Spreadsheet Knowledge: Measuring What User Developers Know

Jane Klobas
Graduate School of Management
University of Western Australia
GPO Box 1164, Nedlands WA 6909, Australia.
and
Bocconi University, Milan, Italy
jane.klobas@uni-bocconi.it

Tanya McGill
School of Information Technology
Murdoch University
South St. Murdoch WA 6150, Australia.
mcgill@murdoch.edu.au

ABSTRACT

Spreadsheets are frequently used to conduct analyses using organizational data. Spreadsheets, however, frequently contain errors and these affect the quality of the analyses performed by the users who develop their own spreadsheets. This paper describes the development of a diagnostic test for spreadsheet knowledge. The test is designed to be used by individuals, teachers, trainers, and organizations to identify the spreadsheet development training needed by spreadsheet user developers and to examine the success of spreadsheet training programs. Reliability and validity of the test are reported.

Keywords: Spreadsheets, Spreadsheet Knowledge, Training, Spreadsheet Quality, End User Development, End User Computing

1. INTRODUCTION

Many important organizational information systems are developed by their users. Employers acknowledge spreadsheet skills to be among the most beneficial information technology skills a new employee can have (Davis, 1997; Ives, Valacich, Watson, Zmud et al., 2002). Often, quite critical organizational data, and analyses based on these data, are entrusted to individuals who produce spreadsheets to conduct their analyses (Govindarajulu, 2003). Spreadsheets, however, frequently contain errors, and these errors affect the quality of the analyses performed by the users who design and build their own spreadsheets (the 'user developers') (Panko and Halverson, 2001).

Educators suggest that spreadsheet quality would be improved by teaching students principles of system design and maintenance as well as how to use the various features of spreadsheets (Teo and Tan, 1999). Improving the quality of university education for spreadsheet user developers should improve the quality of spreadsheets developed by new entrants to the workplace. In addition, appropriate training courses need to be developed for spreadsheet user developers who are already in the workplace (Govindarajulu, 2003; Kreie, Cronan, Pendley and Renwick, 2000).

User developers in the workplace need, however, to be motivated to attend spreadsheet training courses. We cannot assume that user developers are aware either that their spreadsheets contain errors or that the quality of their spreadsheets can be improved (McGill, 2002). Noting that accurate self-knowledge of spreadsheet expertise is rare and difficult to obtain, Hall (1996) called for development of a 'spreadsheet metric'. More recently, Torkzadeh and Lee (2003) reiterated the need for measures of end user computing skills. This paper takes up the challenge and introduces a diagnostic test for spreadsheet knowledge. The test is designed to be used by individuals, teachers,

trainers, and organizations to identify the spreadsheet development training needed by spreadsheet user developers. Our goals, in developing this test, were to:

- measure a range of knowledge that contributes to effective development of accurate spreadsheets,
- develop a test that would validly test the knowledge of user developers, from novices to experienced spreadsheet users and developers, and
- ensure that the test provided reliable results, from a statistical point of view.

1.1 The Nature of Spreadsheet Development Knowledge

Very little research has looked explicitly at defining and measuring spreadsheet development knowledge or skill. The most common approach has been simply to use spreadsheet training or experience as a surrogate for spreadsheet knowledge (Chan and Storey, 1996; Harrison and Rainer, 1992; Janvrin and Morrison, 2000; Panko and Sprague, 1999; Rivard and Huff, 1988). Implicit in this approach is the assumption that experience and training lead to greater levels of end user knowledge and skill. But some research into end user development calls this assumption into question. For example, Chan and Storey (1996) found no relationship between computer training and spreadsheet proficiency and McGill (2002) found no relationship between spreadsheet experience and spreadsheet quality. An instrument that explicitly measures spreadsheet knowledge is required.

User competence has been defined as "the user's potential to apply technology to its fullest possible extent so as to maximize performance of specific job tasks" (Marcolin, Compeau, Munro and Huff, 2000 p. 38). This definition certainly applies to the use of spreadsheet software to develop systems that support a user's work. Two related forms of knowledge appear to be required in order to be a competent user developer: knowledge associated with effective use of the development tool itself (Bowman, 1988; Cheney and Nelson, 1988), and knowledge associated with the development of information systems in general (Bowman, 1988; Janvrin and Morrison, 2000). Thus, competent end user developers require knowledge of the general characteristics of the type of tool being used (e.g. spreadsheet or database management system) and knowledge of the specific features of the package chosen (e.g. Microsoft Excel or Microsoft Access). They also require system development knowledge which includes the ability to model real world problems as well as knowledge of appropriate systems analysis and design techniques. Together these related forms of knowledge are believed to enable an end user developer to produce easy to use, reliable and maintainable applications (Bowman, 1988). We can therefore identify two categories of knowledge required for competent end user development of spreadsheets:

a) Knowledge of the tool: spreadsheet features in general, and the features of the specific spreadsheet packages being used, and
b) System development knowledge, as applied to spreadsheet development.

