TOWARDS A THEORY OF SPREADSHEET
ACCURACY: AN EMPIRICAL STUDY

S. E. Kruck

Dissertation submitted to the faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

in

Business Administration

with a major in Accounting Information Systems

Dr. Larry N. Killough, Chairman

Dr. John J. Maher

Dr. Tarun K. Sen

Dr. Steven D. Sheetz

Dr. Robert C. Williges

August 20, 1998

Blacksburg, Virginia

Key Words: Spreadsheet Accuracy, Spreadsheet Model, Spreadsheet Theory

TOWARDS A THEORY OF SPREADSHEET
ACCURACY: AN EMPIRICAL STUDY

S. E. Kruck

(ABSTRACT)

Electronic spreadsheets have made a major contribution to financial analysis and problem solving. Although professionals base many decisions on the analysis of a spreadsheet model, literature documents the data quality problems that often occur, i.e. underlying formulas and resulting numbers are frequently wrong. A growing body of evidence, gathered from students in academia as well as working professionals in business settings, indicates that these errors in spreadsheets are a pervasive problem. In addition, numerous published articles describe techniques to increase spreadsheet accuracy, but no aggregation of the topics and no model explaining this phenomenon exist.

The research described here develops a theory and model of spreadsheet accuracy and then attempts to verify the propositions in a laboratory experiment. Numerous practitioner articles suggest techniques to move spreadsheets into a more structured development process, which implies an increase in spreadsheet accuracy. However, advances in our understanding of spreadsheet accuracy have been limited due to a lack of theory explaining this phenomenon.

This study tests various propositions of the proposed theory. Four constructs were developed from the theory to test it. The four constructs are planning and design organization, formula complexity, testing and debugging assessment, and spreadsheet accuracy.   From these four constructs three aids were designed to test the relationship between the four constructs. Each of the three aids developed was designed to increase spreadsheet accuracy by addressing a single proposition in the model.

The lab experiment conducted required the participants to create a reusable spreadsheet model. The developed model and theory in this paper appear to represent the spreadsheet accuracy phenomenon. The three aids developed did increase spreadsheet data quality as measured by the number of errors in the spreadsheets. In addition, the formula complexity participants created spreadsheets that contained significantly fewer constants in formulas, and the testing and debugging participants corrected a significant number of errors after using the aid.

DEDICATION

This dissertation is dedicated with love to my parents, who were concerned that I might not finish high school. They gave to me an appreciation for education that I will always treasure. Their love and unwavering belief in me has made this dissertation possible.

ACKNOWLEDGMENTS

I would like to express my gratitude to Professor Larry H. Killough, my chairperson, for all his advice, assistance, and encouragement and to the other members of my committee, Professor John J. Maher, Professor Steven D. Sheetz, Professor Tarun K. Sen, and Professor Robert C. Williges for their suggestions and encouragement throughout this project. I could not have hoped for a better committee and I sincerely thank each one of them.

Additionally, I would like to thank Robert M. Brown, for his support and encouragement throughout my Ph.D. program. This group of gentlemen have supported and encouraged me to reach for the brass ring I wanted, even though I took a different path and I truly appreciate that support. I would also like to thank Bill and my brother and his family for their support during this project.   Without their unselfishness, I could never have focused all my energies into completion of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Electronic spreadsheets have made a major contribution to financial analysis and problem solving. Spreadsheets were originally developed in the late 1970s by two accounting students, Daniel Bricklin and Robert Frankston, who were tired of continually having to re-total accounting worksheets (Kay, 1984). Spreadsheet programs were the first software packages produced for serious business use, and their introduction marked a new era of end-user orientation of computers (Chan, 1987). Spreadsheets are used extensively today in accounting and business for applications such as preparing budgets, forecasting production, financial modeling, workpaper generation, cost/benefit analysis, foreign exchange analysis, assets and liability management, determining rate of return on investments, mathematical modeling, analyzing scientific and engineering data, projecting market penetration, and evaluating the feasibility of divestitures, acquisitions, and mergers. Although many decisions are based on the analysis of a spreadsheet model, documentation shows that there are often data quality problems, i.e. underlying formulas and resulting numbers are frequently wrong. A growing body of evidence indicates errors in spreadsheets are a pervasive problem (Brown & Gould, 1987; Cragg & King, 1993; Davies & Ikin, 1987; Panko & Sprague, 1997).

The research described here develops a theory and model of spreadsheet accuracy and then attempts to verify the propositions in a laboratory experiment. Numerous articles do suggest techniques to move spreadsheets into a more structured development process and imply an increase in spreadsheet accuracy. However, advances in our understanding of spreadsheet accuracy have been limited due to a lack of theory explaining this phenomenon. Additionally, little empirical research has been conducted to determine

1

whether the addition of proposed structure in the form of spreadsheet development aids could be created to increase spreadsheet accuracy.

Kerlinger (1986) defines theory as a set of related propositions that specify relationships among variables. Developing theory is an important part of research, just as important as theory verification (Brewer & Hunter, 1989). Glaser and Strauss (1967) contend that most researchers focus on verification of existing theory, however, they suggest that generating or developing new theory is equally important. Freese (1972) suggests that it is the researcher's responsibility to "define(s) some social phenomenon as problematic," develop models and theory and then test them. A caution is given when generating theory because it is not always "clearly recognized as the main goal of a given research, [and] it can be quickly killed by the twin critiques of accurate evidence and verified hypotheses" (Glaser & Strauss, 1967). Furthermore, Glaser and Strauss (1967) suggest that testing theory will only modify the theory, not destroy it. A starting point of an unexplained phenomenon must occur to develop a new theory and then accumulate evidence about that phenomenon and new theory. Verifying existing theory with accurate evidence is extremely important, but should not overshadow theory generation.

This research project is directed towards the end-user or spreadsheet developer who creates, maintains, and uses the spreadsheet. This research does not cover those spreadsheets that are created by professional programmers or spreadsheet developers because they generally use a structured development process. Typically, professional spreadsheet developers determine the users' needs; plan the application and user interface; create the spreadsheet, formulas, macros and user interface; test the application and make it as "bullet proof" as possible; document the program; distribute the spreadsheet and train the user; and update the application as necessary.

The proposed study will contribute to the literature in several ways. First, it attempts to establish theory and a model where one does not now exist. Second, the study attempts to test the proposed model by investigating if spreadsheet accuracy can be

increased through the introduction of structure by the use of spreadsheet development aids.

The structure for the remainder of this paper follows. Chapter Two provides a description of the importance and uniqueness of the spreadsheet environment, including a discussion of data quality.   A review of previous studies demonstrating high error rates in spreadsheet models illustrates the problem. Chapter Two also provides a discussion of theory and model development and development of the hypotheses. Chapter Three presents the research methodology including the experimental procedures. Chapter Four presents the statistical analysis and results. Chapter Five provides a discussion of the results as well as a discussion of the study's limitations and implications for future research.

CHAPTER 2

LITERATURE REVIEW, THEORY/MODEL DEVELOPMENT AND HYPOTHESES

2.1 Literature Review

Several studies demonstrate that business professionals use spreadsheets extensively to make decisions (Heagy & McMickle, 1988; Heagy & Gallun, 1994; Mingers, 1991; Waller & Gallun, 1985; Zhao, 1997). Spreadsheet accuracy is essential when making decisions. Although several empirical studies have found high error rates in spreadsheet models (Brown & Gould, 1987; Cragg & King, 1993; Davies & Ikin, 1987; Panko & Sprague, 1997), no previous published empirical research has examined techniques to increase the accuracy of spreadsheets by reducing the number of spreadsheet errors. This chapter provides a discussion of the literature relevant to the current study. The second section provides discussion on theory and model development, and then presents a theory and model of spreadsheet accuracy. The final section presents the hypotheses to be tested.

2.1.1 Spreadsheets – An Important Skill

Daniel Bricklin and Robert Frankston, two accounting students, first introduced electronic spreadsheets in 1979 (VisiCalc©). It took only a short time before spreadsheets became the preferred microcomputer application. Several professional organizations have adopted guidelines for incorporating computers into the educational curriculum and are briefly mentioned. Then the following series of studies demonstrate that electronic spreadsheets hit the market quickly and became one of the most frequently used software applications, if not the top software application, for the personal computer.

Various professional organizations and commissions have recommended the inclusion of computers in higher education. The American Accounting Association (AAA) has recommended the use of computers in the educational curriculum process for the past forty years (1959; 1964; 1970; 1986). Similarly, the Accounting Education Change Commission (AECC) has more recently reiterated the importance of incorporating computers into the educational process. Furthermore, the AECC recognizes the need for understanding design and implementation issues. The objectives of the AECC include having students understand how information is identified, measured, communicated, and used. More specifically, skills necessary for a successful accounting career include the ability to use data to "solve real-world problems" (AECC, 1990). Spreadsheets are one way to manipulate data to solve these problems.

Waller and Gallun (1985) surveyed thirty-six companies -- the "Big Eight" accounting firms, fifteen other local and national accounting firms, and thirteen industrial and financial firms -- in order to determine specific software they used and the coverage of spreadsheets they would like to see in a university curriculum. Waller and Gallun discovered that all three surveyed groups wanted an in-depth knowledge of spreadsheets more than any other microcomputer concept included in the university curriculum. Furthermore, all of the survey participants wanted some coverage of spreadsheets and only one firm in the financial group wanted merely an overview of spreadsheets. It should be noted that spreadsheets were relatively new at this time and companies were just starting to use personal computers. It is clear, however, that the companies surveyed wanted their employees to have spreadsheet skills and knowledge.

In the Waller and Gallun survey, practitioners indicate that spreadsheets should be covered at higher education levels. However, practitioner needs and academic offerings are often at odds with one another. A study of 122 practitioners in public accounting (51% of sample), business and industry (44% of sample), and government were contrasted with 172 accounting academics (Heagy & McMickle, 1988). These groups were questioned on fifty-nine accounting information systems topics divided into ten categories.

Academics and practitioners were in agreement on twenty (34%) of the topics, but they differed significantly on thirty-nine (66%) of the topics. Generally, practitioners wanted more emphasis on all programming and application topics, with spreadsheets, database management systems, integrated software, and accounting modules ranking the highest. The top two topics for practitioners were internal controls with which academics agreed ($\chi^2 = 10.30$, p < .01) and spreadsheets, with which the academics did not agree ($\chi^2 = 66.08$, p < .01). Academics ranked spreadsheets only thirty-eight out of fifty-nine. Practitioners' first-hand experience or desires demonstrate that spreadsheet skills are important.

Spreadsheet knowledge is not only an important skill in the United States but a skill desired by other countries as well. A questionnaire sent to sixty-six operational research groups in the UK asked respondents to rank seventy-five different topic areas (Mingers, 1991). The respondents included both those in private and public sectors, as well as consultants. As might be expected, the practical operational research process topics of report writing, presentation skills, and group or teamwork skills were among the top fifteen topic areas. The top skill received a mean ranking of 4.4 out of 5. Basic computing skills, but more specifically, spreadsheets were also in the top fifteen ranking with a mean score of 4.2 out of 5, only .2 lower than the top-desired skill. Looking specifically at computing skills, spreadsheets ranked higher than microcomputer skill, programming, databases, and fourth-generation languages.

Heagy and Gallun (1994) questioned accountants in large and mid-sized firms regarding computer literacy desired of new employees. Two hundred twenty-nine accountants in public practice and 144 accountants from industry were included in the results. The accountants chose spreadsheets over five broad categories of computer knowledge needed by new hires. The other categories were database management systems, telecommunications, accounting systems, and systems development. The specific spreadsheet skills considered most critical were strategy and design of models. The lowest

The most recent survey was directed at companies that recently entered the Fortune 500 list of corporations (Zhao, 1997). The researcher believed that these successful corporations would be considered leaders in using information technology. Zhao wanted to obtain data about end-user skills that business professionals need now, and to determine what skills were expected to be necessary in the year 2000. The study asked about specific hardware skills such as the keyboard, mouse, printer, scanner, and modem. The study also asked about specific software packages by name in five software skill areas: operating systems, word processing, spreadsheets, databases, and desktop publishing. All of the thirty-five skills included in the survey decreased in importance for the year 2000 except for spreadsheet skills, UNIX operating system knowledge, and the ability to use a scanner. Spreadsheet skills were considered the most important application software to understand and apply both now and in the year 2000.   These results indicate that spreadsheets are likely to remain among the top software skills needed for business professionals well into the foreseeable future.

It is evident from the review of the literature that from the time spreadsheets were introduced, they have been one of the top skills needed by business professionals and they will also continue to be so. With such wide use, it becomes all the more crucial to develop techniques to minimize errors in spreadsheets. In the next section, a comparison of computer programmers and spreadsheet developers are contrasted to illustrate the spreadsheet environment.

2.1.2 Spreadsheets – The Problem

Spreadsheets are unique for a variety of reasons, and this uniqueness is part of what has created the problem. This section will provide a description of the uniqueness of the spreadsheet environment and contrast it with the traditional computer programming environment.

Spreadsheet software has a unique environment. It is likened to a visible calculator that allows direct input and output of data using the identical interface, which is not true in

traditional programming environments. This format leads to a natural representation of many problems, but the interactive nature of spreadsheets makes every part of the spreadsheet available to every other part (Alexander, 1994; Carlsson & Stabell, 1986). Another advantage is the small, focused pieces of code in a graphically meaningful arrangement (Johnson, Nardi, Zarmer, & Miller, 1993), but the disadvantage is that the formulas tend to be cryptic and normally only one formula is visible at a time. Spreadsheets permit such rapid development that users easily neglect design, specification, testing, and maintenance. Spreadsheets are not based on concepts such as complex control structures, parameter passing or recursion that traditional computer programming requires. As users continue to use a spreadsheet program, they are not faced with the job of stringing together low-level programming functions, but can concentrate on solving the actual problem. With spreadsheets, users have sufficient means to model their problems at a very small price in terms of programming effort (Nardi, 1995).

User-developed applications lack design review and control procedures that are normally associated with traditional applications developed by information system professionals (Alavi, Nelson, & Weiss, 1987-8; Pierson, Forcht, & Teer, 1990; Sumner & Schultheis, 1990). The user often starts entering information into the spreadsheet with little or no planning (Bissell, 1986; Brown & Gould, 1987; Cragg & King, 1993). The user ignores traditional distinctions between program design, development and testing and generally no testing and debugging period exists.

In addition to the above, the price of spreadsheet software has fallen so that most computer owners can afford to purchase it; regardless of whether they are familiar with its use. Furthermore, advertising claims that the software does all the work, and is a user-friendly program help make the user feel comfortable with the purchase and use of spreadsheet software. The quick response and creation of colorful spreadsheets and graphs breeds confidence and excitement. These inherent properties of spreadsheets can cause data quality issues or errors. Data quality issues are of major concern to those making business decisions using spreadsheet models.

Several advantages of end-user developed applications exist. First, end-user developed software helps with the shortage of system development personnel and the overload of work they receive. Encouraging end-users to develop their own spreadsheet applications instead of assigning work to system development personnel allows the system development personnel to be assigned to more important organizational priorities. Second, end-users who create their own applications eliminate the necessary communication of system requirements to the systems development personnel. Both the end-user and the systems personnel have their own jargon and may find it difficult to communicate. End-user development does not eliminate the need to identify the requirements, but it does eliminate any communication problems (Davis, 1982). Finally, end-user programming transfers the system implementation processes from the technical systems expert to the non-technical user. Systems development personnel can design a more sophisticated system than the users understand, and users may therefore not use the system. End-users have been encouraged to develop applications by both the private (Brancheau & Wetherbe, 1987; Dickson, Leitheiser, & Nechis, 1984; Niederman, Brancheau, & Wetherbe, 1991) and public sectors (Swain, White, & Hubbert, 1995). End-users creating their own spreadsheets relieves some of the problems of traditionally developed applications.

Professional programmers typically learn and apply structured development life-cycle processes when they develop computer programs. These professionals are concerned with reliability, auditability, and control (Ronen, Palley, & Lucas, 1989). However, the majority of spreadsheet developers are not professional programmers, but end-users. These users have obtained their office suite or spreadsheet package and want to apply these tools to help them make better decisions. The research conducted here attempts to focus on this large segment of end-users that usually develop their own spreadsheets using ad hoc approaches, rather than following a formally structured development life-cycle process. Procedures and attitudes that tend to enforce a reasonable level of data quality in centralized data processing operations are often missing in the end-user computing environment (Ballou & Pazer, 1987). Professional computer programmers

know the importance of deep debugging, running test data, and having outside programmers do a detailed, line by line audit. All of these measures enhance data accuracy, but end-users who develop their own spreadsheets usually use an ad hoc approach and do not follow a structured development life cycle process.   The theory, model, and aids developed and tested in this study are intended to help add *some* discipline and structure to the end-user spreadsheet development environment. That is, to apply techniques known to aid systems developers in the creation of better quality programs to the help end-users create better quality spreadsheet models.   All of the end-user development characteristics discussed here create data quality issues which will now be discussed.

2.1.3 Data Quality – Anecdotal Evidence

Business decisions are made every day and the data on which those decisions are based on is often inaccurate. Thus, data accuracy is an important issue. Several recent studies provide evidence that data accuracy is a problem. Knight (1992) surveyed medium and large companies and Arnold (1992) surveyed medium companies with sales in excess of $20 million and both found that more than 60% of the companies reported data quality problems. Laudon (1986) sampled two computerized criminal-history record systems both operated by the FBI. Analyzing the verifiable cases, Laudon found that approximately 26% of the records were complete, accurate, and unambiguous in the criminal-history records, while more than 74% had "significant quality problems."   The United States Secretary of Education reported that data on student education loans was not entered correctly, and as a result, many defaulters got student loan renewals while applications from deserving students were rejected (Knight, 1992).

Accuracy of data is an attribute desired by data consumers. Both professionals working in the data industry and MBA students, when approached as data consumers, indicated that data accuracy was the most important attribute of data (Wang & Strong, 1996). In addition, accuracy was the most cited data quality dimension based on a

literature review (Wand & Wang, 1996; Wang, Storey, & Firth, 1995). The present proposal focuses on data accuracy in spreadsheets.

In order to illustrate the effects of poor data quality, some actual incidents are described here. Relatively few incidents of spreadsheet errors are made public and these are usually not revealed by choice.   For several reasons, spreadsheet errors do not make it into the public eye or the firm involved is not identified. One reason is that the companies are embarrassed and concerned about their image. Reporting one error could raise a frightening thought that many other problems are yet to surface. Customer confidence may be shaken by reports of such an incident. Others fear the legal or financial repercussions of making the error public ("How Personal Computer Can Trip Up Executives," 1984; Edge & Wilson, 1990; Godfrey & Flatau, 1995; Knight, 1992; Krull, 1989). However, there have been a few incidents of spreadsheet errors that did make it to the popular press.

One of the most publicized incidents occurred at James A. Cummings, Inc., a Fort Lauderdale construction company. Using Symphony[©], a spreadsheet package developed by Lotus, Cummings bid $3 million on an office complex for a local utility. The controller added a row to include additional overhead of $254,000 but failed to check whether or not this row was included in the formula that totaled the column. This caused Cummings to underestimate the cost of the entire project. Cummings did win the bid but severely underestimated the cost of the project, resulting in a large financial loss to the firm (Cragg & King, 1993; Edge & Wilson, 1990; Floyd, Walls, & Marr, 1995; Hayden & Peters, 1989; Kee & Mason, 1988; Schultheis & Sumner, 1994; Simkin, 1987; Stone & Black, 1989).   This particular incident was in the public eye because Cummings sued Lotus for the spreadsheet error.   The case was eventually settled out of court.

