



# Effects of error factors and prior incremental practice on spreadsheet error detection: an experimental study

Thompson S.H. Teo \*, Joo Eng Lee-Partridge

*Department of Decision Sciences, Faculty of Business Administration, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore*

Received 3 December 1999; accepted 15 May 2001

---

## Abstract

Previous research has shown that spreadsheet errors are common and are not easily detected. In this paper, an experiment was conducted to examine the rate of detection of both quantitative errors and qualitative errors in two domain-free spreadsheets. A detailed list and explanation of the types of common spreadsheet errors are presented. Results showed that the ability to detect various types of errors appears to be dependent on the type and prominence of errors as well as prior incremental practice with spreadsheet error detection. Implications of the findings are discussed. © 2001 Elsevier Science Ltd. All rights reserved.

*Keywords:* Debugging; Detection of errors; Qualitative errors; Quantitative errors; Spreadsheet errors

---

## 1. Introduction

Spreadsheets are used for a wide variety of tasks including financial analysis, budgeting and forecasting. Such tasks are often critical in facilitating decision-making. Consequently, errors in spreadsheets can often result in incorrect or “less than optimal” decisions being made, which may result in undesirable outcomes. Research has shown that spreadsheet errors can result in losses ranging from hundreds of thousands to millions of dollars [1–3]. Furthermore, the pervasiveness of spreadsheet use in the business world has made the problem of errors a serious matter. It is therefore crucial that spreadsheet errors are eliminated or minimized through proper design and systematic inspection for errors.

Researchers have examined a wide range of issues pertaining to spreadsheets including errors at different life cycle stages [4,5], policies on development and implementation [6,7] and user work practices [8,9]. Field audits and experiments at different life cycle stages, e.g., cell-entry stage [10], development or draft stage [11,4], debugging stage [12,13]

and operational stage [14,15] have generally found spreadsheet errors to be common and difficult to detect. Indeed there is some evidence that between one-third and one-half of all spreadsheets contain errors [16–18].

In fact, Panko and Halverson [18] noted that “every study that has looked for errors has found them and has found them in considerable abundance” (p. 327). One primary reason is that end-users tend to view spreadsheets as simple tools. Consequently, they are often over confident about the error-free nature of their spreadsheets [19]. This leads to the “blind” acceptance of spreadsheet solutions with no thorough examination of the accuracy or the validity of the spreadsheet model. This problem is compounded by poor spreadsheet design which makes error detection and correction difficult.

A substantial amount of research has described how errors are made during spreadsheet development, estimated its effects, and/or prescribed techniques or practices to mitigate them [20–23]. Research has also examined the effect of team size in reducing errors [24], patterns of spreadsheet errors [25,26], cognitive difficulties [27], overconfidence in spreadsheet modeling [28], and the impact of individual factors such as expertise and experience [12] and presentation factors such as screen versus paper and location of

---

\* Corresponding author. Tel.: +65-874-3036; fax: +65-779-2621.

*E-mail address:* fbateosh@nus.edu.sg (T.S.H. Teo).

formulas, on error detection [2]. Galletta et al. [12] emphasized that future research should focus more on the detection of spreadsheet errors since there is “no silver bullet” to prevent errors. Although software is available for auditing common errors, limitations in its features have made automatic detection of all types of errors impossible.

This paper describes two experiments to examine the detection of both quantitative and qualitative spreadsheet errors. In both experiments, a domain-free spreadsheet problem was used in order to rule out threats to validity caused by differences in task domain knowledge among participants in the experiments. Also, the occurrence of poor fit between tasks and domain knowledge is minimized or eliminated through the use of a domain-free spreadsheet problem.

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 (e.g., prominence of different instances of the same type of errors) affects detection rates. The results should enable us to better understand spreadsheet error detection and provide some insights as to the ease or difficulty of detecting certain types of spreadsheet errors.

The paper is organized as follows. First, we introduce the theoretical framework that serves as the basis for this study. Second, hypotheses are postulated based on the theoretical framework. Third, the sample and experimental procedures are described. Fourth, the results are presented and discussed. Fifth, contributions of the study and implications for researchers and practitioners are examined.

## 2. Theoretical framework

The theoretical framework for this study is derived from Galletta et al.’s [2,13] work on spreadsheet error-finding as shown in Fig. 1. Galletta et al. proposed that four main factors affect error-finding performance, namely error factors, individual factors, presentation factors and external factors.

### 2.1. Error factors

Certain errors may be more prominent than others. Such errors may therefore be more easily detected because they are more conspicuous and/or their values are inconsistent with the users’ notion of an appropriate or common value. The complexity of the spreadsheet or formula containing the error is also important because a complex spreadsheet (especially if poorly designed) makes error finding difficult. The quantity and type of errors may make the spreadsheet user more aware of the presence of errors. Since errors can be classified into various types, the next section examines the types of errors in greater detail.

#### 2.1.1. Types of errors

In general, spreadsheet errors may be classified into two main types: quantitative and qualitative (summarized in Table 1). Quantitative errors result in incorrect bottom-line values while qualitative errors are the outcome of poor spreadsheet design (which may potentially lead to incorrect bottom-line values). The distinction between the different types of errors is important as they may require different strategies for error reduction or correction.

Quantitative errors may be classified into three main types: mechanical, logic and omission [4]. Mechanical errors usually result from carelessness, mental overload and distractions, e.g., simple slips such as mistyping a number, pointing to a wrong cell address, or selecting an incorrect range of values or cells. In contrast, logic errors are more complex as they are often related to the cognitive processes of the spreadsheet developer. For instance, in order to determine whether the various steps used to arrive at the solution are appropriate or correct, one may need to understand the cognitive processes of the spreadsheet developer as incorrect formulas caused by incorrect algorithms would affect the solution. An example of a logic error caused by incorrect algorithms is the addition of profit margin to fixed costs only, rather than total costs (fixed costs plus variable costs). As the name implies, omission errors usually result from leaving out something that should be in the spreadsheet model. An example would be to omit the number of laborers per team in the computation of labor costs. As with logic errors, omission errors can be difficult to detect.

Two other categories of quantitative errors were suggested by Galletta et al. [2,12,13] namely, domain errors (e.g., error in an accounting concept) and device errors (e.g., omitting a cell or typing an incorrect cell address). Domain errors are analogous to logic errors while device errors are analogous to both omission and mechanical errors. In this study, we will adopt Panko’s classification of quantitative errors since it is more comprehensive and easier to define.

