Available online at www.sciencedirect.com
ScienceDirect
Omega 37 (2009) 579–590
omega
The International Journal of Management Science
www.elsevier.com/locate/omega

A comparison of spreadsheet users with different levels of experience☆

Barry R. Lawson*, Kenneth R. Baker, Stephen G. Powell, Lynn Foster-Johnson

Tuck School of Business at Dartmouth College, HB 9000, Hanover, NH 03755, USA

Received 20 November 2007; accepted 16 December 2007
Available online 28 December 2007

Abstract

How do experienced spreadsheet users compare to inexperienced ones, and what light can this comparison shed on spreadsheet best practices? This is the question we address in this paper, using the results from a survey of nearly 1600 respondents. This survey was completed by a wide range of spreadsheet users and focused on their significant characteristics and practices. We were interested in their training, experience, collaboration, and quality control methods. We also examined the number of spreadsheet functions they used regularly, the manner in which they created spreadsheets, and the types of tests they used to check results. We compared two subgroups corresponding to two extremes with respect to their self-reported level of experience and skill. Each subgroup was represented by roughly 10% of the total respondents. Our results suggest that there is a substantial difference between these groups, not only in their personal backgrounds and the corporate setting within which they work, but also in their individual spreadsheet skills and practices. We find that the most experienced subgroup exhibits many desirable characteristics and practices.

Keywords: Spreadsheets; End-user computing; Spreadsheet practices; Decision-support systems; Information systems; MIS

1. Introduction

The prolific use of spreadsheets in industry underscores the degree to which individuals and organizations rely on them for record keeping, analysis, prediction, and decision making. Because spreadsheets are easy to learn and capable of sophisticated analyses, they have been accepted by users spanning a broad continuum from beginner to expert. Recent examples on the theme of spreadsheet modeling in this journal include [1–5]. The flexibility of spreadsheets also allows them to be used without great discipline. Poor design practices and errors are too easily introduced and not so easily detected. The resulting risk calls forth the need for improved spreadsheet management.

In the Tuck Spreadsheet Engineering Research Project (SERP), we have examined current organizational practice as it relates to the use of spreadsheets, with the ultimate aim of developing a set of good practices for creators and users. An early step in this research was to document how spreadsheets are currently being used. For this purpose, we developed a detailed questionnaire that could be administered on the internet. The questionnaire was made available to

☆ This manuscript was processed by Area Editor B. Lev.
* Corresponding author. Tel.: +1 802 592 3632; fax: +1 603 646 8711.
E-mail address: lawson384@charter.net (B.R. Lawson).

0305-0483/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.omega.2007.12.004

seven different groups representing three corporations, graduates of two business schools, and affiliates of two software firms. In this paper, we discuss results drawn from this survey, with special attention to two distinct subgroups that arguably represent opposite ends of the spectrum of spreadsheet sophistication. Our main purpose is to compare the results from these two subgroups in order to identify best practices. Although prescriptions for best practice can be found in a few publications, those are not based on systematic empirical research. By basing our conclusions on questionnaire data from the most comprehensive survey of its kind, we hope to legitimize our recommendations for best practice with new evidence from the real world.

In the following section, we review the results of other surveys of spreadsheet users that have appeared in the literature. In Section 3, we describe the SERP questionnaire and the populations to whom the survey was administered. Section 4 explains how the two subgroups were defined, and Section 5 focuses on the differences in their spreadsheet practices. Finally, in Section 6, we discuss the implications of the results and draw conclusions for improving spreadsheet design and use.

2. Literature review

Spreadsheets have been around for over 25 years, but there have been few published surveys that provide a broad-based look at spreadsheet practices. Here are the important surveys that we found in the research literature:

