GREENWICH UNIVERSITY

MSc IN FINANCE & FINANCIAL INFORMATION SYSTEMS

"SPREADSHEET ERROR DETECTION PERFORMANCE:

AN EMPIRICAL EXAMINATION IN GREECE"

By Zournatzidou Efthalia

Supervisor: Dr Dimitrios I. Maditinos

Special thanks

to my supervisor Dr. Dimitrios I. Maditinos

and my family

Index

Abstract

The personal computers era contributed to passing out advanced programming tasks from the specialized professionals to the end users. Spreadsheet models is one of the most widely used applications worldwide that can produce valuable results with minimal training and effort. The errors however that most spreadsheets include may be catastrophic and difficult to detect. In this study, we make an investigation on the influence that experience and presentation on the error finding performance by end users. To reach the desired outcome we use a sample of 123 business and finance students that were submitted to a task of finding errors in a simple free cash flow model. Findings of our research reveal that presentation of the spreadsheet is of major importance as far as the error finding performance is concerned while experience does not seem to affect students on their performance. The confined pool of the sample, the self reported experience and the necessarily simple spreadsheet (unlike spreadsheets in real world) that we used in our experiment, are the main limitations of our study.

Introduction

Spreadsheet programs are widely used to analyse and manipulate advanced and complex numerical data. One can enter numbers into a spreadsheet and perform a large variety of mathematical and economical calculations even with a quite complicated structure. The problem of eliminating errors from software in general and spreadsheets in particular has been bothering literature since the beginning of the computer era. With the advent of the personal computers in the eighties and the rapid development of end-user computing, control of software development passed out from the professionals to the end users. As spreadsheets have diffused throughout industry, evidence has accumulated that many spreadsheets contain errors (Panko and Halverson, 1996; Panko, 1998) and that errors can be costly to the organisations that use them.

Although research has suggested that errors are prevalent in spreadsheets, there is much left to be examined about the types of errors that occur, the reasons that they occur, and the ways to avoid them. Spreadsheets are used very widely and frequently for business decisions. Spreadsheet modelling is nowadays recognised as the most frequently used application in the modern industry, which is also quite easy to use with minimum training. Furthermore, spreadsheets' tabular structure and embedded formulas enable any user to start building a spreadsheet from scratch with very little experience and knowledge. Spreadsheet errors on the other hand are common both in complicated spreadsheets where the amount of the cells is usually large but also in relatively simple ones. Studies have shown that spreadsheets have high error rates (Panko, 1998; Rajalingham, D. R. Chadwick, and B. Knight, 2001).

The most important issue that the industry is alarmed for is that spreadsheet errors can potentially cost large amounts of money (Galletta *et al.*, 1993; Galletta *et al.*, 1996; Galletta *et al.*, 1997). Apart from their wide use in the industry, spreadsheets can serve as major vehicles

for educational purposes. The quite widespread adoption of spreadsheets, creates a greater need for better auditing tools. Even so, spreadsheet auditing requires too much time, effort and expertise (Panko, 1998). According to Panko (1998) only half or fewer errors can be detected using the available auditing approaches. Besides it would be unreasonable to expect users disadvantaged by their background, education, or learning style to carry out effective code inspections and spreadsheet testing with no prior experience in technicalities software systems.

Errors in spreadsheets might arise due to a variety of reasons ranging from the user's lack of the specifications or requirements of the spreadsheet understanding, to errors arising from incorrectly entering the formulas or values. In this study we employ an experiment among business and finance students of the technological institution of Kavala, Greece, to investigate if user experience and the presentation of a spreadsheet can affect the auditors' error finding performance. In this respect we measure error finding rates by everyday users (such as students) and try to identify their failure. Also we try to identify and evaluate their success categorising the factors that cause it. To reach this goal we follow the categorisation that Panko (2000) and Panko (2006) proposed, we classify the students' errors and investigate patterns that ultimately reveal causes of error making.

In the second chapter we present the literature resources that we use as our theoretic background. Specifically we mention the basic findings of several studies dealing with error classification, those which present the prevalent spreadsheet auditing and testing methods and finally those that deal with spreadsheet error performance. Especially at the late years there is a great interest of literature in examining the factors that can help towards the improvement of error finding performance. Galletta *et al.* (1996) and Panko and Sprague Jr. (1998) conducted two studies that opened the way for this thus they are used as our driver in this study. In chapter three we outline the methodology of our research and the hypotheses that we

developed to examine our research framework. Specifically after reviewing the most significant, in our opinion, methodological approaches in the determination of the factors that improve error finding performance we set the scene for our study's methodology. Then we introduce the two factors we examine in our research, spreadsheet presentation and user experience. These factors are slightly adapted from Galletta *et al.* (1996) and Panko and Sprague Jr. (1998) and are often met in literature. Finally we develop a set of hypotheses so as to statistically check if our proposed factors can affect error finding performance.

In chapter four we present the main findings of our research. In this chapter we outline all the outcomes from the previously conducted statistic analysis and discuss their meaning. Also we show whether our findings support our claims, presented in chapter three and analyse the reasons. Ultimately in chapter five we make a discussion of our findings and their practical importance for spreadsheet practitioners as well as how do we thing our study contributes to the literature. Finally we explain our main limitations and our proposals for further research in this field.

2. Literature review

2.1 Introduction

In the first chapter, we outlined the basic concepts of spreadsheet modelling as well as its use in the industry. As spreadsheets play a significant role in the modern economic activity, worldwide literature started to examine issues concerning spreadsheet modelling. A large segment of literature deals with spreadsheet errors and their influence in their purpose. Scholars and stakeholders are strongly interested in examining the causes of errors and the

most efficient ways to resolve them. In this study, we deal with this issue and examine whether spreadsheet presentation and experience affect spreadsheet errors finding performance following the most cited literature. In the next chapter, we present the studies that are used as our drivers for the implemented methodology. Specifically we present the purpose and the main conclusions of each study.

2.2 Error types

The immense need for accuracy, speed and concreteness in assessing information for enterprises needs promoted the use of spreadsheets. Due to their friendly interface, the use of spreadsheets has grown extensively in the last two decades and the importance of their role is highlighted in industry. Moreover, spreadsheets have been employed as a useful tool for business evaluation and management of personal accounting, data analysis (Connors 1983, 1984; Davis 1997; Heagy & McMickle 1988; Heagy & Gallun 1994; Lee 1986; Mingers 1991; Waller 1985, Panko 1998, Kreie *et all* 2000, Pryor 2004, Powell *et al* 2008) and for many other simple or more complicated applications. However, the extensive application of spreadsheets embodies the risk of error and represents a major control risk to organizations (Janvrin and Morrison, 1996).

Many scholars have underlined the fact that spreadsheets enclose errors (Panko, 1998; Kreie et all, 2000; Panko and Sprague, 1998; Panko and Halverson, 1996). As spreadsheets have become a significant tool for businesses and for individuals, researchers suggested a classification of these errors so as to identify errors more easily and quickly (Ronen *et al.* 1989; Cragg and King 1993) categorised spreadsheet errors into: wrong formula, erroneous ranges, omitted factors, incorrect data entry, misuse of spreadsheet functions and duplication of effort. Galletta *et al.* (1996) categorised errors in those caused by the spreadsheet

technology (device errors) and those caused by entering wrong data (domain errors). Furthermore, Saarilouma and Sajaniemi (1994) distinguished them between location errors and formula errors. The most popular classification however is the one proposed by Panko and Halverson (Powell *et al.*, 2008) who classified errors to quantitative and qualitative and introduced the first concrete and general classification of errors. Quantitative errors have to do with incorrect values while qualitative with the formula design of the spreadsheet (Panko and Halverson, 1996).

Panko and Halverson (1997) after having classified spreadsheet errors into quantitative and qualitative, they specified them into three more definitive subcategories; the mechanical errors by entering wrong information in spreadsheets, the logic errors that happen when the user does not apply the appropriate algorithm for a particular formula and the omission errors when information is missed out. On the other hand, there are two types of quantitative errors; those that occurred by insufficient knowledge of implementing spreadsheets (reasoning errors) and those that occur by insufficient real world or mathematic knowledge (accidental errors). Teo and Tan (1997) based on Panko's and Halverson's (1997) error categorisation, introduced two more types of qualitative errors. The first one is duplication errors; errors that arise when one parameter is entered more than once in a spreadsheet. The second type is jamming errors; these errors arise when the user enters more than one parameter in one spreadsheet cell.

2.3 Most common auditing and testing methods

Testing is the controlled execution of a spreadsheet, checking that its function meets its specification. Testing is the only way to tell whether the spreadsheet actually works and is therefore a vital step in the process of obtaining confidence in the results (Pryor, 2004). With

a good testing procedure, tests are easy to run, and the auditor can automatically have a log making it easy for him to see what tests have been run and when, as well as the results of its test. Testing is part of ensuring spreadsheet quality and complements spreadsheet review. Unlike testing, a review consists of looking at the code and trying to spot the errors, while testing a spreadsheet includes running it and look at the results. Both are imperative, and neither is likely to find all the errors on its own (Pryor, 2004). Testing can take place at different stages of the development process.

