SPREADSHEET ERROR DETECTION: AN EMPIRICAL EXAMINATION IN THE CONTEXT OF GREECE

Dimitrios Maditinos[1], Dimitrios Chatzoudes[2], Charalampos Tsairidis[3]

[1]*Kavala Institute of Technology, Agios Loukas, 654 04 Kavala, Greece*
[2, 3]*Democritus University of Thrace, Panagi Tsaldari 1, 69100 Komotini, Greece*
E-mails: [1]dmadi@teikav.edu.gr (corresponding author); [2]dchatzoudes@yahoo.gr;
[3]*xtsairid@pme.duth.gr*

Abstract. The personal computers era made advanced programming tasks available to end users. Spreadsheet models are one of the most widely used applications that can produce valuable results with minimal training and effort. However, errors contained in most spreadsheets may be catastrophic and difficult to detect. This study attempts to investigate the influence of experience and spreadsheet presentation on the error finding performance by end users. To reach the target of the study, 216 business and finance students participated in a task of finding errors in a simple free cash flow model. The findings of the study 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. Further research proposals and limitations of the study are, moreover, discussed.

Keywords: end user computing, spreadsheet error detection, spreadsheet presentation, user experience, Greece.

1. 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 if they incorporate 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 (Chan *et al.* 2000; Kruck, Sheetz 2001). With the advent of the personal computers in the 1980s and the rapid development of end-user computing, control of software development was handed over from the professionals to the end users (Kreie *et al.* 2000; Clermont 2002). As spreadsheets have diffused throughout industry, evidence has underlined that spreadsheets should not contain errors and that errors can be costly to

the organisations that are using spreadsheets (Ayalew *et al.* 2000; Caulkins *et al.* 2006; Panko, Halverson 1996a; Panko 1998).

Nowadays, spreadsheets are used widely and frequently for business decisions. Spreadsheet modelling is recognised as the most frequently used application in the modern industry, while it is also quite easy to use with minimum training (Kruck 2006). 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 all spreadsheets have high error rates (Panko 1998; Rajalingham *et al.* 2001). Although research have 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. The most important issue that alarms industry is the fact that spreadsheet errors can potentially cost large amounts of money (Galletta *et al.* 1993, 1996, 1997).

Errors in spreadsheets may arise due to a variety of reasons, ranging from the user's lack of understanding about the specifications and/or spreadsheet requirements, to errors arising from incorrectly entering the formulas or values (Panko 2006; *Powell et al.* 2008). In the present study, an experiment among business and finance students of the Technological Educational Institute (T.E.I.) of Kavala, Greece, is employed in order to investigate whether (a) user experience in spreadsheet modelling and (b) spreadsheet presentation affect the auditors' error finding performance. Under this framework, the error finding rates of everyday users, such as students, are investigated, while the reasons of the students' failure are being interpreted. Moreover, a categorisation of the factors that lead to students' success in spreadsheet error detection is being attempted.

The following section includes a literature review concerning the main variables of the present study. In the third and fourth section, the proposed hypotheses and the research methodology are being presented. The results, conclusions, study limitations and future research are discussed in the sections 5, 6, 7 and 8 respectively.

2. Literature review

Several studies have dealt 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. They used a sample of first year business undergraduate students who were taking an introductory course on spreadsheets using Microsoft Excel. Results 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 impact 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.

Howe and Simkin (2006) 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 identified 67 percent 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. Howe and Simkin (2006) 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 102 undergraduate MISstudents and 50 postgraduate Business Administration students to develop a model from a word problem that was 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 revealed 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 up a sample of 113 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 percent 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. According to the authors (Galletta *et al.* 1996), paper versus screen effects applies well to spreadsheet error finding.

The studies of Galletta *et al.* (1996) and Panko and Sprague (1998) that are presented above are used as main drivers for the present paper.

3. Hypotheses development

Galletta and Hufnagel (1992) revealed that detection of spreadsheet errors is a quite demanding task. Sometimes, even experienced auditors fail to detect errors in simple spreadsheets. This incompetence is attributed to numerous factors that can hinder the error finding performance (Thompson, Tan 1997; Simkin 2004).

