Association for Information Systems
AIS Electronic Library (AISeL)

2004

Spreadsheet Visualization Effects on Error Correction

Hock Chuan Chan

National University of Singapore, chanhc@comp.nus.edu.sg

Follow this and additional works at: http://aisel.aisnet.org/sighci2004

Spreadsheet Visualization Effects on Error Correction

Hock Chuan Chan
National University of Singapore
chanhc@comp.nus.edu.sg

ABSTRACT

Spreadsheets have been used by organizations for decades. Errors in spreadsheets are commonly found in laboratory and field findings. In recent years, many exciting new visualization techniques have been developed to help users understand spreadsheet models and to check for errors. Two visualization tools were tested in an experiment for their effects on error correction. The first is a simple arrow tool which shows dependencies among cells. The second shows the input-process-output function of cells in addition to the dependency arrows. The experiment shows significantly better error detection with the arrow method than for the plain method (without visualization tools). Wrong data errors took more time to correct than missing data errors.

Keywords

Spreadsheet error detection, spreadsheet visualization, error types.

INTRODUCTION

Spreadsheets have been used in organizations and in diverse areas, such as business, engineering, physical and life sciences (McGill and Klobas, 2004). Many years of laboratory and field studies concluded that spreadsheet errors are very widespread, very hard to avoid and often have serious consequences (Chan and Storey, 1996; Chen and Chan, 2000; Galletta, Abraham, Louadi, Lekse, Pollalis and Sampler., 1993; Galletta, Hartzel, Johnson, Joseph and Rustagi., 1996; Panko, 2000; Panko and Halverson, 2001; Panko, 2004; Teo and Lee-Partridge, 2001; Teo and Tan, 1999). It is very important to try to reduce spreadsheet errors. Strategies have been put forth to reduce errors in spreadsheet during development and checking stages. New visualization tools have been developed to aid users in error detection (Chan, Ying and Peh, 2000; Clermont, 2003; Davis, 1996; Igarashi, Mackinlay, Chang and Zellweger, 1998; Sajaniemi, 2000; Shiozawa, Okada and Matsushita, 1999). Graphics or visualization is expected to be easier compared to printed text, as people are known to have remarkable visual perceptual abilities (Shneiderman, 1998).

Empirical investigation of spreadsheet visualization tool effectiveness has been lacking, with only a few such studies (e.g. Chen and Chan, 2000; Davis, 1996). More studies are needed to assess and provide guidance on whether visualization tools are effective. Such information will be of high value to the millions of spreadsheet users, trainers, researchers and developers.

LITERATURE REVIEW

This section reviews empirical studies on user performance, error types and visualization tools.

Empirical Studies

Many years of research into spreadsheet development have concluded that a large percentage of spreadsheets contain errors (Panko and Halverson, 2001). Spreadsheet development studies are reported by Brown and Gould (1987), Panko and Halverson (1997), Panko and Halverson (2001), Panko and Sprague (1998) and Teo and Tan (1999). Studies on spreadsheet error-checking include Galletta et al. (1993), Galletta et al. (1996), Panko (1999), and Reithel, Nichols and Robinson (1996). Errors have been found in abundance in most of the studies, including many field audits (Panko and Halverson, 2001). Work on reducing spreadsheet errors is especially important as spreadsheets are used throughout all organizational levels.

Reason (1990) presented a comprehensive framework on why humans make errors. Humans have always sacrificed precision for speed. Research has found that cell error rates (cells with error divided by total number of cells) are consistent with general error rates in other work domains (Panko, 2004). Human errors occur in 0.5 to 5% of all actions, increasing with complexity. Established guidelines for inspecting computer programs usually detect errors in about 5% of all program statements (Panko, 2004). Guidelines have been recommended for spreadsheet checking (also known as spreadsheet auditing) but have not been popular. There is a tradition in spreadsheet development that users are not bound by design or development guidelines (Davis, 1996).