The second type of errors, namely qualitative errors, are the outcome of poor spreadsheet design. They may be measured using Panko’s [29] two laws of good spreadsheet modeling. Violations of these two laws have been used successfully by Teo and Tan [30] as measures of qualitative errors. The first law emphasizes that a variable shall be defined in only one of two ways; either as a single number or as a formula not containing any numbers. Violation of the first law results in more than one variable (e.g., length and breadth) being placed in a single cell. This type of error is known as jamming error. Although jamming error, by itself, may not result in incorrect values, it makes it harder for the developer to understand and make changes to the spreadsheet model. This may potentially lead to incorrect changes, thereby resulting in incorrect bottom-line values.

The second law emphasizes that information pertaining to any variable shall not be repeated or duplicated in different

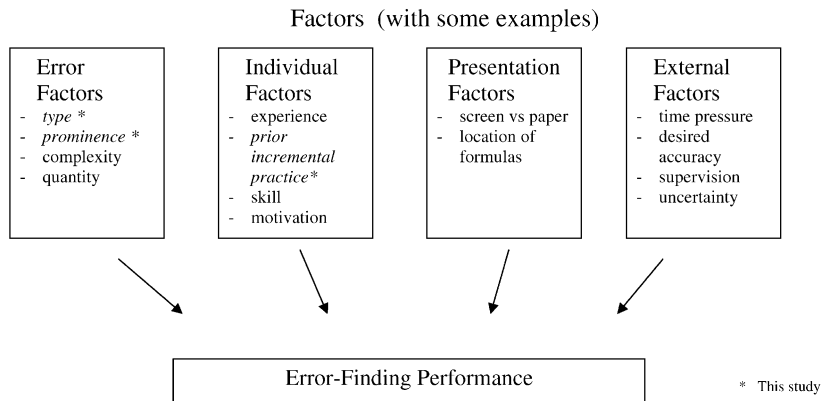


Fig. 1. Theoretical framework (Adapted from [13,14]).

Table 1  
Types of errors

Types of errors	Description
<i>Quantitative</i>	
<i>Classification by Panko [28]</i>	
1. Mechanical	Result from carelessness, mental overload and distractions, e.g., simple slips such as mistyping a number, pointing to a wrong cell address.
2. Logic	Relate to the cognitive processes of the spreadsheet developer, e.g., incorrect algorithms used to solve a problem.
3. Omission	Result from leaving out something that should be in the spreadsheet, e.g., omission of the number of laborers per team in the computation of labor costs.
<i>Classification by Galletta et al. [12,13,2]</i>	
1. Domain errors	Error in a particular domain area, e.g., accounting concept. Related to logic errors.
2. Device errors	Errors involved in use of device (i.e., spreadsheet software) on which spreadsheet model is constructed. Related to mechanical and omission errors.
<i>Qualitative</i>	
Outcome of poor spreadsheet design (which may potentially lead to incorrect bottom-line values).	
<i>Classification by Panko [29]</i>	
1. Jamming	Result of more than 1 variable (e.g., length and breadth) being placed in a single cell.
2. Duplication	Repeated or duplicated information in different cells of the spreadsheet.

cells of the spreadsheet. For example, if the value for length (as a variable used in the spreadsheet) is duplicated in the adjacent cell, it would violate the second law. This type of error is known as duplication error and may result in data inconsistency since if the value of length is changed, the spreadsheet developer may forget to change the value in the duplicated cell. Consequently, bottom-line values may also be affected.

In summary, qualitative errors are violations of proper spreadsheet design maxims, as opposed to quantitative errors, which are errors leading to incorrect results. Qualitative errors, by itself, do not necessarily lead to incorrect results, but they are also important because such errors may potentially lead to incorrect results if a spreadsheet is later

modified and the presence of qualitative errors results in incorrect modification.

### 2.2. Individual factors

Individual factors include a wide range of individual characteristics such as the degree of experience or expertise with the problem domain (e.g., cost accounting in a pricing model), experience or expertise with the device (in this case, the spreadsheet software), and motivation and skill in performing the task of finding errors (e.g., incentives offered).

In this study, we extend the notion of experience by examining the effect of prior incremental practice defined as giving participants a similar spreadsheet problem immediately

after they have worked on a spreadsheet problem. While experience usually refers to expertise gained over a period of time, prior incremental practice refers to “expertise” gained within a relatively short time span as the result of working on a similar problem. Unlike experience, the effect of prior incremental practice has not been examined in past research.

### 2.3. Presentation factors

The third set of factors that may influence error finding pertains to the presentation of information to the spreadsheet user. For example, research in human–computer interaction [31–34] has generally found that reading from paper is generally about 20% to 30% faster than reading on screen. Thus, presentation media (screen versus paper) can influence error-finding performance especially in the presence of time constraints.

Another aspect of presentation factors is the presence of spreadsheet formulas during error finding. For example, although Galletta et al. [2] found no significant differences in error-finding performance for subjects provided with spreadsheet formulas versus those not provided, there was some evidence that providing formulas in an integrated manner (i.e., formulas are provided together with printed spreadsheet) may reduce the number of false positives (correct items flagged as errors) when compared to other treatment conditions.

### 2.4. External factors

The last category of factors that may influence error-finding performance is external factors such as time pressure and desired accuracy which may affect the amount of effort spent in detecting errors. The desired accuracy may be dependent on the criticality of the spreadsheet solution, e.g., a spreadsheet solution involving an expenditure of millions of dollars is likely to be more closely scrutinized compared to say, a spreadsheet solution pertaining to expenses for a small-scale office function.

The extent of supervision may also affect error detection since poor supervision may result in cursory review procedures. Poor supervision may also cause the spreadsheet developer to view the spreadsheet problem as non-critical and trivial, thereby ignoring the need to systematically check for errors. On the other hand, if lack of supervision is due to employee empowerment, it may lead to greater employee confidence and sense of responsibility and more careful review procedures. Finally, overall uncertainty in the problem solved by the spreadsheet can make it difficult to detect incorrect items in the model.

## 3. Hypotheses

Galletta et al.’s [2,13] studies have focused solely on presentation factors. In this paper, we extend Galletta et al.’s

work by examining the effects of error factors (in terms of type and prominence) and individual factors (in terms of prior incremental practice) on error-finding performance. The specific items examined are shown in italics in Fig. 1.

### 3.1. Error factors

#### 3.1.1. Type of errors (*quantitative versus qualitative errors*)

With the exception of Teo and Tan’s [30] study, none of the existing research on spreadsheet error detection examined both quantitative and qualitative errors. Most research focused solely on the detection of quantitative errors. For example Panko and Sprague [26] examined the detection of quantitative errors by undergraduate and MBA students while Galletta et al. [12] examined the detection of domain and device errors (which are also quantitative errors) using CPAs and MBA students.