There are different types of testing of which unit testing is the most detailed type. Individual components are tested in isolation. Unit testing should take place frequently throughout the development process. System testing looks at the system as a whole and tests the final output. It should take place at a minimum when a spreadsheet is released for use, and preferably more often during development. Regression testing compares the results of a new version against those of a previous version. It is a specialised form of system testing, used to check that there are no intended changes have been introduced. Acceptance testing, test allows the spreadsheet users when they receive the spreadsheet from the developer, to determine that it meets their requirements and is fit for use.

Auditors and academics have been busy with two interrelated areas; spreadsheet error rates, and how to generate error-free, trustworthy and reliable spreadsheets. Auditing methods are the different ways of identifying errors in a spreadsheet model. The most common auditing methods that are used in the industry today are classified by Chan, Ying and Peh (2000). Spreadsheet description is the first method they propose and contains the logic specification of the model. Then the authors propose the most traditional methods of spreadsheet auditing: on-screen spreadsheet examination and printed spreadsheet, which are used in many other studies and is employed by our study as well. Findings of the studies that investigated these two auditing methods favour the printed spreadsheets. The following

methods are the more recent on-screen visualisation tools. Comment notes are little comment notes that are attached to individual cells. When the spreadsheet model is being built, the user can insert comments to clarify the meaning of a particular cell. This aids at a later stage the understanding of the model. Auto calculate is a feature that automatically shows the sum (or other functions as selected by the user) of a block of selected cells. Furthermore, modern spreadsheet software allows the user to set a range of valid values for a cell. For example, a cell that contains age may be constrained to a value between 18 and 99. The user can use advanced the invalid data input method to highlight cells that contain values outside the valid ranges.

2.4 Spreadsheet errors finding performance

The impact of spreadsheet errors however is not yet widely examined in literature and the number of studies dealing with this issue is quite limited (Powell *et al.,* 2008). Clermont (2002) stated, "We did not find any tremendous erroneous result values that might have had severe negative effects on the company". The European Spreadsheet Risks Interest Group (EUSPRIG) reports the impact of spreadsheet errors in several cases. Caulkins *et al.* (2006) conducted an investigation on the impact of spreadsheet errors concluding, "Spreadsheet errors are a significant threat to decisions". Moreover, practitioners identify the importance of accuracy and have published many techniques for improving it in their spreadsheets (Brown and Gould, 1987). Their prescriptions and results of research are consistent and they form the basis for spreadsheet accuracy theory. Three propositions describe how the three constructs influence spreadsheet accuracy. Kruck's (2006) study indicated that his proposed Spreadsheet Accuracy Theory is significantly improved by the development of accurate spreadsheets.

The penetration of spreadsheets in various applications has triggered our interest to investigate spreadsheets errors and their effects. Numerous publications over the years have described the seriousness of the spreadsheet errors' problem and the extent to which it has unfavourably affected businesses. Some of the most recent and concrete researches are cited next.

There is a major need for testing in order to reduce spreadsheet errors since spreadsheet systems are the most widely used and the most popular end user systems (Ayalew *et al.*, 2000). The only empirically proved way to significantly reduce errors is testing. Programmers have learned that they can reduce error rates from two per cent to five per cent on average of all code lines to only 0.1 per cent to 0.3 per cent (Putnam & Myers, 1992). However, there are multiple types of testing, most of which are insufficient to reduce errors to the point where significant errors will be highly unlikely. The important criterion for selecting a testing methodology is the method's ability to reduce formula error rates from two per cent to five per cent to a far lower value, detecting sixty per cent to eighty per cent of all errors in each round of testing. Panko and Sprague (1998) suggested testing throughout the development process of a spreadsheet. Furthermore, they suggested beginning testing before code or logic creation, at the requirement documents.

Panko and Sprague (1988) examined the frequency of errors when building a spreadsheet. Their sample under examination was a group of over a hundred MIS undergraduate students and fifty MBA students; all students were familiar to spreadsheet modelling and accounting. Their experimental trial was the students to create a simple model and then to test it for errors. Panko and Sprague (1988) made an error classification in order to determine the error frequency when examining the results. Furthermore, the error classification contributed in realising "who" makes the errors and what kinds of errors are most common. Opposite to other studies, this one required the creation of a simple model;

still errors were not avoided. More particular, thirty five per cent of the models and two per cent of the cells presented errors. However, the errors were met at a lower per cent compared to past studies but still not at a safe percentage for developing spreadsheets. The study concludes that errors cannot be avoided since error is almost natural. Even professional programmers are incapable to identify all the errors in a spreadsheet.

Panko (2000) concluded in his study that errors cannot be completely avoided and he recommended ways to reduce them rather than to eliminate them. When spreadsheets are developed, pure testing is applied (Cragg and King, 1993; Gable, Yap and Eng, 1991; Hall, 1996; Nardi, 1993; Schultheis and Sumner, 1994). In order to decrease spreadsheets' errors Panko (2000) suggested a cell to cell examination for each spreadsheet thorough an inspection code. However, coding inspection however costs in money and time and is yet not effective, since this method is found to detect about eighty per cent of the errors (Panko, 2000). Furthermore, Panko (2000) suggested checking spreadsheet outcome for reasonableness since this practise is easy, effective and does not cost money or time before applying the code inspection. Moreover, Panko (2000) recommended evading errors by entering information twice in order to verify the accuracy of information already entered the second time information is entered. Finally, Panko (2000) suggested using cell protection in order to prevent users from modifying another user's work. Cell protection is easy to apply and helpful.

According to the American Management Association, spreadsheets are employed by the ninety one per two per cent to five per cent of an end-users sample (American Management Association, 1988). Businesses mainly employ spreadsheet models in order to reach financial decisions concerning a new investment project, or to estimate their budget, or about forecasting, therefore businesses should be convinced that spreadsheets are a reliable tool.

Thompson and Tan (1997) also made an investigation on spreadsheet errors and their impact on businesses. Realising the impact of an incorrect decision caused by a possible error in the construction of spreadsheets, Thompson and Tan (1997) examined dissimilar categories of errors in the building of spreadsheets. For this purpose, they investigated the types of errors that may occur even for simple domain-free spreadsheet problems. Additionally they show that spreadsheet errors are difficult to detect during "what-if" analysis when spreadsheets are not properly designed. In this study Thompson and Tan (1997) measure for the first time in literature qualitative errors and thus they consider their paper seminal for the literature. Thompson and Tan (1997) followed Panko's (1998) findings on spreadsheet errors that they are repeated quite often even for applications that do not require high domain experience (Panko, 1998).

Panko (1998) examined the construction and the design of spreadsheets as well as the errors that may arise from them. Furthermore they investigated the types and the frequency of errors in order to acquire information about which kind of errors occur more frequently when changes on the design of spreadsheet are applied. Finally, scholars examined the connection between dissimilar kinds of errors in order to avoid the repeated mistake of correlating dissimilar kind of errors which results into not well-build or advanced spreadsheets. The sample of their investigation was a group of one hundred and seventy-six students studying information systems at the University of Singapore. The students were already familiar with spreadsheet use. The students were assigned two spreadsheet exercises. The first was a homework task concerning a wall construction problem (Panko, 1996) and the second was a classroom task concerning changes that should be applied on the design of spreadsheets. Results from their investigation illustrate that errors are repeatable and result in new errors that mainly arise when rebuilding a spreadsheet.

Thomson and Partridge (2001) also investigated spreadsheet errors following Panko (1998). Their examination sample was a first-year group of one hundred seventy nine students participating in a computing course of a university. The students were already familiar with developing spreadsheets, formulas, data tables, and "what-if analyses". Students were provided with two spreadsheets in order to identify their errors. The spreadsheets were given along with hard copies so as students are able to detect errors both on the printed copy and on screen. Thomson and Partridge (2001) provided the hard copies based on other researches, which concluded that hard copies contribute in detecting errors more easily and quickly (Dillon, 1992; Gould *et al.*, 1987; Oliver, 1994). Results of this study illustrated that only fifty per cent of the errors have been identified. The quantitative errors were more easily identified but still not at a sufficient level. On the other hand, qualitative errors were detected by a relatively low percentage of the students. Thomson and Partridge (2001) reached the conclusion that identifying errors depends on the type of error and on the user experience in identifying errors.

Chan *et al.* (2000) in their review about spreadsheet errors underlined the fact that despite the simplicity of developing a spreadsheet model it is realised that frequent errors occurred during this developing process. Moreover, the fact those spreadsheets errors are difficult to be perceived, it results in misleading decisions for the shareholder. Therefore they considered it crucial to look for useful tools and strategies that contribute to the identification of those kinds of errors. In order to succeed their purpose a suite of new auditing visualisation methods have been implemented. Examples have been given on how each of the methods can be useful. Furthermore, four strategies for detecting errors were proposed. The first strategy covers an overview understanding of the model, with the auditing method helping the user to identify component blocks of the model. The subsequent strategies are targeted specifically at the main types of error: mechanical, logical and omission. These strategies can focus on the

component blocks one at a time, instead of having to deal with the full spreadsheet model. The new visualisation methods enable these strategies to be achieved visually, with less search effort by the user.

