1-1-2010

Applying an error taxonomy to examine inexperienced spreadsheet users' planning and execution errors

Amy Y. Chan
University of Wollongong, amychan@uow.edu.au

Peter Caputi
University of Wollongong, pcaputi@uow.edu.au

Suzanne Hilellis
University of Wollongong

Zi Zhu
University of Wollongong

Rohan Jayasuriya
University of New South Wales, ajayasur@uow.edu.au

Applying an error taxonomy to examine inexperienced spreadsheet users' planning and execution errors

Abstract

This paper presents preliminary findings of an application of the Panko-Aurigemma (2010) error taxonomy to errors generated in a spreadsheet training task. Data from 11 inexperienced trainees were obtained on two spreadsheet training tasks of different complexity and scored by two judges. High levels of inter-rater reliability were obtained for a refined planning and execution error classification. Preliminary findings relating to the likely processes underlying task completion indicate that trainees make more execution than planning errors on easy tasks. Negative associations were found between the frequency of execution and planning errors and task performance when completing easy tasks. For hard tasks, we found a negative association between task performance and frequency of execution errors; however, the association between the number of planning errors and performance was positive. These findings point to the utility of examining error types in IS training strategy research.

Keywords

Applying, error, taxonomy, examine, inexperienced, spreadsheet, users, planning, execution, errors

Disciplines

Arts and Humanities | Life Sciences | Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Chan, A. YC., Caputi, P., Hilellis, S., Zhu, Z. & Jayasuriya, R. (2010). Applying an error taxonomy to examine inexperienced spreadsheet users' planning and execution errors. Proceedings of the 21st Australasian Conference on Information Systems: Defining and Establishing a High Impact Discipline Atlanta, GA, USA: AIS Library.

Applying an Error Taxonomy to Examine Inexperienced Spreadsheet Users' Planning and Execution Errors

Amy Chan
Peter Caputi
Suzanne Hilellis
Zi Zhu
School of Psychology
University of Wollongong
North Wollongong, New South Sales
Email: pcaputi@uow.edu.au

Rohan Jayasuriya
School of Public Health and Community Medicine
University of New South Wales
Sydney, New South Wales
Email: r.jayasuriya@unsw.edu.au

Abstract

This paper presents preliminary findings of an application of the Panko-Aurigemma (2010) error taxonomy to errors generated in a spreadsheet training task. Data from 11 inexperienced trainees were obtained on two spreadsheet training tasks of different complexity and scored by two judges. High levels of inter-rater reliability were obtained for a refined planning and execution error classification. Preliminary findings relating to the likely processes underlying task completion indicate that trainees make more execution than planning errors on easy tasks. Negative associations were found between the frequency of execution and planning errors and task performance when completing easy tasks. For hard tasks, we found a negative association between task performance and frequency of execution errors; however, the association between the number of planning errors and performance was positive. These findings point to the utility of examining error types in IS training strategy research.

Keywords

Spreadsheet, spreadsheet error, taxonomy, planning error, execution error.

INTRODUCTION

The advent of electronic information technology systems has led to the development and widespread use of computer spreadsheet programs in organisational and educational settings, with its goal to facilitate more efficient and accurate information processing. Researchers have used errors in spreadsheet development, to study and develop taxonomies of such errors (Panko & Halverson, 1996; Panko & Aurigemma, 2010). Evidence has accumulated that many spreadsheets contain errors, and that errors can be costly to the organisations that use them (Powell, Baker, & Lawson, 2008). An error is defined as an action that is neither intended nor desired by the actor (Senders & Moray, 1991), and results in a failure to achieve the intended outcomes (Reason, 1990).

Due to the high prevalence of errors in spreadsheet use, it is important to appropriately identify and classify spreadsheet errors to facilitate diagnoses of their causes, the relative frequency of different error types, and their impact on judgement and decision making (Powell et al., 2008). Furthermore, an effective classification system of spreadsheet errors is useful for comparing errors across users and application areas. Hence, since the mid 1990s, many classification schemes for spreadsheet errors have been developed (e.g., Galletta, Abraham, El Louadi, Leske, Pollalis, & Sampler, 1993; Saarilouma & Sajaniemi, 1994; Panko & Halverson, 1996).