In the only study examining both quantitative and qualitative spreadsheet errors, Teo and Tan [30] found that the number of qualitative errors (jamming and duplication errors) made during spreadsheet design for Panko’s wall problem (explained later in the Method section) is about nine times that of quantitative errors (mechanical, logic and omission errors). It seems that spreadsheet developers commonly “jammed” more than one variable in a single cell or duplicate cell information for convenience. This implies that it may be easier to detect quantitative errors rather than qualitative errors since designers may view the latter as less significant than the former (which usually impacts directly on the bottom-line solutions).

Furthermore, since qualitative errors result from poor spreadsheet design, the spreadsheet designer may not actually be aware of good spreadsheet design practices. Consequently, the focus in error detection may be more on quantitative errors that directly affects bottom-line values. We therefore hypothesize that:

**Hypothesis 1.** Quantitative errors are more easily detected than qualitative errors.

#### 3.1.2. Prominence of errors

In a study of spreadsheet error-finding, Galletta et al. [12] found that subjects were able to detect about 55% of the simple, conspicuous errors (defined in terms of domain and device errors) in a given spreadsheet model. Other studies by Galletta et al. [2,13] using different media (screen versus paper) and different presentation of formulas (conventional formulas versus no formulas versus integrated formulas) also found that subjects generally detected about 50% of the errors. Subjects with printed spreadsheets generally found more errors than those with screen-only spreadsheets but they took a longer time to do so. No significant differences were found among groups with different formula treatments. It is conceivable that since the subjects in the formula treatments were exposed to exactly the same seeded

errors, the prominence of the seeded errors is the same, thereby resulting in insignificant differences among treatments.

Different groups of subjects working on the same spreadsheet problem should produce approximately the same result (assuming that groups are equivalent since they were randomly assigned to treatments). However, if the spreadsheet problem was similar (but not exactly the same due to the different nature of seeded errors), there may be some differences in error detection rates between groups. In other words, within each type of error (e.g., mechanical, logic), the actual nature of the error (e.g., different instances of mechanical errors) may influence error detection rates since some errors are more obvious or prominent than others.

Davis and Olson [35] emphasized that the “ability of human information processors to identify differences may be important in detecting errors (i.e., noticing differences between correct and incorrect data) and also in their reactions to variations in data they receive” (p. 246). They also commented that Weber’s law of just noticeable differences (which states that the difference that is noticeable is a constant proportion of the physical dimensions of the stimulus) may be important in explaining how individuals determine whether a particular value is appropriate. For example, an error of \$10 in the appropriate value of \$50 is more likely to be detected as compared to an error of \$10 in an appropriate value of \$500. Hence, in this study, we define prominence of errors in terms of whether the errors are conspicuous and inconsistent with the participant’s notion of an appropriate or common value. This definition also includes errors that may have a greater quantitative impact as well as its position on the spreadsheet. We therefore hypothesize that:

**Hypothesis 2.** The prominence of seeded errors will influence spreadsheet error detection rates.

### 3.2. Individual factors

#### 3.2.1. Effects of prior incremental practice

Research has shown that experience may influence self-efficacy, which has both direct and indirect effects on computer usage [36]. This is important since self-efficacy enables the user to use computers more efficiently and effectively. Such experience or expertise has also been shown to be important in program debugging [37] since it helps programmers to better examine and understand sequences in computer commands and in finding errors. Consequently, experience is also likely to play an important role in the detection of spreadsheet errors. In fact, there is some empirical evidence on the role of expertise or experience in spreadsheet error finding. For example, Galletta et al. [12] reported that subjects with accounting and spreadsheet expertise found the most spreadsheet errors in the shortest time compared to other subjects. However, spreadsheet

experience only facilitated the speed of error detection, and not the quantity of errors detected.

In the context of our study, we extend the notion of experience by examining the effects of prior incremental practice defined as having participants work on a spreadsheet problem, and then immediately given another similar spreadsheet problem to work on. Hence, if the participant is given a spreadsheet error detection practice problem and then immediately given a similar spreadsheet problem (i.e., matched-pair samples), the rate of error detection should improve due to knowledge gained from the earlier incremental practice session. A similar situation would arise if we have independent samples, i.e., if one group of participants is given some incremental practice on spreadsheet error detection, it should do better than another group with no prior incremental practice.

Intuitively, experience leads to better error detection. Of greater interest is the question of whether prior incremental practice also leads to better error detection. We therefore hypothesize that:

**Hypothesis 3.** Subjects with prior incremental practice in error detection will be able to detect more errors than subjects without prior incremental practice.

## 4. Method

### 4.1. Sample and procedures

The sample comprised first year business undergraduate students who were taking an introductory course on spreadsheets using Microsoft Excel. We chose business students because such students are likely to be potential users of spreadsheets later in their working life. An experiment was designed as part of a mid-term test. Before taking the mid-term test, students were taught the fundamentals of spreadsheet design and use for computation and problem solving. Students were not explicitly taught to look for errors in spreadsheet.

Specifically, students had 3 lectures (1 h each) on several aspects of Excel like relative and absolute referencing, the steps to take in designing spreadsheets, formulas, goal seek, data table, and what-if analysis. They also had 4 tutorials of 2 h each on working on building various spreadsheets. Students were made aware that good spreadsheet design should not have jamming and duplication, but they are not told explicitly that these are considered as errors. The reason for doing so is that we wanted to test whether students remember good spreadsheet design principles when they check spreadsheets created by others.

Further, since we did not specifically teach students about the three types of quantitative errors prior to the experiments, there are no compelling reasons to highlight that jamming and duplication are considered errors. Instead, we highlighted that good spreadsheet design should not have

Table 2  
Description of spreadsheet errors

Types of errors	Spreadsheet UNCLE	Spreadsheet DAD
<i>Quantitative</i>		
Mechanical	Number of hours per day = 7 (not 8)	Wall thickness = 0.5 m (not 5 m)
Logic	Incorrect cell reference used in formula for profit IF(C29 < 1600, C29 * B6, Formula should be C29 * B5) IF(C29 < 1600, C29 * B5, C29 * B6)	Number of laborers multiplied twice (in different cells)
Omission	Number of laborers per team is omitted	CPF <sup>a</sup> is omitted
<i>Qualitative</i>		
Jamming	Jamming of height and length in single cell (height*length)	Jamming of days per team and hours per day in a single cell (No. days* No. hours per day)
Duplication	Duplication of labor costs	Duplication of volume in another column

<sup>a</sup>CPF (Central Provident Fund) is a compulsory savings plan where employers are expected to contribute a certain percentage of their employees' salary to this fund.

jamming and duplication. In doing so, we created a “level playing field” to examine which types of errors are more easily detected.