Galletta *et al* (1996) also investigated the reliability and validity of spreadsheets through a research on a group of MBA students. The purpose of the study was to realize how easy and quickly errors can be detected on spreadsheets. The particular study concentrated on the presentation formula that a spreadsheet has and whether the different formulas of spreadsheets contribute in the error detection. Three hypotheses were formed concerning the diverse presentation formula of a spreadsheet and several different kinds of errors were lodged in under the trial examination of spreadsheets. The examination concluded that errors can be more easily and quickly detected when spreadsheets are escorted with hard copies of the spreadsheet.

Ayalew, Clermont and Mittermeir (2000) attempted through their research to overcome the tension between the statements "Spreadsheets are Software too" and "spreadsheet-authors are no Programmers" in order to improve the quality of spreadsheet software. More specific, they presented two different but supplementary approaches in order to detect errors in spreadsheet programmes. They stressed the fact that spreadsheets are programs that are written not only by professionals but also by end-users. Furthermore, since spreadsheet systems are easy to use, they do not require much training in formal methods of designing and programming, and show – in contrast to conventional programs – the results of the effort while the development effort is still in progress, they are also written in a style different from conventional software (Beiser, 1990). However, end-users do not have the capacity that professional have in programming and certainly, their model is not related to programming concepts. In order to overcome end-users lack of capacity in programming they suggested the model visualization and the interval testing spreadsheets. The model visualization provides

the spreadsheet user and programmer more insight into the structure of the spreadsheet, which will help him to shorten the trial and error process of creating the spreadsheet and to understand and debug spreadsheets in use. The other approach, interval testing spreadsheets, tries to overcome the difficulties resulting from the lack of specification of spreadsheets by introducing interval arithmetic as basic device. The combination of these two approaches might contribute in detecting errors of frequently occurring nature.

Howe and Simkin (2006) conducted a study in an attempt to detect not only errors in a spreadsheet but also the aspects that affect identifying errors in a spreadsheet. The aspects that they tested if they affect error detection were the age of the user (maturity), his educational skills (grade), how familiar the user is with spreadsheet and whether the user is male or female. After collecting the above information for their sample they provided them with a simple task in electronic copy only. Howe and Simkin categorised errors into the following four types: a) entering data that does not comply with the principles of the company in a spreadsheet (rule violation errors), b) entering incorrect information in spreadsheets c) accounting errors (clerical and nonmaterial errors) and d) errors in the spreadsheet formulas. According to their error and end-user categorisation the results demonstrated, they concluded that younger user detect errors more easily compared to all users. Moreover, that the educational skill plays a significant role in detecting errors; and finally that ladies identified more errors than gentlemen. However, only sixty seven per cent of the errors have been detected. Furthermore, the types of errors that have been more easily detected are the errors that concerned incorrect data entered. The lower percentage in detecting errors presented those that concerned spreadsheet formulas.

Kruck (2006) investigated ways to enhance spreadsheet reliability and to reduce errors in spreadsheet. Kruck underlined that errors can be reduced if spreadsheets were designed more cautiously .Moreover, numerous errors can be avoided if spreadsheet became less

complicated and multiple test were applied on spreadsheets during the procedure of designing them (Kruck and Sheetz, 2001, Simkin, 2004). However, even if all the above practices were applied still Kruck concluded that the inclination for making errors would not eradicate. The above three topics were tested on a sample of one hundred twenty nine students by testing three hypotheses: a)first hypothesis about the development of the spreadsheet, b)second hypothesis about the simplicity of spreadsheets; formulas and third hypothesis about applying tests on the spreadsheets. The research results were consistent with former researches (Kruck and Sheetz, 2001, Simkin, 2004) that errors were limited.

Purser and Chadwick (2000) investigated the spreadsheet errors concerning the professional spreadsheet end-user group. More specifically the sample of the end-users was consisted by individuals whose everyday work employed the use of Excel and by students of Greenwich University who undertake courses of computing and mathematics. The professionals were the tested group and the students were the control group. The investigation was carried out through internet and the survey was distributed to the shareholders' mail account.

Pryor (2004) has underlined the significance of testing spreadsheets. In order to test a spreadsheet it must be verified the expected output compared to the known input. Testing is very important as we can realise whether spreadsheet provides the correct information. However, reviewing a spreadsheet should not be considered testing, as testing means to check the results, while reviewing means to check the spreadsheet formulas and detect errors. A spreadsheet can be tested at all its development phases. According to Pressman and Ince (2000) there are various types of testing. More particular, there is testing of every unit of spreadsheet separately (unit testing). Furthermore, testing altogether the system (system testing) and testing the results of a spreadsheet before and after implementing changes in it (regression testing).Finally, there is testing that run the shareholder before it accept the

spreadsheet by the developer (acceptance testing).Pryor (2004) concluded that testing even though is difficult to implement in some cases it is feasible and it offers many advantages.

2.5 Summary

In this chapter, we presented the prevailing studies dealing with spreadsheet errors. Particularly, we distinguished the studies dealing with the errors' categorisation and the studies that examine the causes of errors and factors that lead to error. Using these studies as driver, we present the methodology of our study in the next chapter. We base our methodology on Thompson and Partridge (2001), Harry Howe and Mark G. Simkin (2006), Janvrin and Morrison (1996), Galletta *et al* (1996), Panko and Sprague (1998) studies and test the effect of the spreadsheet presentation and experience on error finding performance. The methodologies that we base our research framework are also presented on the next chapter.

3. Methodology

3.1 Introduction

After a thorough examination on the literature concerning spreadsheet error detection, we proceed on the specification of the methodology. Initially, we re-examine of the resembling methodologies in relevant studies upon which we based our approach. Furthermore, in this chapter, we outline the research framework as well as all methodological issues. Specifically, we present a set of factors that may affect error finding performance. Afterwards, we develop a set of research hypotheses in order to test the research framework

and to specify the extent of each factor is related to the error finding performance. This research hypotheses set is submitted to statistical tests using a dataset of 123 finance and business students, presented in the following analysis along with the methodological framework.

3.2 Relevant Methodologies

Several studies are dealing with error finding performance. Thompson and Partridge (2001) conducted an experiment to examine the rate of detection of both quantitative and qualitative errors in two domain-free spreadsheets. Specifically, their study contains two experiments examining detection of both quantitative and qualitative spreadsheet errors. In both experiments they used a domain-free spreadsheet problem in order to rule out threats to validity caused by differences in task domain knowledge among participants. Briefly, this study extends previous studies by examining differences in error detection rates between quantitative and qualitative errors; error detection with and without prior incremental practice; and whether the nature of errors affects detection rates. They used a sample of first year business undergraduate students who were taking an introductory course on spreadsheets using Microsoft Excel. In order to reach the desired outcome they examined the following three hypotheses:

H1: Quantitative errors are more easily detected than qualitative errors.

H2: The prominence of seeded errors will influence spreadsheet error detection rates.

H3: Subjects with prior incremental practice in error detection will be able to detect more errors than subjects without prior incremental practice.

Results of the specific study showed that the ability to detect errors appears to be dependent on the error type (logical, mechanic or omission) and the error prominence as well as prior incremental practice.

In their research Janvrin and Morrison (1996) explore the impacts of different design approaches through two field experiments evaluating the use of a structured design methodology when developing complex spreadsheets. For this purpose the authors used sixty one accounting and business administration students that were assigned several linked spreadsheets with errors to treat. Their methodology is implemented in two phases where in the first phase subjects could work together at relatively simple workbooks while at the second phase subjects were asked to work individually at relatively more demanding workbooks.

Results of their study indicate that subjects using the methodology showed a significant reduction in the number of "linking errors," that is mistakes in creating links between values that must flow from one area of the spreadsheet to another or from one worksheet to another in a common workbook. They also observed that factors such as gender, application expertise, and workgroup configuration influenced spreadsheet error rates as well.

Harry Howe and Mark G. Simkin (2006) in their study for spreadsheet error detection report the results of an experiment in which they investigated the potential determinants of spreadsheet error-detection rates in a sample of 228 participants. On average, participants found 67 per cent of the 43 errors embedded in a test spreadsheet model. In this experiment the scholars also gathered information about a number of factors that might plausibly account for differences in error-detection rates. Their independent variables included gender, age, number of university credits taken, grade, years of prior programming experience, years of prior spreadsheet experience, years of prior Excel experience, user confidence, and a dummy variable for school location. The most important conclusion is a lack of explanatory power for

all the variables they used. Such a finding implies that other, unknown, factors must be at work. Scholars also believe that all spreadsheet errors are not the same, and that it makes sense to categorise them. Therefore they propose a four error type categorisation.

In their study Panko and Sprague (1998) asked from a sample of a hundred and two undergraduate Management Information Systems' students and fifty Master in Business Administration students to develop a model from a word problem that free of domain knowledge. The scholars followed Galletta *et al.* (1996) and tested the error finding performance of their subjects according to their development, auditing, and training experience. Results of their study reveal that inexperienced and experienced spreadsheet developers do about the same number of errors. These results are consistent with Galletta *et al.* (1996) who found that when experienced spreadsheet developers audited models, they did not find a higher percentage of the errors in these models than did inexperienced spreadsheet developers.