Following the context of previews research in the field (Galletta, Hufnagel 1992; Galletta *et al.* 1996; Panko, Sprague 1998), the present study is testing whether (a) spreadsheet presentation and (b) user's experience affect the spreadsheet error detecting performance. Spreadsheet presentation refers to whether the spreadsheet is presented on screen or on paper, as well as to whether formulas or values appear (Kruck, Sheetz 2001; Clermont 2002; Powell *et al.* 2008). Spreadsheet experience refers to the familiarity that each respondent attributes to himself about working with spreadsheets.

Dillon (1992), Gould *et al.* (1987) and Oliver (1994) used two spreadsheet presentation patterns and found out that reading from the screen is generally faster than reading from the paper (printed form of the spreadsheet). Thus, spreadsheet presentation (screen versus paper) can influence the error finding performance. Thus, we hypothesize:

Hypothesis 1: More errors are found when spreadsheet is presented on screen than on paper.

Furthermore, besides testing the average amount of errors found, the average time that each subject spends to complete the task given is, moreover, investigated:

Hypothesis 2: Errors are found less frequently when the spreadsheet is presented on screen rather than on paper.

As said above, the second aspect of spreadsheet presentation 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:

Hypothesis 3: More errors are found when the auditor works with formulas (added next to actual numerical values).

Hypothesis 4: Errors are found more quickly when the auditor works with formulas (added next to actual numerical values).

Finally, following the context developed by Galletta and Hufnagel (1992), the present study is testing whether the subjects' experience in spreadsheets affects the error finding performance:

Hypothesis 5: More errors are found when the auditor has high domain experience.

Hypothesis 6: Errors are found more quickly by experienced auditors.

Hypotheses 1 and 2 study the first aspect of spreadsheet presentation (paper vs screen), while hypotheses 3 and 4 the second (values vs formulas). Finally, hypotheses 5 and 6 study the relationship between spreadsheet experience and error detecting performance.

4. Research methodology

4.1. Variable definition and measurement

4.1.1. Spreadsheet errors

There are various types of errors that appear on spreadsheets and they are examined in a number of relevant studies (Rajalingham *et al.* 2000; Ayalew *et al.* 2000; Caulkins *et al.* 2006). Spreadsheet error detection literature mainly deals with users' errors and not with software errors that have to do, mainly, with programming issues (Rajalingham *et al.* 2000). The most common distinction of users' errors is in quantitative and qualitative errors. Quantitative errors, usually, have the form of numerical entry errors, while qualitative errors take the form of poor spreadsheet design and format and could potentially lead to quantitative errors. 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 (Kruck, Sheetz 2001; Simkin 2004; Panko 2006; Kruck 2006).

Since the main objective of the present study is to investigate the impact of spreadsheet presentation and end user experience on spreadsheet error detecting performance, it is obvious that only quantitative errors will be accounted for. That is because qualitative errors are not, mainly, related to the end user, since they foremost concern the initial designer of the spreadsheet (Rajalingham *et al.* 2000).

Panko and Halverson (1996b), distinguished the three dominant types of quantitative errors (mechanical, logic and omission errors). Mechanical errors are defined as simple mistakes, such as mistyping a number (Salchenberger 1993). Logic errors involve mistakes in reasoning that leads to wrong formula entering and are considered as quite difficult to detect and correct (Allwood 1984). The most critical type of errors, however, are the omission errors that occur when something is left out; this type is considered as the most difficult to detect (Allwood 1984; Bagnara *et al.* 1987; Woods 1984). In the present study, an evaluation of a free cash flow model containing eight errors is being attempted. These errors represent the three error types mentioned above. More specifically, three mechanical errors, three logic errors and two omission errors were seeded to all participants of the experiment.

4.1.2. Spreadsheet experience

User experience is a composite factor that is measured with the use of three indicators: (a) self-reported spreadsheet familiarity, (b) free cash flow concept familiarity and (c) years of experience with spreadsheets. Howe and Simkin (2006), moreover, used respondents' age as an indicator of experience and investigated its impact on the error finding performance. However, age is not incorporated in the present study as an indicator of experience, since it does not represent the actual time that each respondent have spent working with spreadsheets (Randolph *et al.* 2002).

Self-reported spreadsheet familiarity and free cash flow concept familiarity were measured with the use of 3 questions (items) each (the five-point Likert scale was employed), while the years of spreadsheet experience were calculated as the difference between the students' age and the age of his admittance on the Department of Business Administration. At this point, it should be mentioned that students on the Department of Business Administration are being taught spreadsheet modelling since the first semester of their studies.