2. METHOD

Psychometric principles and techniques were used to develop the diagnostic test for spreadsheet development. The detail of the methods used for development and testing is provided in this section.

2.1 Development and Pilot Testing of Spreadsheet Knowledge Questions

The first stage in development of the diagnostic test to measure spreadsheet knowledge was development of a pilot test. The pilot test drew, as far as possible, on existing tests that could be adopted or adapted to address the two sets of knowledge: spreadsheet features and spreadsheet development knowledge.

Kreie's (1998) spreadsheet knowledge instrument was used as the basis for the questions to measure knowledge of spreadsheet features. Nine of Kreie's 17 items were selected for inclusion. Spreadsheet development knowledge was measured using two sorts of questions. Questions to test knowledge of processes of spreadsheet development were developed specifically for this study and drew upon two published methodologies for the development of spreadsheets (Ronen, Palley and Lucas, 1989; Salchenberger, 1993). These questions covered areas such as the need for modeling and planning, and methods of testing. The second source of spreadsheet development questions was material drawn from Rivard et al.'s (1997) instrument to measure the quality of end user developed applications.

Each item on the test was presented as a multiple choice question with 5 options; in each case the fifth option was 'I don't know' or 'I am not familiar with this feature'. The test was examined for content validity (comprehensiveness) by four information technology academics who have been involved in teaching spreadsheet use and design. A few revisions were made on the basis of their suggestions. The resulting 32 item test was piloted with 60 predominantly mature-aged students enrolled in undergraduate business degrees. They reported an average of 4.4 years spreadsheet experience with a minimum of just a few weeks and a maximum of 13 years. The students were recruited during class and completed the knowledge test on the spot. It was stressed that completion of the test was voluntary and that it formed no part of their assessment.

The pilot test was statistically reliable (Cronbach's alpha of 0.77 indicates that scores on the test questions are internally consistent). Seven questions did not, however, discriminate well among the 60 students when the questions were subjected to a Guttman analysis (Guttman, 1950). These questions were removed.

The resulting spreadsheet knowledge test contained 25 multiple choice questions. Nine of the questions test knowledge of spreadsheet features, and 16 questions relate to spreadsheet development knowledge (eight of these concern the spreadsheet development process and eight relate to knowledge of spreadsheet quality). Spreadsheet knowledge is represented by the number of correct questions on the test. The test is reproduced in the Appendix.

2.2 Examining the Quality of the Test

We collected a new sample of 159 spreadsheet user developers to enable examination of the quality of the 25 item spreadsheet knowledge test. Recruitment of this sample is described in the next section.

We examined the test's quality from several points of view. Where the data permitted (more demographic data was gathered in the main study than in the pilot study), we pooled the responses from the pilot and main samples in order to have a larger and more varied sample for psychometric analyses where larger sample size and more variation is preferred to a smaller sample drawn from a single sampling frame. The description of each analysis indicates whether the main sample or the pooled sample was used. In these cases, both individual sample and pooled sample results are reported where appropriate. The analyses conducted were:

- *Content validity*. The test's coverage of the range of knowledge required for spreadsheet development had already been examined by submitting the pilot test to experienced teachers of spreadsheet design and use. We used factor analysis to test for evidence of more than one dimension among the questions included in the test. We pooled the responses of the two groups of participants (students, from the pilot test, and new sample of user developers) for this analysis.
- *Reliability*. We calculated Cronbach's alpha for the new sample to confirm that item scores were internally consistent across different samples.
- *Additivity*. Tukey's test of additivity was used to ensure that scores on the questions in the test could be added to achieve a composite score.
- *Range of difficulty*. We used Rasch item response analysis (Andrich, Sheridan, Lyne and Luo, 1998) to identify the extent to which the test measures spreadsheet knowledge across a range of levels of difficulty. We pooled the results of the two samples for this analysis.
- *Nomographic validity* is the ability of a test to discriminate between groups of participants for whom differences are expected. Based on the assumptions of earlier research, we expected to observe differences in spreadsheet knowledge among people with different levels of training and experience. We pooled the responses of the two groups of participants for this test, and compared the scores of a) those participants who had some formal spreadsheet training with those who had little or none, and b) spreadsheet novices (less than 2 years experience) with participants who had a moderate amount of experience (2 to less than 5 years) and those who were very experienced (more than 5 years).
- *Predictive validity* is the validity of a test to predict performance on a subsequent task. Because the spreadsheet knowledge test was designed to test knowledge associated with development of good quality spreadsheets, we used linear regression to test the extent to which spreadsheet knowledge explained the quality of spreadsheets developed by the study participants.