Two domain-free spreadsheets (named UNCLE and DAD after the main character described in the spreadsheet problems) seeded with errors were given to students for debugging on paper and on screen, respectively. The paper version includes two spreadsheets: one shows a printed copy of the spreadsheet (showing the proposed solution for the business scenario); the other shows the formulas for the various cells in the spreadsheet. For the screen version, students can access both the spreadsheet and the formulas on screen.

Note that the business scenarios for the two spreadsheets (UNCLE and DAD) were also given to students. Basically, the scenarios (see the appendix) were the same but the numerical values (e.g., costs of bricks, number of workers) as well as the actual seeded errors were different. Each spreadsheet contained five different types of errors (mechanical, logic, omission, jamming and duplication) as shown in Table 2.

The two spreadsheet exercises (shown in the appendix) were a close adaptation of the wall problem used by Panko [4]. We chose Panko's wall problem because the exercise is relatively simple and does not require any specific domain knowledge (e.g., accounting expertise). Only basic knowledge such as computation of volume of the wall (given length, breadth and height), labor costs (given manpower and wage rate) and bid amount (given profit margin) are required. Another reason is that Panko's wall problem has been successfully used in several studies (e.g., [5,24,26]), thereby enabling us to build on past research as well as compare our results with previous studies. In order to be consistent with the local terminology for building materials, the types of wall were called brick and concrete (rather than brick and lava as in Panko's experiment). Further, the term “fringe benefits” was changed to “CPF” in order to cater to

the local context. We also modified some of the numerical values used by Panko.

The 382 students taking the introductory course in spreadsheets were randomly split into 17 tutorial groups with an average size of 22 students per group. These tutorial groups were randomly assigned to experiments 1 and 2. Both experiments were given to students as part of a mid-term test. This may motivate students to take the spreadsheet exercises seriously.

*Experiment 1:* Hundred and Seventy-nine students were given a spreadsheet problem (called UNCLE) together with its printed solution on paper containing errors. They were not told the number of errors present but were asked to detect and describe any errors present. After 20 min, the exercise sheet containing their descriptions of errors were collected from them. The students were then given a second exercise containing a similar spreadsheet problem (called DAD) but with the numerical values (e.g., cost of bricks, number of workers) and seeded errors changed. They were asked to retrieve the spreadsheet solution stored on a diskette and look for errors in it on screen. As before, students were asked to write down the errors.

*Experiment 2:* In the second experiment, the same procedure as experiment 1 was followed with the exception that the two spreadsheets (UNCLE and DAD) were interchanged. In other words, the second group of students ( $n_2 = 203$ ) worked on a spreadsheet problem called DAD on paper and the second spreadsheet problem called UNCLE on screen. The results of experiment 2 were intended to give an indication of the generalizability of the findings for experiment 1 as well as an indication of whether the nature of seeded errors (within each type of errors like mechanical, logic, etc.) affects error detection. The experimental procedure is summarized in Fig. 2.

In both experiments, error detection was carried out first on paper, then on screen. The reason was that spreadsheets on paper should result in higher performance in error

Experiment 1

Group 1:	Task on paper followed by (UNCLE)	Task on screen (DAD)
----------	--------------------------------------	-------------------------

Experiment 2

Group 2:	Task on paper followed by (DAD)	Task on screen (UNCLE)
----------	------------------------------------	---------------------------

Fig. 2. Experiments 1 and 2.

detection compared to reviewing spreadsheets on screen. This is in line with past research which showed that reading from screen tends to be slower than reading from paper [31,32,34]. However, Gould et al. [33] found that reading from screen can be as fast as reading from paper if the image quality on screen is comparable to that on paper.

In contrast, Galletta et al. [13] found that the number of errors detected and time taken behaved in opposite directions, i.e., although subjects using paper found more errors compared to subjects using screen, they took a longer time. In any case, past results usually showed that either the type of medium had no effect on error detection or that performance on paper was generally better than on screen. Consequently, if the number of errors detected on screen was equivalent or greater than that of paper, we can perhaps attribute it mainly to the effect of prior incremental practice on paper.

Another reason was that we wanted to simulate real-life spreadsheet usage where printed spreadsheets are first used for checking. Consequently, during error correction or performance of subsequent analysis on screen, further errors may be detected. The experiment also allowed participants to work on similar spreadsheet problems but with different numerical values, thereby simulating to some extent changes in initial conditions of the spreadsheet problem.

## 5. Results and discussion

### 5.1. Characteristics of respondents

The participants of the experiment were first-year business students in a large university taking a course on computing. The age of the students ranged from 17 to 21 with the males in the group generally being older than the females as the males had a two-year stint in national service before entering the university. The distribution of males to females among the participants in experiment 1 was quite even. Of the 179 participants, there were 90 males and 89

females. For experiment 2, there were 88 males and 115 females.

### 5.2. Coding of outcomes

For each spreadsheet exercise, the participants wrote down the errors detected. These materials were used for coding purposes. An independent coder was trained and employed to identify the types of errors detected by the students. To check for reliability, two batches of the experimental material (one for DAD and the other for UNCLE) were randomly selected and used for coding by both the coder and one of the authors. For the first batch of materials involving the UNCLE task, there were a total of 230 detected errors and only 7 discrepancies among the two coders over these. This gives a reliability factor of 0.98 (or a Cohen kappa value of 93.4%). Likewise for the DAD task, the coders differed in 17 aspects over a total of 322 detected errors, giving a reliability of 0.95 (or a Cohen kappa value of 89.3%). In both cases, the reliability is sufficiently high implying that the coder was able to code quite accurately the errors given by the students. The coder proceeded to complete coding all the experimental materials.

### 5.3. Hypotheses testing

The overall results are shown in Table 3. The means and standard deviations for each type of errors were computed and both matched-pair and independent sample *t*-tests were carried out. Since the five types of errors occurred only once in each of the spreadsheets, the mean value of each type of error also indicates the percentage of errors detected. For example, for experiment 1, the mean of 0.84 and 0.56 for mechanical and logic errors, respectively, can be interpreted as 84% and 56% of participants detected the mechanical and logic errors, respectively, for experiment 1.