One of the earliest comprehensive attempts to create a taxonomy of errors in spreadsheet development and use came from Panko and Halverson (1996). Their taxonomy was built upon findings from protocol analyses of students' mathematical problem solving (Allwood, 1984). Of key relevance to the present study, this taxonomy described quantitative and qualitative error categories and features of errors in general. A quantitative error occurs when incorrect knowledge or procedure is applied to a spreadsheet task such as generating graphs or manipulating data. This normally results in a wrong number or an incorrect display in the spreadsheet, although it has been acknowledged more recently that sometimes an erroneous procedure can still end up with the correct outcome (e.g., when a wrong cell that contains the same value as the correct cell on the spreadsheet is selected

for a calculation – see Panko & Aurigemma, 2010). By comparison, qualitative errors are inappropriate or risky practices that do not immediately lead to wrong calculations or displays (e.g., using specific numbers instead of relevant cell references in writing a formula for calculations), with such practice leading to incorrect results in subsequent uses. Notably, the definitions of quantitative and qualitative errors in this taxonomy represent a departure from the earlier notion of errors - that an error is something that has to result in a failure to achieve the intended outcomes (Reason, 1990). Even though the immediate outcome for a task may appear to be correct, it may have been achieved via erroneous procedures or practices. From a practical perspective, observing the process undertaken by a spreadsheet user while completing tasks can be useful in identifying the user's level of knowledge and understanding about the spreadsheet application, as well as the levels of attention and care placed onto the task.

As indicated earlier, spreadsheet error taxonomies are useful in informing and improving practice in spreadsheet use. Hence, a systematic examination of spreadsheet errors can be particularly relevant to understanding the approaches and mechanisms underlying inexperienced users' behaviour as they acquire skills within a spreadsheet environment. Although accuracy in task performance is generally a good indicator of learning success, such that the better the user's task performance, the fewer the number of errors are made, our stance in this research is that errors are more than just the flip side of task performance. In the process of producing a correct outcome in a spreadsheet task, the user may have made and corrected a variety of errors. Thus it will be important and informative to identify appropriate ways to measure the relative frequencies of different types of errors made by the user while attempting a spreadsheet task.

The purpose of the present study is to develop a systematic approach that can reliably examine and assess the prevalence of different types of spreadsheet errors in inexperienced computer spreadsheet users. To this end, we primarily draw upon Panko and Aurigemma's (2010) most recent revision of the Panko-Halverson taxonomy. Focussing on inexperienced users is relevant from a training perspective given the amount of resources invested in IS and computer training needs.

A Case for Using the Panko-Aurigemma (2010) Taxonomy to Examine Inexperienced Spreadsheet Users' Errors

Of particular relevance to the present study is Panko and Halverson's (1996) stance that spreadsheet errors are frequent, and that many such errors are quantitative in nature. In an attempt to characterise the different sources of quantitative errors, Panko and Halverson (1996) further classified such errors into mechanical errors (typing or pointing incorrectly), logic errors (deciding incorrectly on which function or formula is appropriate), and omission errors (misinterpreting the task requirement and therefore skipping part of the required task). This early classification system has formed an important basis for the development of many more recent spreadsheet error taxonomies (e.g., Rajalingham, Chadwick, & Knight, 2000; Purser & Chadwick, 2006).