The quality of each of the spreadsheets was also independently assessed by two experienced information systems developers. The items used to measure system quality were obtained from the instrument developed by Rivard et al. (1997) to assess the quality of user developed applications. For this study, items that were not appropriate for the spreadsheets under consideration (e.g. specific to database applications) were excluded. Minor adaptations to wording were also made to reflect the environment in which spreadsheet development and use occurred. The resulting system quality scale consisted of 20 items, each scored on a Likert scale of 1 to 7 where (1) was labeled 'strongly agree' and (7) was labeled 'strongly disagree'. A typical question item was "Errors in the spreadsheet are easy to identify". The instrument was shown to be reliable with a Cronbach's alpha of 0.94. Agreement between the two independent assessors was high ($r = 0.80$, $p < 0.001$) and averages of the ratings for each item were used to calculate system quality, which was created as a composite variable using the factor weights obtained from measurement model development using AMOS 3.6.

2.3 Recruitment of Participants

The test is designed to be administered to end users (as distinct from information systems experts) who develop their own spreadsheets. To ensure that the test was appropriate for use with end users, a sample of user developers with a broad range of spreadsheet knowledge was needed. We therefore sought study participants from members of the public. As an incentive to participate, we offered a one hour spreadsheet training course on 'Developing Spreadsheet Applications' and $20 to compensate for parking costs, petrol and inconvenience. Recruitment occurred firstly through a number of advertisements placed in local newspapers calling for volunteers. These were followed by e-mails to three large organizations that had expressed interest in the study and finally word of mouth brought forth some additional participants. Whilst being essentially a convenience sample, the participants covered a broad spectrum of ages, spreadsheet experience and training. Of the 159 participants in this sample, 32.7% were male and 67.3% female and their ages ranged from 14 to 77 with an average age of 42.7. They reported an average of 4.5 years experience using spreadsheets (with a range from less than 1 year to 21 years).

2.4 Procedure

Fourteen separate experimental sessions of approximately four hours were held over a period of five months. The number of participants in each session ranged from seven to seventeen, depending on availability.

Each session was conducted in several parts. In the first part, participants completed a short questionnaire about themselves and their previous experience with spreadsheets and then completed the spreadsheet knowledge test. They were then given a problem statement related to making choices between car rental companies, and asked to develop a spreadsheet to solve it using Microsoft Excel. They were encouraged to treat the development exercise as they would a personal or work task, rather than as a test. In the next part, each participant used the spreadsheet to answer a set of questions. The questions ranged from comparison of the three rental firms when no excess kilometer charges are imposed through to questions where excesses are applied and basic parameters are assumed to have changed from those given in the original problem description. A typical question is 'Which rental company is the cheapest if you wish to hire a car for 6 days and drive approximately 1500 kilometers with it?'. After completing this task, they completed a final questionnaire to gather their perceptions of the quality of their spreadsheet and satisfaction with their use of it. The final part of the session consisted of the training course. More detail of the procedure is provided in McGill (2004).

3. RESULTS

3.1 Content Validity and Dimensionality

Confirmatory factor analysis (principal components analysis with oblimin rotation of .25) identified two weak factors which together explained 26% of the variance in scores on the knowledge test. The factors match the two areas of knowledge needed for competent use of spreadsheets:

- Knowledge of the tool (knowledge of spreadsheet features),
- Spreadsheet development knowledge (knowledge of modeling and systems design and evaluation principles which might be applied to spreadsheets).

Table 1 lists questions in each of these categories.

This solution, although suggesting that we have tapped two spreadsheet knowledge domains that reflect the two generic competencies needed by user developers, is statistically weak. It includes only 18 of the 25 questions in the test. The other questions could not be classified. These characteristics suggest that the test is uni-dimensional, that is, that it may be used reliably to measure generic spreadsheet knowledge, which includes the two spreadsheet knowledge domains, but not to measure knowledge of each of the individual domains on its own.