The results are presented as follows. First, we will discuss error factors in terms of types of errors (quantitative and qualitative errors) and prominence of errors. Second, we will discuss individual factors in terms of the effects of prior incremental practice on error detection.

#### 5.3.1. Error factors

*5.3.1.1. Types of errors (quantitative versus qualitative errors)* Consistent with previous studies, the total number of errors detected was generally about 50% or below (Table 3). Matched-pair *t*-tests were carried out to compare quantitative and qualitative errors within each treatment condition. The results in Table 3 indicate that quantitative errors were more easily detected than qualitative errors: paper/UNCLE ( $t = 19.28, p < 0.001$ ); screen/DAD ( $t = 33.62, p < 0.001$ ); paper/DAD ( $t = 38.09, p < 0.001$ ); and screen/UNCLE ( $t = 33.31, p < 0.001$ ). Hence, hypothesis 1 is supported. Overall, mechanical errors were most easily detected followed by either logic or omission errors.

Table 3  
Results of *t*-tests on different types of errors<sup>a</sup>

Types of errors	Experiment 1 ( <i>n</i> <sub>1</sub> = 179)		Experiment 2 ( <i>n</i> <sub>2</sub> = 203)		Matched pair <i>t</i> -tests		Independent sample <i>t</i> -tests			
	1	2	3	4	<i>t</i> <sub>12</sub>	<i>t</i> <sub>34</sub>	<i>t</i> <sub>13</sub>	<i>t</i> <sub>24</sub>	<i>t</i> <sub>14</sub>	<i>t</i> <sub>23</sub>
	Paper (UNCLE) Mean (SD)	Screen (DAD) Mean (SD)	Paper (DAD) Mean (SD)	Screen (UNCLE) Mean (SD)	(Hypothesis 3)		(Hypothesis 2)		(Hypothesis 3)	
<i>Quantitative</i>										
Mechanical	0.84 (0.37)	0.84 (0.36)	0.82 (0.38)	0.94 (0.25)	-0.16	-3.52**	0.38	-2.87**	-3.01**	0.55
Logic	0.56 (0.50)	0.72 (0.45)	0.64 (0.48)	0.78 (0.41)	-3.62***	-3.34**	-1.63	-1.53	-4.76***	1.56
Omission	0.19 (0.39)	0.85 (0.36)	0.64 (0.48)	0.59 (0.49)	-18.11***	1.13	-9.93***	6.00***	-8.72***	4.95***
Mean (SD)	0.53 (0.26)	0.80 (0.25)	0.70 (0.26)	0.77 (0.27)	-11.73***	-2.98**	-6.34***	1.28	-8.81***	1.28
<i>Qualitative</i>										
Jamming	0.07 (0.25)	0.06 (0.23)	0.00 (0.00)	0.04 (0.21)	0.63	-3.06**	3.58***	0.52	0.97	3.25**
Duplication	0.05 (0.22)	0.04 (0.19)	0.00 (0.07)	0.05 (0.22)	0.53	-2.76**	2.65**	-0.48	0.05	2.23*
Mean (SD)	0.06 (0.17)	0.05 (0.17)	0.00 (0.04)	0.05 (0.15)	0.78	-4.19***	4.35***	0.04	0.73	0.04
Total errors detected	1.70 (0.83) 34.0	2.50 (0.82) 50.0	2.10 (0.79) 42.0	2.40 (0.84) 48.0	-10.79***	-4.09***	-4.80***	1.22	-8.10***	4.83***
Matched-pair <i>t</i> -test (Hypothesis 1)	19.28***	33.62***	38.09***	33.31***						

<sup>a</sup>Note: \* *p* < 0.05; \*\* *p* < 0.01; \*\*\* *p* < 0.001.

This is consistent with Teo and Tan’s [30] study which found similar results.

Both jamming and duplication errors had very low detection rates of less than 8%. One possible reason is that participants may not consider jamming or duplication errors to be significant. Rather, their focus may be more in terms of quantitative errors (which directly affect bottom-line values) rather than qualitative errors in overall spreadsheet design (which may or may not affect bottom-line values).

**5.3.1.2. Prominence of errors** Independent sample *t*-tests (*t*<sub>13</sub> and *t*<sub>24</sub>) were conducted to examine differences in error detection using similar spreadsheet tasks but with different seeded errors (UNCLE and DAD), and different groups of participants (*n*<sub>1</sub> = 179 and *n*<sub>2</sub> = 203). For the paper medium, the results showed that it was easier to detect omission errors for DAD than for UNCLE spreadsheet tasks (*t*<sub>13</sub> = - 9.93, *p* < 0.001). In contrast, it was easier to detect jamming (*t*<sub>13</sub> = 3.58, *p* < 0.001) and duplication (*t*<sub>13</sub> = 2.65, *p* < 0.01) errors for UNCLE than for DAD spreadsheet tasks. However, this must be interpreted with caution since the error detection rates for jamming and duplication were very low for both groups, thereby implying that slight differences may give rise to significant results.

From the description of spreadsheet errors in Table 2, and the above results, it appeared that the omission of number of laborers per team (in UNCLE spreadsheet) was much harder to detect than the omission of CPF (in DAD spreadsheet). This was consistent with Teo and Tan’s [30] study which found that omitting the number of laborers per team occurred most frequently among subjects who designed the spreadsheet model based on the wall problem. Since subjects were likely to omit the number of laborers per team when designing the spreadsheet model, by inference, they were also unlikely to detect such error in a given spreadsheet model.

A related interpretation is that the omission of CPF was a more prominent error than the omission of number of laborers per team. The reason is that contribution to CPF is compulsory for all employees and employers in Singapore. In contrast, the number of laborers per team may vary depending on the nature of tasks and time schedule for completion of work.

For the screen medium, the results showed that it was easier to detect mechanical errors for DAD than for UNCLE spreadsheet tasks (*t*<sub>24</sub> = -2.87, *p* < 0.01). Apparently, based on Table 2, it was easier to detect an error in wall thickness of 5 m (instead of 0.5 m) than it was to detect an



error in the number of hours per day of 7 (instead of 8). One likely reason is that the difference between 7 and 8 h per day was not very noticeable. Moreover subjects were more used to the notion of working 8 h per day and their inherent bias for working 8 h per day may have resulted in the failure to detect this mechanical error. In contrast, a wall thickness of 5 m was obviously incorrect since the wall is being built for a house (in fact, 0.5 m is already more than twice that of a normal wall!). Further unlike the US, students in Singapore are taught to always put a zero in front of the decimal point, i.e., 0.5 instead of .5, thereby making the error between 0.5 m and 5 m more obvious). Applying Weber's notion of just noticeable differences, it is obvious that the difference between 7 or 8 h is less noticeable than the difference between 0.5 m and 5 m. This showed that mechanical errors, which were prominent (or inconsistent with the subjects' notion of an appropriate or common value), were more likely to be detected.