Another well-publicized case involves the Fidelity Magellan fund and its year-end distribution. The mistake was the result of an employee entering a plus sign, rather than a minus sign, in front of a net capital loss. The item in question had a value of $1.3 billion and an incorrect sign caused a $2.6 billion swing in earnings. As a result of the fund's incorrect estimated earning, Fidelity Magellan's investment professionals estimated its

year-end payout to be $4.32 per share. This error had several impacts. First, the net asset value of the Magellan fund fell 4.4% from the day they made the mistake until December 5th, the day they caught it. Although Fidelity notes that the mistake did not affect customer accounts, it most likely did influence investor actions and investor confidence. Customers may have sold their investment or waited to purchase the fund until after the year-end payout to avoid the distribution and resulting capital-gains tax consequences (Godfrey & Flatau, 1995; Savitz, 1994).

The most recent incident involves the Beardstown Business and Professional Women's Investment Club, more popularly know as the Beardstown Ladies. This investment club is comprised of fourteen women, with a median age of 70, from a small Illinois town (Rich, 1998). Prompted by a *Chicago Magazine* report, the Beardstown Ladies accounting books were audited by Price Waterhouse LLP at the request of Hyperion, the publisher of the Beardstown Ladies' five books. The audit discovered that their return was actually about 9.1%, far below the 23.4% originally reported. Price Waterhouse attributed the mistake to a computer input error (*The New York Times,* D2). Other sources report that the membership dues were included as investment gains (Kadlec, 1998; Krugman, 1998; Skrzycki, 1998), "nobody double-checked the math" (Kadlec, 1998), "their math shouldn't be a problem" (Wiles, 1998) and "the Ladies now blame it on a 'computer error'" (Skrzycki, 1998). Although this may or may not be specifically a spreadsheet error, it illustrates the data quality problem.

References to other spreadsheet errors exist but the companies have chosen to remain anonymous due to the negative publicity that arises from such errors. In one case, *The Wall Street Journal* reported an unnamed Dallas oil and gas company that fired several executives due to a spreadsheet error that caused the firm to lose millions of dollars in an acquisition deal (Cragg & King, 1993; Creeth, 1985).   In another case, a chief operating officer was working with two related spreadsheets containing 15,000 cells to project the market for computer-aided design equipment for the manufacturing industry. The size of the market was underestimated by $36 million when instructions

were given to the computer to round all spreadsheet figures to the nearest whole number. The inflation rate of six percent (1.06) was rounded to one, thus negating the inflation factor ("How Personal Computer Can Trip Up Executives," 1984; Krull, 1989; Schultheis Sumner, 1994; Watt, 1985). In a final case, a senior consultant for the Price Waterhouse accounting firm was asked to "untangle" an unnamed client's multi-billion dollar spreadsheets. One-hundred-twenty-eight errors were found in four worksheets (Edge & Wilson, 1990; Schultheis & Sumner, 1994; Simkin, 1987).

These real-world errors are unacceptable, but are they isolated incidents? The following section describes a number of empirical studies demonstrating that the previous cases of spreadsheet errors are not isolated incidents.

2.1.4 Data Quality – Empirical Evidence

2.1.4.1 Error Rates

Much evidence documents the existence of data quality problems in spreadsheets. This evidence has been gathered from professionals in business settings as well as from students in academia. Each of the following studies demonstrates unacceptably high error rates in spreadsheets in real-world applications as well as in experimental studies.

An experiment conducted by Brown and Gould (1987) involving experienced spreadsheet users found that 44% of their end products contained errors. Furthermore, it was discovered that little time was spent planning the actual spreadsheet. The subjects were business professionals employed by IBM, each of whom had one to five years of experience and used spreadsheets about eight hours a week. Each subject created three different spreadsheets. These spreadsheets were simpler than those created at IBM as evidenced by the subjects considering ten of the twenty-seven spreadsheets created were simple compared to the participants' real-world work. Brown and Gould observed that the 44% error rate may have actually understated the true error rate because the problems used in the experiment were so well defined, and real-world problems are often ill-defined

13

and unstructured. Although this experiment may not generalize to the spreadsheet models that are created in a real-world setting, the following study overcomes this limitation by investigating "live" spreadsheets – spreadsheets currently in use.

Davies and Ikin (1987) studied nineteen spreadsheets from ten sources. The spreadsheets analyzed were from four major industrial firms, an international firm of chartered accountants, a local government authority, a state government statutory authority, two firms of consultants, and a lecturer. The spreadsheets examined involved project and product costing, budget, payroll, loan calculations, and investment analysis applications. Five (26%) worksheets were considered error free and four (21%) of the spreadsheets contained "major" errors including an error of seven million dollars, different exchange rates for the Australian dollar within the same time period, and negative units in the current stock account. The remaining spreadsheets (53%) were considered "inadequate and extremely prone to accidental errors" in actual real-world usage (Davies & Ikin, 1987). For example, several spreadsheets contained no text or labels but consisted of all numbers and formulas.

A study by Cragg and King (1993) examined real spreadsheets from ten companies and discovered a 25% error rate, even though all but one of the spreadsheets had gone through a formal testing process. Half the spreadsheets had been used at least six months and averaged seven revisions. Two-thirds of the spreadsheets had been revised at least one time. This may have compounded the problem since modified spreadsheets tend to be more error prone than spreadsheets designed from scratch (Brown & Gould, 1987). Although a 25% error rate is substantial, it may actually understate the true error rate of the sample spreadsheets because examination time was limited to two hours per spreadsheet. The actual size of the examined spreadsheets ranged from 150 to 10,000 cells. These two real-world studies of spreadsheets illustrate that even when spreadsheets are created and used by people in the real world, they often have an unacceptably high error level.

The final empirical study conducted used upper-division undergraduate business and MBA students. In addition to creating spreadsheets, Panko and Sprague (1997) had some of the undergraduate students debug their own spreadsheet models. For analysis purposes, they sub-divided the MBA students into inexperienced and experienced groups. Spreadsheet experience was operationally defined as self-reported spreadsheet development, auditing, and training experience. Data entry or other superficial manipulations were not considered experience for this classification. Panko and Sprague used a "domain-free" wall-bidding problem statement. They discovered that 37% of the undergraduate students' spreadsheet models had errors and 35% of the inexperienced MBA students' spreadsheet models had errors. Surprisingly, the experienced MBA students had errors in 24% of their spreadsheet models, which was not statistically different than the inexperienced MBA students ($p = .414$). Furthermore, the undergraduate students were not significantly different than the combined MBA group's error rate of 30% ($p = .223$). Twenty-three of the undergraduate participants were asked to debug their spreadsheet models. First they received a ten-minute lecture about the dangers and propensity of spreadsheet errors, followed by another ten minutes of instruction on how to code-inspect a model by going through it cell-by-cell and how to use Microsoft Excel's[©] auditing tool. Eighteen percent of the errors were discovered and corrected by the participants which completely corrected 13% of the models. One participant actually added an error. Discovering errors in their own spreadsheet was difficult and implies that we need to find techniques to prevent errors.

The above studies on spreadsheet errors are summarized in Table 1. These studies demonstrate that spreadsheet error rates range from 25% to 44%. These high error rates are found under different settings, both real-world and experimental.

TABLE 1

Summary of Studies of Spreadsheet Errors

Authors	Year	Participants	% of Spreadsheets w/Errors
Brown & Gould	1987	IBM employees	44%
Davies & Ikin	1987	Live/real company spreadsheets	
		-major errors	21%
		-inadequate and extremely error prone	53%
Cragg & King	1993	Live/real company spreadsheets	25%
Panko & Sprague Jr.	1997	Undergraduate students	37%
		Inexperienced MBA students	35%
		Experienced MBA students	24%

A further problem that affects the quality of data in spreadsheets is the possibility of significant magnification of errors. An important problem develops because even if only a nominal percentage of errors occur in spreadsheet design, these are often part of a long cascade of cells leading to bottom-line values. Even a small cell error rate can multiply into a large bottom-line error. Magnification of errors can create a major problem for spreadsheet models. Lorge and Solomon (1955) developed a general model for analyzing error cascades:

$$E = 1 - (1 - e)^n$$

where:

E = probability of finding an error at the end of the cascade

e = probability of a cell error[1]

n = number of stages or events in the cascade[2]

---

[1] Cell error—incorrect formula or value in any particular cell or combination of cells

[2] Cascade —sequence of cells where a cell value or formula affects subsequent cell values

For example, assume that the probability of a cell error is 1% and the number of cascading cells (i.e. cells in the sequence) is ten, then the probability of an error in the spreadsheet is 10%. When the probability of a cell error is 3% and the number of cascading cells is ten, then the probability of an error in the spreadsheet is 26%. More examples can be seen in Table 2. These numerical examples are based on small spreadsheets. In more complex spreadsheets, the probability of finding an error is larger due to multiple cascades.

TABLE 2

Probability of Error Due to Cascading Effect

Cell Percentage of Errors	Number of Cascades or Sequences	Probability of Error
1%	10	10%
1%	50	39%
1%	100	63%
3%	10	26%
3%	50	78%
3%	100	95%
5%	10	40%
5%	50	92%
5%	100	99%

What has not been addressed is how we can decrease the error rate. Researchers are just beginning to investigate error detection in spreadsheets. The next section describes two empirical studies that specifically look at error detection in spreadsheets.

2.1.4.2 Error Detection

Two recent studies have tried to determine the effectiveness of error detection. Neither of these studies provided subjects with error detection or debugging techniques prior to the experiment, which may explain why the detection rates are low.

Galletta, Abraham, El Louadi, Lekse, Pollalis, and Sampler (1993) examined the differences in domain knowledge and its interaction with self-reported spreadsheet ability to determine if any significant differences existed in spreadsheet error detection. The domain knowledge was defined as accounting and therefore CPAs were considered experts. Spreadsheet ability was defined as time spent in the creation and or modification of spreadsheets in excess of 250 hours for the experienced group, and less than 150 hours of such experience for the novices. Data entry or other superficial manipulation of spreadsheets was not considered experience for this study. The mean (median) number of hours for spreadsheet experts was 1,548 (925), and for novices was 54 (44). The design was a 2 x 2 with subjects either CPAs or MBA students and level of spreadsheet experience as either experts or novices. The test instrument consisted of ten spreadsheets that had two errors each. One error was a domain error in an accounting concept or principle, for example, a prepaid item listed as a non-cash item. The second error was a device or spreadsheet error, for example an incorrect formula. Care was taken to construct the instrument so that an error could only belong to one of those two categories. Overall participants found 46% of the domain errors and 65% of the device errors. Only the domain knowledge (accounting expertise) was significant in finding the accounting errors ($p = .021$). Spreadsheet expertise was not significant in finding the device errors ($p = .671$), however the time taken was significantly decreased for those with spreadsheet experience ($p = .013$).

This study used computer screens, rather than printouts, for the spreadsheet display to discover errors for the above study. Considerable literature suggests that reading from a computer screen is not as effective or efficient as reading from paper (Gould & Grischkowsky, 1984; Gould, Alfaro, Barnes, Finn, Grischkowsky, & Minuto,

1987a; Gould, Alfaro, Finn, Haupt, & Minuto, 1987b; Haasen & Haas, 1988; Oliver, 1994). To determine whether the difference between electronic and paper display interacted with spreadsheet error detection, Galletta, Hartzel, Johnson, and Joseph (1996-97) conducted a second study in spreadsheet error detection. Participants, MBA students taking a graduate MIS course, worked from either a computer screen or paper copies. The experiment further subdivided groups into those using the traditional formula listing that Lotus 1-2-3[©3] creates and those without formulas. A fifth group worked from a printed listing with formulas spatially integrated with the numbers and results. Participants received a one-page budget worksheet that contained eight errors; however, they were not told how many errors the worksheet contained. The best group found an average of 4.6 or 58% of the errors and the worst group found 3.2 or 40% of the errors. The number of errors found was significantly higher for the paper group ($p = .007$). No significant difference existed in performance based on the presence or absence of formulas. Furthermore, the group that used the integrated formulas and values did not perform significantly better than any of the other groups. Error detection may be improved if researchers can provide participants with some procedure or aid for detecting errors.

Allwood (1984) noted an interesting finding in error detection in statistical problem solving. The subjects' verbal protocol was analyzed in an effort to learn more about the problem-solving process. Allwood found that subjects often did not detect the error even after having verbally noticed the symptom. This is relevant to finding and correcting spreadsheet errors because a participant may notice an inconsistent number or a cell with an error message but fail to correct it. This was also noticed by Carroll (1990) when participants were creating a form to record address information. One participant found he could only enter a four digit zip code, yet continued to enter data until getting a second message when entering a date. At which time a comment was noted that the participant could not understand why so much was being made of "one" error. The

---

[3] Lotus 1-2-3[©] formula listing is a sequential listing in cell order. There are no visual cues or indication of the formulae's location other than cell address.

participant completely missed the problem with the zip code, even though the system sent an error message. To date, this has not been tested with spreadsheets.

2.1.5 Summary of Literature Review

As demonstrated by several surveys, the ability to develop spreadsheets is currently, and will continue to be, a very desirable skill. Developing spreadsheets is also unique in that it gives the end-user development capabilities without the structure common in traditional programming. A number of empirical studies have demonstrated that spreadsheet error rates are high, although errors are not generally discussed openly. A number of practitioner articles suggest ways to decrease these errors, but to date, no published empirical study has tested any of these techniques. The following section develops the theory and a model for use in spreadsheet error reduction.

2.2 Theory/Model Development

Numerous published articles describe techniques to increase spreadsheet accuracy, but there has been no aggregation of the topics and no model explaining this phenomenon. To date, no published studies determine if any of these techniques *actually* increase spreadsheet accuracy. Instead, evidence has been largely anecdotal, having been acquired on a trial-and-error basis in the field by practitioners and users. Practitioners have written about spreadsheet techniques because they know and care about the real-world outcome of spreadsheet modeling. A theory of spreadsheet accuracy is proposed from the practitioner literature and the limited research in this area. Propositions of theory are tested by creating three spreadsheet development aids. Each aid is designed to increase accuracy by addressing issues associated with a single proposition.

As suggested by several researchers in the introduction, theory building is just as important as verification and testing of existing theories. Theory building consists of creating or building new theories to explain known but previously unexplained empirical results. Freese (1972) suggests that it is the researchers responsibility to study some

social phenomenon, then construct and test theories. The research study here is considered applied, meaning that it is the scientific discovery of knowledge having applicability to a specific, identifiable problem (Ashton & Willingham, 1988). The phenomenon of spreadsheet accuracy is examined and a theory developed. In the literature review section of this paper empirical evidence supports the contention that spreadsheets are important and are used in real-world business. Furthermore, evidence shows that they often contain significant errors. The lack of progress in this area is in part due to a lack of existing theory.

This model is driven by those techniques considered relevant by practicing spreadsheet developers who have collectively published in excess of thirty-five articles. These articles are summarized and referenced in Appendix A. This group of end-user developers considers these techniques important enough to take the time and effort to share their knowledge with others. They consider these techniques to be the ones that helped them increase spreadsheet accuracy in their own spreadsheet models.

The first step in creating the theory required analyzing the literature for explicitly stated techniques to increase spreadsheet accuracy. The authors of the examined literature have a common belief that the adoption of their techniques will decrease errors in spreadsheet models. Some of the error reduction techniques were cited only a few times whereas others were mentioned in more than half of the papers. The techniques were classified into like or similar categories or activity relationships to develop the model. For example, using cross-footing techniques cannot be checked until the spreadsheet model is designed and limit proofs cannot be checked until after spreadsheet numbers and formulas have been entered. Therefore, both of these techniques would usually be done after the model is created, often during some type of error checking phase, so they were placed into a single category. After grouping the error reduction techniques, the researcher gave them a construct name and definition.

The constructs developed are planning and design organization, formula complexity, testing and debugging assessment, and spreadsheet accuracy. The planning

and design organization is defined as the degree to which the spreadsheet was laid out into
an orderly and cohesive format. Formula complexity is defined as the degree to which the
formulas required in spreadsheet cells are difficult to understand. Testing and debugging
assessment is the degree to which detecting and correcting errors took place.
Spreadsheet accuracy is defined as the degree to which the spreadsheet is error-free or
accurate. These constructs and relationships were developed into the model shown in
Figure 1. To conduct an initial test of the model, three propositions are developed to test
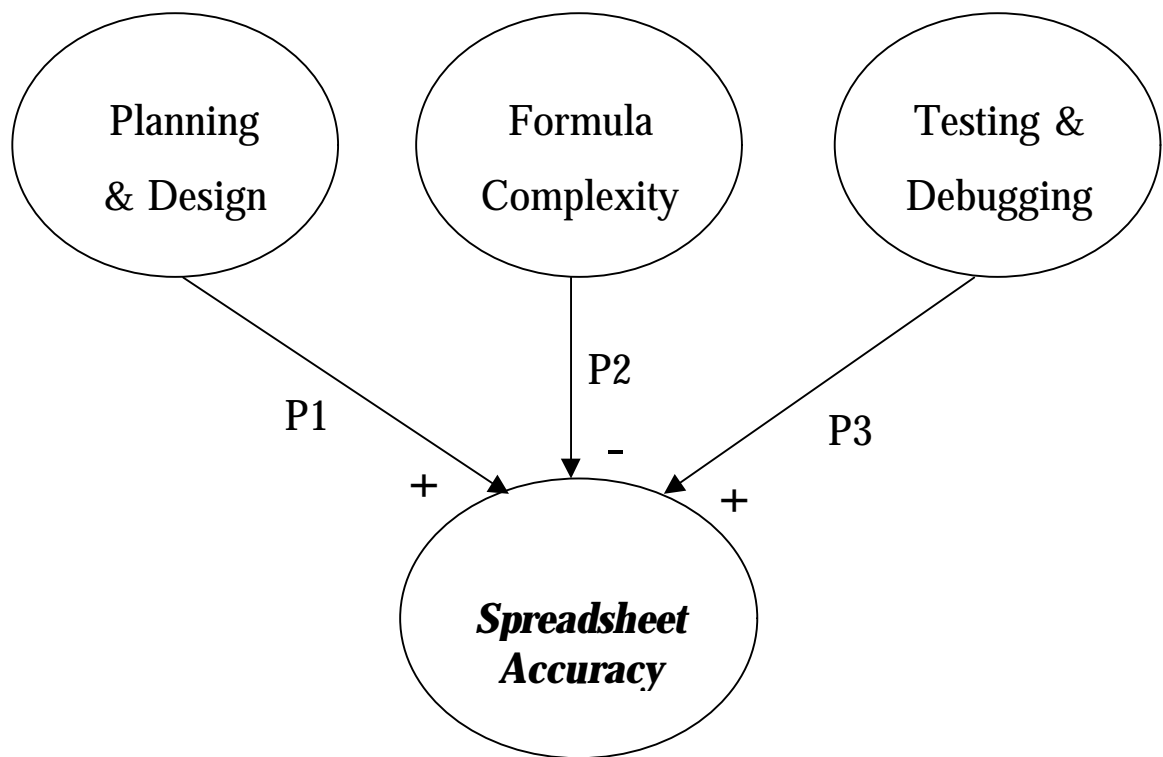the theory.

FIGURE 1

Spreadsheet Error Reduction Model

A short description of each proposition and an example follow.