Studies have indicated that spreadsheet error-checking is very difficult. For example, Brown and Gould (1987), in a study of nine experienced and confident users developing spreadsheets, found that subjects made at least an error each. It was observed that most errors involved formulas. Formulas are usually observed by clicking on cells. Although there is an option in Excel that allows users to view all formulas instead of computed numbers, it is still difficult to detect errors. Another experiment by Galletta

et al. (1993), with thirty certified public accountants and thirty MBA students, found that "both accounting and spreadsheeting expertise contributed to the subjects' error-finding rate" (Galletta et al., 1993, p.79). It was also found that efficiency (speed) was aided by spreadsheet but not accounting expertise. It was concluded that errors are difficult to locate, even for simple obvious errors in well documented spreadsheets. Similar findings on the difficulty of error detection and lack of error awareness were reported by Teo and Tan (1999).

Many different approaches have been studied on how to reduce or detect spreadsheet errors. One approach to reduce errors during development is to have team rather than individual development (Panko and Halverson, 1997, 2001). Other approaches consider the provision of printed worksheets (Galletta et al., 1996; Teo and Lee-Partridge, 2001), and the use of different cell sizes and plain / fancy formats (Reithel et al., 1996). Another approach aims to understand and develop a model of how users debug spreadsheets (Chen and Chan, 2000)

Error Types

Spreadsheets errors are classified as qualitative or quantitative (Galletta et al., 1993; Panko and Halverson, 1996; Teo and Tan, 1997; Teo and Tan, 1999). Qualitative errors are poor spreadsheet design and format. These are excluded from this study. Quantitative errors are errors that lead (usually) to wrong computations. A possible classification is by the way they are committed: e.g. mechanical, omission and logic errors (Panko and Halverson, 1996). This classification is good for identifying errors during spreadsheet development (Teo and Tan, 1999). Where users have to check pre-built models, the distinction may not be relevant. For example, a wrong cell reference can result from mechanical or logic error.

For the purpose of error detection, we propose a new error classification based on cell content. A cell can contain a formula, which contains operations and operands. This is a common distinction in programming. Operations include "+", "-", and predefined functions such as "sum" and "fdist". Operands are values, which can be a data, or a reference to another cell. Thus, the first classification of errors has three levels: operation, data or reference. The second classification is based on whether the content is missing, wrong, or extra. For example, if the correct formula is "=A1+B1", than an example of a missing reference error is "=B1", a wrong reference error is "=A2+B1", and an extra reference error is "=A1+B1+C1". These two classifications are orthogonal. Crossing the two classification levels produces a total of 9 different error classes.

Visualization Tools

Spreadsheets essentially comprise an inter-connected web of cells that reference one another through the use of

formulas. Spreadsheet structures can be classified into two levels (Chen and Chan, 2000; Saariluoma and Sajaniemi, 1989, 1991). Surface structures are represented by values, figures and spatial positions. Deep structures are formed by formulas. Surface and deep structures are often inconsistent and users need to memorize deep structures, which demands heavy memory load and results in errors (Saariluoma and Sajaniemi, 1989). Users "often find the structure of the computations being carried out in a spreadsheet rather obscure – they are not as visible as one might expect" (Hendry and Green, 1994, p. 1045).

Sajaniemi (2000, p. 49) highlights the same problem, "Computations in spreadsheets are hard to grasp … the problem … lies in the invisibility of the structure of calculations." Visualizations aim to provide a visual mapping between the surface and deep structures, and help lighten the memory load (Chan et al. 2000; Sajaniemi, 2000).