In contrast, it was harder to detect omission errors ( $t_{24} = -6.00, p < 0.001$ ) for UNCLE than for DAD spreadsheet tasks. This confirmed previous findings that the omission of number of laborers per team (in UNCLE spreadsheet) is much harder to detect than the omission of CPF (in DAD spreadsheet). The results for logic, jamming and duplication errors were not significant.

Comparison between the means of quantitative errors for the paper medium was statistically significant ( $t_{13} = -6.34, p < 0.001$ ) but not for the screen medium ( $t_{24} = 1.28, p > 0.05$ ). Similar results were obtained for qualitative errors for paper ( $t_{13} = 4.35, p < 0.001$ ) and screen media ( $t_{24} = 0.04, p > 0.05$ ). The above results imply that hypothesis 2 is supported since there appeared to be some differences in error detection between the two spreadsheets (UNCLE and DAD) due to the different nature (or prominence) of seeded errors.

### 5.3.2. Individual factors

**5.3.2.1. Effects of prior incremental practice** In examining the effects of prior incremental practice on error finding, both matched-pair and independent *t*-tests were used. The matched-pair *t*-test results enabled us to examine the effects of prior incremental practice on the same group of subjects using both UNCLE and DAD spreadsheet tasks. In contrast, the independent *t*-test results compared two different groups of subjects on exactly the same spreadsheet tasks (i.e., either UNCLE or DAD). Both results served to reinforce the findings regarding the effects of prior incremental practice on spreadsheet error detection.

For experiment 1, matched-pair *t*-tests revealed that the number of logic and omission errors detected with prior incremental practice was significantly greater than that without prior incremental practice ( $t_{12} = -3.62$  for logic errors and  $t_{12} = -18.11$  for omission errors,  $p < 0.001$ ). The results for mechanical, jamming and duplication errors were not statistically significant. In contrast, for experiment 2 when

the spreadsheet tasks (UNCLE and DAD) were reversed, the matched-pair *t*-test showed significant results for all types of error with the exception of omission errors. For comparison of means of quantitative errors, both experiments 1 ( $t_{12} = -11.73, p < 0.001$ ) and 2 ( $t_{12} = -2.98, p < 0.01$ ) were statistically significant; whereas for qualitative errors only experiment 2 ( $t_{12} = -4.19, p < 0.001$ ) was statistically significant. Comparisons of the total number of errors in both experiments 1 ( $t_{12} = -10.79, p < 0.001$ ) and 2 ( $t_{12} = -4.09, p < 0.001$ ) yielded significant results.

Overall, our results showed that error detection was generally facilitated by prior incremental practice. However, the nature of the errors also played an important role since some errors were more easily detected than others. This is evident in our results for experiments 1 and 2 (since we found significant results for certain types of errors in experiment 1 but not in experiment 2 and vice versa). Overall, there is support for hypothesis 3 which states that subjects with prior incremental practice in error detection will be able to detect more errors than subjects without prior incremental practice.

Independent sample *t*-tests ( $t_{14}$  and  $t_{23}$ ) were also conducted to examine differences in error detection using the *same* tasks (UNCLE or DAD), and different groups of participants ( $n_1 = 179$  and  $n_2 = 203$ ). For the UNCLE spreadsheet task, the results showed that it was easier to detect quantitative errors with prior incremental practice ( $t_{14} = -8.81, p < 0.001$ ). Specifically, for the same task, mechanical ( $t_{14} = -3.01, p < 0.01$ ), logic ( $t_{14} = -4.76, p < 0.001$ ) and omission ( $t_{14} = -8.72, p < 0.001$ ) errors were easier to detect with prior incremental practice. In contrast, comparisons between qualitative errors ( $t_{14} = 0.73, p > 0.05$ ), i.e., jamming ( $t_{14} = 0.97, p > 0.05$ ) and duplication errors ( $t_{14} = 0.05, p > 0.05$ ), were not statistically significant. Overall, the results demonstrated the beneficial effects of prior incremental practice in enhancing quantitative error detection, thereby adding further support for hypothesis 3.

For the DAD spreadsheet task, the results showed that it was easier to detect omission ( $t_{23} = 4.95, p < 0.001$ ), jamming ( $t_{23} = 3.25, p < 0.01$ ) and duplication ( $t_{23} = 2.23, p < 0.05$ ) errors with prior incremental practice. Hence, this lends further support for hypothesis 3. However, it should be noted that the effect of prior incremental practice was also dependent on the nature of errors as different types of errors (even within the same category, e.g., omission errors) can determine the ease or difficulty of error detection.

## 6. Limitations

There are three main limitations in this study. First, the use of undergraduate students in our spreadsheet experiment may not be representative of actual spreadsheet developers in organizations. However, it is practically impossible to persuade a large group of practitioners to participate in spreadsheet experiments. Furthermore, since the students in our sample were studying business administration, they

represented a potential pool of spreadsheet users or developers. In addition, the use of students enabled us better control over extraneous variables such as age and educational level which may have confounded the results.

Second, the spreadsheet exercises were generally simple compared to real-world spreadsheet problems. However, such simplicity was deliberate as it enabled us to control for domain knowledge requirements. In other words, the simple spreadsheets used in this study did not require domain knowledge and consequently helped us to rule out any confounding effects caused by different levels of domain knowledge among participants.

Third, the order of the experiment treatment was the same in both experiments, i.e., paper followed by screen. We were unable to vary the order of treatments because the experiments were carried out as part of a mid-term test. We were concerned that students would feel that the mid-term test was not fair if the order of treatments was varied. Although this treatment sequence may have confounded the results to some extent, sequencing screen after paper mitigated such confounding. Since previous research usually showed either no differences between paper or screen, or that paper usually resulted in better error detection than screen, we could perhaps attribute the better performance on screen (as compared to paper) to other factors (other than type of media). Thus the sequence of activity was closer to real-life spreadsheet usage where printed spreadsheets are first used for checking, with subsequent changes made on screen.

## 7. Contributions and implications for research and practice

This study extends previous work on spreadsheet error finding and helps to build a cumulative tradition in IS research. The theoretical framework (adapted from [2,13]) serves as a useful basis for investigating factors influencing error finding performance. Future research can expand the framework to include other variables and/or examine other factors in the model.