In order to verify the construct validity of the spreadsheet experience variable, (a) exploratory factor analysis and (b) reliability analysis with the use of the statistical package S.P.S.S. 14.0 was being conducted. Seven items were used in the appropriate analysis and all relative indexes were within the acceptable range (Kaiser-Mayer-Olkin, Bartlett's Test of Sphericity, Eigenvalue, Variance, Cronbach Alpha). After successfully completing the control for the construct validity of spreadsheet experience, the final score of the variable was calculated using the mean of the 7 items used.

4.2. Data collection (experimental method)

A thorough experiment was implemented in order to test the proposed hypotheses. The experiment was conducted with the use of a simple free cash flow evaluation spreadsheet and an answer sheet, where respondents were asked (a) to identify the cells containing errors and (b) to describe each error. Respondents were, also, asked to include some basic demographics (age, genre, school) and answer relative questions about their spreadsheet and cash flow concept familiarity.

The evaluation spreadsheet was delivered in five presentation forms. The first two forms were electronic, specifically a Microsoft Excel file, where only values were presented and the auditor had to point to each cell to see the formula on the formula bar and a Microsoft Excel file, where the respondent could see only formulas. The other three presentation patterns were printed spreadsheets; specifically a printed spreadsheet where the subject (auditor) could see values but no formulas at all, a printed spreadsheet where the subject could see only formulas and no values and, finally, a printed spreadsheet where the subject could see both formulas and values. The five forms of the test were named in a descriptive manner: (a) Excel with values, (b) Excel with formulas, (c) PDF with values, (d) PDF with formulas, (e) PDF with values and formulas.

The experimental material was addressed to students of the undergraduate and the postgraduate programme of the Business Administration Department of the Technological Educational Institute (T.E.I.) of Kavala, Greece. All participants had taken at least one spreadsheet modelling course using Microsoft's Excel, as well as one investment evaluation course and they were familiar with both Microsoft's Excel and the free cash flow evaluation concept.

The participants were 123 and they were split into five random groups. Each group was given one of the five spreadsheet forms discussed above. No time constraints were set, but participants were informed that they were to detect exactly eight errors. The process allowed to record the time each student needed in order to complete the task. Participants were asked not to collaborate with each other or use another kind of external help and were discretely supervised as to do so.

After the completion of the task, 123 valid responses were collected. The members of the research team recorded the total number of valid errors that each student was able to identify, as well as the time each one spent in finding them. In addition, after coding the collected data into an Excel spreadsheet, the errors each student found were classified according to the three error types mentioned above. Furthermore, the time each student used per found error was calculated. Using these data, an index of overall performance for each student was extracted. This index is expressed as the average time each student spent to find a valid error and algebraically calculated by the quotient: Total time spent / Valid errors found.

5. Results

5.1. Sample demographics

Fifty two percent of the students participated in the experiment (64 students) were undergraduate students in the Department of Business Administration, while 48 percent (59 students) were postgraduate students (23.5 percent in the Finance Stream and 24.5 percent in the Financial Information Systems stream). Furthermore, 70.7 percent (87 students) were male and 29.3 percent (36 students) female. Finally, 64.2 percent (79 students) were less than 24 years, while the rest 35.8 percent (44 students) were 25 years or older.

5.2. Spreadsheet presentation and detection performance

As mentioned above, the free cash flow model of the experiment was delivered in five distinct forms, each including eight errors. Students were separated into five random groups and each group was provided with the spreadsheet in one of these forms.

In order to conclude which of the five forms is the most efficient in terms of error detection, an examination of (a) the errors found, (b) the time spent for the completion of the experiment and (c) the error detection rate was conducted. Means, standard deviation and skewness were calculated for all cases.

When it comes to the valid errors found by auditors (students), the empirical analysis revealed that the group that worked with "Excel with values" found 70.63 percent of the errors seeded in the spreadsheet. The group that worked with "Excel with formulas" found 55.5 percent of the total errors. The group that worked with the "PDF with values" found 30 percent of the total seeded errors, which is the lowest performance. The other two groups both found half of the total seeded errors (50.5 and 52 percent respectively).

Standard deviation in all cases is low and, therefore, acceptable, while it seems to be relatively constant for all groups of students. Moreover, the skewness scores (less than 1 in all cases), show that all data follow the normal distribution.