3.2 Reliability

To test that the 25 questions in the test together formed a single, reliable scale, we calculated Cronbach's alpha for the sample of 159 spreadsheet users who developed the car rental application. In this sample, alpha was .78, similar to that of the pilot sample. Alpha for the two samples, pooled, was .80, and therefore above the threshold for a satisfactory scale.

Table 1: Test questions representing domains of spreadsheet development knowledge

Question	Factor loading
Knowledge of spreadsheet functions	
Q1. Identification of the row in a cell reference	-0.50
Q6. Currency feature	-0.54
Q9. COUNT function	-0.52
Q11. Meaning of absolute cell reference	-0.51
Q12. Order of arithmetic operations	-0.50
Q16. IF function	-0.56
Q17. Formula errors	-0.47
Q21. MAX function	-0.43
Q22. Spreadsheet protection features	-0.65
System development knowledge	
Q3. Criteria for effective spreadsheets	0.44
Q4. Characteristics of high quality spreadsheet	0.36
Q13. Determining the necessary input data for a spreadsheet	0.41
Q15. Methods for testing spreadsheets	0.35
Q18. Reasons for spreadsheet documentation	0.58
Q19. Reasons for planning on paper	0.61
Q23. Aspects of spreadsheet documentation	0.60
Q24. Characteristics of a well-designed spreadsheet	0.54
Q25. Improving a given spreadsheet	0.36

3.3 Additivity

The test can be scored, as proposed, by counting the number of correct answers. The test questions were additive when they were administered to the new sample (Tukey's power for additivity = 1, $p = .98$).

3.4 Range of Difficulty

The questions in the test ranged from easy to difficult, although there was a slight imbalance toward easier questions. The simplest item was question 1 (identification of reference to a row), while the most difficult were questions 20 (use of reference fields) and 8 (operation involving an absolute reference).

Rasch analysis also ranks participants in order of score. The participant ranking suggested that it may be possible to develop additional questions that discriminate more finely by level of difficulty. In particular, some more moderately easy questions would help to discriminate better among the knowledge of those participants with moderately low scores, while some moderately difficult questions would help to discriminate better among those

participants whose scores are clustered around and above the mid-point.

3.5 Nomographic Validity: Differences by Level of Training and Experience

Scores on the test vary, as expected, with formal spreadsheet training and years of spreadsheet experience. These differences are summarized in Table 2. Across the pooled sample (n=219), it was possible to observe that user developers with little or no training received significantly lower scores on the test (F=25.15, df=2,213, p<0.001). Those with novice levels of spreadsheet experience also performed significantly worse than did more experienced or expert participants (F=8.11, df=1,213, p<0.001).

Table 2: Differences in test scores among groups of participants

	n	Mean	sd
Formal training			
Little or none	103	11.1	4.4
More than a little	116	15.4	4.3
Total	*219*	*13.3*	*4.8*
Experience			
Novice (< 2 years)	63	10.2	4.3
Experienced (2 to < 5 years)	71	13.8	4.1
Expert (5 or more years)	85	15.3	4.6
Total	*219*	*13.3*	*4.8*

The differences in spreadsheet knowledge scores by training and experience are illustrated in Figure 1. This figure shows how scores increased with experience, as well as how participants who had more than a little spreadsheet training performed better than those with little or no training.

3.6 Predictive Validity: Use of Spreadsheet Knowledge to Predict Spreadsheet Quality

Spreadsheet knowledge was correlated with independently rated spreadsheet quality (r=.44, n=159, p<.001). Spreadsheet quality could be explained partially (19%) by spreadsheet knowledge as measured by the spreadsheet knowledge test (F=36.99, df=1,157, p<0.001).

4. DISCUSSION

The spreadsheet knowledge test described in this paper will help organizations to diagnose the level of spreadsheet knowledge among their employees; it will also help instructors to determine the level of spreadsheet knowledge of their students. It can also be used by individual spreadsheet users and developers to test their own spreadsheet knowledge. The test is easily administered and scored. It meets statistical standards for additivity and reliability. It is able to distinguish between the spreadsheet knowledge of users with different levels of training and experience. Importantly, scores on the test can be used to predict the quality of spreadsheets developed by users. Figure 2 illustrates these relationships.