There are many different visualization tools in the literature (Ballinger, Biddle and Noble, 2003; Chen and Chan, 2000; Clermont, 2003; Hendry and Green, 1993; Igarashi et al., 1998; Sajaniemi, 2000; Shiozawa et al., 1999). This paper focuses on two fundamental methods. Information about cell connections is one of the most fundamental information that users need to know (Davis, 1996). For example, if cell B1 contains "=A1+A2", then cell B1 is called a dependent cell of A1 (and also of A2), while A1 and A2 are precedent cells of B1. Visualization help is very much needed since locating the dependents and precedents by "manually finding cell after cell may be frustrating, time-consuming, and error-prone" (Davis, 1996, p.432). Hendry and Green (1994) state "understanding how a formula works often requires the user to recursively track down the meaning of cell references" (p. 1053), and "even for simple problems spreadsheet formulae are not always easy to create or understand" (p. 1062). Thus, the first method is the arrow method, which shows cell connections with blue arrows. For example, if cell B4 contains "=A1+A2+A3", arrows will go from A1, A2 and A3 to B4, as shown in figure 1. This method is available in Microsoft Excel, e.g. version 2003, as a formula auditing toolbar.

The second method aims to label cells for their roles in the model (Chan et al., 2000; Davis, 1996; Hendry and Green, 1993; Sajaniemi , 2000). "The concept of the cell, its purpose and applications should be emphasized" (Teo and Tan, 1999, p.157). Cells with different roles are given different colors. Colors enable further visualization of deep structures. One implementation colors cells according to the input-process-output computational model (Chan et al., 2000). Cells are classified as input cells, processing cells, output cells and stand alone cells. Input cells are referred to by other cells, processing cells refer to other cells and are referred to by other cells, output cells refer to other cells, but are not referenced, and finally, stand alone cells have no links to other cells. As

shown in figure 2, input cells are colored light gray (these cells are in column A), process cells are light blue (these are in column B), the one output cell (cell D12) is dark gray, and the one stand alone cell (cell D3) is bright pink.
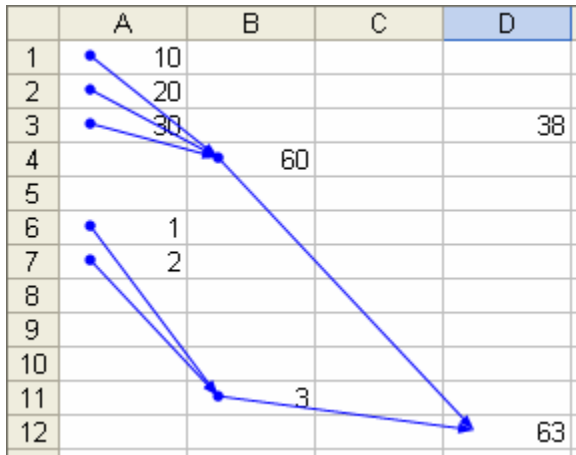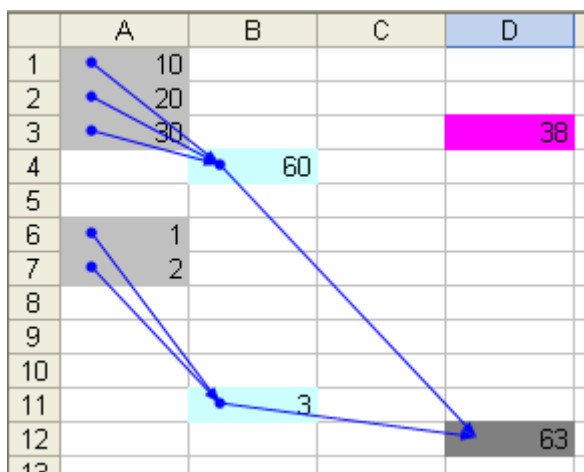

Figure 1. Arrow Tool

Figure 2. Combined Tool

There are very few empirical studies of spreadsheet visualizations. One experiment by Davis (1996) compared two tools, the data-dependency tool and arrow tool. The first tool produced a flowchart that shows cells as inputs, outputs, decision variables, parameters or formulae. It is similar to the input-process-output tool (Chan et al., 2000), but it generates a separate chart instead of superimposing on the spreadsheet. Performance (dependent cell identification) showed no significant differences between the tools, but significantly better than without tools. The second experiment by Davis (1996) compared arrow tools with Excel 3.0's method of listing precedent and dependent cells. There was no significant difference in debugging performance.