- Cragg and King [6] investigated spreadsheet practices in 10 firms. They concluded that spreadsheets are normally created in an "informal and iterative manner", and that there is a need for more training and enforcement of design and use guidelines in organizations.
- Schultheis and Sumner [7] noted that while some controls were being applied in organizations they researched, spreadsheet developers tended to use more controls in high-risk spreadsheet applications.
- Floyd et al. [8] studied management policies in four large organizations and found that few formal policies were in place to govern spreadsheet model development and use.
- Chan and Storey [9] surveyed members of a Lotus mailing list in 1992, sending out 1000 questionnaires and receiving 256 returns from business analysts in various functional specialties. The respondents were distributed broadly over several industries. The survey described their training and the most frequent types of analyses they did, along with an indication of the frequency with which they used nine prominent spreadsheet features. The main part of the Chan–Storey article describes a model for the (statistical) relationship among analytic tasks performed, spreadsheet proficiency, use of specific spreadsheet features, use of other software packages, and satisfaction with these software packages. The strongest relationship linked spreadsheet proficiency with the performance of specific tasks. In an expanded version of the model, spreadsheet proficiency and the importance of decisions made were found not to be significantly related.
- Hall and Johnstone [10] surveyed spreadsheet developers in Australia in late 1991. They sent out 268 questionnaires, received 106 returns, but only 82 of those completed the questions on controls (good practices), which were the focus of the study. The respondents answered questions about a specific spreadsheet project of their choice. One major finding was the low awareness of any kind of spreadsheet quality control among those surveyed.
- Pemberton and Robson [11] surveyed part-time students (who were working full time) at the University of Northumbria Business School. Of the 227 students surveyed, about 30 did not use spreadsheets, so the effective sample size was 197. The average age was 29. About half the sample (48%) used spreadsheets three or more times a week. The software used was Excel (94%), Lotus (5%), and QuattroPro (1%). The survey suggested that most spreadsheet use was unsophisticated, perhaps due to limited amounts of training.
- Caulkins et al. [12] surveyed 45 executives and senior managers/analysts about their experiences with spreadsheets. Not only did almost all respondents report that errors are common, but most attributed losses or bad decisions to such errors, even though it was unclear whether the ultimate consequences were severe. Many of these executives expressed an opinion that error checking and quality control procedures can be informal and will detect gross errors. Others thought that more formal quality control processes could be beneficial.

One conclusion from this literature review is that there have not been many recent surveys undertaken. A majority of the surveys listed were carried out before Excel became the dominant software. Moreover, the sample sizes in these surveys have not been large. At a size of 256, the Chan–Storey sample appears to be the largest. Our survey reflects a sample of spreadsheet

users virtually all of whom use Excel, and our sample size of 1597 is capable of providing a much broader picture than previous surveys.

A related literature deals with studies (not surveys) of spreadsheet use in practice. These articles deal with rather small samples, and they address narrowly focused research questions. They also contribute to our understanding of spreadsheet practice, but in very specialized ways:

- Conway and Ragsdale [13] considered the value of using structured rules to achieve goals of reliability, auditability, and modifiability. They concluded with several guidelines they believed to be helpful for creating spreadsheet modules especially for optimization problems.
- Edwards et al. [14] developed a set of guidelines for "do-it-yourself" spreadsheet creators and a set of best practices for verifying spreadsheets and improving logic and data management. Some of these practices are reflected in the SERP survey used as a basis for this research.
- Kreie et al. [15] studied 66 end users, contacted over the internet, to investigate the question of whether the quality of end-user computing applications could be improved by training end users in analysis and design methods. (Their answer: yes.)
- Lawrence and Lee [16] presented a report to the Financial Services Forum. They provided a framework for the analysis of project financing and presumed that the accompanying analysis could apply as well to spreadsheets. In the appendix, the report summarized the experience of Mercer Finance and Risk Consulting, profiling the 30 largest spreadsheet models they studied during the preceding year. Their statistical results provide a benchmark for some of our findings.
- McGill and Klobas [17] studied 159 end users to test hypotheses related to a multifaceted model of the relationship between the quality of designed spreadsheets and the extent of developer and user spreadsheet knowledge. Developer knowledge was found to be closely related to the perceived quality of the application, whereas user knowledge was found to be closely related to the impact of the application on decision making. Their work is also important because it advocates objective mechanisms to assess levels of spreadsheet knowledge.
- Croll [18] described interviews with about 20 auditors, accountants, bankers, insurers, analysts, and the like, showing how spreadsheets play a critical role in London's financial community. Croll concluded that the awareness and control of risk are uneven, with banking, professional services, and private finance being the most aggressive at dealing with the potential for spreadsheet errors. His findings provide a useful backdrop for the portion of our results that deal with risk.
- Grossman et al. [19] conducted field interviews to identify spreadsheets that were vital to the companies that use them. They identified five classes of such spreadsheets: application software, financial risk management tools, executive information systems, business process infrastructure, and complex analytical tools. In each category, they describe one or more spreadsheets in use. In general, they observe a misalignment between the importance of these spreadsheets and the resources devoted to creating and maintaining them.