Galletta *et al.* (1996) picked a sample of one hundred thirteen MBA students and assigned them the task to find eight errors planted in a single-page spreadsheet. Their purpose was to discover if differences in the presentation format would facilitate error-finding performance. Totally they used five presentation formats. Spreadsheets were presented on the screen, both with and without formulas' presentation. Spreadsheets were also presented on paper with a list of formulas attached, or without formulas. An integrated formula paper treatment was introduced, with formulas presented in each cell directly under each calculated value. The participants found, on average, only about 50 per cent of the errors across all presentation formats. The on-screen treatments were clearly inferior to the paper treatments, regardless of the presentation of formulas. Their study showed that users who attempt to find errors in spreadsheets are not aided by formulas, but are aided by paper copies of a spreadsheet. Paper versus screen effects according to the authors applies well to spreadsheet

error finding. We consider the studies of Galletta *et al.* (1996) and Panko and Sprague Jr. (1998) which resembles the first, seminal for our study and so these two studies are used as the main drivers for our research.

3.3 Research Framework

There are various types of errors in spreadsheets, examined in a number of studies. Spreadsheet error finding literature mainly deals with users' errors unlike software errors that have to do more with programming issues (Rajalingham *et al.,* 2000). The most common distinction of users' errors is in qualitative and quantitative. Quantitative errors usually has the form of numerical errors when designing a spreadsheet while qualitative take the form of poor spreadsheet design and format and could potentially be the cause to quantitative errors. According to Rajalingham *et al.* (2000) qualitative errors may be due to formatting, update, hard-coding and semantic factors. Quantitative errors may immediately distort the spreadsheet result; qualitative errors on the other hand may lead to quantitative errors later when the spreadsheet is in use and refer to errors by the poor spreadsheet design.

Panko and Halverson (1996) distinguished the three dominant types of quantitative errors. Mechanical errors are defined as simple mistakes such as mistyping a number. In this category belong errors that have to do with overwriting or misleading data input (Salchenberger, 1993). Logic errors involves mistakes in reasoning that leads to wrong formula entering and are considered as quite difficult to detect and correct (Allwood, 1984). Errors in enabling and planning skills are representative examples. The most critical type of errors however are the omission errors occurring when something is left out; this type is considered as the most difficult to detect (Allwood, 1984; Bagnara *et al.*, 1987; Woods, 1984). Any cell that may contain an error can be either referenced or unreferenced

(Rajalingham *et al.,* 2000). In the second case things are rather easy for an auditor who has just to correct the specific error. Referenced cell that contains errors however may cause a series of other errors in the referenced cells and make the auditors' task rather hard.
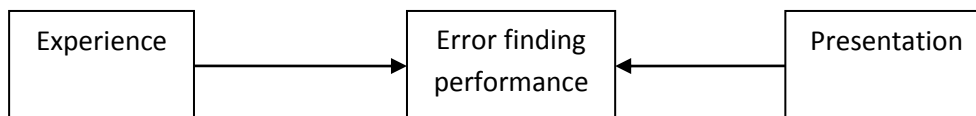
In this study we employ a common evaluation free cash flow model in which we seeded eight errors that represent the three error types mentioned above and asked the respondent body to identify them. Specifically we seeded three mechanical errors, three logic errors and two omission errors as they are considered the most difficult to detect. Before respondents' start the activity of error finding they were asked to fill in a questionnaire about their personal demographic profile. This questionnaire contained questions regarding age, years coping with the Free Cash Flow concept, years coping with spreadsheet activities, gender and the program of studies followed. Respondents were also asked to self evaluate their spreadsheet familiarity using a five-point Likert scale. Using these data and their performance on finding errors we test a set of hypotheses concerning the factors that affect the spreadsheet error finding performance. The exact form of the research model is presented in graph 1 that follows.

Galletta and Hufnagel (1992) revealed that detection of spreadsheet errors is a quite demanding task. Sometimes even experienced auditors cannot detect errors in simple spreadsheets. This incompetence is attributed in numerous factors that can hinder the error finding performance. Following Galletta *et al.* (1996) we use two factors that are tested for affecting spreadsheet error detecting performance. Specifically the factors we use are *spreadsheet presentation and user's experience.* Presentation refers to whether the spreadsheet is presented on screen or on paper as well as to whether formulas or values appear (Galletta *et al.*, 1996). Spreadsheet experience refers to the familiarity that each respondent attributes to himself about working on spreadsheets and about the concept of free cash flow. Howe and Simkin (2006) tested respondents' age for directly affecting error finding

performance as an indicator of experience. We believe that age is not indicative of a respondent's experience because it does not represent the time that each respondent spent working with spreadsheets (Randolph *et all*, 2002). For this reason experience for this study is a factor occurring from three different parameters; self-reported spreadsheet familiarity, free cash flow concept familiarity and years of study. As the purpose of this study is to identify whether the two literature-based factors affect error finding performance, we only employ quantitative errors for our research. Besides qualitative errors, as proposed by Rajalingham *et al.* (2000) can only potentially be the cause of any form of quantitative errors. Thus qualitative errors are a subject beyond the present analysis.

Graph 1-Research

```
┌─────────────┐        ┌─────────────┐        ┌─────────────┐
│ Experience  │───────▶│ Error finding│◀───────│ Presentation│
│             │        │ performance │        │             │
└─────────────┘        └─────────────┘        └─────────────┘
```

3.4 Hypotheses Development

For every tested hypothesis, we perform a twofold test; the first concerns the average amount of errors found and the second the average time that each subject spends to complete the task. Dillon (1992), Gould *et al.* (1987) and Oliver (1994) used two presentation patterns and found that reading from screen is generally faster than reading paper. Thus, presentation (screen versus paper) can influence error finding performance. After this the first two hypotheses are the following:

H1: More errors are found when spreadsheet is presented on screen than on paper.

H2: Errors are found at less time when the spreadsheet is presented on screen than on paper.

Another aspect of presentation pattern is the appearance of spreadsheet formulas during the error finding process. Although Galletta *et al.* (1996) found no significant differences in error finding performance for subjects provided with spreadsheet formulas versus those who weren't, there is evidence that providing formulas in an integrated manner (i.e., both formulas and values are provided in the same spreadsheet) may reduce the number of invalid errors found, compared to other treatments. The following two hypotheses refer to the appearance of formulas and are the following:

H3: More errors are found when the auditor works with formulas.

H4: Errors are found more quickly when the auditor works with formulas.

In this study, we also examine the subjects' overall experience retrieved from Galletta and Hufnagel (1992) although they did not directly test it for affecting error finding performance. Thus, hypotheses 5 and 6 are formed as follows:

H5: More errors are found when the auditor has high domain experience.

H6: Errors are found more quickly by experienced auditors.

3.5 Method

The experiment we implemented to test the set of hypotheses contain a simple free cash flow evaluation spreadsheet and an answer sheet where respondents were asked to identify the cells containing errors as well as to describe each error. Respondents were also asked to include some basic demographics. The evaluation spreadsheet was delivered in five presentation forms. The first two forms was electronic, specifically a Microsoft excel file

where only values are presented and the auditor must point each cell to see the formula on the formula bar and a Microsoft excel file where the respondent can see only formulas. The remaining three presentation patterns are printed spreadsheets; specifically a printed spreadsheet where the subject can see values but no formulas at all, a printed spreadsheet where the subject can see only formulas and no values and a printed spreadsheet where the subject can see both formulas and values.

The experimental material was delivered to the students of two business administration undergraduate classes and two MSc in Finance and Financial Information Systems' classes of the Technological Institution of Kavala, Greece. Participants amounted 123 and were split into four different groups. All participants have taken at least one spreadsheet modelling using Microsoft's excel course as well as one investment evaluation course and they are familiar with both Microsoft's excel and the free cash flow evaluation concept. Next, we split the total amount of participating students in five groups randomly and delivered each one of the different five spreadsheet forms in each group. We set no time constraints but informed students that they were asked to detect exactly eight errors. This process allowed us to record the time every student needed to complete the task. Participants were asked not to collaborate with each other or use another kind of external help and were discretely supervised to maintain this condition.

After the task was completed, 123 valid responses were collected and we recorded the total valid errors each student was able to find as well as the time each one spent to find them. In addition, after coding the data into an excel spreadsheet we classified the errors each student found into the three error types already mentioned before. Furthermore, we calculated the time each student used per found error. Using these data, we extracted an index of overall performance for each student. This index is expressed as the average time each student spends

to find a valid error and algebraically calculated by the quotient: Total time spent / Valid errors found

Sample

As already mentioned the material was distributed in four different classes of the technological institution of Kavala and in particular to the students of two undergraduate business administration classes and two postgraduate finance and financial information systems classes. Specifically 52.03 per cent (64 students) were undergraduate students of business administration while 47.97 per cent (59 students) were postgraduate students of which 23.58 per cent are in the finance stream and 24.39 per cent are in financial information systems stream. Furthermore 70.73 per cent (87 students) of students are male and 29.27 per cent (36 students) are female. Additionally 64.23 per cent (79 students) are younger than 24 years old while the rest 35.77 per cent (44 students) are 25 years old or older. Finally, given that each student is admitted at the age of eighteen years for undergraduate studies, we consider the years of experience each student have in spreadsheet modelling and business courses as the difference between their age and the age of the admittance. The total experience years of the participants per stream are presented in the following table 1.