Proposition 1 (P1): Increasing the degree that the spreadsheet is planned and designed will increase accuracy of the spreadsheet model. For example, explicitly planning to round cell values should result in rounding at the proper time so that results are correct. An illustration can be found in the case discussed earlier where the spreadsheet developer inadvertently rounded all cells after the model was complete and as a result lost the effect of the inflation factor. Planning a rounding strategy may have prevented this error. Another benefit, or example, of the planning and design aid would be decreasing the number of rows and columns that would be inserted or deleted because of prior planning of the data placement. Inserting rows or columns can create disaster as illustrated by the Cummings case discussed previously. Designing the spreadsheet so that rows or columns could be inserted without disastrous results could have prevented this error.

Proposition 2 (P2): Decreasing the degree of complexity in formulas will increase the accuracy of a spreadsheet model. For example, splitting a long formula into smaller parts should make a formula less complex and, therefore, increase the accuracy of a spreadsheet. A formula that adds and subtracts twenty different cells could be split into two simple formulas. One of the formulas could total the cells to be added and the other formula could total the cells to be subtracted; then the final formula would just subtract the second formula from the first. Another example would be to use range names instead of cell addresses because range names are easier to understand. A formula that states "sales – expense" would be less error prone than a formula that states "G58 – G109." Formulas written clearly or broken into smaller pieces increases the accuracy of the spreadsheet model.

Proposition 3 (P3): Increasing the degree of testing and debugging assessment will increase the accuracy of a spreadsheet model. In this stage, the spreadsheet developer deliberately looks for and corrects errors. Most end-user developers feel that the spreadsheet they have just created is correct and looks good, so they often ignore or skip this stage. However, spreadsheet developers should use some testing and debugging

techniques just as traditional computer programmers do. For example, testing the model with actual historical data or re-computing all the calculations by hand will help increase spreadsheet accuracy. Another example would be to use the built-in Auditing Tool in Excel©.

To test these propositions, three aids were developed. The aids attempt to influence a construct in the theory and changes in spreadsheet accuracy. The planning and design aid to test Proposition 1 is a short list of reminders including a brief description of some proper spreadsheet development techniques. The formula complexity aid to test Proposition 2 is designed to reduce formula length which is part of what makes a formula complex, and will remind the spreadsheet developer of techniques to use along with a short description of each. The testing and debugging aid developed to test Proposition 3 is a listing and description of techniques to help ensure that any errors will be discovered and corrected. These aids are contained in Appendix B.

The specific techniques for each aid were based on two criteria. First, the literature reviewed indicated that some techniques were more general than others. For example, techniques that referred to formulas were more universal than those relating to graphs because graphs are not included in spreadsheet models as often as formulas. Each spreadsheet contains a multitude of formulas, whereas graphs exist only in a few of the real world spreadsheets. Second, due to the exploratory nature of the experiment, the techniques had to be reasonable to implement during 1-1½ hours experiment time. For example, developing a table of contents or map of the worksheets is not a reasonable exercise to conduct in the computer lab with a small spreadsheet.

Checklists and decision aids have been used and empirically studied in accounting audit programs. The use of decision aids has the potential to improve and facilitate auditor judgement (Anderson, Jennings, Kaplan, & Reckers, 1997), however empirical results have been mixed. Decision aids have demonstrated a beneficial effect in both auditing and non-auditing environments (Ashton & Willingham, 1992; Anderson, Jennings, Kaplan, & Reckers, 1995; Anderson et al., 1997; Benbasat & Schroeder, 1977;

Butler, 1985; Klein, Goodhue, & Davis, 1997; Libby & Libby, 1989; Sharda, Barr, & McDonnell, 1988), also they have demonstrated a dysfunctional effect (Ashton & Willingham, 1990; Fischhoff, Slovic, & Lichtenstein, 1978; Kachelmeier & Messier, 1990; Pincus, 1989).

Accounting firms have used decision aids for internal control questionnaires and various other types of support and documentation (Elliott & Kielich, 1985). Several studies have looked at the benefit of decision aids in auditing. These studies have found that such decision aids generally decreased auditor liability when used correctly (Anderson et al., 1995) [4], and checklists have decreased the variability of sample sizes (Butler, 1985). Butler used a checklist technique in the assessment of sampling risk. Prior to making risk assessment, eleven of the auditors answered the four decision aid questions in the demographic questionnaire. These four questions were intended to remind the auditors of factors relevant to an assessment of sampling risk. Butler found evidence that participants exposed to the decision aid made risk assessments that were closer to the normative criterion and, also, made more correct accept/reject decisions about account balances than the auditors not exposed to the decision aid. It is hoped that a similar aid intended to remind the spreadsheet developer of specific error reduction techniques will yield similar improvements in spreadsheet accuracy. Klein et al. (1997) conducted two studies on error detection and found evidence that when error detection was an explicit goal, subjects were able to find significantly more errors. This may suggest that having an aid to remind spreadsheet developers of the goal may work as an explicit statement to reduce errors. This study includes the checklist and explicit goal techniques that Butler (1985) and Klein et al. (1997) utilized in the aids developed to test the theory.

---

[4] The quality of decisions made to determine if they were better was not evaluated.

2.3 Hypotheses

While many documented studies demonstrate the existence of high error rates in spreadsheet construction, no published studies test why they occur or how to increase spreadsheet accuracy. This is a relatively new area for research because spreadsheets have only existed for approximately two decades (VisiCalc© was first released in 1979). Most of the spreadsheet-related research has occurred in the past decade and has been confined primarily to documenting the existence of high error rates in spreadsheet models developed by novices and experienced users in either lab or real world settings, not increasing accuracy.

The purpose of this research is to propose and examine the theory of spreadsheet accuracy. The theory and model drive all hypotheses. Three aids have been developed to test the constructs: one for the planning and design stage, a second one for the formula creation stage, and a third one for the testing and debugging stage. These aids will be evaluated on the basis of outcomes. Outcome oriented evaluations will focus explicitly on the accuracy of the spreadsheet made with the benefit of the aid.   Four groups comprise the study: a control group which will not use any aid, a group that will use a planning and design aid, a group that will use a formula complexity aid, and a fourth group that will use the testing and debugging aid. The intended outcome is for the aids to increase spreadsheet accuracy, therefore, the following hypotheses, in null form, are tested:

> H_{1A}*: The participants that use the planning and design aid will have no significant increase in spreadsheet accuracy over the control group.*

> H_{1B}*: The participants that use the formula complexity aid will have no significant increase in spreadsheet accuracy over the control group.*

> H_{1C}*: The participants that use the testing and debugging aid will have no significant increase in spreadsheet accuracy over the control group.*

The planning and design aid includes those techniques that practitioners believe increase the accuracy of spreadsheets. This represents the relationship between the design and organization of the spreadsheet model and the accuracy of the spreadsheet. Accordingly, the following hypothesis, in null form, is tested:

> H_2: *The participants that use the planning and design aid will have no significant relationship between spreadsheet accuracy and the quality of their design and organization.*

The formula complexity aid includes those techniques that practitioners and users believe will increase the accuracy of spreadsheets by decreasing formula complexity through helping to decrease formula length. This implies that there should be a relationship between accuracy and formula complexity in the spreadsheet models. Therefore, the following hypothesis is tested, in null form:

> H_3: *The participants that use the formula complexity aid will have no significant relationship between spreadsheet accuracy and the complexity of the formulas.*

The final hypothesis examines only the spreadsheets of the testing and debugging group. The testing and debugging aid includes techniques that practitioners believe will increase the accuracy of spreadsheet models by directing the spreadsheet developer in discovering and eliminating errors. This implies that a significant difference exists for the number of errors corrected for the participants using the testing and debugging aid. The final hypothesis, in null form, is tested:

> H_4: *The participants in the testing and debugging group will not correct a ny more of the spreadsheet errors after using the testing and debugging aid.*

The goal of this research is to advance spreadsheet theory and understand the components of spreadsheet accuracy so that increases can be achieved.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter presents the research methodology employed to examine the hypotheses presented in Chapter Two. One independent variable was a between-subjects variable. This research study manipulated the variable at four levels: 1) no aid, 2) planning and design aid, 3) formula complexity aid, and 4) testing and debugging aid. Tests of the research hypotheses were accomplished through the use of four dependent variables. Hypotheses H_{1A}, H_{1B}, and H_{1C} were tested using the number of spreadsheet errors. Hypothesis H_2 was tested using two expert's ratings of the "goodness" of the spreadsheet and the number of spreadsheet errors. Hypothesis H_3 was tested using the length of the longest formula in each spreadsheet and the number of spreadsheet errors. Hypothesis H_4 was tested using the number of spreadsheet errors before using the aid and again after using the aid.

The remaining sections of this chapter provide detailed discussions of the research methodology. The first section presents the experimental design, independent and dependent variables. The second section describes the study's participants and section three describes the experimental procedures.

3.1 Experimental Design

Participants were assigned randomly to one of four groups. The experiment was conducted in the computer lab. The task was relatively uncomplicated and required only knowledge of basic formulas for completion. To test the reliability of the instrument, a pilot test was run using eight graduate students. No major modifications were made after discussion with the participants and analysis of the data.

3.1.1 Independent Variables

The independent variable is the aid used.   The aid used was manipulated at four levels: no aid, planning and design aid, formula complexity aid, or testing and debugging aid. Each aid listed four techniques. Each technique had a short heading followed by a one sentence description. Attached to each aid was a detailed description, and examples, including screen shots where appropriate, to help the participant with its use. Appendix B includes the three aids.

3.1.2 Dependent Variables

Table 3 at the end of this section summarizes the hypotheses and variables. The first hypothesis required testing for a difference in spreadsheet accuracy between the treatment and control groups. The dependent variable for these hypotheses was the number of spreadsheet errors. The operational definition of a spreadsheet error is any formula or number that causes an incorrect calculation. Each spreadsheet was compared to the solution. If an error existed in the bottom-line value, the formulas and values were reviewed and changed until the correct bottom-line value was obtained.   The number of changes required to obtain the correct bid was recorded. A typographical error in a label was not considered an error for this experiment.

The second hypothesis required testing for a relationship between the "goodness" of the spreadsheet model and spreadsheet accuracy. The dependent or response variables were the number of spreadsheet errors and the rating/rank of "goodness" for each spreadsheet. The expert ratings/ranks of "goodness" were operationally defined by spreadsheet experts and will be discussed in the next chapter. Two spreadsheet experts were asked to rank or rate each spreadsheet model in both the control group and the planning and design group. Each expert received fifty-eight printouts (thirty from the control group and twenty-eight from the planning and design aid group) in random order. The experts were instructed to use their expertise to develop a four-point scale from good to bad, place the printouts into the four groups, and then define and describe each group.

No discussion of what constituted the four groups was provided to avoid a demand effect until completion of the task.

Hypothesis 3 required testing for a significant relationship between the formula complexity and spreadsheet accuracy. The dependent or response variables were the number of spreadsheet errors and the length of the longest formula. Thus, formula complexity was operationalized as the length of the longest formula. The length of the formula was determined by the number of cell references and constants in the longest formula. For example, the formula =A1+B1 is considered a length of two.

Hypothesis 4 determined whether the participant correctly fixed a significant number of spreadsheet errors using the testing and debugging aid. Each spreadsheet model was duplicated before the participant used the testing and debugging aid. The pre-aid and post-aid spreadsheet were compared to determine the total number of errors fixed or corrected after the participant uses the testing and debugging aid.

TABLE 3

Summary of Hypotheses

Hypotheses	Independent Variables (Levels)	Dependent Variables
H_{1A}: The participants that use the planning and design aid will have no significant increase in spreadsheet accuracy over the control group.	Aid used (four levels)	Number of spreadsheet errors
H_{1B}: The participants that use the formula complexity aid will have no significant increase in spreadsheet accuracy over the control group.	Aid used (four levels)	Number of spreadsheet errors

TABLE 3 (con't)

H_{1C}: The participants that use the testing and debugging aid will have no significant increase in spreadsheet accuracy over the control group.	Aid used (four levels)	Number of spreadsheet errors
	Dependent or Response Variable	**Dependent or Response Variable**
H_2: The participants that use the planning and design aid will have no significant relationship between spreadsheet accuracy and the quality of their design and organization.	Expert rating/rank of "goodness" (four levels)	Number of spreadsheet errors
H_3: The participants that use the formula complexity aid will have no significant relationship between spreadsheet accuracy and the complexity of the formulas.	Longest formula length	Number of spreadsheet errors
	Dependent Variables	**Dependent Variables**
H_4: The participants will not correct any more of the spreadsheet errors after using the testing and debugging aid.	Number of errors before using aid	Number of spreadsheet errors corrected

3.2 Participants

The participants in the experiment were graduate and undergraduate students at Virginia Polytechnic Institute and State University who were randomly assigned to one of the four groups. The four groups were: 1) control, 2) use of the planning and design aid, 3) use of the formula complexity aid, and 4) use of the testing and debugging aid. Students

were compensated in two ways. First, motivation to participate was encouraged by rewarding the students with state lottery tickets or extra credit. Hunton (1996) used the lottery ticket technique after participants said that lottery tickets would be more valuable than an equivalent amount of money. Hunton discovered that a lottery ticket incentive program appeared to work because differences in performance were noted between the treatment conditions when lottery tickets were used. However, when only cash was used there was no significant difference in the performance between the treatment conditions. For this study, there was no significant difference in performance by those who received extra credit and those who received lottery tickets (F = .606, p = .438). Second, motivation to perform well was induced with a monetary drawing. The participants performing the best (i.e., creating the highest spreadsheet accuracy) in each of the four groups were entered into a drawing for a monetary reward.

Supplemental data were collected after the participants created their spreadsheet model. Participants were asked to rate their spreadsheet experience as beginner, moderate, or expert. Most participants (95%) considered themselves to have moderate or expert spreadsheet experience. They were also asked to rate their spreadsheet knowledge on a four-inch scale[5] anchored at "Very Low" and "Very High." The mean (standard deviation) spreadsheet knowledge of 3 (.7) based on a four-inch scale. The participants reported mean (standard deviation) spreadsheet experience of 493 (666) hours. Only about 5% considered themselves beginner spreadsheet users. Sixty males and fifty-five females participated[6]. Over 90% were business students and 77% were either accounting or accounting information systems majors[7]. Eighty-three participants were undergraduate students and thirty-three were graduate students. Additional descriptive demographic data are contained in Table 4.

---

[5] All scales were four-inch line segments without any markings. The participant response was measured to the nearest sixteenth of an inch.

[6] One student did not provide gender.

[7] Five students did not provide major.

TABLE 4

Demographic Data

Panel A. Discrete Measures		
Attribute	Level	N
Group	Control	28
	Planning & Design	26
	Formula Complexity	29
	Testing & Debugging	33
Spreadsheet	Beginner	6
Experience	Moderate	90
	Intermediate	20
Number of	1	47
Spreadsheet	2	54
Programs Used	3 or more	15
Academic	Undergraduate	83
Rank	Graduate	33
Gender[6]	Male	60
	Female	55
Major[7]	Accounting	86
	Other-Business	16
	Other-Non-Business	9

[6] One student did not provide gender.

[7] Five students did not provide major.

TABLE 4 (con't)

Panel B. Continuous Measures

Attribute	N	Minimum	Maximum	Mean	Standard Deviation
Spreadsheet Knowledge 0-very low 4-very high	116	0.56	4.00	3.04	0.74
Spreadsheet Experience Hours	95	5.00	3,120.00	492.57	665.51
Hours Work per Week	76	1.00	50.00	20.26	13.23

3.3 Experimental Procedures

The experiment was conducted in campus computer labs. Participants signed up on a volunteer basis. They were told that they would be asked to create a spreadsheet model in the computer lab using Microsoft Excel©. They were also told that they must be able to create a spreadsheet using addition, subtraction, multiplication and division and they did not need to know how to use functions or write macros. The groups ranged between two and fourteen participants. Participants were sent an e-mail reminder of their agreement to participate, which also included the lab location. To eliminate any confounding effects of different instructions, all instructions were read from a script included in Appendix C. First the students were thanked for their participation, reminded of their compensation, and assured anonymity. The researcher then outlined the process, gave them a consent form, and reminded them not to discuss the experiment until the end of the month.

After each student agreed to participate, they were given some background on the importance of spreadsheets. If they were in one of the three treatment groups, they were then given the appropriate aid. Each aid included four techniques and detailed instructions for each of the four techniques. In a lecture format, each technique was covered and examples shown on the overhead computer projector. An example of a blank Bidder spreadsheet was described and shown so that the participants would understand to enter their social security number in the message box. This was used to track the analysis of their spreadsheet data and later matched with their demographic data at which time it was also matched with their experimental group. This was to eliminate any bias in the analysis of the data as the researcher did not know which spreadsheet was in which group. After matching all participants information, the social security number was removed and replaced with observation numbers to ensure privacy.

Each participant was then given a diskette with the Bidder.xls file and asked to open the file and enter their social security number. Finally they were given the task which is included in Appendix D. The task required participants to calculate two bids or estimates to remodel a kitchen. The task had several options for materials at different costs. Approximately 30% of the participants asked about linear feet. The researcher explained that "linear feet was the area around the kitchen as though you were to paint a line around the wall." Approximately 45% of the participants did not calculate linear feet correctly for either bid; however, one participant calculated correctly for one bid but not the other. After the participants finished creating their spreadsheet model, the final step was to complete a demographics/post-test questionnaire. The following day a final e-mail was sent thanking them for their participation.

CHAPTER 4

RESULTS

Six sections comprise this chapter. The first section provides the results of the preliminary analysis including manipulation and reasonableness checks that were included to ensure the effective manipulation of the independent variables. This section also presents and discusses the tests of assumptions related to the statistical procedures used in the analysis sections. The second through fifth sections present the analyses of the data associated with each of the research hypotheses. This chapter's final section presents supplemental data analysis.

4.1 Preliminary Analysis

The preliminary data analysis presented in this section addresses several issues. First, the validity and reliability of this study are discussed. Second, manipulation checks performed to assess the effectiveness of the manipulation of the independent variable are presented. Finally, the assumptions underlying the statistical analysis were tested to provide justification for the performance and interpretation of the results of the tests of hypotheses.

4.1.1 Validity and Reliability

To test the reliability and validity of the instrument, a pilot study was conducted using eight graduate students. The participants were randomly assigned to one of the four experimental groups. These participants completed the study as described in Section 3.3, Experimental Procedures. After debriefing the pilot participants about the use of the instrument, no major modifications were made. The participants believed the instrument to be realistic and understandable. Manipulation checks and statistical analysis were

performed on the pilot study data. The results of the statistical analysis were significant. Reliability of the post-test questionnaire was tested by calculating Cronbach's alpha for several questions. The Cronbach's alpha for the confidence questions was .90 and for the accuracy questions was .67. High alpha values imply high interrelations among the questionnaire items, because internal consistency measures are products of the number of items and the intercorrelations among them.

4.1.2 Manipulation and Reasonableness Checks

A total of 129 students participated, however, six participants turned in incomplete spreadsheets and their data were dropped from the preliminary analysis. As part of the post-test questionnaire, the researcher asked participants several questions to ensure that they understood the task and that the technology used in the experiment was not a problem.