Thus, field work in end-user computing has supplemented broad surveys with specialized portraits of spreadsheet use, within the bounds of narrow research questions posed in experiments and interviews. Such efforts complement surveys by exploring the dimensions of organizational behavior that influence the use of spreadsheets.

A third segment of current literature provides guidance on "recommended" spreadsheet practices. Three of these sources, due to Read and Batson [20], Raffensperger [21], and BPM [22], are good examples of detailed recommended practices. Several of their recommendations are reflected in the SERP questionnaire used in this research.

We studied 1597 responses from seven different sources to determine general spreadsheet practices and what level of experience and types of personal characteristics are associated with the use of those practices. We believe that this survey is the first of its kind in terms of scope and detail. In general, we observe significant differences in practices, and the evidence suggests that these differences can often be explained by such factors as training, experience, spreadsheet complexity, and the importance of spreadsheets in the organization. Most of our observations relate to user characteristics, but some also reflect the posture of their organizations toward good practices for spreadsheet design and use.

3. The SERP Survey

Our questionnaire was based on a seven-stage model for the evolution of a typical spreadsheet. The stages are designing, testing, documenting, using, modifying, sharing, and archiving. We developed this model based

on field visits to several firms and discussions with spreadsheet users at those firms. We used the seven stages as a mechanism for organizing the contents of our survey.

The questionnaire itself addressed each of the seven areas; in addition, it contained questions on training, quality control, and risk as they relate to spreadsheets. We circulated a draft of the questions among several leading researchers who have been writing about spreadsheet risk, and we incorporated their suggestions into the final questionnaire. By using a variety of editors, we exposed ourselves to possible criticism that our questions were sometimes inconsistent or that our measurements were vague, but we valued the input of the active research community. The content of our questionnaire was also informed by our reading of the substantial literature on spreadsheets. That literature is reflected in the SERP bibliography. Finally, we included questions that described the respondents themselves. In all, the questionnaire contained 67 items, some of which allowed for open-ended answers, and it took about 15–20 min to complete. The questionnaire is available on the project website (http://mba.tuck.dartmouth.edu/spreadsheet/index.html).

Seven populations were invited to fill out the questionnaire. Two populations (referred to later as II and III) consisted of people on mailing lists of (and contacted independently by) two software companies that specialize in spreadsheet-related software for optimization and simulation. Together, these two samples contained 568 respondents. We anticipated that these two samples would represent a relatively sophisticated and technical set of spreadsheet users because of their association with operations research applications. Two of the populations (VI and VII) were MBA alumni from two business schools, one in the United States and the other in Europe, together accounting for 846 responses. The other three populations (I, IV, and V) came from private corporations and accounted for 183 total responses.

The response rate in these surveys is difficult to determine, but in two of the private cases, the respondents represent over 50% of the number possible. These were companies where a concerted effort was made to have members of specific departments respond to the survey. In one of the MBA alumni surveys, the response rate is estimated at about 12%; the other is unknown but lower. For respondents on the mailing lists of the two software companies, the response rates are difficult to estimate, given the manner in which availability of the survey was announced and distributed.

We limit our discussion to those survey questions that shed light on the characteristics of the respondents, their work settings, and their spreadsheet practices. For a detailed summary of results, interested readers can visit our project's website.

Our main interest in this paper is the relation between respondent characteristics and spreadsheet practices. To achieve our purposes, we identified two subgroups of respondents: one composed of those who appeared to have the lowest level of capability (Group A), the other those who were the most advanced (Group B). We wanted to determine what differences in practices occur between the respondents in the two groups. We assumed that the most advanced users are more experienced, have greater expertise, work on larger spreadsheet models, and find spreadsheets to be more important to their jobs than their counterparts in the less advanced subgroup. We hypothesized that these users would be likely to use a range of good practices. Our main goal was to identify the practices associated with membership in the more advanced subgroup of users.

In order to select two subgroups for comparison, we used responses to three of the survey questions:

1. What level of importance do spreadsheets have in your job?
2. Please classify your experience with spreadsheets.
3. How large are the models you normally create?