Table 1. Errors found in each form – mean, standard deviation, skewness

	Errors' Mean	Standard Deviation	Skewness
Excel With Values	5.65 (70.63%)	1.22	–0,22
Excel With Formulas	4.44 (55.5%)	1.04	0,17
PDF With Values	2.40 (30%)	1.00	0,43
PDF With Formulas	4.04 (50.5%)	1.42	0,29
PDF With Values and Formulas	4.16 (52%)	1.24	0,65

A primary conclusion that can be drawn from the data presented in Table 1 is that the students' capability to point the cursor at each cell and see the formula is decisive in their performance to find valid errors. "Excel with values" seems, therefore, to be the most efficient form of error detection.

The above conclusion is, moreover, supported when the time that students spent in order to find exactly eight errors (valid or not valid) is being calculated. As shown in Table 2, the "Excel With Values" form is the one that helps students to identify eight errors faster that any of the other forms.

Additionally, it seems that standard deviation is relatively high when it comes to the last two forms of the experiment (PDF with formulas, PDF with values and formulas). The dispersion of time each student spent in order to complete the task in these two cases is significant (4.43 and 3.92 respectively). This means that there were students who completed the error finding task in a very short period of time and others that spent much more time doing so. Finally, skewness scores are acceptable in all cases, showing that all data follow the normal distribution.

Table 2. Time spent in each form – mean, standard deviation, skewness

	Time Mean	Standard Deviation	Skewness
Excel With Values	21.61 min.	1.16	0.10
Excel With Formulas	22.4 min.	1.41	–0.31
PDF With Values	30.6 min.	3.15	0.73
PDF With Formulas	26.4 min.	4.43	1.05
PDF With Values and Formulas	29.16 min.	3.92	–0.11

Table 3. Error detection rate – mean, standard deviation, skewness

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

Table 3 presents the error detection rate for each of the five spreadsheet forms (time spent in order to find a valid error). The data presented in Table 3 support the claims made after observing Tables 1 and 2, clearly pointing out that working with "Excel with values" is the most efficient method when it comes to error detection. When students are using an Excel file with values, they are able to find a valid error every 4.01 minutes, far less than with any other spreadsheet form. Furthermore, the dispersion of values (0.94) shows that the above performance is consistent. Finally, the skewness index (0.35) identifies a normal distribution, unlike in the case of "Excel with formulas" and "PDF with values". For the case of "PDF with values", dispersion is extremely high (8.56 minutes), showing that students had quite contradictory performances.

5.3. Error type analysis

The quality of all valid errors identified in the experiment is very interesting to be analysed. As mentioned earlier, the relevant literature makes three main error distinctions (mechanical errors, logic errors and omission errors) (Panko, Halverson 1996). Table 4 presents an analysis of the valid errors found, according to their category. The group of students that worked with the "Excel with values" form performed extremely well in finding logic errors (70 percent). The other 4 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 concluded that the "Excel with values" form is quite convenient for finding more logic errors.

Table 4. 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%

5.4. Hypotheses verification

In order to test the four first hypotheses of the present study, (a) cross tabulation analysis, (b) mean comparison and (c) paired samples tests were conducted. The verification of hypotheses 5 and 6 was made with the use of linear correlation. The presentation of the hypotheses testing is being attempted in three steps.

5.4.1. Hypotheses 1 and 2

As shown in Table 5, 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 (3.53). More specifically, working with the on screen presented spreadsheet led students to detect 1.85 more errors than the ones they detected when working on paper. This difference is, moreover, statistically significant ($t = 6.389$, *sig.* $= 0.000$).

Furthermore, the 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, when working with on screen presentation students needed on average 22 minutes to complete their task, while when working with the paper version they needed about 29 minutes. The average difference of 6.72 minutes is, also, statistically significant ($t = -10.428$, *sig.* $= 0.000$).

Finally, when dealing with the on screen version, students needed 4.68 minutes on average in order to detect a valid error, while when dealing with the paper version they needed 10.24 minutes on average. That means that students that worked with the on screen spreadsheet used on average 7.11 minutes less to find each valid error. This mean difference is, moreover, statistically significant ($t = 6.25$, *sig.* $= 0.000$).

Thus, it is concluded that hypotheses 1 and 2 are fully supported by the empirical data.