The first of such questions asked if the participant had any problems with the technology in the lab or the spreadsheet package itself. Six participants indicated that they had a problem with the technology. Two participants indicated that they either did not like the computer because it was too slow or they "didn't like the border formatting." Those two participants' data were not dropped. However, four participants reported problems that could affect their spreadsheet solution. For example, one participant reported that he/she did not "know some debugging messages," another participant reported a problem "knowing/finding all the functions to utilize," and another reported "not sure that I completely understood the problem." These individuals were dropped from subsequent analyses.

As a check on this question, the researcher also asked participants to indicate the difficulty they had using the spreadsheet program itself on a four-inch scale anchored "None" and "A lot." An additional participant was dropped because he/she indicated "A lot" even though he/she did not indicate the specific problem.

As an additional check, the researcher also asked participants if their knowledge of the spreadsheet program mechanics was more than adequate, adequate, or inadequate to do the task. One participant indicated that his/her knowledge of the spreadsheet mechanics was inadequate; this participant was dropped.

Beyond assessing the participants' understanding of the technology and software, a final question asked if the participant understood the problem. The question asked if the participant possessed the knowledge of the logic required in the bidding process. One participant indicated that his/her skill was inadequate and was dropped. Participants also indicated which spreadsheet packages they had used in the past. All remaining participants indicated that they had previously used Excel[©]. The final set of data contains 116 observations and the deletions are summarized in Table 5. The sample contained twenty-eight participants in the control group, twenty-six participants in the planning and design group, twenty-nine participants in the formula complexity group, and thirty-three participants in the testing and debugging group.

TABLE 5

Summary of Deleted Participants

Summary of Participants	
Number of original participants	129
Incomplete spreadsheets	6
Technology problems	4
Trouble with spreadsheet program	1
Inadequate knowledge of the logic required	1
Inadequate knowledge of spreadsheet mechanics	1
Total participants used in analysis	116

4.1.3 Tests of Statistical Assumptions

The data were then subjected to tests of the assumptions of independence, normality, and homogeneity of variance. The first assumption, independence, may be satisfied by randomly assigning participants to experimental conditions. Accordingly, all participants were randomly assigned to either the control group or one of the three treatment groups (planning and design aid, formula complexity aid, and testing and debugging aid). Randomization tests were performed using the four experimental groups as the independent variable. Table 6 presents the results of those tests. As can be noted from Table 6, it appeared that academic rank was not random.

TABLE 6

Tests of Randomization

Independent Measure - Treatment Group*		
Dependent Variable	Test	Significance Level
Spreadsheet Knowledge	ANOVA	0.922
Spreadsheet Experience Hours	ANOVA	0.816
Hour Work per Week	ANOVA	0.911
Gender	Chi Square	0.614
Academic Rank	Chi Square	0.000^{+}

* control, planning and design, formula compleixty, testing and debugging

[+] see further statistical tests

Composition of the groups is shown in Table 7. It can be seen that the control group was 50% graduate students whereas in the other three experimental groups, there were only 21% to 23% graduate students.

TABLE 7

Frequency Counts - Academic Rank

Group ▶ Academic Rank ▼	Control	Planning & Design	Formula Complexity	Testing & Debugging
Undergraduates	14	20	23	26
Graduate	14	6	6	7

The three treatment groups did not appear to be different from one another, whereas the control group appeared to be different based on academic rank. The next step was to calculate a t-test for the graduate and undergraduates in each of the four experimental groups to determine if differences in task performance existed. The study cannot reject the null hypothesis that there is no significant difference in performance for the two academic ranks in any of the four groups (control group $t = 1.36$, $p = .186$; planning and design group $t = .83$, $p = .413$; formula complexity group $t = .31$, $p = .759$; testing and debugging group $t = .02$, $p = .985$). Therefore, statistically, it appears that the groups are sufficiently randomized.

Graphical and statistical methods were used to test the assumption of normality. First, normal probability plots were generated to allow a visual inspection of the data. Several of these plots were slightly non-normal in appearance, so further analysis was conducted through the use of both the Kolmogorov-Smirnov Normality Test with a Lilliefors Significance Correction for testing normality, and the Shapiro-Wilks statistic which is used for small samples (Norusis, 1997). The test statistics confirmed the visual information from the plots. Thus, the hypothesis of normality is rejected and as a result both parametric and nonparametric procedures were calculated. In cases where the data

are non-normal, the nonparametric procedures are often more robust. Only the parametric results are reported where they are consistent with the equivalent nonparametric test, otherwise the nonparametric results are reported.

The homogeneity of variance assumption was evaluated using the Levene's Test of Homogeneity. Researchers consider Levene's better when the data come from continuous, but not necessarily normal distributions, across any number of groups (Hair, Anderson, Tatham & Black, 1995). Other procedures commonly used to test homogeneity of variance (e.g., Hartley, Cochran, and Bartlett) may have insufficient power in the presence of non-normality (Keppel, 1991). The Levene statistic of .8575 (p = .4655) demonstrates that there is homogeneity of variance and that the unequal cell sizes should not impact the sensitivity of the statistical tests.

4.2 Hypothesis One

This study tests hypotheses $H1_A$, $H1_B$, and $H1_C$ by comparing each treatment group, which received an aid, to the control group, which did not receive an aid, to determine whether a significant increase in the quality of the spreadsheet models occurs. This was accomplished by the t-test.[8] An ANOVA was not performed because the a priori interest is not between the different treatment groups but between the control and each individual treatment group. In each case, the treatment group had significantly fewer errors than the control group[9]. Figure 2 contains a bar chart of the mean number of errors per spreadsheet model in each of the four experimental groups. Table 8 contains the statistical results.

---

[8] The Wilcoxon Rank Sum test and Mann-Whitney test are the equivalent nonparametric procedures. The Mann-Whitney, is equivalent to the rank sum test except that it uses an exceedences method to compute the test statistic (Schlman, 1992). These two nonparametric tests give the same significance level.

[9] There were no significant difference between any of the three treatment groups.

Number of Spreadsheet Errors
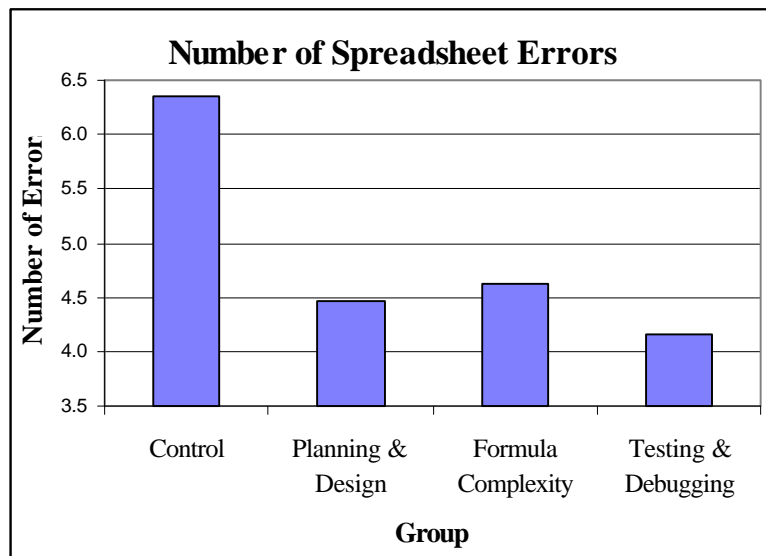

FIGURE 2

Mean Number of Spreadsheet Errors

TABLE 8

Summary of Hypothesis One

Hypothesis One - Results

Group Compared to Control	t-value	Significance Level
Planning and Design	1.69	0.048
Formula Complexity	1.76	0.042
Testing and Debugging	2.36	0.022

These results support hypotheses H1$_A$, H1$_B$, and H1$_C$: a significant increase in the data quality as measured by a decrease in spreadsheet errors occurs for the groups of participants using an aid. The next three sections further investigate these results by presenting the results for the remaining three hypotheses followed by some supplemental analyses.

4.3 Hypothesis Two

Hypothesis H2 was tested by comparing the treatment group, which received the planning and design aid, with the control group, which did not receive the aid, to determine whether there was a significant relationship between the quality of a spreadsheet and the "goodness" in terms of the design and organization of the model. Determining the "goodness" of the spreadsheet was accomplished by the use of two spreadsheet experts, referred to as Expert A and Expert B.[10] They worked independently and ranked the spreadsheets into four groups. Then they named and defined each group in their own words.

The question of interest here is whether or not a better spreadsheet design, as defined by the experts, is associated with higher data quality, as defined by the number of errors. In looking at the details of the ranking, there is a consistent pattern in both the rankings, definitions, and descriptions provided by the Experts. Expert A consistently ranked the spreadsheets more stringently than Expert B. Expert A ranked thirty of the fifty-four spreadsheets one or two levels below that of Expert B. In no case did the experts differ by more than two ranks. This suggests that the ordering of the spreadsheets is consistent between the two experts. Using an average of the two experts' ratings, the relationship was tested using the Kendall's modified Theil-Sen correlation coefficient.[11]

---

[10] Reference of Expert A and Expert B are for convenience only.

[11] The nonparametric equivalents to the Pearson's correlation coefficient are the Kendall's modified Theil-Sen and the Spearman correlation coefficient. The results of the parametric and nonparametric tests were significantly different and as a result the nonparametric results are presented.

The results indicate that no relationship exists, and thus, hypothesis H2 is not supported. Chapter 5 includes a detailed discussion of the experts and additional analysis.

4.4 Hypothesis Three

Hypothesis H3 was tested by comparing the formula complexity treatment group, which received the formula complexity aid, to the control group, which did not receive the aid, to determine whether there was a significant relationship between the level of complexity of a formula and the quality of a spreadsheet model. Formula complexity was defined by the length of the longest formula in each spreadsheet model. The formula length is a count of the variables (or cell references) and constants in the formula. This number ranged from three to eighteen for the control group and from two to nineteen for the formula complexity group. The mean (standard deviation) of formula length for the control group was 6.29 (3.25) and for the treatment group was 5.93 (4.15). No significant difference existed between these two groups ($t = .36$, $p = .990$). The Kendall's modified Theil-Sen correlation coefficient was calculated to determine if there was a relationship between the longest formula and the number of spreadsheet errors[12]. The results indicate that no relationship exists, and thus, does not support hypothesis H3. Although this hypothesis is not supported as operationally defined, Chapter 5 discusses post-hoc analysis which suggests this operational definition of the construct may be inadequate.

4.5 Hypothesis Four

Hypothesis H4 was tested by comparing the treatment group, which received the testing and debugging treatment aid, to the control group, which did not receive the aid, to determine whether there were any more errors corrected after the participants used the aid.[13] The two variables, the number of errors before and the number of errors after, are not independent. The hypothesis test was, therefore, accomplished by using the one-tailed

---

[12] ------.

[13] One participant actually created one error during the testing and debugging stage.

paired t-test.[14] The participants corrected a significant number of errors as determined by the results of the one-tailed paired t-test (t = 1.87, p = .035). Hypothesis H4 is supported. The use of the testing and debugging aid does significantly decrease the number of spreadsheet errors.

In summary, this study supports hypotheses $H1_A$, $H1_B$, and $H1_C$, and H4. All three aids increase spreadsheet quality as defined by decreased errors when compared to a control group; and participants who used the testing and debugging corrected a significant number of spreadsheet errors. These results will be further discussed in Chapter 5 and these results are reported in Table 15 at the end of Chapter 5.

4.6 Supplemental Analysis

The supplemental analysis uses various items taken from the post-test questionnaire. Confidence measures, spreadsheet experience, several systems courses, and a few miscellaneous demographic items were examined for their relationship to the dependent variable.

4.6.1 Confidence and Task Difficulty

The demographics questionnaire asked several questions about confidence and the perceived quality achieved in the participants' spreadsheet model. This information was captured both as a discrete variable and as a continuous variable. Three different questions captured the discrete data. One question asked what level participants considered their spreadsheet solution quality; very high, somewhat high, neither high nor low, somewhat low, or very low. Over 90% answered either very high or somewhat high quality. Another question asked participants about their confidence of the accuracy of their solution using the same 5 choices. Eighty-seven percent of the participants rated their confidence in the accuracy of their solution as either very high or somewhat high.

---

[14] The Wilcoxon matched-pairs signed-rank test, which uses information about the size of the difference between the pairs (Norusis, 1997) is the equivalent nonparametric procedure.

The third question asked participants the grade they deserved using the traditional grading scale of A, B, C, D, and F. Over 96% thought they deserved an A or B, three participants thought they deserved a C, none thought they deserved a D or F, however, one participant felt their model was incomplete. These three attitudinal measures are correlated as demonstrated by the Pearson's correlation coefficients ranging from .62 to .82 (p = .000). A composite variable was created using the average of these three questions. No significant relation existed between each individual variable or the composite variable and the number of spreadsheet errors.

The researcher also collected confidence data with two questions measured on continuous scales. Participants were asked to rate how error-free they believed their solution was on a four-inch scale anchored "Not At All" and "Definitely." Most participants considered their models to be error free with a mean (standard deviation) of 2.40 (1.07). The second question asked if the participants believed they set forth a high-quality solution on a four-inch scale anchored "Not At All" and "Definitely." Most participants considered their solution to be of high-quality with a mean (standard deviation) of 2.77 (.74). No significant relation existed between either of these confidence variables and the number of spreadsheet errors.

As a check on the confidence questions, the researcher also asked participants to rate how difficult they felt the "Bidder" task was on a four-inch scale anchored "Not Difficult" and "Very Difficult." Most participants rated the task as not difficult with a mean (standard deviation) of 1.19 (.89). Again, no significant relation existed between assessed difficulty and the number of errors in the spreadsheet. Section 5.1.4 further discusses these results.

4.6.2 Experience

Experience was captured as both a discrete variable and as a continuous variable. The discrete variable was captured as beginner, moderate or intermediate spreadsheet experience. These data were analyzed using a 3 x 4 ANOVA with three levels of

experience by the four experimental groups.[15]   No significant differences existed in the interaction or the main effects of experience when measured as a discrete variable.

The continuous variable was captured by asking the participant the length of time in years and months and the number of hours per week they used a spreadsheet. These data were analyzed using the Pearson's correlation coefficient. No significant relation exists between the number of spreadsheet errors and the hours of spreadsheet experience. Categorical data were created using the number of hours of experience broken into five categories (<100, 100-249, 250-499, 500-999, and ≥ 1000). This test determined whether an interactive effect of spreadsheet experience and errors occurred between any one of the created groups. The correlation analysis may have masked any interactivity if it was only at the extremes. A 4 x 5 ANOVA with four experimental groups and five levels of experience was used. Again, no significant relationship existed between spreadsheet experience expressed as hours and the number of spreadsheet errors for any of the groups.

4.6.3 Systems Courses

Several different types of systems courses were also analyzed. Systems development courses teach students to approach a problem in a more structured manner. Over 80% of the participants had completed one or more collegiate courses in structured systems development. These data were analyzed using a Oneway ANOVA with five levels (1-no; 2-yes, one course in high school; 3-yes, more than one course in high school; 4-yes, one course at a college or university; 5-yes, more than one course at a university). This ANOVA showed marginal significance ($p = .06$). However, post hoc testing revealed that this was obtained by one cell having only one observation. Participants may not have considered spreadsheet development as systems development and therefore, may not have applied the same processes.

---

[15] The Kruskal-Wallis test (which is like the Mann-Whitney test except that there are more groups (Norusis, 1997)) is the equivalent nonparametric procedure.

Over 80% of the participants had completed one or more collegiate courses in database development and design; and one or more collegiate courses in spreadsheet development. The Oneway ANOVA with five levels (1-no; 2-yes, one course in high school; 3-yes, more than one course in high school; 4-yes, one course at a college or university; 5-yes, more than one course at a university) found no significant differences for database or spreadsheet courses. Programming language courses also require students to develop a structured methodology to developing systems applications. The most used programming language was Visual Basic, which is the macro language for Excel[©]. Over 60% of participants indicated that they knew how to use Visual Basic. A t-test failed to find significant differences between any systems course and the number of spreadsheet errors. Other languages included in the demographics questionnaire were C, C++, COBOL, Fortran, Pascal, and Basic, none of which were indicated as used by more than one-third of the participants.

4.6.4 Demographics

Over 45% of the participants were not working during the semester the study was conducted, and almost 40% worked less than twenty hours a week. Over 90% of the students took over ten credit hours and 15% took over sixteen credit hours. Approximately 52% of the participants were male. Each of these attributes was captured as a discrete variable and analyzed using a Oneway ANOVA. Again, none of these attributes were significantly associated with the number of spreadsheet errors. Approximately 17% of the participants indicated that English was not their primary language. This attribute was analyzed using the Oneway ANOVA with the number of errors as the dependent variable and primary language as the independent variable. This attribute was found to be significant overall, however, when it was further analyzed to see if it interacted with any of the experimental groups, it did not. Primary language did not affect the number of spreadsheet errors.

CHAPTER 5

DISCUSSION, IMPLICATIONS, AND LIMITATIONS

This chapter contains a discussion of the results of the data analysis presented in the previous chapter and the conclusions drawn from those results. The first section contains a discussion of the results of the analysis of each aid individually and a few comments on participant confidence. The second section contains the implications of the present study. The limitations of the study are reviewed in the third section. The concluding section provides recommendations for future research.

5.1 Discussion of Results of Hypotheses Testing

Spreadsheet models have been used extensively today in accounting and business applications. Unfortunately mounting anecdotal and empirical evidence indicate that these spreadsheets often contain data quality problems, i.e. the underlying formulas and resulting numbers are frequently wrong. Learning about, and understanding, ways to increase data quality will help decision-makers because these decision-makers often base decisions on incorrect numbers. Providing a decision-maker with a more accurate tool should yield better decisions.

Research that has examined the data quality problem was discussed in Chapter Two. This research indicates that there is a problem with the data quality in spreadsheets and provides some anecdotal examples. Chapter Two also presents the model and theory of spreadsheet accuracy.

The purpose of this study was to develop a theoretical model for spreadsheet accuracy. The study achieved this purpose by proposing a theory from a detailed review of the available literature and operationalizing this through the creation of four aids and

examining the number of spreadsheet errors found in spreadsheet models that participants created in a laboratory setting using those aids. The reliability and validity, manipulation checks, and statistical assumptions were described. Then this paper reported the results of the tests of the four hypotheses, each comparing the control group to one of the following: planning and design aid, formula complexity aid, or testing and debugging aid. Following is a discussion on each hypotheses and aid results.

5.1.1 Planning and Design Aid

Hypothesis $H1_A$ predicted that the number of spreadsheet errors would decrease when participants used the planning and design aid. Those participants that used the planning and design aid had, on average, 1.9 fewer errors per spreadsheet model (p = .048).

5.1.1.1 Experts

Hypothesis H2 required testing for a relationship between the "goodness" of a spreadsheet model, as defined by the spreadsheet experts, and the number of errors. Two experts rated/ranked each of the spreadsheets in both the control and the planning and design group. As indicated in Chapter Four, no relationship exists between the expert rankings and the number of spreadsheet errors. One possible reason for the insignificance is that the inter-rater agreement between the experts was not high. The experts who ranked the spreadsheets also provided names and definitions for each group. Table 9 shows the names and definitions provided by the experts. Expert A was more concerned with the functionality of the spreadsheets, whereas Expert B's top grouping refers to cosmetic type attributes. The instructions to the experts may have induced this effect when participants used colors or fonts as indicators. The instructions to the experts explained that on screen background colors were displayed as different shades of gray on the printouts and on screen font colors were displayed as different font styles on the printouts. The instructions to the experts are contained in Appendix E. Expert A, after completing the task, said that segmenting a spreadsheet so that the input area is very clear

and separate was the most important design criteria and furthermore, he/she would only consider employing individuals who created spreadsheets in the first grouping. This is in agreement with the practitioner literature; the functionality of segmentation reduces the possibly of accidentally over-writing formulas.