Figure 1: Training and experience differences in spreadsheet knowledge, pooled sample (n=219)

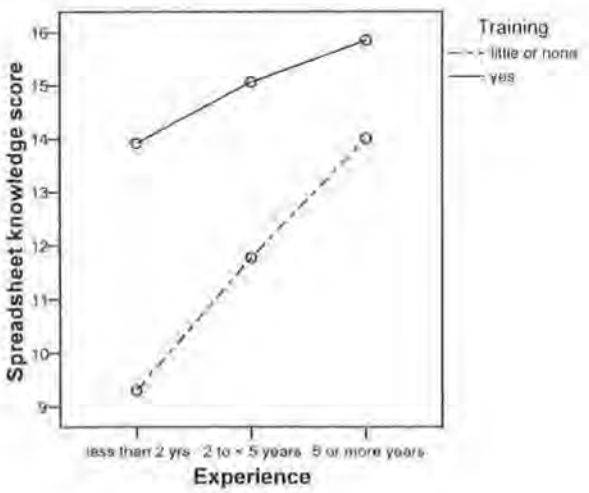
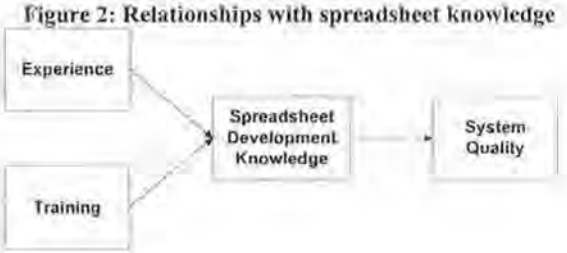


Figure 2: Relationships with spreadsheet knowledge



4.1 Improving Spreadsheet Quality

The relationships summarized in Figure 2 suggest that spreadsheet training that develops knowledge of spreadsheet features and spreadsheet development practices should contribute to improvements in spreadsheet quality. However, spreadsheet quality is only partially explained by spreadsheet knowledge. Other factors that may contribute to spreadsheet quality include the general intelligence of the user developer and their motivation for the task (Marcolin et al., 2000), and possibly, the user's knowledge of the problem domain. Several authors have explored the role of domain knowledge in the success of end user development, but have concluded that the limiting factor is usually knowledge of the development tool rather than knowledge of the domain for which the application is being developed (Agboola, 1998; Galletta et al., 1993; Mackay and Elam, 1992). Nonetheless, future research could usefully examine the relative influence of training, the general ability of the user developer, user developer motivation and user developer knowledge of both spreadsheets and the problem domain on spreadsheet quality.

4.2 Opportunities for Further Research on Spreadsheet Knowledge

Although the items in our diagnostic test are designed to measure spreadsheet knowledge along the two dimensions needed for competent end user development – knowledge of spreadsheet features and spreadsheet development knowledge – we do not recommend use of this test to measure knowledge on each of these dimensions, separately. The factor analysis described in this paper confirmed that, while measurements on the two underlying dimensions could be identified, the scale is most effective when it is used in its entirety. We do not know if this is because the interplay between these two dimensions of spreadsheet knowledge is so important that spreadsheet knowledge on either dimension cannot be separated from spreadsheet knowledge on the other, or if a different form of test would be required to make such a distinction. Future research into the nature and measurement of spreadsheet knowledge should examine each of these possibilities.

While the test is capable of distinguishing between different levels of spreadsheet knowledge, and the Rasch analysis confirmed that questions on the test ranged from simple to difficult, it may be possible to develop a test that is even more sensitive to differences in spreadsheet knowledge for use in situations where greater sensitivity is required. A test based on the current test, but including some additional difficult items, could be developed, and the procedure described in this paper followed to confirm that it had the desired qualities.

5. CONCLUSION

The diagnostic test for spreadsheet knowledge described in this paper meets the need for a 'spreadsheet metric'. Not only is it valid and reliable, it is also useful. It distinguishes different levels of spreadsheet knowledge both among university students and among users who develop spreadsheets at home and in the workplace. Spreadsheet knowledge, as measured with the test, acts as a link between training and experience and spreadsheet quality. The data gathered in this study provide evidence that it is worth training people to develop spreadsheets because spreadsheet knowledge varies with training as well as with experience, and better spreadsheet knowledge results in higher spreadsheet quality. The test could therefore be used both to identify the need for spreadsheet training and to examine the success of spreadsheet training programs.

6. REFERENCES

Agboola, Issac O. (1998), "An Experimental Study of the Effects of Application Domain Knowledge and Data Modeling on the Quality of Database Implementations by Non-Experts," PhD, University of Maryland College Park.