Since this study confirms previous research that spreadsheet errors are difficult to detect, it implies that more attention should be paid to the types of errors (qualitative versus quantitative) as well as the prominence of errors since different errors may require different error-reduction or detection strategies. Obviously, there is a need for the teaching of proper spreadsheet design to be incorporated in the curriculum rather than the current emphasis on the mechanics of spreadsheets.

This study contributes to existing research in four ways. First, to the authors' knowledge, this is the first study that requires the subjects to look for both quantitative and qualitative errors using a spreadsheet model developed by someone else. Previous research that examined quantitative and qualitative errors involved error detection in subjects' own spreadsheet model [4,5] or detection of errors without distin-

guishing among quantitative and qualitative errors [12,13]. Future research can build on this study by examining different strategies for reducing or detecting errors.

Second, this study was conducted as part of a mid-term test. There was therefore some assurance that students would try their best in detecting the seeded errors since their performance constituted a certain percentage of the final grade. Previous studies commonly used student volunteers who were either paid a token sum for participation or were eligible for prizes if they were in the top 33% or 50% of the subjects. In such cases, the incentive to persevere in looking for errors in the spreadsheet model may be weaker compared to experiments conducted as part of a graded test which would motivate students to take the spreadsheet exercises seriously. Further, the pressure of an exam might simulate to some extent the work pressure in the real world where spreadsheet solutions are expected within a certain time constraint.

In addition, using a volunteer sample may result in some sample bias compared to administering the experiments to a large group of students. Hence, this study also demonstrates that it is possible to incorporate experiments as part of mid-term test, thereby leading to less sample bias compared to using volunteers. Future research can replicate this study in a non-test environment. Comparisons can then be made as to whether a test environment makes any difference in error detection. Such results would be useful to practitioners as it may provide some measures to encourage conscientious error detection.

Third, this study provides some empirical evidence that prior incremental practice on error detection can improve subsequent error detection. This is reassuring as it emphasizes that error detection can improve with practice. Although this result is obvious and expected, what is significant here is our narrow definition of prior incremental practice (as having worked on a spreadsheet problem and then immediately working on a similar spreadsheet problem in a single session). In other words, although previous results have shown that experienced spreadsheet users are generally able to detect more errors than inexperienced spreadsheet users, what we have shown in this study is that even our narrow definition of prior incremental practice (as opposed to long term experience) does help in error detection.

Hence, there is a need to emphasize not only the teaching of good spreadsheet design in IS courses, but also practice in detecting the different types of spreadsheet errors. For practitioners, the implication is that it is important to have the spreadsheet model reviewed periodically since prior practice does help in detecting errors. For researchers, the experiment can be repeated by implementing a series of spreadsheet exercises and determining whether improvements in error detection rates tend to reach a peak or achieve a steady state. In other words, is there a point where any incremental practice would result in diminishing return, where improvements in error detection decreases or reaches a steady state?

Fourth, this study demonstrates that while prior incremental practice may help to detect errors, error detection is also dependent on the nature of errors as some errors are more prominent than others. Attention should be focused on logic and omission errors, as these errors are generally much harder to detect than mechanical errors. For the reduction of qualitative errors, emphasis should be given to good spreadsheet design practices that eliminate the occurrences of jamming or duplication errors. Users need to realize that although jamming and duplication may be commonly viewed as idiosyncratic preferences of the spreadsheet developer, such practices make errors harder to detect and are likely to result in further errors when modifications are made to the initial model. In contrast with our study where there are no direct computational errors resulting from qualitative errors, future research can replicate this study by building spreadsheets with computational errors that do result from jamming and/or duplication.

It is important that practitioners realize that proper design procedures similar to programming and systems development should be adopted for spreadsheet development. For example, the systems development life cycle (incorporating structured system design) can be used as a systematic approach to spreadsheet development [38,39]. Analogous to structured design for software applications, spreadsheets should be developed in well-documented, logically related blocks. This will make error detection and checking of spreadsheets easier and more systematic. The result is not only a reduction of the occurrences of errors (through proper design) but also the facilitation of error detection (through proper walkthroughs and reviews).

## Appendix A.

### A.1. UNCLE

Your uncle is a contractor. He is interested in a job that requires him to build a wall for a house in Binjai Park. As he knows that you know Excel, he asked you to build a spreadsheet model to help him create a bid for building the wall.

He provided you with the following information:

There are two options for the wall—concrete or brick.

The wall will be built by teams of three laborers. Each team will take four 7-h days to build either type of wall.

The wall will be 6 m long, 1 m tall, and 0.5 m thick.

Wages will be \$15 per hour per laborer. You will have to add 22% to wages to cover employer's contribution to CPF.

Concrete wall will cost \$20 per m<sup>3</sup>. Brick wall will cost \$30 per m<sup>3</sup>.

Your bid must add a profit margin of 15% if your expected cost is less than \$1600, and 25% if your expected cost is greater than or equal to \$1600.

### A.2. DAD

Your dad is a contractor. He is interested in a job that requires him to build a wall for a house in Binjai Park. As he knows that you know Excel, he asked you to build a spreadsheet model to help him create a bid for building the wall.

He provided you with the following information:

There are two options for the wall—concrete or brick. The wall will be built by teams of two laborers. Each team will work three 8-h days to build either type of wall.

The wall will be 20 m long, 2 m tall, and 0.5 m thick.

Wages will be \$10 per hour per laborer. You will have to add 20% to wages to cover employer's contribution to CPF.

The material cost for the concrete wall is \$3 per m<sup>3</sup> while the cost for the brick wall is \$2 per m<sup>3</sup>.

In his bid, he wants to add a profit margin of 20% if your expected cost is less than \$1000, and 30% if your expected cost is greater than or equal to \$1000.

*Instructions for error detection on paper:* On the following page, you will find two specimens; one the spreadsheet as shown and the other the spreadsheet showing the cell formulas. Look at the spreadsheet. In the spaces provided below highlight any errors that you find in the spreadsheet by providing the cell reference and the reason(s) why you think it is an error.

*Instructions for error detection on screen:* On the diskette provided, you will find an Excel file named UNCLE.xls (or DAD.xls). Open the file; look at the spreadsheet. Highlight any errors on the spreadsheet by citing the cell reference and the reason(s) for the errors in the spaces provided. Correct the errors on the spreadsheet and save the file as xxxxW1.xls (where xxxx is the 4 digit number on your matric number; e.g. if your matric number is 97-1234a-07, save the file as 1234W1.xls). Remember to put your matric number in Cell A1.