TABLE 9

Names and Definitions Given By Experts

Expert Opinion			
Rank		Expert A	Expert B
1	Definition	Understandable, segmented & reusable	Most professional
	Elaboration	Have most required elements, well organized	Made good use of spreadsheet fonts, shading and appears reusable
2	Definition	Readable format, something missing	Good ideas - needs more work
	Elaboration	Something noteworthy, be it format or understandability	Marginal use of spreadsheet capabilities/more design needed
3	Definition	Convey information, but format and segmentation need work	Maybe a start - lot more work
	Elaboration	Some required elements, not executed effectively	First step in something that could be useful
4	Definition	Not reusable, poor segmentation, poor understandability	Not much - start over
	Elaboration	Few required elements	Used spreadsheet as a word processor - not reusable

In further analyzing this problem, it appears that one of the experts may not have been a spreadsheet expert. Both experts were referrals and unknown to the researcher. Both considered themselves spreadsheet experts. Expert A worked for a large multi-national firm. This expert supervised several employees and worked on reporting various financial performance data which was conveyed to others both inside and outside the firm

using spreadsheets. Expert B started a computer consulting business and regularly participated in a local computer club's help session for other end-users, answering their questions. Since conducting this portion of the experiment, it was learned that this individual was considered an expert by very novice users. Expert B also did not know how to write spreadsheet macros. Figures 3 and 4 show Boxplots that were created for each experts' ratings and the number of spreadsheet errors. Rating group 1 is considered the best and rating group 4 is considered the worst. Boxplots are formed from "boxes", which contain 50% of values falling between the 25th and 75th percentiles. Vertical lines that extend from the box to the highest and lowest values, called "whiskers," exclude outliers (Norusis, 1997). Outliers are displayed as points in the diagram. The horizontal line across the box indicates the median. Boxplots also convey information about spread and skewness (Norusis, 1997). Visually Expert A's ratings appeared to be related to the number of spreadsheet errors (Figure 3). Expert B had a random visual appearance (Figure 4).
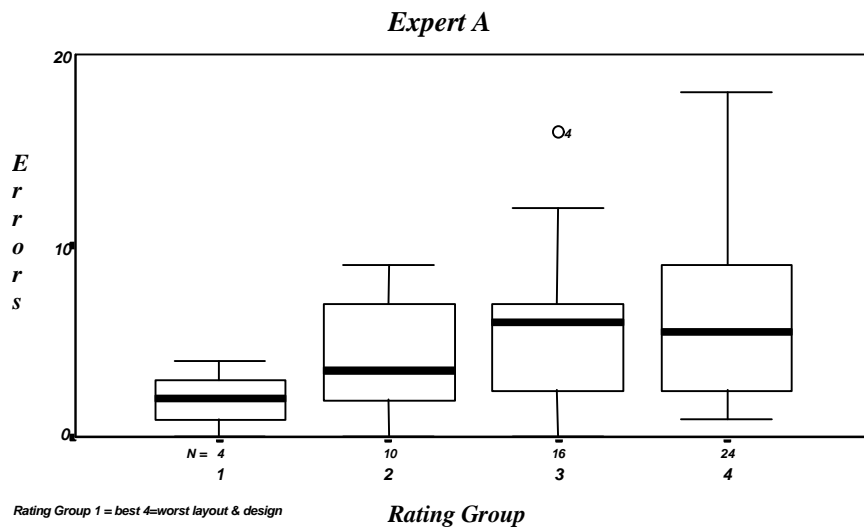

FIGURE 3

Boxplot of Expert A's Ratings

Expert B

Rating Group 1 = best 4=worst layout & design
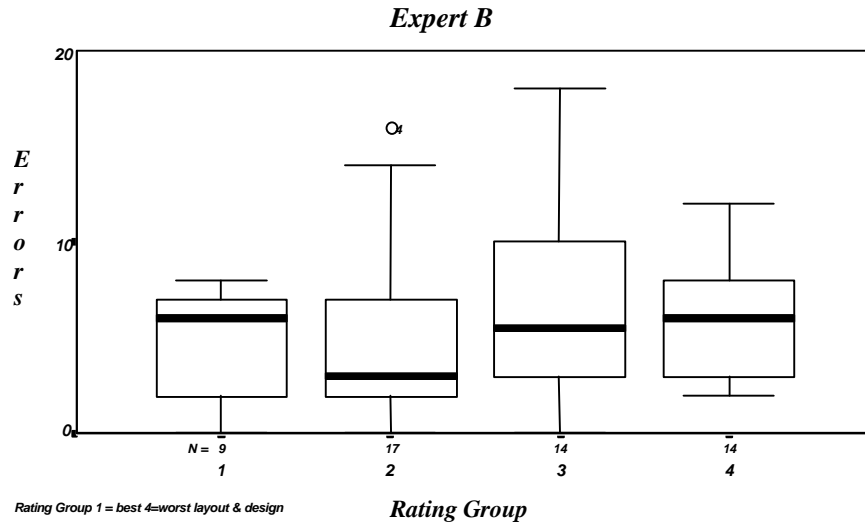
FIGURE 4

Boxplot of Expert B's Ratings

Based on this qualitative evaluation, Expert B was dropped. Correlation measures were calculated using Expert A's ratings only. The correlation was low (.3) yet significant (p = .014). This area needs further research and possible rewording of the instructions to the experts to provide stronger support for this proposition.

5.1.1.2 Time

An additional question to be answered was if the additional time required to use the aid was worth the reduction of approximately two errors. During the post-test questionnaire, the researcher asked participants to report the amount of time they estimated it took to plan their spreadsheet. The spreadsheet task also recorded both the time the participant opened the spreadsheet and left the opening screen. This was verified by having participants open the blank bidder task spreadsheet before receiving the task statement. Each participant pressed a button when ready to enter data. This time interval was recorded internally by the computer and later compared to the participants estimate of time. Previous studies had noted that participants rarely spend time planning the

spreadsheet before beginning to enter data (Brown & Gould, 1987; Panko & Sprague Jr., 1997). During an examination of the data, two outliers became apparent. A participant in the control group who reported sixty minutes of planning had only thirty-three seconds recorded and a participant in the planning and design group who reported ninety minutes of planning had less than twenty-seven minutes recorded. Excluding these two observations, the planning time statistics are reported in Table 10.

TABLE 10

Planning Time Statistics

Planning Time Statistics in Minutes (1 outlier removed per group)					
Group	N	Minimum	Maximum	Mean	Standard Deviation
Control					
Reported by participants	26*	1.00	19.00	8.73	5.14
Recorded by spreadsheet	27	0.03	23.63	5.95	5.54
Planning & Design					
Reported by participants	25	1.45	45.00	15.36	10.67
Recorded by spreadsheet	25	0.07	29.83	11.21	7.97

* one participant did not answer this question

The amount of time to use the planning and design aid, on average, regardless of whether reported or recorded, was approximately 5 1/2 to 6 1/2 minutes more than the control group. In both the control group and the planning and design group, the participants reported spending more time than the spreadsheet recorded. The additional time required to reduce the errors is approximately 3% of the total development time. Nonetheless, a significant difference existed in the amount of planning time between the two groups. This is illustrated in Table 11. The end-user will have to perform a

cost/benefit analysis to decide if the reduction of approximately two errors is worth the additional time required to use the planning and design aid.

TABLE 11

Statistical Tests of Planning Time

Significance Tests for Planning Time Between Groups					
	Test		Significance		Significance
	Statistic	Value	Level	Value	Level
		Self-Reported Time		Recorded Time	
All Observations	t-test	1.86	0.069	3.15	0.003
Without 2 Outliers	t-test	2.84	0.006	2.78	0.008

5.1.2 Formula Complexity

Hypothesis $H1_B$ predicted that the number of spreadsheet errors would decrease when participants used a formula complexity aid. This hypothesis was supported. Those participants that used the formula complexity aid had, on average, 1.7 fewer errors per spreadsheet model ($p = .042$) than the control group.

Hypothesis H3 required testing the complexity of formulas to determine if there was a significant relationship between formula complexity, as defined by length, and spreadsheet data quality. This hypothesis was not supported, as operationally defined. The average of all formulas in each spreadsheet model was also analyzed to determine if it was a better operational definition. Again there was no significant difference. The operational definition of formula length as the longest formula or average formula length are inaccurate because both the longest formulas and the average formula length in the

55

control group and the formula complexity group were similar. The formula length and average formula length should be similar because the participants are creating the models from the same task statement and therefore would require similar if not the same calculations. Upon closer examination of the spreadsheets, an interesting difference between the control group and the formula complexity group emerged. The formulas of the formula complexity group used more cell references or variables instead of constants or hard-coded numbers. This may be a more appropriate measure of the formula complexity construct.

The following explanation and research studies support use of the number of constants as the operational definition of formula complexity. Constants are more complex to use in formulas. A constant used in a formula will not allow automatic updating when a number is changed. For example, if cell A1 contained 2 and cell B1 contained 4, cell C1 would display 6 if the formula in the cell were = A1 + B1, or if the formula were = 2 + 4. The first formula has two variables or cell references (not constants) and if the value were changed in either cell A1 or cell B1, the value displayed in cell C1 would also change. In the second case, the numbers are constants or hard-coded and if the value in either cell A1 or cell B1 were changed, the value displayed in cell C1 would not change. This defeats one of the main benefits of spreadsheets, furthermore, using constants in formulas is essentially using the spreadsheet program as a word processor. One of the main functions of a spreadsheet program is the ability to change any parameter and have it reflected throughout the spreadsheet.

Te'eni (1990) conducted a study to compare two types of interfaces and making a decision. Participants created graphs using either a command-based (DOS -- typing) or a direct manipulation interface (Windows -- point and click). The users that directly entered the data made fewer logical errors. In addition, these participants were more efficient in terms of time and cognitive effort required. Morgan, Morris, and Gibbs (1991) compared the direct manipulation interface with the command-based interface. Participants were computer novices but familiar with the keyboard. Participants made half as many errors

and were more satisfied when they used the direct manipulation interface. Davis and Bostrom (1993) recently studied two computer interfaces to determine if using a direct manipulation interface was different than a command-based interface. These researchers found that there was a significant difference between the performance of the direct manipulation subjects and the command-based subjects. Furthermore, there was a significant difference in the perceived ease of use of the direct manipulation interface over the command-based interface. Participants in the direct manipulation interface group found the computer easier to use.

This generation of students/participants have been raised on computers with the Windows operating system and are intimately familiar with using the mouse for point and click techniques. Formulas that utilize variables are less complex because the participant can use a point and click method to create the formula. For example to write a formula to add together cell A1 and cell B1 would begin either by typing the equal sign (=) or by selecting an appropriate built-in function. As an example, using the equal sign, they would first start the formula by typing the equal sign, then point and click on cell A1, type a plus sign (+), point and click on cell B1 and hit the enter key to finish. If a built-in function were used instead, the spreadsheet creator would select the function, then fill in the arguments needed by pointing to their cell locations. Using the previous example, the spreadsheet developer would point to the *fx* button, select the Sum function, and then point and click on each cell or the range of cells to be added. They do not need to remember cell addresses to write the formula nor do they need to remember the numbers. If the spreadsheet developer decided to use constants, they would be required to type in the correct number, which increases the likelihood of an error because they must remember and type it correctly (Shneiderman, 1992). Using the point and click method makes creating spreadsheets with variables less complex than using constants.

The two variables that were most often hard-coded by the participants in this experiment were the 30% profit margin and the 20% additional worker benefits. Table 12 contains the descriptive statistics on the number of constants used and significance level

for both the control group and the formula complexity group. Panel A includes all observations from both groups. One spreadsheet in the formula complexity group contained three times as many constants as the next highest one. Upon further inspection, this participant may have misunderstood the instructions because all the material costs and units were hard-coded along with the profit margin and the additional worker benefits. Panel B reports the same information but excludes this one outlier. The formula complexity participants used significantly fewer constants in their formulas.

TABLE 12

Constants Used by Formula Complexity Group

Panel A: Constants Used In Formulas (all observations)

Group	N	Mean	Standard Deviation	Test Statistic	Value	Significance Level
Control	28	11.357	13.287			
				t-test	2.18	0.017
Formula Complexity	29	4.690	9.339			

Panel B: Constants Used In Formulas (1 outlier removed)

Group	N	Mean	Standard Deviation	Test Statistic	Value	Significance Level
Control	28	11.357	13.287			
				t-test	3.11	0.002
Formula Complexity	28	3.143	4.301			

The operational definition of formula complexity as the number of constants contained in each spreadsheet model is significantly different between the control group and the formula complexity group. It also appears to have captured the formula complexity construct. The use of the formula complexity aid did affect the number of constants in the spreadsheet model, however, there is no relationship between the number of constants and the number of spreadsheet errors.

5.1.3 Testing and Debugging

Hypothesis $H1_C$ predicted that the number of spreadsheet errors would decrease when participants used the testing and debugging aid. This hypothesis was supported. Those participants that used the testing and debugging aid had, on average, 2.2 fewer errors per spreadsheet model ($p = .022$) than the control group.

Hypothesis H4 required testing for a significant decrease in the number of errors after using the testing and debugging aid. This testing necessitated making a copy of the spreadsheet model prior to the participant using the testing and debugging aid. Due to experimental constraints and consideration for the participant's time, the training on the use of the aid took place prior to creation of the spreadsheet. This may have masked the before and after effects of the aid because the participants may have used the testing and debugging techniques as they created their model. However, since the participant assignment to groups was random, there is no reason to believe that the groups were significantly different prior to the experiment, therefore, the results can be attributed to the aid. The participants corrected only nine errors (which was statistically significant), yet there was approximately 2.2 fewer errors per spreadsheet when compared to the control group. Even though this masking may have taken place, the results are still profound. This aid supported the end result of a better quality spreadsheet, thus the conclusion that the aid significantly reduced spreadsheet errors in the final model is consistent.

5.1.4 Confidence

One set of post-test questions appears to be conflicting because a spreadsheet model not considered error-free was considered to be of high quality. The researcher asked participants to rate how error-free they felt their solution was on a four-inch scale anchored "Not At All" and "Definitely." Overall, participants considered their models to be error free with a mean (standard deviation) of only 2.77 (.74). Had the participants believed that their models were completely error-free this mean would be closer to four. Only four participants rated their solution at the extreme of "Definitely" error-free and only four participants rated their solution at the other extreme of "Not At All" error-free. This leaves 93% of the responses in between. The quartiles from not error-free to definitely error-free (1-14%, 2-19%, 3-30%, and 4-37%) show that approximately one-third of the participants indicated that they considered their models to contain errors, yet over 96% of the participants thought they deserved an A or B. Furthermore, over 90% considered their solution quality as very high or somewhat high. In addition, the participants did not consider the task difficult, as the mean rating was 1.19 on a four-inch scale anchored "Not Difficult" and "Very Difficult." In other words, the participants thought a solution that was not error-free was also one of high quality. It is of that students feel they deserve an A or B, and they constructed a solution they considered to be of high quality, yet they don't think their work is error-free.

In summary, the model and theory developed in Chapter Two appear to represent the phenomenon of spreadsheet accuracy. The three aids developed to test the theory and model did increase the data quality of spreadsheets. Furthermore, the formula complexity participants used fewer constants in their formulas, utilizing the power of spreadsheets; the testing and debugging group did find and correct a significant number of errors. However, what is unknown at this point is specifically which technique in each aid has led to the increase in data quality, or the use of fewer constants and more variables. Table 13 contains a summary of the research results.

TABLE 13

Summary of Research Results

Hypotheses	Significant	Meaning
H$_{1A}$: The participants that use the planning and design aid will have no significant increase in spreadsheet accuracy over the control group.	✓ p = .048	Spreadsheet accuracy increased
H$_{1B}$: The participants that use the formula complexity aid will have no significant increase in spreadsheet accuracy over the control group.	✓ p = .042	Spreadsheet accuracy increased
H$_{1C}$: The participants that use the testing and debugging aid will have no significant increase in spreadsheet accuracy over the control group.	✓ p = .022	Spreadsheet accuracy increased
H$_2$: The participants that use the planning and design aid will have no significant relationship between spreadsheet accuracy and the quality of their design and organization.	✓ p = .014[16]	No relationship between spreadsheet design and errors
H$_3$: The participants that use the formula complexity aid will have no significant relationship between spreadsheet accuracy and the complexity of the formulas.	p = .220	No relationship between the number of constants used and errors
H$_4$: The participants will not correct any more of the spreadsheet errors after using the testing and debugging aid.	✓ p = .035	Significant number of errors were corrected

---

[16] Although the p-value is significant, the Pearson's correlation coefficient is only .3008.

5.2 Implications

The findings in this study are applicable to researchers, practitioners, and educators. This study has developed a model and theory that explains the phenomenon of spreadsheet accuracy.   Researchers should take this study as the first in a stream of research to advance our understanding of spreadsheet accuracy and further develop the model. This research may have implications for systems development research because in most cases the end-user is the system developer.

Spreadsheet data quality is a problem which in many cases affects the bottom line. Practitioners may want to incorporate techniques studied here into the development of spreadsheets. Practitioner use of these aids should increase the data quality of their spreadsheet models and reduce costly errors.

Educators may want to consider teaching all business students some structured spreadsheet development. As discussed in the literature review section, spreadsheet skills are highly desired by business professionals. Students in systems degree programs do receive training in systems development. It is somewhat apparent in this study, however, that it did not impact the quality of the spreadsheets created by systems students. This result may have been obtained because students do not consider spreadsheet development the same as systems development. The educational program should specifically point this out to students so that they will develop spreadsheets using a structured process and thus create a higher quality spreadsheet.

5.3 Limitations

These results are subject to several limitations. First, participants were students, and from a single institution, which limits the generalizability of the results.   Second, this was a laboratory experiment. The laboratory experiment provided greater control, but was limited because it did not allow for the participants' natural interaction with colleagues. Thus, a laboratory setting limits the external validity of the study. Finally, the task may have been too difficult, as indicated by participants' difficulty in the linear feet

calculation; forty-four percent of the participants miscalculated the linear feet. The researcher could not determine if the miscalculations were a logical or formula error.

The reduction in errors may have been a result of the use of any aid and not any particular technique. The use of an aid may have helped participants become more aware and focus their attention on the quality of their spreadsheet. This should be tested at some future date.

5.4 Recommendation for Future Research

This study highlighted several areas of interest for future research. A larger study should be undertaken to add support to the one hypothesis that was not supported. This hypothesis (formula complexity) showed that the construct was correctly manipulated, however, did not attain significance.

Although the theory was supported, each technique in each aid should be tested to determine which specific techniques created the increase in data quality. The techniques included in each aid were not considered to be all-inclusive. Additional techniques should also be investigated and tested. Also, significant techniques in each of the three areas (planning and design, formula complexity, and testing and debugging) could be combined and tested. Regarding the model and theory, additional nodes could be developed, tested, and then added to the model. Testing of the interaction of the constructs should also be considered for future research.

REFERENCES

Accounting Education Commission (1990). Objectives of education for accountants: Position statement number 1. Issues in Accounting Education, 5 (2): 307-312.

Alavi, M., Nelson, R. R. & Weiss, I. R. (1987-88). Strategies for end-user computing: An integrative perspective. Journal of Management Information Systems, 4 (3), 28-49.