Group A consisted of those who said:

(a) the level of importance spreadsheets have in their job is either "unimportant" or "moderately important",
(b) their experience with spreadsheets was "little or no experience", "some experience; still a beginner", or "extensive experience; some expertise", and
(c) the sizes of models they normally create are under 1000 cells.

Group B consisted of those who said:

(a) the level of importance spreadsheets have in their job is "critical",
(b) their experience with spreadsheets could be characterized as "very experienced; high expertise", and
(c) the sizes of models they normally create exceed 10,000 cells.

We would guess that spreadsheets are somewhat important to most of those who responded to the survey in the first place, so our contrasts probably do not involve people who are oblivious to spreadsheets. In fact, less than 1% of the respondents indicated that spreadsheets were unimportant, so we mainly contrasted users who considered spreadsheets moderately important with those who

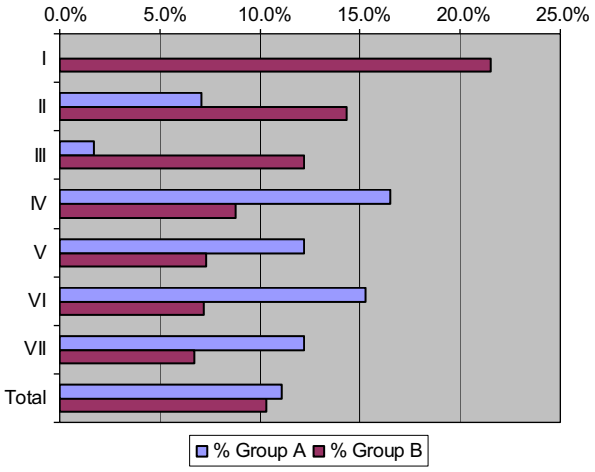

Fig. 1. Percentage of each of seven survey audiences represented in Groups A and B.

considered spreadsheets to be critical. Secondly, we did not attempt to measure the complexity of the spreadsheets used by respondents; rather, we took spreadsheet size to be a proxy for spreadsheet complexity. We might also wonder about the difference between the classification of "extensive experience–some expertise" and that of "very experienced–high expertise". However, some of our survey editors pointed out that there are many instances of spreadsheet users who use the same type of spreadsheet over and over. Someone who has a lot of experience with one spreadsheet is different from someone who has experience with many different spreadsheets. The wording in the two answers was a coarse attempt at capturing this distinction.

There is obviously some subjectivity in the classification, partly due to the fact that the respondents were allowed to self-classify. However, by using the intersection of the answers to three different questions, we hoped to overcome much of this subjectivity and ultimately identify two groups that would represent relative extremes within our sample. Of the 1597 total respondents, 175 (10.9%) were in Group A, 165 (10.3%) in Group B. "All" responses refer to those from the total of 1597 respondents.

A confirming measure of the expertise in the two groups is reflected in the responses to an open-ended survey question soliciting information on "practices particularly helpful to you or your organization in improving the quality of spreadsheets". Three times as many Group B respondents offered such practices (43.0–13.7%); and the practices covered over a dozen general categories from planning, standard formats, and modularization to version control, documentation, and testing.

Fig. 1 shows the distribution of Groups A and B among the seven participating survey audiences ordered by the percentage of respondents in Group B. The number of respondents varied considerably among these audiences as did the populations from which these sample respondents came. A mixture of countries was represented among the respondents in almost all of the seven audiences.

We believe that the two groups, A and B, represent opposite ends of a spectrum regarding experience, importance, and complexity. The substantial differences we discuss in the following section represent a partial validation of that belief. To the degree that we were successful in identifying contrasting groups, the differences (or similarities) in spreadsheet practices between these two groups can be enlightening. Identifying these differences is the principal intent of this paper.

We make no claim that the respondents to the survey (including Groups A and B) represent "typical" spreadsheet designers and users. Most come from the business world, have advanced degrees, and can point to considerable work experience. We can assume that the differences between the typical spreadsheet user and our survey respondents are even greater than the differences portrayed between Groups A and B.

4. Characteristics of Groups A and B

In order to contrast the members of the groups, we selected several questions in addition to the three used initially to define the two groups.