Table 5. Screen vs Paper (hypotheses 1 and 2)

	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	0.000
Time	22.02	1.34	28.72	4.20	–6.72	–10.428	0.000
Rate	4.68	1.24	11.79	6.64	–7.11	–6.25	0.000

5.4.2. Hypotheses 3 and 4

As shown in Table 6, the students that worked with spreadsheets showing formulas added next to actual values found about the same number of errors with those who worked with spreadsheets showing values only (4.21 and 4.02 respectively). However, the mean difference of 0.19 in favour of spreadsheets showing formulas is not statistically significant ($t = -0.823$, *sig.* $= 0.413$). Thus, it is concluded that the third hypothesis is rejected by the empirical data.

Moreover, students who worked with spreadsheets showing formulas completed the task in 25.99 minutes on average, while those who worked with spreadsheet showing values only needed 1.28 minutes more. Since the mean difference is not statistically significant ($t = 1.332$, sig. $= 0.187$), the fourth hypothesis is, also, rejected by the empirical data.

Despite the above, when taking into consideration the rate of finding each valid error, results seem to present a different point of view. More specifically, students who worked with spreadsheets showing values needed 2.43 more minutes in order to detect each valid error. The mean difference in that case is statistically significant ($t = 2.74$, *sig.* $= 0.008$), something that implies that working with spreadsheets showing formulas actually helps students to perform better in terms of errors found per minute. Although the fourth hypothesis cannot be fully supported according to the empirical data presented above, one could suggest that it could be perceived as partially supported.

Table 6. Values vs Formulas (hypotheses 3 and 4)

	Values		Formulas		Mean difference	T statistic	Sig.
	Mean	Standard Deviation	Mean	Standard Deviation			
Number of errors found	4.02	1.75	4.21	1.24	–0.19	–0.823	0.413
Time	27.27	1.75	25.99	1.24	1.28	1.332	0.187
Rate	9.23	1.75	6.80	1.24	2.43	2.74	0.008

5.4.3. Hypotheses 5 and 6

In order to test hypothesis 5 (more errors are found when the subject has high domain experience), a correlation between the number of valid errors detected and the experience of the user was conducted. Results failed to support the hypothesis, since the correlation coefficient and the significance index were non satisfactory (0.086 and 0.342 respectively). The same case was for hypothesis 6 (errors are found more quickly by experienced auditors) (0.030 and 0.741 respectively). Therefore, hypotheses 5 and 6 are rejected by the empirical data (it could be advocated that experience do not at all influence the error detection performance).

6. Conclusions

The present paper is based on the relevant studies conducted by Galletta and Hufnagel (1992), Galletta *et al.* (1996) and Panko and Sprague (1998). The main objective of the study is to test whether spreadsheet presentation and overall user experience influence the error detecting performance. The tests conducted are bilateral, since they do not only check whether spreadsheet presentation and overall experience influence the number of errors found, but also the time needed for these errors to be found.

The first two hypotheses are supported by the empirical data, showing that on screen spreadsheet presentation is more convenient in the error detection process. Additionally,

students tend to find errors less frequently when they work with spreadsheets on screen rather than on paper.

The second two hypotheses tested the error detection performance when working with spreadsheets containing values and spreadsheets containing values and formulas, but they were both rejected by the empirical data. More specifically, it was proved that the spreadsheet form presented with formulas failed to 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, the hypothesis can be supported when investigating the empirical data concerning the performance rate (average time in detecting each error). 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 spent needless time to find errors that were not valid. Thus, these students took much more time to complete the whole task. The findings regarding the second set of hypotheses are consistent with the studies of Galletta and Hufnagel (1992) and Galleta *et al.* (1996).

The third pair of hypotheses regarding the experience of each student is not at all supported by the empirical data. The empirical examination showed that experience is not at all related with the error detection performance. These findings contradict Thompson and Partridge (2001) who extended the study of Galleta *et al.* (1996) and showed that prior incremental practice increased error detection performance.

Generally, it is empirically evident that on screen presentation of spreadsheets in combination with the appearance of formulas is the most useful way for error detection. Additionally, the above combination presents the best results as far as the types of errors found. Using this method, auditors can detect a large amount from all three prevailing types of errors; even omission ones, which are considered as the most difficult to detect (Panko, Sprague 1998).

Finally, the satisfactory performance of on screen spreadsheets with value presentation in detecting logic errors should be underlined. Specifically, students of the specific group found about seventy percent of the planted logic errors. The phenomenon of higher performance in logic errors is, however, also observed, at a smaller degree, in the other four 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 errors of the specific type.