In line with their original taxonomy, the Panko and Aurigemma (2010) revised taxonomy broadly classifies errors into qualitative and quantitative errors. Within the subcategory of quantitative errors, this taxonomy distinguishes between planning and execution errors. This distinction is useful for gaining insights into the point at which an error is made. Planning errors occur before the user begins typing a formula or selecting a tool from the spreadsheet application. As such, planning errors reflect impoverished knowledge or consideration of concepts required for choosing the correct options and using the spreadsheet application competently. This is contrasted with execution errors, which occur after the user has commenced typing or has made a selection of tools or menu options, and are reflected in simple mechanical slips (e.g., mistyping and incorrectly pointing at a wrong cell) and lapses (e.g., forgetting to carry out a task, or repeating the same incorrect task procedure despite already having received error messages that indicate the incorrectness of the procedure).

In assessing their revised error taxonomy, Panko and Aurigemma (2010) report a high inter-rater reliability (96.6%) in their classification of errors in 39 spreadsheets produced by students. They called for further application of this taxonomy to other corpuses, especially those with different levels of complexity. Of central interest to the present study, these researchers report that planning errors are much more prevalent (82%) than execution errors (18%) in the data they examined. This pattern is contrasted with those reported by Powell, Baker, and Lawson (2009), who found that across a representative cross-section of completed, operational spreadsheets generated from a wide variety of sources (e.g., consulting companies, banking sector, tertiary education sector, government agency), only 32% of spreadsheet errors they classified were logic errors in using formulae, which were conceptually similar to mistakes in planning as per Panko and Aurigemma's definition. This discrepancy in the observed relative frequencies of different error types may partly reflect the role of varying levels of experience and expertise on the nature of errors, with less experienced spreadsheet users (students in the case of Panko & Aurigemma's research) more likely to make more planning errors than more experienced, professional users. Another possible issue inherent in is the use of completed spreadsheets to trace and classify different types of errors by users, and this warrant further consideration and investigation.

It is possible that the effectiveness of code inspection, as well as spreadsheet auditors' knowledge of the spreadsheet development context, are two factors that may contribute towards error detection and classification (Panko & Aurigemma, 2010). Furthermore, it is also possible that the relative prevalence of different error types may vary according to task difficulty, and the pattern observed by Powell et al. (2009) may reflect a culmination of different error patterns across tasks of varying levels of complexity. These possibilities need to be examined empirically. However, there is yet another inherent difficulty in examining completed spreadsheets for the purpose of correctly identifying and classifying spreadsheet errors: auditors typically have to rely on *observable* errors in the completed spreadsheets, to *infer* the nature of the process involved in producing each error. This adds to the difficulty with which different auditors can consistently classify errors into the same categories, which, in turn, complicates the task of assessing and refining spreadsheet error taxonomies.

In this paper, we address the issue of the difficulty of identifying errors post-hoc (after the data is entered), by examining spreadsheet users' actions recorded via a computer screen-capturing software. This permits auditors to independently gain insight into the steps and processes used by spreadsheet users while they were attempting a set of required tasks in Microsoft® Excel (Microsoft Corporation, 2003). We propose that this novel approach will circumvent many ambiguities associated with classifying errors as planning versus execution errors. Moreover, to the extent that a spreadsheet user has attempted different methods to tackle a task within the spreadsheet application environment before settling on a specific action, those different methods (and corresponding spreadsheet errors) can all be identified and scored accordingly. This is a further benefit of this approach to assessing spreadsheet errors, as compared to other approaches suggested in the literature.

We report on findings to date from a study currently in progress, based on Panko and Aurigemma's (2010) taxonomy for classifying planning and execution errors in assessing errors made by inexperienced spreadsheet users. Two judges were trained to become familiar with the context and requirements of two training tasks provided to inexperienced users of the spreadsheet application. The Panko-Aurigemma (2010) taxonomy for classifying planning and execution errors was refined and contextualised for identifying and classifying specific actions as belonging to different types of errors in this study. The judges then examined and classified the quantitative spreadsheet errors made by inexperienced undergraduate student users on two tasks during an introductory spreadsheet training session. In addition, we introduced the difference in complexity using two tasks (a Hard task and an Easy task) which varied in the level of spreadsheet knowledge required to successfully complete the task (see Method section for details).