Alexander, R. A. (1996). Teaching good systems design for spreadsheet projects. Journal of Accounting Education, 14 (1), 113-122.

Allwood, C. M. (1984). Error detection processes in statistical problem solving. Cognitive Science, 8 413-437.

American Accounting Association. (1959). Report of committee on accounting instruction in electronic data processing. Accounting Review, 215-220.

------. (1964). Report of committee on accounting systems instruction. Accounting Review, 715-720.

------. (1970). Report of the 1968-69 committee on the role of the computer in accounting education. Accounting Review (supplement), 29-43.

------. (1986). Future accounting education: Preparing for the expanding profession (Committee on the Future Structure, Content, and Scope of Accounting Education). Issues in Accounting Education, 168-195.

Anderson, J. C., Jennings, M. M., Kaplan, S. E., & Reckers, P. J. J. (1995). The effect of using diagnostic decision aids for analytical procedures on judges' liability judgements. Journal of Accounting and Public Policy, 14 33-62.

------. (1997). The effects of interference and availability from hypotheses generated by a decision aid upon analytical procedures judgements. Behavioral Research in Accounting, 9 1-25.

Arnold., S. E. (1992). Information manufacturing: the road to database quality. Data Base 15 (5) 32.

Ashton, R. & Willingham, J. (1988). Using and evaluating audit decision aids. In Touche Ross/University of Kansas Audit Symposium Proceedings, University of Kansas.

------. (1990). Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback and justification. <u>Journal of Accounting Research, 28, (Supplement)</u> 148-180.

------. (1992). Effects of justification and a mechanical aid on judgment performance. <u>Organizational Behavior and Human Decision Processes, 52</u> 292-306.

Ballou, D. P., & Pazer, H. L. (1987). Designing inspection strategies for uncertain environments. <u>Decision Sciences Journal, 18</u> (2), 217-233.

Beardstown ladies had the wrong results. (1998, March 18). <u>The New York Times</u>, D2.

Benbasat, I., & Schroeder, R. G. (1977). An experimental investigation of some MIS design variables. <u>MIS Quarterly, 1</u> (1), 37-50.

Bissell, J. L. (1986). Spreadsheet planning and design. <u>Journal of Accountancy, 161</u> (5), 110-120.

Brancheau, J., & Wetherbe, J. (1987). Key issues in information systems management. <u>MIS Quarterly, 11</u> (1), 23-45.

Brewer, J., & Hunter, A. (1989). <u>Multimethod research: A synthesis of styles.</u> Newbury Park: Sage library of social research.

Brown, P., & Gould, J. (1987). An experimental study of people creating spreadsheets. <u>ACM Transactions on Office Information Systems, 5</u> (3), 258-272.

Butler, S. A. (1985). Application of a decision aid in the judgmental evaluation of substantive test of details samples. <u>Journal of Accounting Research,</u> 513-526.

Carlsson, S. A., & Stabell, C. B. (1986). Spreadsheet programs and decision support: A keystroke-level model of system use. <u>Decision Support Systems: A Decade in Perspective.</u> Amersterdam: North-Holland, 113-128.

Carroll, J. M. (1990). <u>The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skills.</u> Cambridge: MIT Press.

Chan, W. (1987). Sorting out spreadsheets. <u>Australian Accountant, 57</u> (11), 52-56.

Cragg, P. G., & King, M. (1993). Spreadsheet modeling abuse: An opportunity for OR? <u>Journal of Operational Research Society, 44</u> (8), 743-752.

Creeth, R. (1985). Microcomputer Spreadsheets: Their uses and abuses. <u>Journal of Accountancy, 159</u> (6), 90-93.

Davies, N., & Ikin, C. (1987). Auditing spreadsheets. <u>Australian Accountant,</u> 54-56.

Davis, G. B. (1982). Caution: User-development systems can be hazardous to your organization. Proceedings of the Seventeenth Annual Hawaii International Conference on System Sciences, 750-763.

Davis, P. (1997). What computer skills do employers expect from recent college graduates? Technological Horizons in Education Journal, 74-78.

Davis, S.A., & Bostrom, R. P. (1993). Training end users: An experimental investigation of the roles of the computer interface and training methods. MIS Quarterly, 17, (1), 61-86.

Dickson, G., Leitheiser, R. & Nechis, M. (1984). Key information issues for the 1980s. MIS Quarterly, 8 (3), 135-159.

Edge, W. R., & Wilson E. J. G. (1990). Avoiding the hazards of microcomputer spreadsheets. Internal Auditor, 35-39.

Elliott, R. K., & Kielich, J. A. (1985). Expert systems for accountants. Journal of Accountancy, 126-134.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representations. Journal of Experimental Psychology, 4 330-344.

Floyd, B. D., Walls, J., & Marr, K. (1995). Managing spreadsheet model development. Journal of Systems Management, 46 38-43, 68.

Freese, L. (1972). Cumulative sociological knowledge American Sociological Review, 37 (37), 472-482.

Galletta, D. F., Abraham, D., El Louadi, M., Lekse, W., Pollalis, Y. A., & Sampler, J. L. (1993). An empirical study of spreadsheet error-finding performance. Accounting Management Information Technology, 3 (2), 79-95.

Galletta, D. F., Hartzel, K. S., Johnson, S., & Joseph, J. L. (1996-1997). Spreadsheet presentation and error detection: an experimental study. Journal of Management Information Systems, 13 (3), 45-63.

Glaser, B. G., & Strauss, A. (1967). The discovery of grounded theory. Chicago: Aldine.

Godfrey, K. & Flatau, A. (1995). Computing error at Fidelities' Magellan Fund. The Risks Digest, 16 (72).

Gould, J. D., Alfaro, L., Barnes, V., Finn, R., Grischkowsky, N., & Minuto, A. (1987a). Reading is slower from CRT displays than from paper: Attempts to isolate a single-variable explanation. Human Factors, 29 (3), 269-299.

Gould, J. D., Alfaro, L., Finn, R., Haupt, B., & Minuto, A. (1987b). Reading from CRT displays can be as fast as reading from paper. Human Factors, 29 (5), 497-517.

Gould, J., & Grischkowsky, N. (1984). Doing the same work with hard-copy and CRT terminals. Human Factors, 26 323-337.

Haasen, W. J., & Haas, C. (1988). Reading and writing with computers: A framework for explaining differences in performance. Communications of the ACM, 31 1080-1089.

Hair, Jr., J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). Multivariate Data Analysis. Upper River Saddle: Prentice-Hall.

Hayden, R. L., & Peters, R. M (1989). How to ensure spreadsheet integrity. Management Accounting, 30-33.

Heagy, C. D., & Gallun, R. A. (1994). Recommended microcomputer knowledge for accounting graduates: A survey. Journal of Accounting Education, 12 (3), 205-210.

Heagy, C. D., & McMickle, P. L. (1988). An empirical investigation of the accounting systems course: Academic practice versus practitioner needs. Issues in Accounting Education 3 96-107.

How personal computers can trip up executives. (1984, September 24). Business Week, 94-102.

Hunton, J. E. (1996). User participation in defining system interface requirements: An issue of procedural justice. Journal of Information Systems, 10 (1), 27-47.

Johnson, J. A., Nardi, B. A., Zarmer, C. L., & Miller, J. R. (1993). ACE: building interactive graphical applications. Communications of the ACM, 36 (4), 41-55.

Kachelmeier, S. J., & Messier, Jr., W. F. (1990). An investigation of the influence of a nonstatistical decision aid on auditor sample size decisions. The Accounting Review, 65 (1), 209-226.

Kadlec, D. (1998, March 30). Jail the Beardstown Ladies! Time, 54.

Kay, A. (1984). Computer software. Scientific American, 251 (3), 52-59.

Kee, R. C., & Mason, Jr., J. O. (1988). Preventing errors in spreadsheets. <u>Internal Auditor, 45</u> (1), 42-47.

Keppel, G. (1991). <u>Design and Analysis: A Researcher's Handbook.</u> Upper River Saddle: Prentice-Hall.

Kerlinger, F. N. (1986). <u>Foundations of behavioral research.</u> New York: Holt, Rinehart and Wilson.

Klein, B. D., Goodhue, D. L., & Davis, G. B. (1997). Can humans detect errors in data? Impact of base rates, incentives and goals. <u>MIS Quarterly, 21</u> (2), 169-194.

Knight, B. (1992). The data pollution problem. <u>ComputerWorld, 26</u> (39), 81-83.

Krugman, P. (1998, March 30). There'll always be soros. <u>Fortune,</u> 31.

Krull, A. (1989). Computer rip-offs and foul-ups: Is management to blame? <u>EDPACS, 16</u> (11), 10-14.

Laudon, K. C. (1986). Data quality and due process in large interorganizational record systems. <u>Communications of the ACM, 29</u> (1), 4-18.

Libby, R. & Libby, P. A. (1989). Expert measurement and mechanical combination in control reliance decisions. <u>The Accounting Review, 64</u> 729-747.

Lorge, I., & Solomon, H. (1955). Two models of group behavior in the solution of eureka-type problems. <u>Psychometrika, 20</u> (2), 139-148.

Mingers, J. (1991). The content of MSc operational research courses: results of a questionnaire of OR groups. <u>Journal of Operational Research Society, 42</u> (5), 375-385.

Morgan, K., Morris, R. L., & Gibbs, S. (1991). When does a mouse become a rat? Or…Comparing the performance and preferences in direct manipulation and command line environment. <u>Computer Journal, 34</u> 105-125.

Nardi, B. A. (1995). <u>A Small Matter of Programming.</u> Cambridge: The MIT Press.

Niederman, F., Brancheau, J., & Wetherbe, J. (1991). Information systems management issues for the 1990s. <u>MIS Quarterly, 15</u> (4), 475-495.

Norusis, M. J. (1997). <u>SPSS Guide to Data Analysis.</u> Upper River Saddle: Prentice-Hall.

Oliver, R. (1994). Proof-reading on paper and screens: The influence of practice and experience on performance. <u>Journal of Computer-Based Instruction, 20</u> (4), 118-124.

Panko, R. R., & Sprague, R. H., Jr. (1997). Hitting the wall: Errors in developing and code inspecting a "simple" spreadsheet model. <u>Decision Sciences Journal,</u> forthcoming.

Pierson, J. K., Forcht, K. A., & Teer, F. P. (1990). Determining documentation requirements for user-developed applications. <u>Information and Management, 19</u> 21-31.

Pincus, K. (1989). The efficacy of a red flags questionnaire for assessing the possibility of fraud. <u>Accounting, Organizations and Society, 14</u> 153-163.

Rich, F. (1998, March 21). The liars' club. <u>The New York Times.</u>   A15.

Ronen, B., Palley, M. A., & Lucas, H. C., Jr. (1989). Spreadsheet analysis and design. <u>Communications of the ACM, 32</u> 84-93.

Savitz, E. J. (1994). Magellan loses its compass. <u>Barron's, 84</u> (50), 35.

Schulman, R. S. (1992). <u>Statistics in Plain English with Computer Applications.</u> Great Britain: Chapman & Hall.

Schultheis, R., & Sumner, M. (1994). The relationship of application risks to application controls: A study of microcomputer-based spreadsheet applications. <u>Journal of End User Computing, 6</u> (2), 1-18.

Sharda, R., Barr, S. H., & McDonnell, J. C. (1988). Decision support system effectiveness: A review and an empirical test. <u>Management Science, 34</u> (2), 139-159.

Shneiderman, B. (1992). <u>Designing the User Interface: Strategies for Effective Human-Computer Interaction.</u> Reading Ma: Addison-Wesley.

Simkin, M. G. (1987). Micros in accounting: How to validate spreadsheets. <u>Journal of Accountancy, 164</u> (5), 130-138.

Skrzycki, C. (1998, March 18). Stock recipe needs a grain of salt; Beardstown Ladies miscalculated investment returns. <u>The Washington Post,</u> A01.

Stone, D. N., & Black, R. L. (1989). Building structured spreadsheets. <u>Journal of Accountancy,</u> 131-142.

Sumner, M., & Schultheis, R. (1990). Managing the risks of user-developed database applications. <u>Journal of Manufacturing, 1</u> (4), 30-38.

Swain, J. W., White, J. D., & Hubbert, E. D. (1995). Issues in public management information systems. <u>American Review of Public Administration, 25</u> (3), 279-296.

Te'eni, D. (1990). Direct Manipulation as a source of cognitive feedback: a human-computer experiment with a judgement task. <u>International Journal of Man-Machine Studies, 33</u> 453-566.

Waller T. C., & Gallun, R. A. (1985). Microcomputer competency requirements in the accounting industry: A pilot study. <u>Journal of Accounting Education, 3</u> (2), 31-40.

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. <u>Communications of the ACM, 39</u> (11), 86-95.

Wang, R. Y., Storey, V. C., & Firth, C. P. (1995). A framework for analysis of data quality research. <u>IEEE Transactions on Knowledge and Data Engineering, 7</u> (4), 623-640.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. <u>Journal of Management Information Systems, 12</u> (4), 5-33.

Watt, P. (1985). The power - and pitfalls - of electronic spreadsheets. <u>Government Finance Review, 1</u> (3), 32-33.

Wiles, R. (1998, March 32). Their math shouldn't be problem. <u>Los Angeles Times,</u> D11.

Zhao, J. J. (1997). Computer end-user skills needed by business professionals now and toward 2000. <u>Journal of Computer Information Systems, 37</u> (4), 24-29.

APPENDIX A

PRACTITIONER LITERATURE

Summary of Techniques Mentioned

		Technique Mentioned by Aid		
Authors	Number of Techniques Mentioned	Design Aid	Formula Complexity	Testing & Debugging
AICPA	15	X	X	X
Alexander, R. A.	7	X	X	X
Amoroso, D.	11	X	X	X
Anderson, K., & Bernard, A.	12	X	X	X
Bissell, J. L.	10	X		X
Bromley, R. G.	16	X	X	X
Carlberg, C.	1			X
Chan, W.	10	X	X	X
Dhebar, A.	5	X	X	X
Edge, W. R., & Wilson, E. J. G.	13	X		X
Fleenor, W. C., & Crain, J. L.	14	X	X	X
Floyd, B. D., Walls, J., & Marr, K.	3	X		X
Freeman, D.	8	X	X	X
Grupe, F.	7	X	X	X
Grushcow, J.	5		X	X
Hassinen, K.	5	X	X	X
Hassinen K., Sajaniemi, J., & Vaisanen, J.	3	X	X	
Hayden, R. L., & Peters, R. M.	10	X	X	X
Jones, J. M.	1	X		
Kee, R. C., & Mason, Jr, J. O.	9	X		X
Kiely, T.	4	X		X

71

Krull, A.	1			
Marcella, A., Jr.	6	X		X
Marchand, M. G.	1			X
Miller, S. E.	7	X	X	X
Pearson, R.	4			X
Pratt, M. J., & Coy, D.	10	X	X	X
Ronen, B., Palley, M. A., & Lucas, Jr. H. C.	7	X	X	X
Savage, H. M.	7	X	X	X
Savage, S.	3	X		X
Schultheis, R., & Sumner, M.	5	X		X
Simkin, M. G.	14	X		X
Stang, D.	14	X	X	
Watt, P.	8	X	X	X
Wittig, G. R.	1	X		
Yoon, Y.	<u>5</u>	<u>X</u>	<u>X</u>	
Total Techniques	262	31	21	30

Practitioner Literature References

AICPA (1993). Assisting clients in developing policies and procedures for electronic spreadsheet applications. American Institute of Certified Public Accountants.

Alexander, R. A. (1996). Teaching good systems design for spreadsheet projects. Journal of Accounting Education, 14 (1), 113-122.

Amoroso, D. (1994). Decision making using lotus 1-2-3 for windows. New York: McGraw-Hill, Inc..

Anderson, K., & Bernard, A. (1988). Micros in accounting: Spreading mistakes. Journal of Accounting and EDP, 3 (4), 42-45.

Bissell, J. L. (1986). Spreadsheet planning and design. Journal of Accountancy, 161 (5), 110-120.

Bromley, R. G. (1985). Template design and review: How to prevent spreadsheet disasters. Journal of Accountancy, 160 (6), 134-142.

Carlberg, C. (1996). Spreadsheet quirk keeps PC users going in circles. Denver Business Journal, 48 (14), 2B.

Chan, W. (1987). Sorting out spreadsheets. Australian Accountant, 57 (11), 52-56.

Dhebar, A. (1993). Managing the quality of quantitative analysis. Sloan Management, 69-75.

Edge, W. R., & Wilson E. J. G. (1990). Avoiding the hazards of microcomputer spreadsheets. Internal Auditor, 35-39.

Fleenor, W. C., & Crain, J. L. (1989). Standardizing spreadsheet designs. CPA Journal, 59 (10), 81-84.

Floyd, B. D., Walls, J., & Marr, K. (1995). Managing spreadsheet model development. Journal of Systems Management, 46 38-43, 68.

Freeman, D. (1996). How to make spreadsheets error-proof. Journal of Accountancy, 75-77.

Grupe, F. (1985). Tips for better worksheet documentation. Lotus, 68-70.

Grushcow, J. (1985). Avoid these common spreadsheet errors. Lotus, 59-62.

Hassinen, K. (1988). An experimental study of spreadsheet errors made by novice spreadsheet users. Department of Computer Science, University of Joensuu, PO Box 111, SF-80101 Joensuu, Finland.

Hassinen K., Sajaniemi, J., & Vaisanen, J. (1988). Structured spreadsheet calculation. Proceedings of the IEEE Workshop on Languages for Automation, 129-133.

Hayden, R. L., & Peters, R. M (1989). How to ensure spreadsheet integrity. Management Accounting, 30-33.

Jones, J. M. (1986). Decision analysis using spreadsheets. European Journal of Operational Research, 26 385-400.

Kee, R. C., & Mason, Jr, J. O. (1988). Preventing errors in spreadsheets. Internal Auditor, 45 (1), 42-47.

Kiely, T. (1996). Spreadsheet risk: How and why to build a better spreadsheet. Harvard Business Review, 10-12.

Krull, A. (1989). Computer rip-offs and foul-ups: Is management to blame? EDPACS, 16 (11), 10-14.

Marcella, A., Jr. (1989). Application reviews of end-user spreadsheet design. EDPACS, 17 10-14.

Marchand, M. G. (1995). Successful spreadsheets. Supervisory Management, 40 (1), 3.

Miller, S. E. (1989). 8 ways to avoid worksheet errors. Lotus, 5 (2), 50-53.

Pearson, R. (1988). Lies, damned lies, and spreadsheets. Byte, 13 (13), 299-304.

Pratt, M. J., & Coy, D. (1988). 'TIMADS' to avoid spreadsheet model grief. Accountant's Journal, 67 (10), 34-36.

Ronen, B., Palley, M. A., & Lucas, H. C., Jr. (1989). Spreadsheet analysis and design. Communications of the ACM, 32 84-93.

Savage, H. M. (1991). Document that spreadsheet! Management Accounting, 52-57.

Savage, S. (1997). Weighing the pros and cons of decision technology in spreadsheets. ORMS Today Online Edition, 24 (1).

Schultheis, R., & Sumner, M. (1994). The relationship of application risks to application controls: A study of microcomputer-based spreadsheet applications. Journal of End User Computing, 6 (2), 1-18.

Simkin, M. G. (1987). Micros in accounting: How to validate spreadsheets. Journal of Accountancy, 164 (5), 130-138.