Table 1- Years of experience

		Years of experience				
		less than 2	3 or 4	5 or 6	7 or 8	9 or 10
stream	Finance	0,00%	0,81%	14,63%	6,50%	1,63%
	Financial Information Systems	0,00%	0,81%	13,82%	6,50%	3,25%
	Business Administration	13,01%	27,64%	9,76%	1,63%	0,00%
	Total	13,01%	29,27%	38,21%	14,63%	4,88%

Apart from the total years of experience, every participant was asked to state how he/her feels about his familiarity with excel spreadsheets and the concept of free cash flow evaluation. In the following table 2, we present answers per stream regarding Excel spreadsheets' familiarity.

Table 2- Spreadsheet Familiarity

		Stream			Total
		Finance	Financial Information Systems	Business Administration	
Spreadsheet Familiarity	Excellent	0,00%	6,50%	6,50%	13,01%
	Very good	8,94%	13,01%	21,95%	43,90%
	Good	12,20%	4,07%	17,07%	33,33%
	Fair	2,44%	0,81%	6,50%	9,76%
	Not so good	0,00%	0,00%	0,00%	0,00%
	Total	23,58%	24,39%	52,03%	100,00%

In table 3 that follows, we present answers per stream regarding the free cash flow concept familiarity.

Table 3- Free Cash Flow Concept Familiarity

		Stream			Total
		Finance	Financial Information Systems	Business Administration	
Free Cash Flow Concept Familiarity	Excellent	22,76%	23,58%	11,38%	57,72%
	Very good	0,00%	0,00%	8,13%	8,13%
	Good	0,81%	0,81%	17,89%	19,51%
	Fair	0,00%	0,00%	14,63%	14,63%
	Not so good	0,00%	0,00%	0,00%	0,00%
	Total	23,58%	24,39%	52,03%	100,00%

Based on these data we extracted a factor score for each participant regarding his/her experience. This was achieved by performing an initial solution, principal components factor analysis with no rotation using the SPSS 10.00 software. This procedure produced the overall experience factor score by the three experience items we used to measure users experience (self reported spreadsheet familiarity, free cash flow concept familiarity and years of experience). Both Kaiser-Meyer-Olkin and Bartlett's tests of sampling adequacy and sphericity respectively, produced accepted scores.

3.6 Summary

In this section, we presented the methodology of our study. Specifically after presenting the most prevalent studies' methodologies, we outlined our methodological approach and research framework. In addition, we presented the method of our research and basic descriptive statistics of our dataset. In chapter four, we are going to outline the findings of our study regarding the effect of domain experience and spreadsheet presentation on error finding performance. For this purpose, we implement a statistical analysis on the dataset that helps us to extract useful conclusions.

4. Empirical Research

4.1 Introduction

After presenting all features of our methodological approach, we proceed with the analysis of findings. In this section, we investigate whether spreadsheet presentation and user's experience affects overall error finding performance after reviewing the results of the

conducted experiment. Specifically, we perform a statistical analysis to realise whether the hypotheses presented in chapter three are valid. Furthermore we present all the statistical tests we performed and state their practical implication.

4.2 Empirical Research Results

As mentioned in chapter three the experimental free cash flow model was delivered in five distinct forms. Students were separated in five groups and each group was provided the spreadsheet in one of the five different forms. The group that worked with "excel with values" form found 56.50 per cent of the total errors seeded in the spreadsheet. The group that worked with the "excel with formulas form found 44.40 per cent of the total errors. The group that worked with the "pdf with values form" found 24 per cent of the total seeded errors, which is the lowest performance. The other two groups both found 41.6 per cent of the total seeded errors. For each form, we extracted means, standard deviation and skewness both for the number of valid errors each student found and the time he spent to find them. Then we calculated the rate of how much time on average each student spent for every mistake and extracted means, standard deviation and skewness for each of the five forms. The purpose of this analysis is to find out which spreadsheet form is more convenient for the subjects to perform best in their error finding task. The following tables show results of this analysis.

Table 4- Errors' mean, standard deviation and Skewness

	Errors' Mean	Standard Deviation	Skewness
Excel With Values	5.65	1.22	-0.22
Excel With Formulas	4.44	1.04	0.17
PDF With Values	2.4	1	0.43
PDF With Formulas	4.04	1.42	0.29
PDF With Values and Formulas	4.16	1.24	0.65

In table 4 we can see that the students who on average performed better on error finding performance, are those who were given the excel spreadsheet with values. In this form, the student should point each cell to see the formula it is based on. The worst performance is observed to those who were given a printed pdf, which presented only values of each cell. At this case, students couldn't see formulas at all, only values. Standard deviation in all cases is low and acceptable showing the dispersion of the errors amount is low and relatively constant for all students. Finally, with skewness scores less than the unit in all cases data for all cases surround corresponding means normally (follow normal distribution). A primary conclusion from data in table 4 is that the capability of students to point each cell and see the formula is decisive for their performance to find valid errors. By this fact alone however we cannot determine the overall performance because so far we only presented the amount of errors found without incorporate the time that each student spends to find an error.

In table 5 we present the average time that students spent to find exactly eight errors (valid or not valid) as well as the standard deviation and skewness of these data. As shown in table 5, standard deviation is relatively high when it comes to the last two groups of the experiment, which is pdf with formulas and pdf with both values and formulas. That is that the dispersion of time each student spent to complete the task in these cases is high. Thus, there were students who completed this task in a very short time and others that spent much more time. Skewness scores are acceptable in all cases showing that the all data follow normal distribution. Table 5 shows that excel with values form is the one that helps students to find the eight errors more quickly than any other form. Since this is the form, which enables students to find more errors that are valid (table 4), we conclude that it enables students to be more efficient in the error finding task.

Table 5- Time spent in each form; means, standard deviation and Skewness

	Time Mean	Standard Deviation	Skewness
Excel With Values	21.61	1.16	0.1
Excel With Formulas	22.4	1.41	-0.31
PDF With Values	30.6	3.15	0.73
PDF With Formulas	26.4	4.43	1.05
PDF With Values and Formulas	29.16	3.92	-0.11

The above claim is confirmed from the calculation of the error rate, which represents the time a student takes to find a valid error. In table 6 we present statistics for the error finding rate. It is clear that when students are given an excel file with values they can find a valid error in 4,01 minutes, far less than with any other spreadsheet form. Furthermore, dispersion (0.94) of answers in this case shows that this performance is consistent to all students. Skewness of 0.35 clearly identifies a data normal distribution unlike spreadsheet with formulas and pdf with values. Specifically for the case of pdf with values, dispersion is extremely high which in combination with a relatively accepted skewness score shows that students had quite contradictory performances.

Table 6- Error Rate statistics

	Error rate Mean	Standard Deviation	Skewness
Excel With Values	4.01	0.94	0.35
Excel With Formulas	5.27	1.16	1.40
PDF With Values	15.62	8.56	1.19
PDF With Formulas	7.45	3.29	0.31
PDF With Values and Formulas	7.64	2.53	0.35

After studying tables 4,5,6 it is clear that the best error finding performance is achieved when students are provided an excel file showing only values.

The quality of all valid errors is likewise analysed. As mentioned in chapter three there are three main error distinctions prevailing in literature. These are mechanical errors, logic errors and omission errors classified by Panko and Halverson (1996). In table 7 we present an analysis of the errors found. Particularly, the group that worked with the "excel with values" form performed extremely well in finding logic errors. The other four groups seem to have a somewhat same performance in finding mechanical and logic errors, while none of the five groups' exhibit good performance with omission errors. It is realised that "excel with values" form is quite convenient for finding much more logic errors.

Table 7- Error type analysis

	Mechanic Errors	Logic Errors	Omission Errors
Excel With Values	26.15%	70.00%	3.85%
Excel With Formulas	34.23%	54.95%	10.81%
PDF With Values	46.67%	46.67%	3.33%
PDF With Formulas	52.48%	42.57%	4.95%
PDF With Values and Formulas	42.57%	55.45%	4.95%
Total	38.97%	55.47%	5.77%

4.3 Hypotheses test

In this section, we test the set of hypotheses we proposed in chapter three. To do so we conduct a cross tabulation analysis and then compare means. Furthermore, we perform a paired samples test to determine whether the differences of means are statistically significant or not. The first pair of hypotheses is the following:

H1: More errors are found when spreadsheet is presented on screen than on paper.

H2: Errors are found at less time when the spreadsheet is presented on screen than on paper.

Table 8- Screen vs. Paper

Screen VS Paper							
	Screen		Paper		Mean difference	T statistic	Sig.
	Mean	Standard Deviation	Mean	Standard Deviation			
Number of errors found	5.02	1.28	3.53	1.46	1.85	6.389	.000
Time	22.02	1.34	28.72	4.20	-6.72	-10.428	.000
Rate	4.68	1.24	11.79	6.64	-7.11	-6.25	.000

As shown in table 8, the mean of errors found when the spreadsheet is presented on screen is 5.02 significantly higher than the case of presenting the spreadsheet on paper. Furthermore average time for students to complete the task they were assigned is much lower when the spreadsheet is presented on screen than when it is presented on paper. Specifically working with on screen presentation of the spreadsheet students needed on average 22 minutes to complete their task; working with the paper version, they needed about 29 minutes. That is the average valid error took students working with the on screen version 4.68 minutes on average to find and 10.24 minutes on average for students working with the paper version. In order to determine if these differences are statistically significant we performed paired samples tests between the corresponding means.