To the extent that the refined Panko-Aurigemma (2010) taxonomy is useful for classifying and quantifying inexperienced spreadsheet users' planning and execution errors, we hypothesised that the number of errors reported by two independent judges (i.e., spreadsheet auditors) should show a substantial degree of consistency. Secondly, we hypothesised that this consistency would be maintained across task difficulty and error type.

METHODS

In this section, we first outline the corpus of spreadsheets used in the present study. We then describe the development of our reliability protocol for refining and applying the Panko-Aurigemma (2010) spreadsheet taxonomy.

Background to the corpus of spreadsheet data

The spreadsheets used in the present study were taken from another study conducted by three of the present authors on introductory spreadsheet training (Chan, Caputi, Jayasuriya, De Blasio, & Lewis, 2010). Participants were undergraduate psychology students who were inexperienced in using spreadsheet applications. In this study, students were provided with a single introductory training session that lasted approximately 120 minutes. Each student undertook all relevant tasks in the training session via a personal computer. They initially watched a short demonstration video that oriented them to the spreadsheet application environment. This was followed by a series of short tasks that sought to provide trainees with increased familiarity and experience in using the spreadsheet application. Of particular relevance to the present study, however, was that in the final two tasks completed by trainees (the aforementioned Easy task and Hard task), no assistance or feedback was available from the instructor on any aspect of the tasks. In the Easy task, only a basic level of understanding of the principles of data entry and using formulas within the spreadsheet environment was needed. In contrast, the Hard task additionally required the user's ability to identify and use appropriate formulas and tools for computation and formatting purposes. The computer screen-capturing software Camtasia Studio 3 (TechSmith Corporation, 2005) was used to record each trainee's observable actions on the computer screen throughout the entire training session (e.g., mouse clicks, pointing, selecting, typing, and deleting unwanted/incorrect actions). The Camtasia recordings of trainees' behaviour during the Easy and Hard tasks provided the basis for examining and classifying errors by judges in the present study.

Developing and testing a Reliability Protocol for Applying the Panko-Aurigemma (2010) Taxonomy

As pointed out earlier, spreadsheet auditors' knowledge of the spreadsheet development context may be a key determinant of their ability to classify and count errors appropriately (Panko & Aurigemma, 2010). Hence, two judges who were not involved in the spreadsheet training study (the third and fourth authors of the present study) initially examined all materials used in the training protocol, to become familiar with the training context and task requirements. Next, the two judges watched together the Camtasia recordings generated from two randomly chosen trainees during their training sessions. The purpose was for the judges to begin developing insights into the specific types of errors that one might expect to observe in the corpus of spreadsheets in general. Subsequently, the judges collectively generated a list of scoring criteria for identifying quantitative errors, and for classifying those errors into the planning and execution subcategories.

Afterwards, the judges independently applied their interim scoring criteria to examine the Camtasia data for another randomly chosen trainee. Each judge tallied the planning and execution errors made by the trainee in each of the Hard and Easy tasks. However, interrater reliability was found to be dissatisfactory (< 70% agreement). Hence, the judges examined the Camtasia recording together and discussed the sources of discrepancy. This process was instrumental for finetuning the definitions of planning and execution errors, and for identifying specific behaviours that would fit into each category. The judges underwent four iterations of this process with Camtasia data from four trainees before they arrived at the final scoring criteria. At the end of the final iteration of this developmental process, interrater agreement across those data sets, training tasks, and error types was 97%.

A departure of our taxonomy from that of Panko and Aurigemma (2010) is in how the same error being repeated within the same cell is to be counted. In the Panko-Aurigemma (2010) taxonomy, the same error occurring multiple times within a cell is counted once only. However, in our refined taxonomy, when a trainee initially produces an error within a cell because they do not know what to do, we counted the error as a planning error. However, if they repeat the same action or option within that cell, then this second erroneous behaviour is to be counted as an execution error. Our rationale is that the first time a spreadsheet user makes a mistake due to incorrect logic or conceptual knowledge, the error is justifiably a planning error. However, if the user repeats a previous planning error *despite* having already received feedback from the spreadsheet application on it (e.g., error messages), the error is likely to be due to a lapse (e.g., forgetting that the specific action used is incorrect) or a slip (e.g., engaging in the same sensory-motor action by accident). Such repeated errors are meaningful and informative, and therefore we considered that they should not simply be discounted.