## References

- [1] Ditlea S. Spreadsheets can be hazardous to your health. *Personal Computing* 1987;11(1):60–9.
- [2] Galletta DF, Hartzel KS, Johnson S, Joseph J, Rustagi S. Spreadsheet presentation and error detection: an experimental study. *Journal of Management Information Systems* 1996–97;13:45–63.
- [3] Kee RC, Mason Jr. JO. Preventing errors in spreadsheets. *Internal Auditor* 1988;45(1):42–7.
- [4] Panko RR. Hitting the wall: errors in developing and debugging in a “simple” spreadsheet problem. In: *Proceedings of the 29th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 4–7, 1996. p. 356–63.
- [5] Teo TSH, Tan M. Quantitative and qualitative errors in spreadsheet development. *Proceedings of the 30th Annual Hawaii International Conference on System Sciences*, Wailea, Hawaii, 7–10 January, 1996. p. 149–55.

- [6] Cale Jr. EG. Quality issues for end-user developed software. *Journal of Systems Management* 1994;45:36–9.
- [7] Cragg PG, King M. Spreadsheet modelling abuse: an opportunity for OR? *Journal of the Operational Research Society* 1993;44:743–52.
- [8] Hendry DG, Green TRG. Creating, comprehending and explaining spreadsheets: a cognitive interpretation of what discretionary users think of the spreadsheet model. *International Journal of Human-Computer Studies* 1994;40(6):1033–65.
- [9] Schultheis R, Sumner M. The relationship of application risks to application controls: a study of microcomputer-based spreadsheet applications. *Journal of End User Computing* 1994;6(2):11–8.
- [10] Olson JR, Nilsen E. Analysis of the cognition involved in spreadsheet interaction *Human Computer Interaction* 1987–88;3:309–49.
- [11] Janvrin D, Morrison, J. Factors influencing risks and outcomes in end-user development. In: *Proceedings of the 29th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 4–7, 1996. p. 346–55.
- [12] Galletta DF, Abraham D, El Louadi M, Leske W, Pollalis YA, Sampler JL. An empirical study of spreadsheet error-finding performance. *Accounting, Management and Information Technologies* 1993;3:79–95.
- [13] Galletta DF, Hartzel KS, Johnson S, Joseph J, Rustagi S (1996). An experimental study of spreadsheet presentation and error detection. In: *Proceedings of the 29th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 4–7, 1996. p. 336–45.
- [14] Davies N, Ikin C. Auditing spreadsheets. *Australian Accountant* 1987;57(11):54–6.
- [15] Hall MJJ. A risk and control oriented study of the practices of spreadsheet application developers. In: *Proceedings of the 29th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 4–7, 1996. p. 364–73.
- [16] Brown PS, Gould JD. An experimental study of people creating spreadsheets. *ACM Transactions on Office Information Systems* 1987;5:258–72.
- [17] Creeth R. Microcomputers' spreadsheets: their uses and abuses. *Journal of Accountancy* 1985;159(6):90–3.
- [18] Panko RR, Halverson Jr. RP. Spreadsheets on trial: a survey of research on spreadsheet risks. *Proceedings of the 29th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 4–7, 1996. p. 326–35.
- [19] Floyd BD, Walls J, Marr K. Managing spreadsheet model development. *Journal of Systems Management* 1995;46(1):38–43,68.
- [20] Anderson K, Bernard A. Spreading mistakes. *Journal of Accounting and EDP* 1988;4(2):42–5.
- [21] Ballou DP, Pazer HL, Belardo S, Klein B. Implications of data quality for spreadsheet analysis. *Data Base* 1987;18:13–9.
- [22] Pearson R. Lies, damned lies, and spreadsheets. *Byte* 1988;299–304.
- [23] Stang D. Spreadsheet disasters I have known. *Information Center* 1987;3:26–30.
- [24] Panko RR, Halverson Jr. RP. Are two heads better than one? (at reducing errors in spreadsheet modeling? *Office Systems Research Journal* 1997;15(1):21–32.
- [25] Panko RR, Halverson Jr. RP. Patterns of errors in spreadsheet development I: quantitative errors. Working Paper, Department of Decision Sciences, College of Business Administration, University of Hawaii, March 1995.
- [26] Panko RR, Sprague Jr. RH. Hitting the wall: errors in developing and code inspecting a 'simple' spreadsheet model. *Decision Support Systems* 1998;22:337–53.
- [27] Lerch FJ. Computerized financial planning: discovering cognitive difficulties in model building. Unpublished Ph.D. dissertation, University of Michigan, Ann Arbor, 1988.
- [28] Reithal BJ, Nichols DL, Robinson RK. An experimental investigation of the effects of size, format, and errors on spreadsheet reliability perception. *Journal of Computer Information Systems* 1996;54–64.
- [29] Panko RR. *End user computing: management, applications, and technology*. New York: Wiley, 1988.
- [30] Teo TSH, Tan M. Spreadsheet development and "what-if" analysis: quantitative versus qualitative errors. *Accounting, Management and Information Technologies* 1999;9:141–60.
- [31] Dillon A. Reading from paper versus screens: a critical review of the empirical literature. *Ergonomics* 1992;35(10):1297–326.
- [32] Gould JD, Alfaro L, Barnes V, Finn R, Grischkowsky N, Minuto A. Reading is slower from CRT displays than from paper: attempts to isolate a single-variable explanation. *Human Factors* 1987a;29(3):269–99.
- [33] Gould JD, Alfaro L, Finn R, Haupt B, Minuto A. Reading from CRT displays can be as fast as reading from paper. *Human Factors* 1987b;29(5):497–517.
- [34] Oliver R. Proof-reading on paper and screens: the influence of practice and experience on performance. *Journal of Computer-based Instruction* 1994;20(4):118–24.
- [35] Davis GB, Olson MH. *Management information systems: conceptual foundations, structure and development*, 2nd edn. Singapore: McGraw-Hill, 1984.
- [36] Igarria M, Iivari J. The effects of self-efficacy on computer usage. *Omega* 1995;23(6):587–605.
- [37] Young EA. Human errors in programming. *International Journal of Man-Machine Studies* 1974;6:361–76.
- [38] Goss E, Dillon T, Kendrick J. Bittersweet spreadsheets: application development needs control. *The Woman CPA* 1990;51(3):20–4.
- [39] Janvrin D, Morrison J. Using a structured design approach to reduce risks in end user development. *Information and Management* 2000;37:1–12.