Moreover based on table 8, working with the on screen spreadsheet students perform better finding on average 1.85 errors more than working on paper. This difference is statistically significant (t=6.389, sig. = .000). Students working with the on screen spreadsheet also perform better than those working with the paper version since they completed the task on average 6.72 minutes faster. This difference is also statistically significant (t=-10.428, sig. = .000). Finally students that worked with the on screen spreadsheet used on average 7.11 minutes less to find one valid error than the students who worked with the paper spreadsheet.

This difference is statistically significant (t=6,25, sig. = .000). Data of table 8 validates both first and second hypothesis.

The second pair of hypotheses that we proposed in chapter three, concerns the performance of the students working with spreadsheets showing values versus those showing formulas. The proposed pair is the following:

H3: More errors are found when the subjects work with formulas.

H4: Errors are found more quickly when subjects work with formulas.

The same approach as before is followed in order to determine if students perform better when they are presented the values of the spreadsheet or its formulas. As shown in table 9 the students that worked with spreadsheets showing formulas found about the same amount of errors with those who worked spreadsheets showing values. Actually, the difference is marginal in favour of students who worked with spreadsheets showing formulas who found on average 4.21 errors unlike those who worked with values and found on average 4.02 errors. This difference however is not statistically significant (t = -0.823, sig. = .413) which means that the third hypothesis is rejected.

Students who worked with spreadsheets showing formulas completed the task in 25.99 minutes on average while those who worked with spreadsheet showing values needed 1.28 minutes more on average. This difference is also not statistically significant (t = 1.332, sig. = .187). Thus, the fourth hypothesis is also rejected.

The performance rate for students who worked with spreadsheets showing formulas reveal that these students need less time to find a valid error. On the other hand students who worked with spreadsheets showing values need on average 2.43 minutes more to find a valid error. This difference is statistically significant (t = 2.74, sig. = .008) which means that

working with spreadsheets showing formulas helps students to perform better. Although we cannot extract valid conclusion for the support of the fourth hypothesis, we can use the performance rate to provide support for this hypothesis. That is because the performance rate, as mentioned in chapter three, represents the time it takes to find a valid error. The rate can support the hypothesis for each error separately and not for the whole task so if students can perform more quickly for each error, they can perform more quickly in the whole task.

Table 9- Values vs. formulas

	Values vs. Formulas						
	Formulas		Values		Mean difference	T statistic	Sig.
	Mean	Standard Deviation	Mean	Standard Deviation			
Number of errors found	4.21	1.24	4.02	1.75	-0.19	-0.823	.413
Time	25.99	1.24	27.27	1.75	1.28	1.332	.187
Rate	6.80	1.24	9.23	1.75	2.43	2.74	.008

Finally, there is one more hypotheses' pair that is proposed in chapter three that refers to the relationship of the users' experience with the overall performance. The pair of hypotheses is the following:

H5: More errors are found when the subject has high domain experience.

H6: Errors are found more quickly by experienced subjects.

In order to determine the experience of the students we factored three components, the spreadsheet familiarity, the free cash flow concept familiarity and the total years of experience. For this purpose, we performed a principal components factor analysis that produced the experience factor for each student. Then we calculated the Pearson correlation coefficient to observe whether the two variables have a linear correlation pattern. At first we

calculated the correlation coefficient for experience and the total amount of errors found in order to check fifth hypothesis. This test however is fruitless since the correlation yields a score of 0.086 (sig. = .342) indicating almost no correlation at all. The same disappointing result is produced for the experience and total time of task correlation; coefficient in this case yields a 0.030 score (sig. = .741) also indication no correlation at all. This correlation analysis provides evidence for the rejection of both fifth and sixth hypotheses and advocates that experience do not at all influence the error finding performance.

To sum up the first set of hypotheses introduced in chapter three containing hypotheses one and two is validating by the data. The second set containing hypotheses three and four is not validated by the data; however, the performance rate, measuring the time needed to find one valid error validates fourth hypothesis. Finally, the third set of hypotheses containing hypotheses five and six is rejected since it has no support from the data.

4.4 Summary

In this section, we presented the results of our research. Specifically we outlined the performance of our subjects in the experiment they were submitted. After reviewing the average performance, we used a combination of statistical tests to check the validity of the hypotheses introduced in chapter three. In general, chapter four is the implementation of the research framework presented in third chapter. In the next chapter we draw the conclusions from the experiments and make a discussion about the findings of this research. Furthermore we propose the limitations of the research as well as new insights and challenges for further research.

5. Conclusions

Our research is based on the study of Galletta and Hufnagel (1992) later followed by Galletta *et al.* (1996) and Panko and Sprague Jr. (1998). In our research we only tested two of the factors that Galletta and Hufnagel (1992) first tested for affecting error finding performance. Here we primary check a set of hypotheses regarding the influence of the spreadsheet presentation on error finding performance; also we test whether overall experience affects error finding performance. These tests are bilateral because we not only checked whether spreadsheet presentation influences the amount of errors found but also the time needed for errors found. As stated in chapter four, the first pair of hypotheses is supported by the data showing that on screen spreadsheet presentation is more convenient for the error finding process; additionally students tend to find errors at less time when they work with spreadsheets on screen than on paper.

The second pair of hypotheses that we proposed in chapter three is about the error finding performance when working with formulas or with values. Following Galletta and Hufnagel (1992) we tested the performance of our subjects working both with values and formulas. This spreadsheet form however, did not help students to perform better on average than using a spreadsheet showing only values or only formulas. Although students working with formulas do not find significantly more errors or complete the whole task on considerable less time than those who work with values we can extract evidence to support the hypotheses by their performance rate. Specifically, students working with spreadsheets showing formulas use on average significantly less time per valid error found. It seems that students who worked with spreadsheets showing only values especially presented on paper, got confused and made them spent needless time to find errors that are not valid. Thus, these students took much more time to complete the whole task. These findings regarding our

second set of hypotheses are consistent with Galletta and Hufnagel (1992) as well as with Galleta *et al.,* (1996).

The third pair of hypotheses regarding the experience of each student is not at all supported by the data. This examination shows that experience is not at all consistent with the error finding performance, regardless the spreadsheet presentation. These findings contradict Thompson and Partridge (2001) who extended Galleta *et al.* (1996) and showed that prior incremental practice increased performance although they did not directly link performance with the spreadsheet presentation.

Generally, it is empirically evident, from this research that on screen presentation of spreadsheets in combination with the appearance of formulas is the most useful way for error finding. Additionally this combination presents the best results as far as the types of errors found is concerned. That is, using this kind of spreadsheets, auditors can find a large amount from all three prevailing types of errors even omission ones which are considered as the most difficult to reveal (Panko and Sprague Jr., 1998). Finally, we can highlight the great performance in finding logic errors when using the on screen with values presentation spreadsheets. Specifically students of this group found about seventy per cent of the planted logic errors. The phenomenon of higher performance in logic errors however is also observed, at a significantly smaller degree, in the other working groups. This is attributed to the relatively high familiarity of the students with the free cash flow concept, intuitively leading them to find more valid ones of the specific type.

Limitations

There are some limitations in our study that could potentially cause research biases. First of all the number participants is somewhat limited and extreme observations cannot be tolerable. This is because we wanted to address students of business administration and finance studies to maintain high domain experience. The confined pool of students relative to the spreadsheet's subject, in the specific institution, combined with the nature of the experiment do not allow us to refer to more students and normalise possible extreme observations.

Another serious limitation is the simplicity of the spreadsheet unlike real world spreadsheets. We note that the seeded errors into the experimental spreadsheet model are limited in number, somewhat simplistic in scope, and do not represent the spreadsheet errors in any particular application domain .The low number of errors planted in the spreadsheet and the spreadsheet simplicity however is necessary for the experiment because we needed to preserve the same domain knowledge and skills for all participants. With a simple spreadsheet we can be assured that all the participants possess the minimum required understanding not only of the concepts of the spreadsheet but also of the use of spreadsheets' structure. This way the sample is limited to a small and somewhat homogeneous population of university students who are not necessarily representative of the spreadsheet developers in the real world.

We also recognise that the variable referring to the experience and familiarity are self reported therefore potentially biased by the participant's subjective judgement and possible misunderstanding of what is required. Similarly, we realise that simplistic answers to such questions as "years of experience" mostly capture the quantity of such experience but do reflect the quality of such experience.

Further Research

With this study we extend previous work on spreadsheet error finding. The theoretical framework serves as a useful basis for investigating factors influencing error finding performance. Future research may expand the framework to include other variables and examine more complicated factors in the model. Larger sample sizes as well as more specialised respondents would enable future researchers to incorporate more variables and extract high validity of their results. Unfortunately, it is quite difficult to refer to participants with more complicated and real world spreadsheets retrieved from the industry as challenging as this could be. Additionally, this study highlights the attention future researchers should pay to the types of errors and their classification since different types of errors may require different error detection strategies. Furthermore research on qualitative errors, which is a rather complex activity, is proposed as we believe it would lead to illuminating results on error finding performance both for quantitative and for qualitative errors. Finally, we consider imperative that the spreadsheet development is based on good guidelines for students since they most likely will make frequent use and development in their working life.