Stang, D. (1987). Spreadsheet disasters I have known. Information Center, 3 (11), 26-30.

Watt, P. (1985). The power - and pitfalls - of electronic spreadsheets. <u>Government Finance Review, 1</u> (3), 32-33.

Wittig, G. R. (1990). Spreadsheet design: Role of introductory documentation in creating spreadsheets. <u>Library Hi Tech News, 8</u> (2), 61-67.

Yoon, Y. (1995). Applications-oriented spreadsheet instruction in financial model building: An extension. <u>Financial Practice & Education,</u> 143-147.

APPENDIX B

THE AIDS

Each group was trained on the aid assigned to the group. Each aid listed four techniques. For each technique a page with instructions, examples or screen shots provided detailed explanations to help during spreadsheet creation.

<u>Planning and Design Aid</u>

        The items identified and described here are intended to help you create a well designed spreadsheet. This first sheet is the aid, however, additional instructions and explanations are attached.   Please read through the items carefully and use those items that you feel will help you create an error-free spreadsheet.

I. Plan your spreadsheet.	
	Use a planned layout so that movement (copy, cutting, moving or dragging) of the data after entering it will be minimized.
II. Use an organized layout to isolate data and computation areas.	
	Use some type of layout, such as, a stair-step format or modular layout on different worksheets to isolate data and computation areas so that row or column insertions and deletions will not create errors.
III. Use descriptive labels.	
	If the original data or assumptions change so that the user will know what each data item is, use labels that clearly indicate the input data or assumptions to be changed.
IV. Repeat the input data (assumptions) near the output (analysis).	
	Using a formula, repeat the input data or assumptions near the area that you will use for computations so that it is more efficient/effective to reference the correct cells.

I. Plan your spreadsheets:

Plan the spreadsheet layout so that movement of the data after entering will be minimized. Instead of just starting to enter data and formulas, take a few minutes to plan the spreadsheet layout. Using separate areas for the input or assumptions and output or analysis will reduce the possibility of accidentally overwriting one of the formulas.

Determine where the input area should be and what information will be necessary. For example, if you were creating a spreadsheet model to determine monthly payments, what data or assumptions would you need. Some data or assumptions that would be necessary are the amount borrowed, the interest rate, and the length of the loan. In what ways would you utilize the information to calculate the monthly payments. For example, how should the user enter the term: months or years. If the users look at the previous loan calculated using the model, it may not help them determine if the number is in months or years. For example, if the term of the loan was thirty does that mean thirty years or thirty months; if the interest rate is 2% is that annual, semi-annual, monthly, or bi-monthly. In planning the worksheet there should be enough room to tell the user the proper format necessary for correct calculations.

During this planning stage, consider specifically what information should be provided to the user and how the model should impart it. Continuing with our example, should the model give the monthly payments, total payments or maybe compare monthly and yearly payments. The correct number of rows and columns would need to be allowed to display the information, and allow for eventual printing. In addition, enough space to allow headings and footers on each page or screen should be planned. Each page or screen should contain a reasonable amount of information.

Additional considerations necessary depend on the individual spreadsheet model. A model calculating monthly or yearly loan payments would have very different needs than one that tracks a golf handicap. While both still requiring planning, the required input and output would be very different. Planning the spreadsheet helps eliminate movement of data by reducing copying and pasting, inserting and deleting which often creates errors by changing the locations of cells.

II. Use an organized layout to isolate data and computation areas:

Isolating the input or assumption area from the computation or analysis area is a good practice and will help reduce errors. If the spreadsheet model has any intermediate steps not needed for the output, they should also be put in a separate area. Keeping the input data or assumptions together in one area in the model will facilitate changes in these values. It will also explicitly show all the assumptions to the users so they will not have to hunt for specific ones. In addition, it will serve as a visible reminder of all the assumptions in case the user may overlook one that needs to be changed. Furthermore, inserting and deleting rows and columns will not destroy an organized layout. Displaying all the formulas in one area will help decrease the number of errors because the chance of accidentally overwriting a formula will be reduced.

Data can be isolated in a number of different ways. Two of the easiest ways are the stair-step method and the modular layout on different worksheets in the same workbook. The stair-step method contains no areas that have overlapping rows or columns. The modular layout consists of a different worksheet for each area (To add worksheets using the menu bar: Insert, Worksheet. To move a worksheet: point to its tab, click and hold down the mouse button, then drag it to the new location among the other worksheet tabs.) Below are examples of isolating the areas of the spreadsheet.

Stair-Step Method

Input area		
	Formulas & calculations	
		Print or display area
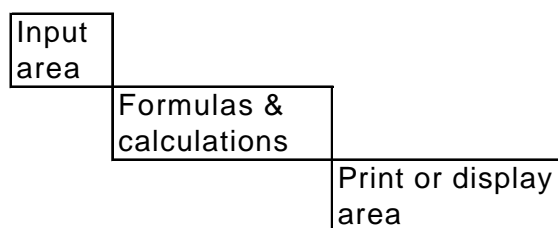
Modular Layout on Different Worksheets
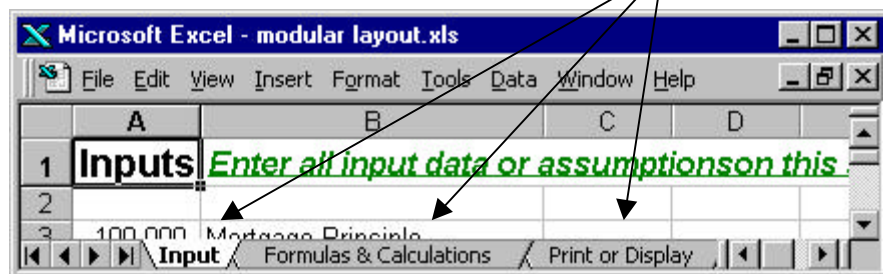
III. Use descriptive labels:

Use labels that clearly indicate all values and formulas whether input, computations, or analysis. The input data or assumptions should be clearly labeled so that the original data or assumptions can easily be changed if necessary. The output or printed data should also be clearly labeled; so the person analyzing the spreadsheet will understand all the information provided.
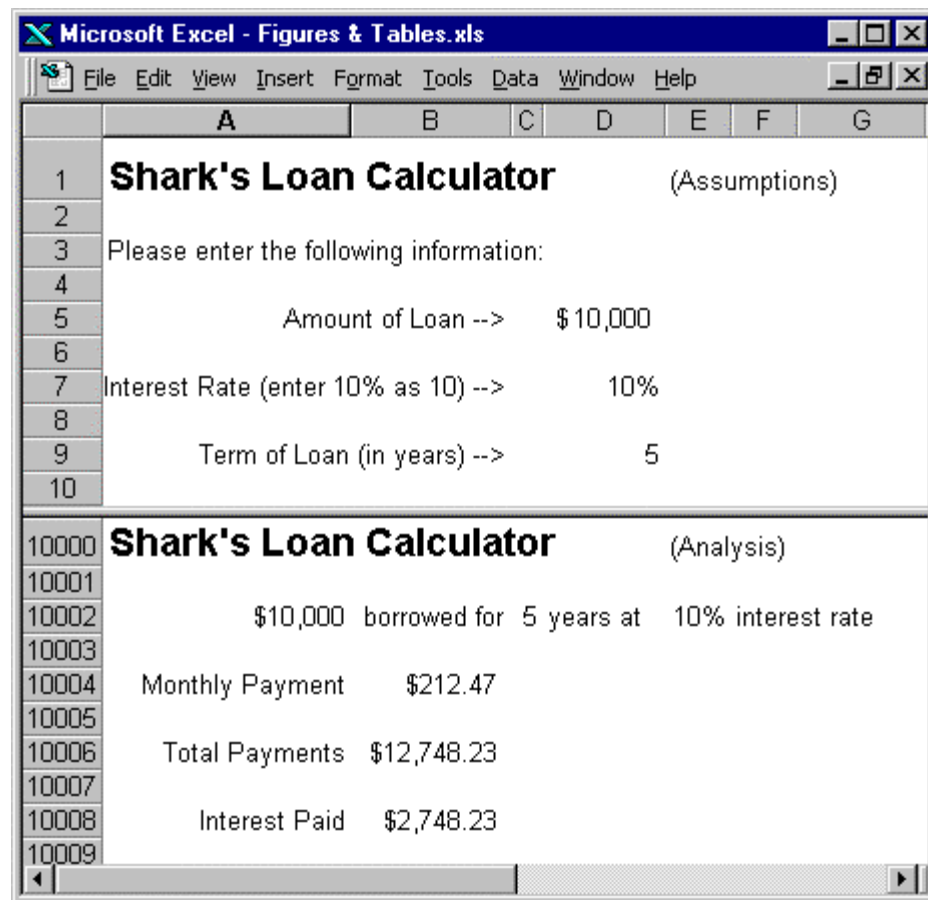
Note in the example below that the values of 18% and $22,000 in rows fifteen and sixteen and the table starting in row eighteen are not labeled making it very difficult to determine their purpose.

Microsoft Excel - Figures & Tables.xls

File Edit View Insert Format Tools Data Window Help

	A	B	C	D	E	F
1	**Bubbles Scuba Shop**					
2	Income Statement					
3	For the Year Ended 12/31/98					
4						
5						
6	Sales				$ 18,000	
7	Cost of Goods Sold				12,600	
8	Income (Loss) Before Expenses				$ 5,400	
9						
10	Expenses Excluding Interest				$ 3,400	
11	Interest				3,960	
12						
13	Net Income (Loss)				$ (1,960)	
14						
15	18%					
16	$ 22,000					
17						
18		16%	14%	12%	10%	
19	$ 22,000	-1520	-1080	-640	-200	
20	$ 20,000	-1200	-800	-400	0	
21	$ 18,000	-880	-520	-160	200	
22	$ 16,000	-560	-240	80	400	
23	$ 14,000	-240	40	320	600	
24	$ 12,000	80	320	560	800	
25						

IV. Repeat the input (assumptions) data near the formulas:

Repeat the input data or assumptions in the analysis or printed region of the spreadsheet to allows the spreadsheet user to see which assumptions were used and to more easily understand the worksheet. This is important whether the user is looking at the computer screen or using a printed copy and is especially important if there are several printouts using different assumptions. Note in the example below if the analysis area were printed, the assumptions would appear so that anyone looking at the worksheet would see that amount of the loan, the interest, and the term.

Formula Complexity Aid

       The items identified and described here are intended to help you create formulas that are less complex and therefore less error prone. This first sheet is the aid, however, additional instructions and explanations are attached.   Please read through the items carefully and use those items that you feel will help you create an error-free spreadsheet.

I. Formulas should contain only cell references.	
	Use only cell references in formulas so that any number can be changed without having to redo any formulas.
II. Formulas can be used to repeat data in other locations.	
	Use formulas to repeat data from the input area to the formula area so that it is easier to use the correct cell reference/data in the formula.
III. Split complex formulas.	
	Use multiple formulas as intermediate steps for any complicated calculations to make formulas less complex.
IV. Relative and absolute cell addressing should be used where appropriate	
	Use relative cell addresses when you want the cell addresses in the copied formula to re-adjust, and use absolute cell addresses when you do not want the referenced cell to change.
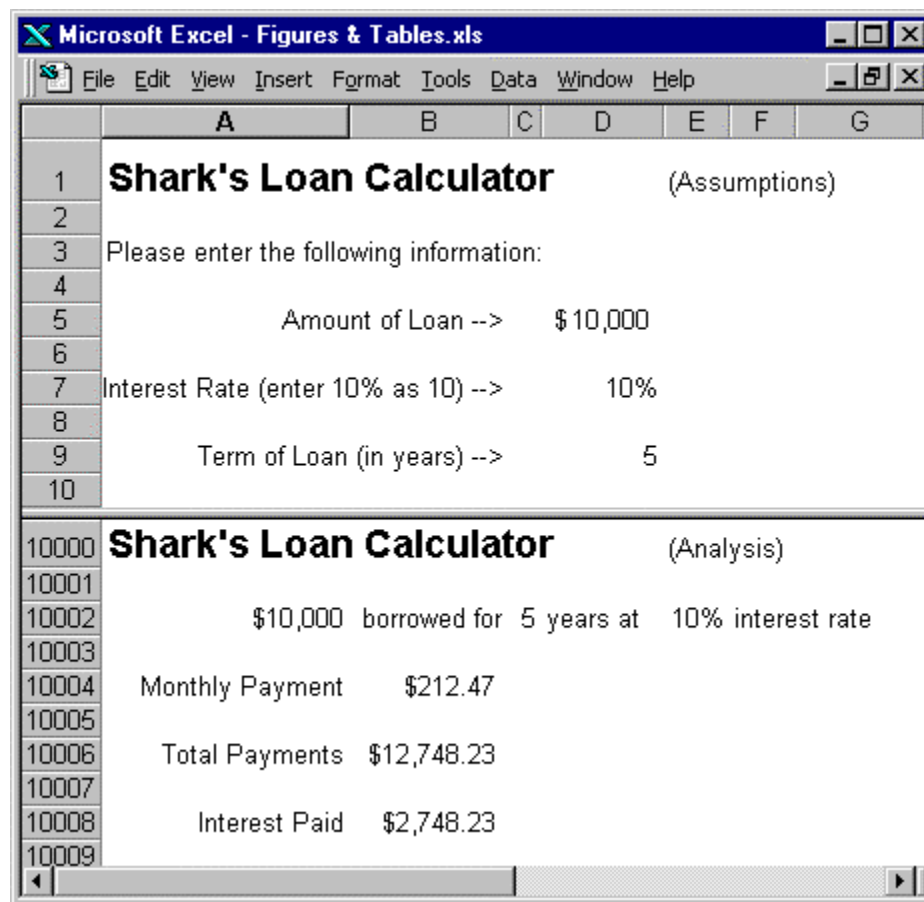
I. Formulas should contain only cell references:

Using only cell references allows changes to any of the assumptions in your spreadsheet model. For example, even if your model includes an inflation factor or a rate of growth you may not expect to change in the future, use these cell references. Include these numbers in your input or assumptions for the model because they may change in the future and it is much easier and less error-prone to include cell references in the formulas than to start trying to locate and change all the places that specific values were used instead of cell references. This means that you can change the number in one place and the formulas will update all locations that use that value. Errors can be created by missing all the locations of the original number or by incorrectly replacing the original number in several locations.

Formulas that contain only cell references simplify trying 'what-if' scenarios. For example, including the number as cell reference allows you to do analysis by varying the inflation rate without having to search for every place the number is used. This is true with any number in the spreadsheet model.

II. Formulas can be used to repeat data in other locations:

Repeating the data using formulas makes it easier to use the correct cell references in the formula. Numbers that are close to the area in which they are used require less searching to find the correct cell to reference. In the example below, note that the data are entered into cells D5, D7, and D9. Using formulas the data are repeated in cells A10002 (formula: =D5), C10002 (formula: =D9), and E10002 (formula: =D7). The formulas use these latter cells to calculate monthly payment, total payment, and interest paid. Having the values repeated close to the formulas requires less searching for the correct cell reference. This would also hold true if there were several worksheets, one of which the user entered data, and then using formulas to repeat that data on the worksheet with the formulas.

	A	B	C	D	E	F	G
1	**Shark's Loan Calculator**				(Assumptions)		
2							
3	Please enter the following information:						
4							
5	Amount of Loan -->			$10,000			
6							
7	Interest Rate (enter 10% as 10) -->			10%			
8							
9	Term of Loan (in years) -->			5			
10							
10000	**Shark's Loan Calculator**				(Analysis)		
10001							
10002	$10,000 borrowed for 5 years at				10% interest rate		
10003							
10004	Monthly Payment	$212.47					
10005							
10006	Total Payments	$12,748.23					
10007							
10008	Interest Paid	$2,748.23					
10009							

III. Split complex formulas:

Splitting a long, complex formula into smaller parts will increase the accuracy of a spreadsheet model. For example, a formula that adds and subtracts twenty different cells could be split into three simpler formulas. In the example below, the formula for the left side (=A2+A3-A4-A5-A6+A7-A8+A9+A10-A11-A12+A13-A14+A15+A16+A17-A18+A19-A20-A21) is long and complicated. One of the formulas could total the cells to be added and the other formula could total the cells to be subtracted; then the final formula would just subtract the second formula from the first. The columns should be created as in Columns E & F so that you just add both columns; Column E, =SUM(E2:E21) and F, =SUM(F2:F21), then the third formula would take the difference =E22-F22. The three formulas are less complicated and much less prone to error than the formula in Column A.

	A	B	C	E	F	G	H
1	**Number**	**Sold or Returned**		**Number Sold**	**Number Returned**		
2	1	Sold		1			
3	8	Sold		8			
4	3	Returned			3		
5	7	Returned			7		
6	3	Returned			3		
7	6	Sold		6			
8	10	Returned			10		
9	7	Sold		7			
10	10	Sold		10			
11	2	Returned			2		
12	2	Returned			2		
13	1	Sold		1			
14	8	Returned			8		
15	1	Sold		1			
16	4	Sold		4			
17	8	Sold		8			
18	2	Returned			2		
19	8	Sold		8			
20	3	Returned			3		
21	9	Returned			9		
22	5	Total Sold or (Returned)		54	49	Subtotals	
23							
24					5	Total Sold or (Returned)	
25							

IV. Relative and absolute cell addressing:

Use relative and absolute cell addressing to aid spreadsheet creation. Relative cell addresses readjust in relation to the cell containing the formula. Copying a correct formula instead of re-entering each and every formula will reduce the number of spreadsheet errors. The formula =A1, entered in cell B1, actually means "the contents of this cell (B1) are equal to the contents of one cell to the left." When the contents of cell B1 are copied it will always mean that it is equal to the contents of one cell to the left automatically. For example, if the formula in cell B1 is copied to Z100, the formula will change to reference the value in cell Y99. However, if you cut and paste a cell with relative references, the references will not change.

Absolute cell references are denoted by dollar signs ($) in front of the cell reference. For example, A1 is an absolute reference to cell A1. An absolute reference does not change when copied to another cell. If formula =A1 is copied to Z100, it will still reference cell A1. A mixed reference contains one absolute coordinate and one relative coordinate. Examples of mixed references are $A1 or A$1. The first reference, $A1 as a formula (=$A1) in cell B1 means that column A will be constant no matter where it is copied. For example, if it is copied to Z1, the value in Z1 will be the same as A1 because the column does not change due to the absolute reference. However, if it is copied to B100, the reference will now point to A100. Note that in both cases the column of A remains the same. The second reference, A$1 as a formula (=A$1) in cell B1 means that row 1 will be constant no matter where it is copied. For example, if it is copied to B100 the reference will now point to A1 because the row does not change due to the absolute reference. However, if it is copied to Z1, the value in Z1 will be the same as Y1. Note that in both cases the row of 1 remains the same.

To change a cell reference to absolute, the F4 acts as a toggle key. The first press creates absolute (A1), the second press makes the row absolute (A$3), the third press makes the column absolute ($A3), and the fourth press changes the cell reference to relative. There are several ways to create an absolute reference. After pointing or typing in the cell address (prior to hitting enter), tap the F4 key as needed. After creating the formula, in the edit mode, position the cursor in any part of the cell reference that is desired to change, and tap the F4 key as needed. Correct use of relative and absolute references can reduce errors copied formulas.