7. Study limitations

A limitation of the present study is the relatively small number of the participants in the experiment. This may lead in recording extreme observations that cannot be tolerable by the statistical analysis. The main reason for not extending the pool of subjects (students)

was due to the need to maintain high domain experience. Therefore, addressing only to students of business administration and financial studies was a one way road. Thus, the confined student pool did not allow referring to more students and normalising possible extreme observations.

Another limitation of the study is the simplicity of the spreadsheet provided for evaluation in comparison with real world spreadsheets. One could argue that the seeded errors are limited in number, somewhat simplistic in scope, and do not represent the spreadsheet errors in any particular application domain. However, the low number of errors planted in the spreadsheet and the spreadsheet simplicity are necessary for the experiment, because the same domain knowledge and skills for all participants need to be preserved. With a simple spreadsheet the research team 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. In 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.

Finally, it is recognised that the variable referring to the experience and familiarity are self reported, therefore, potentially could be biased by the participant's subjective judgement and possible misunderstanding of what is required.

8. Further research

These study extents previous surveys 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. Finally, 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.

References

Allwood, C. M. 1984. Error Detection Processes in Statistical Problem Solving, *Cognitive Science* 8(4): 413–437. http://dx.doi.org/10.1207/s15516709cog0804_5

Ayalew, Y.; Clermont, M.; Mittermier, R. 2000. Detecting errors in spreadsheets, *Spreadsheet Risks, Audit and Development Methods* 1: 51–62.

Bagnara, S.; Stablum, F.; Rizzo, A.; Fontana, A.; Ruo, M. 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.

Caulkins, J.; Morrison, E.; Weidemann, T. 2006. Spreadsheet errors and decision making: evidence from field interviews, *Journal of End User Computing* 40(3): 115–138.

Chan, H. C.; Ying, C.; Peh, C. B. 2000. Strategies and visualization tools for enhancing user auditing of spreadsheet models, *Information and Software Technology* 42(8): 1037–1043.
http://dx.doi.org/10.1016/S0950-5849(00)00130-0

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

Dillon, A. 1992. Reading from paper versus screens: a critical review of the empirical literature, *Ergonomics* 35(10): 1297–1326. http://dx.doi.org/10.1080/00140139208967394

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

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

Galletta, D. F.; Hartzel, K. S.; Johnson, S.; Joseph, J. L.; Rustagi, S. 1996. An Experimental Study of Spreadsheet Presentation and Error Detection, in *Proceedings of the 29th Annual Hawaii International Conference on System Sciences*. http://dx.doi.org/10.1109/HICSS.1996.495417

Galletta, D. F.; Hufnagel, E. 1992. A Model of End-User Computing Policy: Determinants of Organizational Compliance, *Information and Management* 22(1): 1–18.
http://dx.doi.org/10.1016/0378-7206(92)90002-W

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

Howe, H.; Simkin, M. G. 2006. Factors affecting the ability to detect spreadsheet errors, *Decision Sciences Journal of Innovative Education* 4(1): 101–122. http://dx.doi.org/10.1111/j.1540-4609.2006.00104.x

Janvrin, D.; Morrison, J. 1996. Factors Influencing Risks and Outcomes in End-User Development, in *Proceedings of the 29th Annual Hawaii International Conference on System Sciences.*
http://dx.doi.org/10.1109/HICSS.1996.495418

Kreie, J.; Cronin, T.; Pendley, J.; Renwick, J. 2000. Applications development by end-users: can quality be improved?, *Decision Support Systems* 29(5): 143–152.
http://dx.doi.org/10.1016/S0167-9236(00)00068-3

Kruck, S. E. 2006. Testing spreadsheet accuracy theory, *Information and Software Technology* 48(12): 204–213. http://dx.doi.org/10.1016/j.infsof.2005.04.005

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

Oliver, R. 1994. Proof-reading on paper and screens: the influence of practice and experience on performance, *Journal of Computer-based Instruction* 20(4): 118–124.

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

Panko, R. 2006. What we know about spreadsheet errors, *Journal of End User Computings* 10(2): 15–21.

Panko, R. R.; Halverson, R. P. 1996a. Spreadsheets on Trial: A Framework for Research on Spreadsheet Risks, in *Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences,* Maui, Hawaii.

Panko, R.; Halverson, R. 1996b. Spreadsheets on trial: a survey of research on spreadsheet risks, in *Proceedings of the 29th Annual Hawaii International Conference on Systems Sciences* 14(5): 326–335. http://dx.doi.org/10.1016/S0167-9236(97)00038-9

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

Powell, R.; Kenneth, B.; Lawson, B. 2008. A critical review of the literature on spreadsheet errors, *Decision Support Systems* 17(9): 118–127.