References

AMA. Report on end-user and departmental computing. New York: American Management Association, 1988.

Allwood, C. M. (1984), "Error Detection Processes in Statistical Problem Solving", *Cognitive Science*, **8**(4), pp. 413-437.

Ayalew, M. Clermont, R. Mittermier, (2000), "Detecting errors in spreadsheets", *Proceedings of the European Spreadsheet Risks Interest Group,* Annual Conference University of Greenwich, pp. 51–62.

Bagnara, S., F., Stablum, A., Rizzo, A., Fontana and M., Ruo (1987), "Error Detection and Correction: A Study on Human-Computer Interaction in a Hot Strip Mill Planning and Production System", *Preprints of the First European Meeting on Cognitive Engineering Approaches to Process Control*, Marcoussis, France.

Beiser, B., (1990), *Software Testing Techniques*, Van Nostrand Reinhold, New York, second edition.

Brown, P.S. and J.D., Gould (1987), "An experimental study of people creating spreadsheets", *ACM,* **5**(1), pp.258-272.

Caulkins, J., E., Morrison and T., Weidemann (2006), "Spreadsheet errors and decision making: evidence from field interviews", *Journal of End User Computing*, **40**(3), pp. 115-38.

Chan, H.C, C., Ying and C.B., Peh (2000), "Strategies and visualization tools for enhancing user auditing of spreadsheet models", *Information and Software Technology*, **42**(8), pp. 1037-1043.

Clermont, M., (2002) "A spreadsheet auditing tool evaluated in an industrial context," *Proceedings of the European Spreadsheet Risks Interest,* Group Conference Cardiff Wales, pp. 35–46.

Connors, S. (1984), "NAAA research", *Management Accounting*, **21**(3), pp. 16-65.

Connors, S. (1983), "NAAA research", *Management Accounting*, **8**(2), pp. 62-63.

Cragg, P.G. and M., King (1993), "Spreadsheet Modelling Abuse: An Opportunity for OR?" *Journal of the Operational Research Society*, **44**(8), pp. 743-752.

Davis, P. (1997), "What computer skills do employers expect from recent college graduates?" *Technological Horizons in Education Journal*, **31**(7), pp. 74-78.

Dillon, A. (1992) "Reading from paper versus screens: a critical review of the empirical literature", *Ergonomics*, **35**(10), pp.1297–326.

Gable, G., C.s., Yap and M.N., Eng (1991) "Spreadsheet Investment, Criticality, and Control," *Proceedings of the Twenty-Fourth Hawaii International Conference on System Sciences*, Los Alomitos, CA: IEEE Computer Society Press, **17**(3), pp. 153-162.

Galletta, D. F., Hartzel, K.S., Johnson, S., and Joseph, J.L. (1997), "Spreadsheet Presentation and Error Detection: An Experimental Study", *Journal of Management Information Systems*, **13**(2), pp. 45-63.

Galletta, D.F. and E., Hufnagel (1992), "A Model of End-User Computing Policy: Determinants of Organizational Compliance", *Information and Management*, **22**(1), pp.18.

Galletta, D.F., D., Abraham, M., El Louadi, Y.A., Pollailis and J.L., Sampler (1993), "An empirical study of spreadsheet error-finding performance", *Journal of Accounting, Management, and Information Technology*, **3**(2), pp. 7-95.

Galletta, D.F., K., Hattzel, S., Johnson, J., Joseph and S., Rustagi (1996), "An Experimental Study of Spreadsheet Presentation and Error Detection", *Proceedings of the 29th Annual Hawaii International Conference on System Sciences*.

Gould, JD., Alfaro, L., Barnes, V., Finn, R., Grischkowsky, N. and Minuto, A. (1987), "Reading is slower from CRT displays than from paper: attempts to isolate a single-variable explanation", *Human Factors*, **29**(3), pp.269–99.

Hall, M.J.J. (1996), "A Risk and Control Oriented Study of the Practices of Spreadsheet Application Developers", *Proceedings of the Twenty-Ninth Hawaii International Conference on Systems Sciences*, **2**(5), Kihei, Hawaii,Los Alamitos, CA: IEEE Computer Society Press, pp. 364-373.

Heagy, C.D. and P.L., McMickle (1988), "An empirical investigation of the accounting systems course: Academic practice versus practitioner needs", *Issues in Accounting Education*, **44**(8), pp. 96- 107.

Heagy, C.D. and R.A., Gallun (1994), "Recommended microcomputer knowledge for accounting graduates: A survey", *Journal of Accounting Education*, **17**(8), pp. 205-210.

Howe, H. and M.G., Simkin (2006), "Factors affecting the ability to detect spreadsheet errors", *Decision Sciences Journal of Innovative Education*, **4**(1), pp. 101–122.

Janvrin, D. and J., Morrison (1996), "Factors Influencing Risks and Outcomes in End-User Development", *Proceedings of the 29th Annual Hawaii International Conference on System Sciences*

Kreie, J., T., Cronin, J., Pendley, J., Renwick (2000), "Applications development by end-users: can quality be improved?", *Decision Support Systems*, **29**(5), pp. 143–152.

Kruck, S.E. (2006), "Testing spreadsheet accuracy theory", *Information and Software Technology*, **48**(12), pp. 204–213.

S. Kruck, S. Sheetz, (2001) "Spreadsheet accuracy theory", Journal of Information Systems Education **12**(2) pp. 93–108.

Lee, D.M.S. (1986), "Usage pattern and sources of assistance for personal computer users." *MIS Quarterly*, **32**(3), pp. 313-325.

Mingers, J. (1991), "The content of MSc operational research courses: results of a questionnaire of OR groups", *Journal of Operational Research Society*, **42**(4), pp. 375-385.

Nardi, B.A. (1993), *"A Small Matter of Programming: Perspectives on End User Computing"*, MIT Press, Cambridge, Massachusetts.

Oliver, R. (1994), "Proof-reading on paper and screens: the influence of practice and experience on performance", *Journal of Computer-based Instruction*, **20**(4), pp.118–24.

Panko, R. (1996), "Hitting the wall: Errors in developing and debugging in a 'simple' spreadsheet problem", *Proceedings of the 29th Hawaii International Conference on System Sciences* pp. 356–363.

Panko, R. (1998) "What we know about spreadsheet errors", *Journal of End-User Computing*, **10**(3), pp.15–21.

Panko, R. (2000), "Spreadsheet Errors: What We Know. What We Think We Can Do", *Proceedings of the Spreadsheet Risk Symposium*, Greenwich, England.

Panko, R. (2006), "What we know about spreadsheet errors", *Journal of End User Computing's*, **10**(2), pp. 15-21.

Panko, R. and R., Halverson (1996), "Spreadsheets on trial: a survey of research on spreadsheet risks", *Proceedings of the 29th Annual Hawaii International Conference on Systems Sciences*, **14**(5) pp. 326–335.

Panko, R. and R., Sprague (1998), "Hitting the wall: errors in developing and code inspecting a 'simple' spreadsheet model", *Decision Support Systems*, **22**(9), pp.337–353.

Panko, R. and R.P., Halverson, Jr. (1996), "Spreadsheets on Trial: A Framework for Research on Spreadsheet Risks", *Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences*, Maui, Hawaii.

Panko, R. and R.P., Halverson, Jr. (1997) "Are Two Heads Better than One? (At Reducing Errors in Spreadsheet Modeling?", *Office Systems Research Journal*, **15**(1), pp. 21-32.

Panko, R.R. and R.P., Jr. Halverson (1996), "Spreadsheets on Trial: A Framework for Research on Spreadsheet Risks", *Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences,* Maui, Hawaii.

Powell, R., B., Kenneth and B., Lawson (2008), "A critical review of the literature on spreadsheet errors" *Decision Support Systems*, **17**(9), pp. 118-127

Pressman, R., and D., Ince (2000), "A Practitioner's Approach, A European Adaptation" *Software Engineering,* **2**(3), pp.23-32

Pryor, L. (2004), "When, why and how to test spreadsheets", *Proceedings of European Spreadsheet Risks Interest Group (EuSpRIG).*

Purser M., and D. Chadwick (2000), "Does an awareness of differing types of spreadsheet errors aid end-users in identifying spreadsheet errors?" *Proceedings of the European* Spreadsheet Risk Interest Group Annual Conference, Cambridge, UK, pp. 185–204.

Rajalingham, K., D., Chadwick, B., Knight and D., Edwards (2000), "Quality Control in Spreadsheets: A Software Engineering-Based Approach to Spreadsheet Development" *Proceedings of the 33rd Hawaii International Conference on System Sciences*

Rajalingham, K., D.R., Chadwick, and B., Knight (2001), "Classification of spreadsheet errors", *European Spreadsheet Risks Interest Group (EuSpRIG).*

Randolph, N., M., John and L., Gareth (2002), "A Generalised Spreadsheet Verification Methodology" *Australian Computer Society*

Ronen, B., M.A., Palley and H.C., Lucas (1989), "Spreadsheet analysis and design", *ACM*, **32**(1), pp.47-55

Saariluoma, P., Sajaniemi, J. (1989), "Visual information chunking in spreadsheet calculation", *International Journal of Man-Machine Studies,* **30**(2), pp.475-488.

Salchenberger, L. (1993), "Structured development techniques for user-developed systems," *Information and Management*, **24**(3), pp.41-50.