RESEARCH QUESTION AND METHODOLOGY

There is an urgent need to identify methods that can lead to better user performance with spreadsheets (Davis, 1996), and in guiding the development of future tools. According to Galletta et al. (1996) and Teo and Lee-Partridge (2001), error factors and presentation factors are important for error detection. Based on the reviews, we are interested to investigate the effects of visualization tools and error types on user ability to find and correct errors. The following hypotheses are proposed: (H1): Visualization tools will affect error detection performance, in terms of accuracy and time. (H2): Data/reference error type will affect error detection performance, in terms of accuracy and time. (H3): Wrong/missing error type will affect error detection performance, in terms of accuracy and time.

The visualization methods tested in this experiment are the two basic methods in the review. The experiment had three subject groups. The arrow group used the arrow method. The combined group uses the arrow plus cell coloring based on the input-process-output model, which aims to inject an additional perspective on the cell roles. The plain group did not have any visualization.

Of the error classifications described in section 2, four are selected for the experiment (missing data, wrong data, missing reference and wrong reference), primarily to make the experiment more manageable. Each group goes through all four error types. Each subject in each group has to analyze four spreadsheet models: an income statement from Panko and Halverson (1997), a bid-a-wall project from Panko and Sprague (1998), an annual student budget from Galletta et al. (1996) and a final grade computation from Chen and Chan (2000). Each model fits into one screen. Each model is shown four times, once with one of the four errors.

The performance measurements are accuracy in correcting errors with one point awarded for each error corrected, and time in seconds taken to work on each error. Participants were first and second year undergraduate computing students. They were paid for participation and performance. The arrow group had 30 subjects, and the other groups had 33 each. Presentation sequence of error types is randomized. Earlier studies have used groups of roughly the same sizes (Brown and Gould, 1987; Galletta et al., 1993; Panko and Halverson, 2001; Reithel et al., 1996). User training was provided to ensure subjects were familiar with the experiment before they began. There was no time limit for the subjects. Before each model was presented, subjects were given a screen containing the spreadsheet model's problem description (this was also given to them in printed form).

RESULTS AND DISCUSSION

Statistical analysis was performed through a general linear model using SPSS. The independent variable is

group (plain, arrow and combined). The repeated measures are for data/reference and wrong/missing type of errors. Both accuracy and time are analyzed in one test. Instead of sticking rigidly to a p-value of 5%, we also consider the p-value of 0.062 to be sufficiently small. "Logically, of course, there is no justification for a sharp line between a 'significant' and a 'non-insignificant' indifference. (Rosenthal, Rosnow and Rubin, 1999, p.5). The significant effects are: group (p=0.039), data/reference (p=0.001), wrong/missing (p=0.062), and the interaction between data/reference and wrong/missing (p=0.005). To understand the specific effects better, these are analyzed with univariate tests. The significant results are shown in table 1.

Source	Measure	F	Sig.
Within-subject			
wrong/missing	time	6.604	.012
data/reference * wrong/missing	time	12.275	.001
Between Subject			
group	accuracy	6.121	.003

Table 1. Univariate Tests

Time for error correction shows more complication. Wrong data errors took longer time than missing data errors. Wrong and missing reference errors took about the same time. Post-hoc tests show that the arrow group is significantly more accurate than the plain group. The combined group is not significantly different from the others. For all errors, the arrow group shows higher accuracy than the plain group. The combined group varies from the highest (wrong reference) to the lowest (missing reference).