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

Rajalingham, K.; Chadwick, D.; Knight, B.; Edwards, D. 2000. Quality Control in Spreadsheets: A Software Engineering-Based Approach to Spreadsheet Development, in *Proceedings of the 33rd Hawaii International Conference on System Sciences*. http://dx.doi.org/10.1109/HICSS.2000.926736

Randolph, N.; John, M.; Gareth, L. 2002. A Generalised Spreadsheet Verification Methodology, *Australian Computer Science Communications* 24(1): 215–222.

Salchenberger, L. 1993. Structured development techniques for user-developed systems, *Information and Management* 24(3): 41–50. http://dx.doi.org/10.1016/0378-7206(93)90046-V

Simkin, M. G. 2004. Ferret Out Spreadsheet Errors, *Journal of Accountancy* 2(2): 62–67.

Thompson, S. H.; Partridge, L. 2001. Effects of error factors and prior incremental practice on spreadsheet error detection: an experimental study, *Omega* 29(5): 445–456. http://dx.doi.org/10.1016/S0305-0483(01)00037-8

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

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

FINANSINIŲ SKAIČIUOKLIŲ KLAIDŲ APTIKIMAS: EMPIRINIS TYRIMAS GRAIKIJOJE

D. Maditinos, D. Chatzoudes, Ch. Tsairidis

Santrauka

Asmeninių kompiuterių eros pažanga leido eiliniam vartotojui atlikti sudėtingus finansinius skaičiavimus ir programavimo užduotis. Finansinės skaičiuoklės yra viena iš plačiausiai naudojamų kompiuterinių programų, kurios leidžia pasiekti sudėtingus skaičiavimo rezultatus po minimalių apmokymų ir su minimaliomis pastangomis. Tačiau daugumoje skaičiuoklių randamos klaidos, kurios gali būti katastrofiškos ir jas sunku aptikti eiliniam vartotojui. Šis tyrimas siekia įvertinti ryšį tarp skaičiavimo klaidos nustatymo efektyvumo, eilinio vartotojo patirties ir skaičiuoklės rezultato pateiki-mo. Tyrime dalyvavo 216 verslo ir finansų studentų, studijuojančių Graikijos aukštosiose mokyklose. Tyrimo metu buvo analizuojama laisvųjų pinigų srautų modelio finansinė skaičiuoklė. Tyrimo išvados rodo, kad klaidos nustatymo efektyvumui yra svarbus skaičiuoklės pateikimas ir klaidų paieškos kokybė, o vartotojų patirtis neturi tokio stipraus poveikio. Straipsnyje aptariami tyrimo apribojimai ir pateikiami tolesnių tyrimų pasiūlymai.

Reikšminiai žodžiai: eilinis finansinės skaičiuoklės vartotojas, skaičiuoklės klaidų aptikimas, skaičiuoklės pateikimas, vartotojo patirtis, Graikija.

Dimitrios MADITINOS is associate Professor of Information Technology, Finance and Financial Modelling in Kavala Institute of Technology, School of Business and Economics, Greece. He holds degrees in Business Administration with specialisation in Information Technology from Lund University, Sweden, and a PhD in Finance from the Business School, Greenwich University, UK. Before becoming a full academic he was working in senior position for Greek Productivity Centre as responsible for professional training in Information Technology and Management. His research interests are in financial modelling, performance measurement systems, investors' behaviour, financial information systems, ERPs, electronic commerce and e-business.

Dimitrios CHATZOUDES holds a Bachelor Degree in Business Administration from the Technological Educational Institute (T.E.I) of Kavala, Greece and a Postgraduate Degree in International Economics from the Democritus University of Thrace, Greece. He is currently a PhD Candidate at the Department of Production and Management Engineering of the Democritus University of Thrace, Greece. His research interests include Production Management, Operations Management and Research Methodology.

Charalampos TSAIRIDIS is an Assistant Professor at the Department of Social Administration of Democritus University of Thrace, Komotini, Greece. He studied Mathematics and completed his PhD thesis at the University of Ioannina, Ioannina, Greece. Topic of thesis: Statistical Information Theory and censoring. His main research interests are in the area of Statistical Information Theory, Statistical Data Analysis and Social Statistics.