Table 1
Time normally spent on spreadsheets per week, by group

	Group A (%)	Group B (%)	All (%)
0–25%	92.6	11.5	44.7
26–50%	6.9	24.2	30.4
51–75%	0.6	37.6	17.8
76–100%	0.0	26.7	7.2

Table 2
Number of spreadsheets normally used per week, by group

	Group A (%)	Group B (%)	All (%)
0–1	21.1	0.0	5.8
2–5	64.0	21.8	40.2
6–10	13.1	16.4	25.6
More than 10	1.7	61.8	28.3

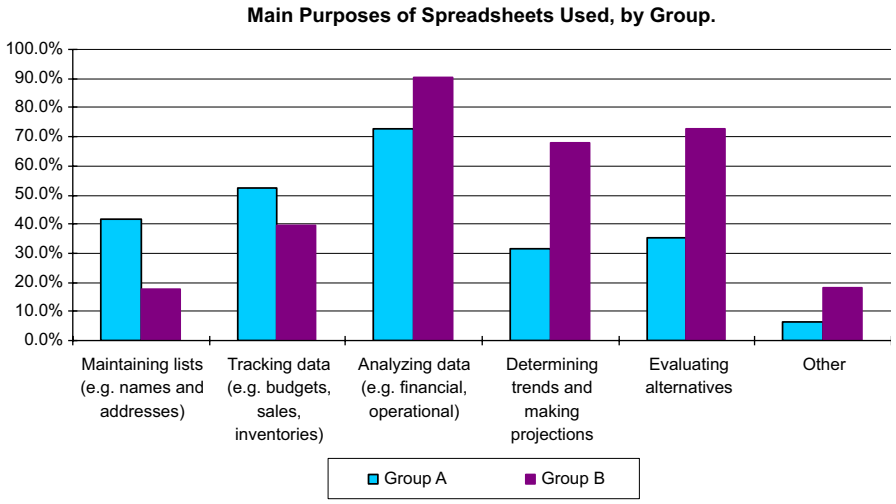

Fig. 2. Main purposes of spreadsheets used, by group.

Table 3
How the spreadsheets you create are used by others, by group

	Group A (%)	Group B (%)	All (%)
My spreadsheets are for my personal use	25.4	4.2	11.5
My spreadsheets are shared with 1 to 2 others	53.2	17.6	42.0
My spreadsheets are used by more than 2 others	19.1	30.9	30.9
My spreadsheets often become permanent assets	2.3	47.3	15.7

4.1. Time spent with spreadsheets

As Table 1 illustrates, nearly 93% of Group A spends *less* than 25% of their time on spreadsheets, while over 88% of Group B spends *more* than 25% of their time on spreadsheets (and over 25% devotes more than 75% of their time). Hence, there is a wide gap between the two groups in the commitment of time and the experience gained from that commitment.

4.2. Number of spreadsheets used in a week

Similarly, Group B respondents use many more spreadsheets in a normal week. While 85.1% of Group A uses five or fewer spreadsheets, 78.2% of Group B uses more than five (and 61.8% uses more than 10 spreadsheets). Table 2 confirms the intensity of spreadsheet use in Group B as compared to Group A.

4.3. Main purposes of spreadsheets

Fig. 2 portrays the differences between these two groups in the uses of their spreadsheets. Analyzing data is the major purpose for both groups, but Group B respondents indicate that analyzing data, determining trends, and evaluating alternatives are main purposes much more frequently than those in Group A.

4.4. Users of spreadsheets

Table 3 presents some data on how spreadsheets are used. Importantly, responses to how spreadsheets are used indicate that 47.4% of Group B (but only 2.3% of Group A respondents) creates spreadsheets that often become "permanent assets" for their organization.

4.5. Risks and risk-avoidance strategies

There are reasons why Group B respondents are as committed to spreadsheet design and use as the survey results suggest. As illustrated in Table 4, a total of 74.0% report that there is a medium to high risk posed by spreadsheets in their organizations compared to only 30.4% for Group A. In both groups, however, awareness of risk is less than we might expect (see Table 5). Although 84.9% of Group B respondents report "some" or

Table 4
Percentage of respondents perceiving levels of risk, by group

	Group A (%)	Group B (%)	All (%)
High risk	5.3	31.6	16.6
Medium risk	25.1	42.4	38.3
Low or no risk	69.6	25.9	45.1

Table 5
Percentage of respondents whose organizations are aware of spreadsheet risk, by group