Schultheis, R. and 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), pp.11-18.

Simkin, M.G., (2004) "Ferret Out Spreadsheet Errors". *Journal of Accountancy*, **2**(2), pp. 62-67.
Teo, T. and M., Tan (1997), "Quantitative and qualitative errors in spreadsheet development", *Proceedings of the 30th Hawaii International Conference on Systems Sciences*.

Thompson, S.H., L., Partridge (2001), "Effects of error factors and prior incremental practice on spreadsheet error detection: an experimental study", *Omega*, **29**(5), pp. 445–456.

Thompson, T.S.H and M., Tan (1997), "Spreadsheet development and 'what-if' analysis: Quantitative versus qualitative errors", *Accounting Management and Information Techonologies*, **9**(2), pp.141–160.

Woods, D.D. (1984), "Some Results on Operator Performance in Emergency Events", *Institute of Chemical Engineers Symposium Series*, **90**(3), pp. 21-31.

Appendix

Table 10-The free cash flow spreadsheet

(Shaded cells are the places where errors are seeded)

DATA					
Cost of debt	10%		EBIT		6,00
Tax rate	40%		Investment		12,00
After-tax cost of debt	6%		Amount borrowed		4,00
Cost of equity	12%		*Subsidized interest rate*		*5%*
			Market value of debt		**2,00**
CASH FLOWS TO SHAREHOLDERS					
Year	**0**	**1**	**2**	**3**	**Goes on forever**
EBIT		6,00	6,00	6,00	
Less interest		(0,20)	(0,20)	(0,20)	
Earnings before taxes		5,80	5,80	5,80	
Taxes		(2,32)	(2,32)	(2,32)	
Net income		3,48	3,48	3,48	
Investment by shareholders	(8,00)	-	-	-	
Net cash flow to shareholders	(8,00)	3,48	3,48	3,48	
NPV at cost of equity	21,00				
Market value of equity	29,00				

WEIGHTED AVERAGE COST OF CAPITAL					
		Proportion	Weighted		
	Cost of	of Market	Cost of		
	Capital	Value	Capital		
Debt (after-tax)	6,0%	6,5%	0,4%		
Equity	12,0%	93,5%	11,2%		
Weighted average cost of capital			**11,6%**		
FREE CASH FLOW					
Year	0	1	2	3	Goes on forever
EBIT		6,00	6,00	6,00	
Taxes		(2,40)	(2,40)	(2,40)	-
EBIAT		3,60	3,60	3,60	
Investment	(12,00)	-	-	-	-
Free cash flow	(12,00)	3,60	3,60	3,60	
NPV at WACC	19,00				
AFTER-TAX CASH FLOWS FROM LENDERS					
Year	0	1	2	3	Goes on forever
Interest		(0,20)	(0,20)	(0,20)	
Interest tax shield		0,08	0,08	0,08	
After-tax interest		(0,12)	(0,12)	(0,12)	
Principal receipt (repayment)	4,00	-	-	-	
Principal plus after-tax interest	4,00	(0,12)	(0,12)	(0,12)	
NPV at after-tax cost of debt	2,00				

CONCLUSION					
Free cash flow	(12,00)	3,60	3,60	discount at WACC	
After-tax cash flows from lenders	4,00	(0,12)	(0,12)	discount at after-tax ordinary cost of debt	
Cash flows to shareholders	(8,00)	3,48	3,48	discount at cost of equity	
NPV of cash flows to shareholders = NPV of free cash flows					
+ NPV of after-tax cash flows from lenders					

Table 11-The free cash flow spreadsheet showing values

DATA					
Cost of debt	10%		EBIT		6,00
Tax rate	40%		Investment		12,00
After-tax cost of debt	6%		Amount borrowed		4,00
Cost of equity	12%		*Subsidized interest rate*		*5%*
			Market value of debt		2,00

CASH FLOWS TO SHAREHOLDERS					
Year	0	1	2	3	Goes on forever
EBIT		**6,00**	**6,00**	**6,00**	
Less interest		**(0,20)**	**(0,20)**	**(0,20)**	
Earnings before taxes		**5,80**	**5,80**	**5,80**	
Taxes		**(2,32)**	**(2,32)**	**(2,32)**	
Net income		**5,60**	**11,80**	**5,80**	
Investment by shareholders	**(8,00)**	**-**	**-**	**-**	
Net cash flow to shareholders	**(8,00)**	**3,28**	**11,80**	**5,80**	
NPV at cost of equity	**27,33**				
Market value of equity	**54,67**				
WEIGHTED AVERAGE COST OF CAPITAL					
	Cost of capital	Proportion of Market Value	Weighted Cost of Capital		
Debt (after-tax)	6,0%	3,5%	0,2%		
Equity	12,0%	96,5%	11,6%		
Weighted average cost of capital			**12,2%**		
FREE CASH FLOW					
Year	0	1	2	3	Goes on forever
EBIT		**6,00**	**6,00**	**6,00**	
Taxes		**(2,40)**	**(2,40)**	**(2,40)**	**.**
EBIAT		**3,60**	**3,60**	**3,60**	

Investment	**(12,00)**	-	-	-	.
Free cash flow	(12,00)	3,60	3,60	3,60	
NPV at WACC	**(12,00)**				
AFTER-TAX CASH FLOWS FROM LENDERS					
Year	**0**	**1**	**2**	**3**	**Goes on forever**
Interest		(0,20)	(0,20)	(0,20)	
Interest tax shield		0,08	0,08	0,08	
After-tax interest		(0,12)	(0,12)	(0,12)	
Principal receipt (repayment)	4,00	-	-	-	
Principal plus after-tax interest	4,00	(0,12)	(0,12)	(0,12)	
NPV at after-tax cost of debt	**2,00**				
CONCLUSION					
Free cash flow	**(12,00)**	**3,60**	**3,60**	**discount at WACC**	
After-tax cash flows from lenders	**4,00**	**(0,12)**	**(0,12)**	**discount at after-tax ordinary cost of debt**	
Cash flows to shareholders	**(8,00)**	**3,28**	**11,80**	**discount at cost of equity**	
NPV of cash flows to shareholders = NPV of free cash flows					
+ NPV of after-tax cash flows from lenders					

Table 12-The free cash flow spreadsheet showing formulas

DATA					
Cost of debt	0,1		EBIT		6
Tax rate	0,4		Investment		12
After-tax cost of debt	=B3*(1-B4)		Amount borrowed		4
Cost of equity	0,12		*Subsidized interest rate*		*0,05*
			Market value of debt		=F6*F5/B3
CASH FLOWS TO SHAREHOLDERS					
Year	0	1	2	3	Goes on forever
EBIT		=$F$3	=$F$3	=$F$3	
Less interest		=-$F$5*$F$6	=-$F$5*$F$6	=-$F$5*$F$6	
Earnings before taxes		=C14+C15	=D14+D15	=E14+E15	
Taxes		=-C16*$B$4	=-D16*$B$4	=-E16*$B$4	
Net income		=C16+C15	=D16+D14	=E16+E13	
Investment by shareholders	=-(F4-F5)	0	0	0	
Net cash flow to shareholders	=SUM(B18:B19)	=SUM(C17:C19)	=SUM(D18:D19)	=SUM(E18:E19)	
NPV at cost of equity	=B2+C20/B6				
Market value of equity	=C20/B5				

WEIGHTED AVERAGE COST OF CAPITAL

	Cost of Capital	Proportion of Market capital	Weighted Cost of Capital		
Debt (after-tax)	=B5	=F7/(F7+B22)	=B28*C28		
Equity	**=B6**	**=1-C28**	**=B29*C29**		
Weighted average cost of capital			=D28+C29		

FREE CASH FLOW

					Goes on forever
Year	0	1	2	3	
EBIT		**=$F$3**	**=$F$3**	**=$F$3**	-
Taxes		**=-C37*$B$4**	**=-D37*$B$4**	**=-E37*$B$4**	
EBIAT	-	**=C37+C38**	**=D37+D38**	**=E37+E38**	-
Investment	=-F4	0	0	0	
Free cash flow	**=B40**	**=C39+C40**	**=D39+D40**	**=E39+E40**	
NPV at WACC	=B41+C4/D30				

AFTER-TAX CASH FLOWS FROM LENDERS					
Year	**0**	**1**	**2**	**3**	**Goes on forever**
Interest		=-$F$6*$F$5	=-$F$6*$F$5	=-$F$6*$F$5	
Interest tax shield		=-C49*$B$4	=-D49*$B$4	=-E49*$B$4	
After-tax interest	-	=C49+C50	=D49+D50	=E49+E50	
Principal receipt (repayment)	=F5	0	0	0	
Principal plus after-tax interest	=B52+B51	=C52+C51	=D52+D51	=E52+E51	
NPV at after-tax cost of debt	=B53+C53/B5				
CONCLUSION					
Free cash flow	=B41	=C41	=D41	discount at WACC	
After-tax cash flows from lenders	=B53	=C53	=D53	discount at after-tax ordinary cost of debt	
Cash flows to shareholders	=B20	=C20	=D20	discount at cost of equity	
NPV of cash flows to shareholders = NPV of free cash flows					
+ NPV of after-tax cash flows from lenders					