Andrich, David, Barry Sheridan, Alan Lyne, and Guanzhong Luo (1998), Rumm: A Windows-Based Item Analysis Program Employing Rasch Unidimensional Measurement Models, Murdoch University, Perth.

Bowman, Brent (1988), "An Investigation of Application Development Process Controls," Ph.D., University of Houston.

Chan, Yolande E., and Veda C. Storey (1996), "The Use of Spreadsheets in Organizations: Determinants and Consequences." Information & Management, Vol 31, pp. 119-134.

Cheney, Paul H., and R. Ryan Nelson (1988), "A Tool for Measuring and Analyzing End User Computing Abilities." Information Processing and Management, Vol 2 No 2, pp. 199-203.

Davis, P. (1997), "What Computer Skills Do Employers Expect from Recent College Graduates." Technological Horizons in Education Journal, pp. 74-78.

Galletta, Dennis F., Dolphy Abraham, Mohamed El Louadi, William Lekse, Yannis A. Pollalis, and Jeffrey L. Sampler (1993), "An Empirical Study of Spreadsheet Error-Finding Performance." Journal of Accounting, Management and Information Technology, Vol 3 No 2, pp. 79-95.

Govindarajulu, Chittibabu (2003), "End Users: Who Are They?" Communications of the ACM, Vol 46 No 9, pp. 152-159.

Guttman, Louis (1950), "The Basis for Scalogram Analysis." Measurement and Prediction, S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarfeld, and J. A. Clausen, eds., Wiley, New York, 1950, pp. 60-90.

Hall, Maria J. J. (1996), "A Risk and Control Oriented Study of the Practices of Spreadsheet Application Developers." Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences, pp. 364-373.

Harrison, Allison W., and Kelly Rainer (1992), "The Influence of Individual Differences on Skill in End-User Computing." Journal of Management Information Systems, Vol 9 No 1, pp. 93-111.

Ives, Blake, J. Valacich, R. Watson, R. Zmud, et al. (2002), "What Every Business Student Needs to Know About Information Systems." Communications of the Association for Information Systems, Vol 9, pp. 467-477.

Janvrin, Diane, and Joline Morrison (2000), "Using a Structured Design Approach to Reduce Risks in End User Spreadsheet Development." Information & Management, Vol 37 No 1, pp. 1-12.

Kreie, Jennifer (1998), "On the Improvement of End-User Developed Systems Using Systems Analysis and Design," PhD, University of Arkansas.

Kreie, Jennifer, Timothy Paul Cronan, John Pendley, and Janet S. Renwick (2000), "Applications Development by End-Users: Can Quality Be Improved?" Decision Support Systems, Vol 29 No 2, pp. 143-152.

Mackay, Jane M., and Joyce J. Elam (1992), "A Comparative Study of How Experts and Novices Use a Decision Aid to Solve Problems in Complex

Knowledge Domains." Information Systems Research, Vol 3 No 2, pp. 150-172.

Marcolin, Barbara L., Deborah R. Compeau, Malcolm C. Munro, and Sid L. Huff (2000), "Assessing User Competence: Conceptualization and Measurement." Information Systems Research, Vol 11 No 1, pp. 37-60.

McGill, Tanya (2004), "The Effect of End User Development on End User Success." Journal of End User Computing, Vol 16 No 1, pp. 41-58.

McGill, Tanya (2002), "User Developed Applications: Can End Users Assess Quality?" Journal of End User Computing, Vol 14 No 3, pp. 1-15.

Panko, Raymond R., and Richard P. Halverson (2001), "An Experiment in Collaborative Development to Reduce Spreadsheet Errors." Journal of the Association for Information Systems, Vol 2 Article 4.

Panko, Raymond R., and Ralph H. Sprague (1999), "Hitting the Wall: Errors in Developing and Code Inspecting a 'Simple' Spreadsheet Model." Decision Support Systems, Vol 22 No 4, pp. 337-353.

Rivard, Suzanne, and Sid L. Huff (1988), "Factors of Success for End-User Computing." Communications of the ACM, Vol 31 No 5, pp. 552-561.

Rivard, Suzanne, Guylaine Poirier, Louis Raymond, and Francois Bergeron (1997), "Development of a Measure to Assess the Quality of User-Developed Applications." The DATA BASE for Advances in Information Systems, Vol 28 No 3, pp. 44-58.