CONCLUSION

Spreadsheet models are very widely used and are very likely to contain errors (Panko and Halverson, 2001). It is recognized that traditional methods of checking for errors, such as by scanning on screen or by looking at printed worksheets, are not very effective. Many visualization techniques have been developed to help users understand spreadsheet models, and to find and correct errors (Ballinger et al. 2003; Chen and Chan, 2000; Hendry and Green, 1993; Igarashi et al., 1998; Sajaniemi, 2000; Shiozawa et al., 1999), but there have been few empirical studies on their effectiveness.

This experiment has shown that a very fundamental visualization method (the arrow method) can lead to significantly better accuracy in error correction than the plain method. Furthermore, the arrow method is more accurate than the plain method for all the error types tested. The experiment also shows that the combined method is not clearly better than the plain method, nor

worse than the arrow method. On the one hand, the experiment shows the promise of visualization tools. On the other hand, piling on visualizations may not definitely lead to better error correction. Researchers should be encouraged to develop more visualization tools, and just as importantly, to empirically test them. The experiment also shows the usefulness of the proposed error classification in understanding the difficulties of error correction. Wrong data/reference errors took very much more time to correct than missing data/reference errors..

The practical implication is that spreadsheet users should be trained and be familiar with the arrow tool. This is a much neglected part of user training. A check of many Microsoft Excel books and training courses shows little or no attention to the arrow tool.

There is need for many more studies in this area. For example, future tests can be on more visualization tools, different spreadsheet sizes and complexities, different subject characteristics, different error types, and different tasks, such as model development, error detection, or model comprehension.

ACKNOWLEDGMENTS

REFERENCES

1. Ballinger D., Biddle R. and Noble J. (2003) Spreadsheet Visualisation to Improve End-user Understanding, *Proceedings of the Australian Symposium on Information Visualizatio*n, 24, 99-109.

2. Brown, P.S. and Gould, J.D. (1987) An Experimental Study of People Creating Spreadsheets, *ACM Transactions on Office Information Systems*, 5, 3, 258-272.

3. Chen, Y. and Chan, H.C. (2000) An Exploratory Study of Spreadsheet Debugging Processes, *The Pacific Asia Conference on Information Systems*, 143-155.

4. Chan, H.C., Ying, C. and Peh, C.B. (2000) Strategies and Visualization Tools for Enhancing User Auditing of Spreadsheet Models, *Information and Software Technology*, 42, 5, 1037-1043.

5. Chan, Y.E. and Storey, V.C. (1996) The Use of Spreadsheets in Organizations: Determinants and Consequences, *Information & Management*, 31, 3, 119-134.

6. Clermont, M. (2003) Analyzing Large Spreadsheet Programs, *10th Working Conference on Reverse Engineering, WCRE*, 306-315.

7. Davis, J.S. (1996) Tools for Spreadsheet Auditing, *International Journal of Human Computer Studies*, 45, 4, 429-442.

8. Galletta, D. F., Abraham, D., Louadi, M.E., Lekse, W., Pollalis, Y. A. and Sampler, J. L. (1993), An Empirical Study of Spreadsheet Error Performance, *Accounting, Management, and Information Technology*, 3, 2, 79-95.

9. Galletta, D.F., Hartzel, K.S., Johnson, S., Joseph, J.L. and Rustagi, S. (1996), Spreadsheet Presentation and Error Detection: An Experimental Study, *Journal of Management Information Systems*, 13, 3, 1996-1997, 45-63.

10. Hendry, D. G. and Green, T. R. G. (1993) CogMap: A Visual Description Language for Spreadsheets, *Journal of Visual Languages and Computing*, 4, 1, 35-54.

11. Hendry, D.G. and Green, T.R.G. (1994) Creating, Comprehending and Explaining Spreadsheets: A Cognitive Interpretation of What Discretionary Users Think of the Spreadsheet Model, *International Journal of Human-Computer Studies*, 40, 6, 1033-1065.