Testing and Debugging Aid

The items identified and described here are intended to help you test and debug your completed spreadsheet. This first sheet is the aid, however, additional instructions and explanations are attached. Please read through the items carefully and use those items that you feel will help you create an error-free spreadsheet.

|		
I. Use Excel's auditing tools to verify formulas.		
	Use Excel's auditing tool (Tools, Auditing, Show Auditing Toolbar) to trace all dependent and precedent cells to determine that the correct cell locations are used in formulas.	
II. Review worksheet for error messages.		
	Look for error messages and correct, such as:	
		##### - indicates cell is to narrow
		#DIV/0! - indicates a division by 0
		#VALUE! - indicates an incorrect type of argument in a function
		CIRC - indicates a reference to itself
III. Determine if the numbers look reasonable.		
	Use your understanding of the problem to decide if the results are reasonable (e.g. income is less than expenses, there should be a loss).	
IV. Verify all calculations with a calculator or test with known models.		
	Use a calculator to confirm that the formulas are correct or confirm the model by using real or historical numbers.	

I. Use Excel's auditing tools to verify formulas:

The auditing tool (to activate choose Tools, Auditing, Show Auditing Toolbar) can trace dependent and precedent cells to determine if the correct values/cell references are used in each formula.

Instructions
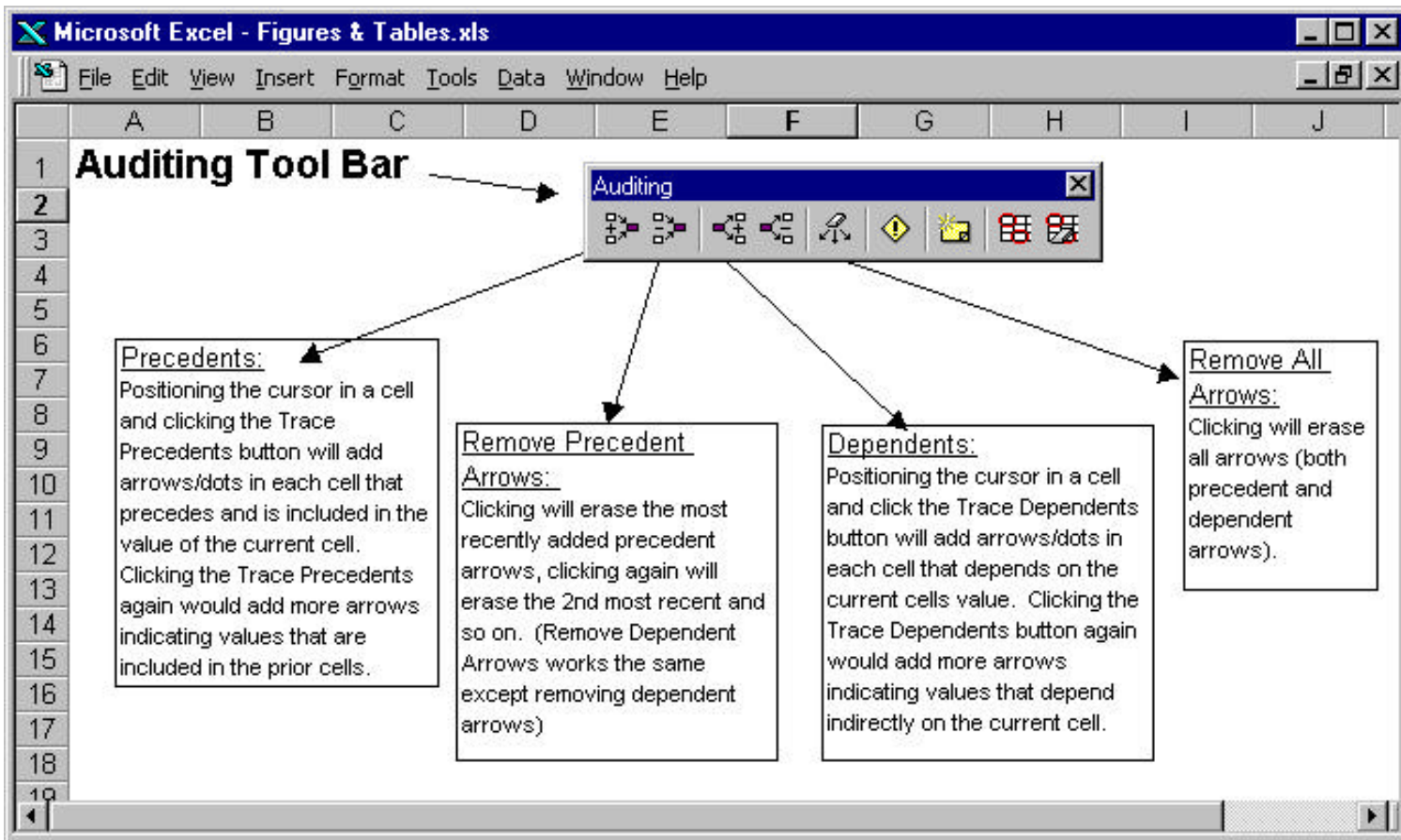
Example
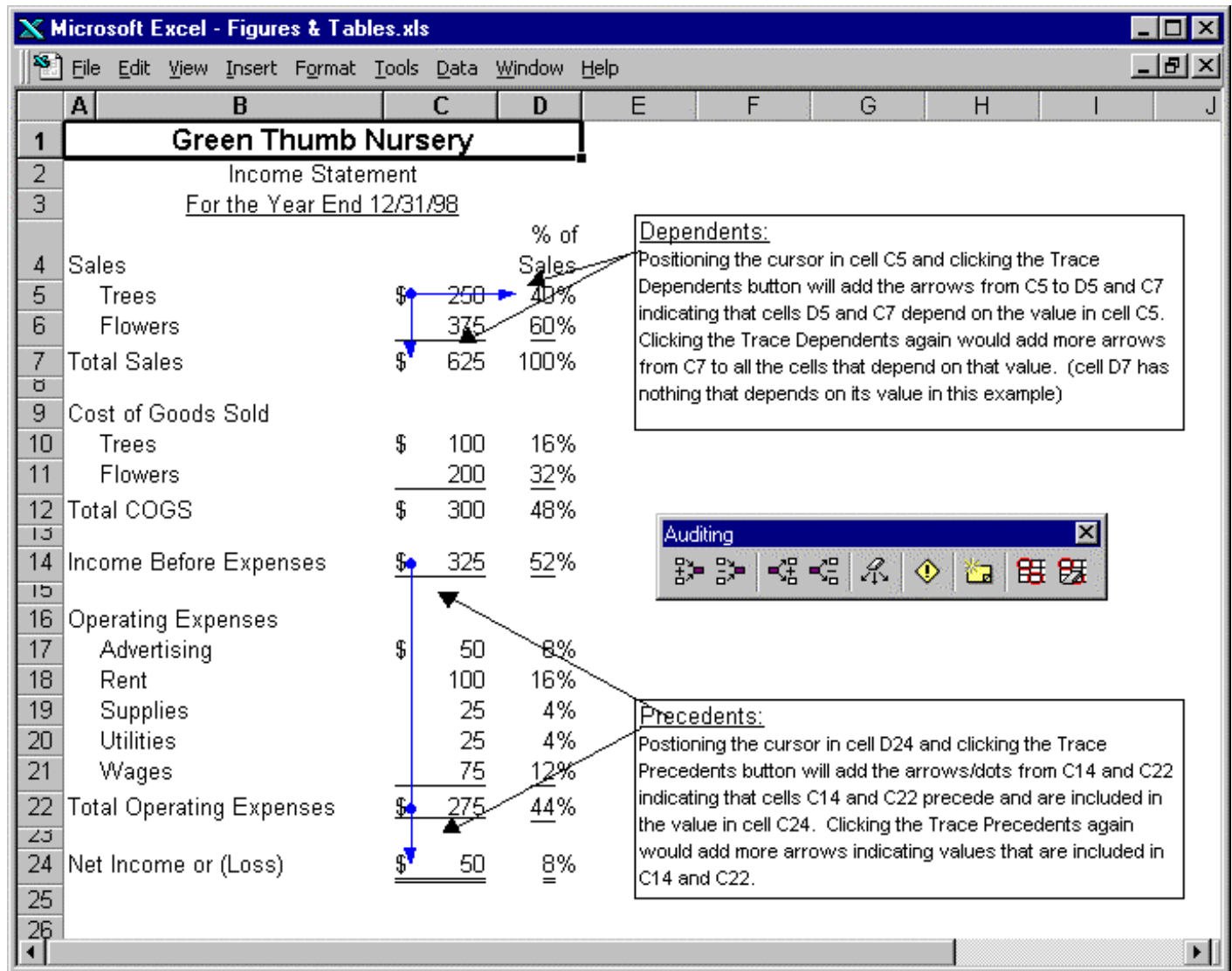
II. Review worksheet for error messages:

Look for error messages in your worksheet. The following items are the most common errors, with an example and appropriate corrections for each. In most cases the Auditing Tool may be of considerable assistance in locating the incorrect references.

Microsoft Excel - Figures & Tables.xls

File Edit View Insert Format Tools Data Window Help

	A	B	C	D	E
1	Values used for the examples below:				
2	0	3			
3	1	4			
4	2	5			
5					
6	Error Message	Example cell references	Problem	Solution	
7	###################	=B3*B2*1000000000	Cell is too narrow	Widen the column	
8	#DIV/0!	=A4/A2	Dividing by zero	Correct the formula	
9	#VALUE!	=B4-A2:A4	Incorrect argument (e.g. value or cell reference): use of range (A2:A4)in formula	Correct the formula (=B5-(A2+A3+A4))	
10	"Circular A11" displayed on Status Bar		Circular reference - refers to its own cell, either directly or indirectly	Cell with circular reference is indicated on the status bar and the formula should be corrected	
11	0	=A3+A4+A11			
12					
13					
14					

Ready   Circular: A11

90

III. Determine if the numbers look reasonable:

Using your knowledge of the problem, determine if the values look reasonable. In the financial statement below, notice that the income statement has comparable numbers except for "Income Before Expense" and "Net Income or (Loss)." If the intervening numbers are comparable, then the final results should also be comparable. Upon closer inspection of the example, it appears that in 1998 the sales and cost of goods sold were added rather than subtracted.

Bubbles Scuba Shop
Income Statement
For the Years Ended 12/31/98 & 12/31/97

	12/31/98	12/31/97
Sales		
Equipment	$ 1,000	$ 875
Travel	375	400
Total Sales	$ 1,375	$ 1,275
Cost of Goods Sold		
Equipment	$ 500	$ 450
Travel	200	225
Total COGS	$ 700	$ 675
Income Before Expenses	$ 2,075	$ 600
Operating Expenses		
Advertising	$ 50	$ 50
Rent	100	100
Supplies	50	25
Utilities	25	25
Wages	75	50
Total Operating Expenses	$ 300	$ 250
Net Income or (Loss)	$ 1,775	$ 350

IV. Verify all calculations with a calculator or test with known numbers:

Use the following two techniques to verify spreadsheet models and will also decrease the number of errors found in spreadsheet models.

Verify calculations with a calculator. All calculations should be redone from scratch, not recomputed from the same formulas in the worksheet. The spreadsheet has calculated exactly what has been entered, but this step is to look for formulas that are mis-specified. For example, the parentheses used in the spreadsheet may be incorrect, but when the formulas are verified it may be determined that they are incorrect. For example, an equation to determine a 10% growth of the $500 sales is written in the spreadsheet as $(500 + 500)* 110\%$, but when it is recalculated it is found that the number should be 550, not 1100.

Test the spreadsheet model with known numbers. If available, use real or historical numbers to determine if the spreadsheet model is correct. For example, the model created is for the next year budget of a new division: use numbers from a current division to ensure that all the formulas are correct.

APPENDIX C

EXPERIMENT SCRIPT

Thank you for agreeing to participate – please stop me if you have any questions. First, a few preliminaries: compensation will be two lottery tickets OR extra credit if your instructor decided to offer extra credit. There will also be a drawing based on performance which will become clear in a few minutes. Each group will have approximately twenty people and those that create spreadsheets based on the best performance will be entered into the drawing. The winner will receive an additional $25 in cash.   Second, everything you do today will be anonymous. I will not match names to spreadsheets, nor will your professor know of any individual results from this study. The only time I will match a name with a social security number, is for the four or five winners of the additional $25 cash, so I won't even know who's model I am analyzing.

Let me give you an overview of the things we will do today. First, I will pass out and have you sign a consent form. Then I will give you a short session on the background of spreadsheets and the problem I am trying to solve. Then you will actually create a spreadsheet from a word problem. After that there is a demographics and post-test questionnaire. I appreciate your participation today.

[Handout the consent form]

Please sign the consent form. Please note that in the consent form it says that you cannot discuss this study until after April 30. If you have any questions about the consent form, please ask. If you have any specific questions about the study, I will be more than happy to answer them, but please do not discuss this study with anyone else until the end of the month.

Background information on spreadsheets:

Spreadsheets have only been around for about twenty years and several studies have shown that spreadsheets are a well-used skill.

For example, in 1985 a study of the accountants in the then "Big Eight," local and national accounting firms, as well as those in industrial and financial firms wanted new hires to have an in-depth spreadsheet knowledge more than any other microcomputer application.

In 1988, a study found that practicing accountants rank spreadsheets in the top two categories out of fifty-nine.

In 1991, operational research groups ranked spreadsheets in top 20% of skills necessary for analysts. The other skills were directly related to their work requirements.

And the most recent study was in 1997. This study demonstrated that 83 new Fortune 500 corporations considered spreadsheets the most important application software to understand now and in the year 2000. Furthermore, it was the only software that increased in popularity.

So because spreadsheets have been found to be so important, I am conducting a study about how spreadsheets are created.

[**Training**: Not applicable for control group. NOTE: read directly from aid and show appropriate examples on computer overhead.]

Now for the spreadsheet model you are asked to create: Before you begin, I want to show you the *Bidder* worksheet. When you open the worksheet, a message box will open [Open example]. Please enter your social security number, which I use only to match up your questionnaire with your worksheet. After pressing OK, a screen will appear. Please read the problem statement which I will be handing out. The problem statement asks you to create a reusable model to make bids on a kitchen using two

different prices for materials. When you are ready to create your model, click the button to move to a blank sheet (click to show). Click this button only when you are ready to start entering your solution.

Any questions?   [Handout the diskettes, handout task individually when each student has the Bidder.xls open.]

Please open the spreadsheet on your diskette. Good luck.

APPENDIX D

THE TASK

"Bidder"

Please open the file "Bidder" contained on your disk before proceeding. Put your social security in the text box that appears when you open "Bidder."

You are starting a new business to remodel kitchens for residential houses. Your task is to build a spreadsheet template model that will calculate two bids for your customers. You should determine the cost so that the customer can make modifications and you can easily determine a new price. (Plus, you will want to reuse this model for all your customer bids.) Base your model on the following information. (You may use a calculator or the back of any of these sheets for scratch paper.) Your performance will not be graded on formatting, only on logic.

Your current customers would like to remodel their kitchen. They are looking at 2 different options and would like a bid for both. Your model should include bids for "Kitchen A" and "Kitchen B." Any donations or sale of leftover material will reduce the cost of that room. Your bid must add a profit margin of 30% to your net expected cost (expected total cost less any reduction for donated or sold items).

The kitchen is (21 X 15), large enough to include a breakfast nook. The flooring for either plan will cost $16.50 per square foot.

The kitchen will require chair molding which will cost $5.35 per linear foot for Kitchen A and $10.45 per linear foot for Kitchen B. You will have 11 feet leftover (due to the way it will need to be cut) which you will sell to another builder for ½ the cost.

The cabinets will cost $47.50 per cabinet for Kitchen A and $65.00 per cabinet for Kitchen B and each kitchen requires 15 cabinets.

Both Kitchens require 25 boxes of tile for the counter. Kitchen A tile will cost $12.75/box and Kitchen B tile will cost $9.50/box.

 The refrigerator will cost $1,000 and require a door to match the cabinets for Kitchen A or Kitchen B. This means that you can sell the door at the flea market for $50, however

you need to purchase an <u>additional</u> cabinet for the matching door.   But you can sell the additional cabinet backing (everything but the door that you will need for the refrigerator) for ½ of the cost to one of your workmen.

The stove/oven combination will cost $650 and comes with an exhaust hood. Kitchen B will use a gas stove at the same price. However, a gas line will be required and the materials will cost $285. You will put in a microwave over the stove and can donate the exhaust hood to the Habitat for Humanity house for $25. The oven does not come with racks so you will need to purchase 2, however, they come in sets of 4 for $100 so you will sell the extra racks for $35 to one of your workmen.

The microwave will cost $250 for Kitchen A and $355 for Kitchen B. Kitchen B will require a special lighting and fan fixture that will cost $250.

Your crew of 5 will need 4 days to complete the Kitchen A but 5 days to complete Kitchen B. Each person of your crew will be paid $10/hour for 8 hours/day plus 20% additional for benefits.

You will remove all old appliances and sell them for $295.

Your performance will be graded on how well you use your assigned aid, not on formatting. [Control group: Your performance will be graded on how well you create the spreadsheet model]

Please be sure to save your spreadsheet model. The file name should be "Bidder.xls."

APPENDIX E

INSTRUCTIONS TO THE EXPERTS

Thank you for agreeing to help me with my dissertation. After we complete this exercise I will be happy to tell you all about my experiment.

Step 1 -

These spreadsheets were created by the study's participants. Their instructions: You are starting a new business to remodel kitchens for residential houses. Your task is to build a **reusable** spreadsheet template model that will calculate two bids for your customers. You should determine the cost so that the customer can make modifications and you can easily determine a new price. (Plus you will want to reuse this model for all your customer bids.) Base your model on the following information. [ask if they want to read the full task]

Please do not worry if the bids are correct.

Please use your expertise to rank or rate the spreadsheets into 4 groups from good to bad. Each spreadsheet is printed on one sheet, if possible. If the workbook was created on more than one worksheet, they are stapled together. If the spreadsheet area is too large to show on one piece of paper, the top page has a miniature overview and each area is

printed and attached. Any cell coloring is shown as shades of gray. Any font coloring is

shown as different fonts. Please use your best judgement to group them into 4 categories.

(NOTE: page break here is intentional)

Step 2 -

Please name each of the four groups:

1. _____

2. _____

3. _____

4. _____

Please give a short definition or a few words to elaborate on the group names:

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

Thank you for your time. May I contact you if I have any additional questions?

VITA

Susan E. Kruck was born in North Carolina on April 5, 1956. She holds a Bachelor's degree in Business Administration with a double major in Accounting and Computer Information Systems with a minor in Decision Support Systems from James Madison University, Harrisonburg, Virginia. She also holds a Master of Business Administration from James Madison University. Her work experience includes consulting for a small certified public accounting firm. Her most current projects involved evaluating client systems, recommending, implementing the requested software, and training clients on new systems. These systems often require hardware additions or software installation to local area networks. Susan also worked in private industry as a corporate financial manager for a multi-corporation organization that included consulting, service and retail corporations. She is also a certified instructor of Scuba Diving and Medic First Aid.

Susan's research interests include end-user computing and education, individual differences in end-users, data quality, cognitive aspects of information systems users, WEB application to small businesses, and social responsibility and ethical issues in information technology.

She is a certified public accountant licensed by the state of Virginia. She is a member of American Accounting Association, American Institute of Certified Public Accountants, American Society for Training & Development, Association for Information Systems, Decision Science Institute, and International Association for Computer Information Systems.