This reliability protocol development resulted in the Panko-Aurigemma (2010) taxonomy being contextualised for examining errors that occurred within the task environments for the present study; but the definitions of different error types were only slightly modified. For the purposes of this study, we did not separate planning errors into domain and spreadsheet subtypes, nor did we separate execution errors into slips and lapses when we scored the data to be reported in the ensuing Results section. We also only focussed on errors displayed as a result of overt behaviours, and did not include omission errors. Our rationale for this non-separation approach was driven by theory as well as by practical consideration. At the theoretical level, our decision has been guided by our broad aim to assess the merits of Panko and Aurigemma's (2010) and Powell et al.'s (2009) different positions regarding the relative frequencies of planning and execution spreadsheet errors. At the practical level, given that the present paper reports on our work in progress, our focus is on examining the interrater reliability across broader error types and task difficulty. For the reliability analysis reported in the ensuing sections, we applied the error scoring criteria developed and outlined above to another randomly selected subsample of 11 trainees from the same spreadsheet training study (see next section for details).

Table 1. Definitions and Examples of Planning versus Execution Errors
in the Refined Panko-Aurigemma (2010) Taxonomy

Error Type	Definition	Examples
Planning	Mistakes (i.e., logic or mathematical errors)	Selecting an unnecessary tab or option; inserting an incorrect formula
Execution	Slips (i.e., sensory-motor errors) or Lapses (i.e., errors caused by memory overload)	Trying the same action that has already been shown to be incorrect; typing a wrong letter or wrong word

Table 1 provides a summary of the final taxonomy used in the present study, and examples of actions that are included in each error type.

Spreadsheets Used for Assessing the Reliability Protocol

To assess the inter-rater reliability of the final scoring criteria developed for the present study, we utilised the Camtasia recordings from an independent set of spreadsheets generated by 11 inexperienced spreadsheet users in the aforementioned introductory spreadsheet training study. This sample was randomly chosen from the full data set (N=106) for this study, and comprised 2 male and 9 female undergraduate students with a mean age of 19 years (SD = 1.84, range = 17 to 24 years). The corpuses examined contained a total of 768 errors identified by the first judge, and 789 errors identified by the second judge. The two judges independently applied the final scoring criteria to examine the frequency of planning and execution errors in each user. Errors were scored separately for the Hard task and the Easy task.

RESULTS AND DISCUSSION

Descriptive statistics for error type (planning or execution) and task difficulty for each judge are presented in Table 2.

Table 2. Descriptive Statistics on the Number of Errors by Task Difficulty and Error Type

		Judge 1	Judge 2
Error Type	Task Difficulty	Mean (SD)	Mean (SD)
Planning	Easy	14.00 (8.05)	15.00 (11.11)
	Hard	28.36 (13.63)	29.00 (15.30)
Execution	Easy	5.64 (5.94)	5.64 (6.45)
	Hard	21.82 (7.01)	22.09 (8.13)

An examination of mean number of errors reported by each judge show a substantial degree of consistency. This consistency is maintained across task difficulty and error type. For the Easy task, both judges reported fewer execution errors than planning errors (t[10] = -5.39, p = .000, t[10] = -5.03, p = .001for Judge 1 and 2 respectively). For the Hard task, although the mean number of execution errors was less than planning errors was reported by each Judge, these differences were not statistically significant (t[10] = -1.77, p = .107, t[10] = -1.70, p = .120 for Judge 1 and 2 respectively).