12. Igarashi, T., Mackinlay, J. D., Chang, B.-W. and Zellweger, P. (1998) Fluid Visualization for Spreadsheet Structures, *Proceedings of the 14th IEEE Symposium on Visual Languages*, 118-125.

13. McGill, T.J. and Klobas, J.E. (2004) The Role of Spreadsheet Knowledge in User-Developed Application Success, *Decision Support Systems*, In Press.

14. Panko, R.R. (1999) Applying Code Inspection to Spreadsheet Testing, *Journal of Management Information Systems*, 16, 2, 159-176.

15. Panko, R.R. (2000) Spreadsheet Errors: What We Know, What We Think We can do, Proceedings of the Spreadsheet Risk Symposium, European Spreadsheet Risks Interest Group (EuSpRIG), 7-17.

16. Panko, R.R. (2004) Human error website (http://www.cba.hawaii.edu/panko/papers/ss/humanerr.htm) (accessed 2004)

17. Panko, R.R. and Halverson, R.P.Jr. (1996) Spreadsheets on Trial: A Framework for Research on Spreadsheet Risks, *Proceedings of the 29th Hawaii International Conference on System Sciences*, 324-325

18. Panko, R.R. and Halverson, R.P., Jr. (1997) Are two heads better than one? (At Reducing Errors in Spreadsheet Modeling), *Office Systems Research Journal*, 15, 1, 21-32.

19. Panko, R.R. and Halverson, R.P., Jr. (2001) An Experiment in Collaborative Development to Reduce Spreadsheet Errors, *Journal of the Association for Information Systems*, 2, 4, 1-29.

20. Panko, R.R. and Sprague, Jr., R.H. (1998) Hitting the Wall: Errors in Developing and Code Inspecting a 'Simple' Spreadsheet Model, *Decision Support Systems*, 22, 4, 337-353.

21. Reason, J. (1990) *Human Error*, Cambridge University Press, U.K.

22. Reithel, B.J., Nichols, D.L. and Robinson, R.K. (1996) An Experimental Investigation of the Effects of Size, Format and Errors on Spreadsheet Reliability Perception, *Journal of Computer Information Systems*, 36, 54-64.

23. Rosenthal, R., Rosnow, R.L. and Rubin, D.B. (1999) *Contrasts and Effect Sizes in Behavioral Research : A Correlational Approach*, Cambridge University Press.

24. Saariluoma, P. and Sajaniemi, J. (1989) Visual Information Chunking in Spreadsheet Calculation, *International Journal of Man-Machine Studies*, 30, 2, 475-488.

25. Saariluoma, P. and Sajaniemi, J. (1991) Extracting Implicit Tree Structures in Spreadsheet Calculation, *Ergonomics*, 34, 8, 1027 – 1046.

26. Sajaniemi, J. (2000) Modeling Spreadsheet Audit: A Rigorous Approach to Automatic Visualization, *Journal of Visual Languages and Computing*, 11, 49-82.

27. Shneiderman, B. (1998) *Designing the User Interface: Strategies for Effective Human Computer Interaction*, 3rd Edition, Addison Wesley Longman, Inc, Massachusetts.

28. Shiozawa H., Okada K. and Matsushita Y. (1999) 3D Interactive Visualization for Inter-Cell Dependencies of Spreadsheets, *IEEE Symposium on Information Visualization*, 79-82.

29. Teo, T.S.H. and Tan, M., (1997) Quantitative and Qualitative Errors in Spreadsheet Development, *Proceedings of the Thirtieth Hawaii International Conference on System Sciences*.

30. Teo, T.S.H. and Tan, M. (1999) Spreadsheet Development and 'What-if' Analysis: Quantitative versus Qualitative Errors, *Accounting, Management and Information Technologies*, 9, 3, 141-160.

31. Teo, T.S.H. and Lee-Partridge, J.E. (2001) Effects of Error Factors and Prior Incremental Practice on Spreadsheet Error Detection: An Experimental Study, *Omega*, 29, 5, 445-456.