Ronen, Boaz, Michael A. Palley, and Henry C. Lucas (1989), "Spreadsheet Analysis and Design." Communications of the ACM, Vol 32 No 1, pp. 84-93.

Salchenberger, Linda (1993), "Structured Development Techniques for User-Developed Systems." Information & Management, Vol 24, pp. 41-50.

Teo, Thompson S H, and Margaret Tan (1999), "Spreadsheet Development and 'What-If' Analysis: Quantitative Versus Qualitative Errors." Accounting, Management & Information Technology, Vol 9, pp. 141-160.

Torkzadeh, Gholamreza, and Jungwoo Lee (2003), "Measures of Perceived End-User Computing Skills." Information & Management, Vol 40, pp. 607-615.

AUTHOR BIOGRAPHIES

Jane Klobas is a research professor at Bocconi University, Milan, Italy and Professorial Fellow at the Graduate School of Management, University of Western Australia. She has a PhD in the psychology of information system and information resource use. Her research concerns evaluation and use of information systems and information resources, and incorporates elements of psychometrics and social and economic impact studies. She has published widely on information management and applications of educational technology.

Tanya McGill is a senior lecturer in the School of Information Technology at Murdoch University in Western Australia. She has a PhD from Murdoch University. Her major research interests include end user computing and information technology education. Her work has appeared in various journals including the Journal of Research on Computing in Education, European Journal of Psychology of Education, Journal of the American Society for Information Science, and Journal of End User Computing.

APPENDIX

Diagnostic Test for Spreadsheet Knowledge

For each of the following questions, please circle the answer that you think is correct. If you are not sure of the answer to a question, don't worry, just circle option **e.** (I am not familiar with this feature OR I don't know).

1. For cell B6, the B refers to the:
   a. Row.
   b. Column.
   c. Cell.
   d. Address.
   e. I am not familiar with this spreadsheet feature.

2. A spreadsheet that is user-friendly:
   a. Is easy to use even if you haven't used it for a long while.
   b. Could be used in other organisations without major modifications.
   c. Does not contain errors.
   d. Is small enough to see all of it on one screen.
   e. I am not familiar with the term user-friendly.

3. Which of the following is NOT a criterion for an effective spreadsheet?
   a. It is small.
   b. It is accurate.
   c. It is easy to change.
   d. It is standardised and consistent.
   e. I don't know.

4. Which of the following is **NOT** a characteristic of a high quality spreadsheet?
   a. Ease of use.
   b. Complexity.
   c. Informativeness.
   d. Modularity.
   e. I don't know.

5. When you need to create a new spreadsheet, the FIRST thing you should do is:
   a. Plan the layout of the spreadsheet on paper.
   b. Work out exactly what the spreadsheet has to do.
   c. Start up your spreadsheet program.
   d. See if you have a previous spreadsheet that you could adapt.
   e. I don't know.

6. If you want the numbers in your spreadsheet to appear as currency (that is with $ signs, etc), you would use the:
   a. Edit feature.
   b. Data feature.
   c. Format feature.
   d. Label feature.
   e. I am not familiar with this spreadsheet feature.

7. A spreadsheet is more likely to be useful over a long period of time if:
   a. Errors are easy to identify.
   b. It is easy to understand the calculations it uses.
   c. It has detailed documentation.
   d. All of the above are true.
   e. I don't know.

8. If you copied the formula =$A$1*B1 from cell C1 to cell C2, the formula in cell C2 would be:
   a. =$A$1*B2.
   b. =$A$2*B2.
   c. =$A$1*B1.
   d. =$A1*B2.
   e. I am not familiar with this spreadsheet feature.

9. Suppose your spreadsheet contains student names and test scores. You can quickly determine how many students on your list with the ___ function.
   a. CALCULATE.
   b. QUERY.
   c. COUNT.
   d. MODAL.
   e. I am not familiar with this spreadsheet function.

10. Dividing your spreadsheet into sections is important because it:
   a. Makes it look more professional.
   b. Enhances the compatibility.
   c. Makes it easier to use and change.
   d. Increases the data storage capacity.
   e. I don't know.

11. An absolute cell reference:
   a. Means you used a cell name rather than the column letter and row number.
   b. Defines what default cell format the spreadsheet uses.
   c. Displays only absolute values.
   d. Always points to the same cell.
   e. I am not familiar with this spreadsheet feature.