Notably, averaged across the two judges, the ratio of planning versus execution errors is *higher* for the Easy task (2.57 to 1) than that for the Hard task (1.31 to 1). This finding is not surprising though, given that the corpus of spreadsheets in the present study was produced by inexperienced spreadsheet users, who might conceivably make more errors - regardless of error type – when faced with a harder task. However, this finding also raises the issue that task characteristics, as well as user characteristics, are both highly relevant determinants of the relative frequencies of different types of spreadsheet errors.

Table 3. Measures of Inter-Rater Agreement

Error Type	Task Difficulty	Intraclass Correlation
Planning	Easy	.85
	Hard	.97
Execution	Easy	.97
	Hard	.96

Inter-rater agreement was further examined by determining the intraclass correlation for the number of errors reported by judges. The results of this analysis are reported in Table 3. These findings show a high level of inter-rater agreement by error type and task difficulty.

Performance scores were also calculated for each respondent in terms of how accurately participants completed steps on each task. A maximum score of 28 could be obtained for the Hard task, while a maximum score for 26 was possible for the Easy task. Task performance was then defined as percentage correct. Performance scores were then correlated with the number of errors determined by each judge for error type and task difficulty. Given the sample size Spearman rank-order correlations were computed. These correlations are reported in Table 4. None of the correlations was significant. As expected, most correlations were negative indicating better performance with fewer errors. Interestingly, performance was positively related to number of planning errors on the Hard task.

Table 4. Spearman Correlations

Error Type	Task Difficulty	Judge 1	Judge 2
Planning	Easy	-.22	-.24
	Hard	.25	.24
Execution	Easy	-.44	-.14
	Hard	-.21	-.17

Although the present study is primarily concerned with errors made by inexperienced spreadsheet users, the general pattern of findings shown in Table 4 may provide some important clues for future research directions. The data suggest that when dealing with a hard spreadsheet task, a greater number of planning errors is associated with *better* task performance. Whilst this finding may initially appear counterintuitive, one possible explanation for it is that users who are prepared to explore more widely within the spreadsheet environment – and hence make more planning errors – are more likely to figure out the correct way to complete steps on the task. This position is consistent with that of error management training (Keith & Frese, 2008), which advocates the benefits of framing errors as a natural and positive by-product of learning and skills acquisition.

CONCLUSION

In this paper, we have argued for the importance of identifying and classifying spreadsheet errors in understanding their causes, the processes underlying task completion and the consequences of different error types in decision making. A number of error taxonomies are described in the literature. One error classification scheme that has received recent attention has been proposed by Panko and Aurigemma (2010).

We used the principles of Panko and Aurigemma's (2010) spreadsheet error taxonomy to examine the relative frequencies of planning and execution errors in inexperienced spreadsheet users' spreadsheet task performance. An aim of our research was to examine the utility of applying this taxonomy to such a task. To this end, we developed a systematic approach to examine the reliability of the classification scheme. The study methodology used two judges who were familiar with the problem context in applying an error scoring criteria, and could trace back to the actual actions and processes utilised by the spreadsheet user (via Camtasia). The findings showed that trained judges can independently score errors with very satisfactory reliability. This finding has an important analytical implication. We have preliminary supporting evidence for a reliable scheme that can be used in assessing spreadsheet errors. This practical implication sets the scene for additional development and research on the use of reliable protocols in future research.

A second aim of this study was to understand the processes underlying completion of spreadsheet task. Across different levels of task difficulty, our data indicate that planning errors are more prevalent than execution errors, although the difference was not statistically significant for difficult tasks. This finding is consistent with that reported by Panko and Aurigemma (2010) and not those reported by Powell et al. (2009). Theoretically, this finding is interesting because it points to potentially different higher-order cognitive processes in assessing and making use of available information (Reason, 1990). However, the observation of differences in error type also has implications for training. The adage that we learn from our mistakes in encapsulated in error management training (Keith & Frese, 2008). It follows that understanding the impact of different error types on learning is of some consequence in error management training. A future direction in this area of research may be to develop an

improved understanding of the relationship between these characteristics and the relative frequencies of different types of spreadsheet errors. This understanding will, in turn, inform instructional designs that target appropriate training strategies to capitalise on ways in which making errors during learning can become beneficial to the learner further down the track.