12. For the spreadsheet formula =B11+B12+B13/A8+A9, which arithmetic operation is performed FIRST?
   a. The values in cells B11 and B12 are added together.
   b. The values in cells B11, B12 and B13 are added together.
   c. The values in cell B13 is divided by the value in A8.
   d. The values in A8 and A9 are added together.
   e. I am not familiar with this spreadsheet feature.

13. In order to determine what input data is required for a spreadsheet you need to:
   a. Know what problem the spreadsheet will be used to solve.
   b. Know what questions the spreadsheet will be used to answer.
   c. Know what outputs are required from the spreadsheet.
   d. All of the above.
   e. I don't know.

Book2

	A	B	C
1	Projected S		
2			
3		Estimated growth	0.1
4			
5		1999	2000
6	Tennis	55380	=B6+B6*$C$3
7	Golf	176000	=B7+B7*$C$3
8	Camping	255650	=B8+B8*$C$3
9	Fitness	300175	=B9+B9*$C$3
10	Total	=SUM(B6:B9)	=SUM(C5:C9)
11			

14. Which of the following ISN'T a section that spreadsheets should normally include:
   a. Documentation section.
   b. Input section.
   c. Development section.
   d. Output section.
   e. I don't know.

15. Which of the following is a method for testing spreadsheets:
   a. Check the logic of your calculations.
   b. Calculate some results by hand.
   c. Verify input values.
   d. All of the above are methods for testing spreadsheets.
   e. I don't know.

16. What is the function that carries out an evaluation (e.g. Is C1 = 10?) and executes either a 'true' or a 'false' action based on the outcome of the evaluation? (Assume the function is preceded by the appropriate symbol for Lotus 1-2-3 or for Microsoft Excel).
   a. BRANCH.
   b. SELECT.
   c. COMPARE.
   d. IF.
   e. I am not familiar with this spreadsheet feature.

17. How many errors does the spreadsheet below have in its formulas?
   a. 0.
   b. 1.
   c. 2.
   d. 5.
   e. I don't know.

Book2

	A	B	C
1	Projected Sales		
2			
3		Estimated growth	10%
4			
5		1999	2000
6	Tennis	$ 55,380	$ 60,918
7	Golf	$ 176,000	$ 193,600
8	Camping	$ 255,650	$ 281,215
9	Fitness	$ 300,175	$ 330,193
10	Total	$ 787,205	$ 867,926
11			

18. Which of the following is **NOT** a reason for documenting a spreadsheet:
    a. It helps other people to understand how to use the spreadsheet.
    b. It helps other people to understand what the spreadsheet does.
    c. It saves other people from having to use your spreadsheet.
    d. It helps you to remember what the spreadsheet does.
    e. I don't know.

19. Which of the following is **NOT** a reason for planning your calculations on paper:
    a. It allows you to make sure you understand the calculation before trying to create a formula for it in your spreadsheet package.
    b. It makes it easier to get someone else to check your logic.
    c. It reduces the likelihood of making errors.
    d. It saves computer processing time.
    e. I don't know.

20. Values that are referred to in more than one formula should be:
    a. Checked carefully to make sure they are the same in each formula.
    b. Avoided whenever possible.
    c. Referenced using relative references.
    d. Stored in a separate section.
    e. I don't know.

21. If you have a long column of test scores and you want to know the highest test score, you could use:
    a. The IF function.
    b. The SCORE function.
    c. The MAX function.
    d. The HIGH function.
    e. I am not familiar with this spreadsheet feature.

22. If you want to prevent changes from being made to a spreadsheet, you would use:
    a. The sheet and worksheet protection features.
    b. The input restrictions in the tools menu.
    c. The autofilter.
    d. The restrict option in the worksheet setup.
    e. I am not familiar with this spreadsheet feature.

23. Which of the following is an important aspect of a spreadsheet's documentation:
    a. The purpose of the spreadsheet.
    b. The information needed to use the spreadsheet.
    c. The name of the author of the spreadsheet.
    d. All of the above are important.
    e. I don't know.

24. Which of the following is **NOT** a characteristic of a well-designed spreadsheet?
    a. Each section of the spreadsheet has a unique function.
    b. It can be printed out on one page.
    c. Corrections are easy to make.
    d. All headings and labels provide clear information about the data they relate to.
    e. I don't know.

25. What would **MOST** improve the quality of the spreadsheet below?
    a. Naming the worksheet.
    b. Adding information about spreadsheet purpose.
    c. Increasing the column widths.
    d. Removing the blank line.
    e. I don't know.
    f.

Information Systems & Computing
Academic Professionals

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the Journal of Information Systems Education have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 1055-3096