Hence, the present study contributes to IS research by identifying a starting point for an alternative, on-line approach to examining errors during training. We envisage that this approach can offer useful insights into how different IS training strategies (e.g., error management training) may influence trainees' behaviour during and beyond training.

From a training perspective, our data also challenge the conventional view that errors are just the flip side of successful task performance. Examining the nature and frequency of different types of spreadsheet errors can provide researchers with rich information on what users do, and how they do it, to lead up to their subsequent objective levels of task performance.

Our research, to date, points to the utility and applicability of Panko and Aurigemma's (2010) error taxonomy. The preliminary findings provide empirical support for a reliable classification protocol that is easy to implement after minimal training. Our secondary findings have implications for understanding task completion and type and frequency of error. These findings point to further research in the application of meta-cognitive training strategies such as error management in IS settings.

REFERENCES

Allwood, C.M. 1984. "Error detection processes in statistical problem solving," *Cognitive Science* (8:4), October-December, pp 413-437.

Chan, A.Y.C., Caputi, P., Jayasuriya, R., De Blasio, F., & Lewis, K.L. 2010. *Computer self-efficacy development during software application training: Influence of general training environment versus specific training strategies.* Unpublished data, University of Wollongong.

Galletta, D.F., Abraham, D., El Louadi, M., Leske, W., Pollalis, Y.A., & Sampler, J.L. 1993. "An empirical study of spreadsheet error-finding performance," *Journal of Accounting, Management, and Information Technology* (3:2), April-June, pp 79-95.

Keith, K., & Frese, M., 2008. "Effectiveness of Error Management Training: A Meta-Analysis," *Journal of Applied Psychology* (93:1), January, pp 59-69.

Microsoft Corporation 2003. *Microsoft® Office Excel 2003.* California, USA: Microsoft Corporation.

Panko, R.R. & Aurigemma, S. 2010. "Revising the Panko-Halverson Taxonomy of Spreadsheet Errors," *Decision Support Systems* (49:2), May, pp 235-244.

Panko, R.R. & Halverson, R.P. 1996. "Spreadsheets on Trial: A Framework for Research on Spreadsheet Risks," *Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences,* Volume II, pp 326-335.

Powell, S.G., Baker, K.R., & Lawson, B. 2008. "A Critical Review of the literature on Spreadsheet Errors," *Decision Support Systems* (46:1), December, pp 128-138.

Powell, S.G., Baker, K.R., & Lawson, B. 2009. "Errors in Operational Spreadsheets," *Journal of Organizational and End User Computing* (21:3), July, pp 23-36.

Purser, M. & Chadwick, D. 2006. "Does an awareness of differing types of spreadsheet errors aid end-users in identifying spreadsheet errors?" *Proceedings of the European Spreadsheet Risk Interest Group Annual Conference,* pp 185-204, Cambridge, UK.

Rajalingham, K., Chadwick, D.R., & Knight, B. 2000. "Classification of Spreadsheet Errors," *Symposium Proceedings EuSpRIG 2000 European Spreadsheet Risks Interest Group,* July, pp 23-34.

Reason, J.T. 1990. *Human Error*. Cambridge, England: Cambridge University Press.

Saarilouma, P. & Sajaniemi, J. 1994. Transforming verbal descriptions into mathematical formulas in spreadsheet calculation," *International Journal of Human-Computer Studies* (41:6), December, pp 915-948.

Senders, J.W. & Moray, N.P. 1991. *Human Error: Cause, Prediction, and Reduction*. Hillsdale, NH: Lawrence Erlbaum.

TechSmith Corporation 2005. *Camtasia Studio Version 3.00*. Okemos, MI: TechSmith Corporation.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Frances De Blasio, Kathryn Baudinette, and Keren Wolstencroft, for their assistance in collecting the Camtasia data used in the present study.