	Group A (%)	Group B (%)	All (%)
Full awareness	12.9	29.6	19.5
Some awareness	42.7	55.3	54.2
No awareness	44.4	15.1	26.3

Table 6
Distribution of respondents, by function, by group

Function	Group A (%)	Group B (%)	All (%)
Sales and distribution	8.4	2.5	4.4
Marketing	18.6	3.8	10.9
Operations/manufacturing	10.2	9.6	9.5
Engineering and research	13.2	18.6	19.8
Finance	10.8	46.8	30.2
Human resources	4.2	1.3	1.3
Other	34.7	17.3	23.9

"full" risk awareness, only 29.6% indicate "full" awareness. For Group A the comparable data are 55.6% and 12.9%.

4.6. Functional area of job

The functional area within organizations is likely to influence the practices used by respondents. For example, Table 6 illustrates that over 46% of Group B respondents indicate that finance is their functional area, while only 10.8% fit into that category in Group A. The type of spreadsheet functions used, the commitment to spreadsheet quality, awareness of risk, and other factors could be largely influenced by these relative proportions.

4.7. Demographic characteristics

Finally, in characterizing the two groups, we used two demographic parameters: gender and age (see Table 7). Group A tends to be more heavily populated by females and slightly older than those in Group B. While 56.9% of Group A is over age 40, 56.1% of Group B is of age 40 or younger.

Table 7
Gender and age characteristics of Groups A and B

	Group A (%)	Group B (%)	All (%)
Gender			
Male	70.1	91.4	83.3
Female	29.9	8.6	16.7
Age			
20–30	15.5	16.0	13.7
31–40	27.6	40.1	38.5
41–50	26.4	30.9	26.2
51–60	20.7	9.3	14.7
Over 60	9.8	3.7	6.9

4.8. Summary of characteristics

Groups A and B differ considerably in several characteristics.[1] Group B individuals spend more time working on a larger number of spreadsheets, work more on spreadsheets designed to serve analytical and evaluative functions, and tend to collaborate with more people. This group is also more likely to have some formal training on spreadsheet creation and use. Moreover, the individuals in Group B tend to be younger and more likely to work in financial functions, where there may be significant concern with skill, accuracy, and advanced practices. The following section considers various types of practices used by individuals in these two groups.

5. Spreadsheet practices

The next focus of analysis is to probe differences between these two groups regarding their spreadsheet design and use practices. Several questions included in the survey relate to such practices as how spreadsheets are created, tested, documented, and shared.

5.1. Spreadsheet design

Several survey questions addressed how spreadsheets are created by respondents. The responses to three of these questions are presented in Table 8. These concern how often spreadsheets are created from scratch, how

[1] Using a standard one-tailed statistical test of hypothesis for differences in proportions, percentage differences between Groups A and B are significant whenever those differences exceed 7%. The differences we observed are generally much larger than this threshold.

Table 8
Selected spreadsheet design practices, by group

	Group A (%)	Group B (%)	All (%)
Do you create spreadsheets from scratch?			
Always	37.1	48.8	36.3
Sometimes	60.6	49.4	62.1
Never	2.3	1.8	1.5
Do you divide your spreadsheets into integrated modules?			
Always	2.3	51.8	20.4
Usually	32.4	35.4	42.6
Sometimes	51.4	12.2	32.7
Never	13.9	0.6	4.2
Do you separate data inputs from formulas in spreadsheet?			
Always	8.1	43.6	22.3
Usually	32.4	40.0	41.4
Sometimes	48.0	14.5	31.1
Never	11.6	1.8	5.2

Table 9
First step in creating spreadsheets, by group

	Group A (%)	Group B (%)	All (%)
Enter the data and formulas directly into a computer	54.9	37.5	48.7
Borrow a design from another spreadsheet	23.4	25.0	22.8
Sketch the spreadsheet on paper	14.9	20.6	17.4
Write the fundamental relationships using algebra	2.9	8.8	5.8
Other	4.0	8.1	5.3

often spreadsheet models are divided into separate, integrated modules, and how frequently data inputs are separated from formulas. The responses to these questions demonstrate that Group B individuals more often create spreadsheets from scratch and are more likely to use good design practices as well.

Table 9 presents the typical first step in creating a spreadsheet. The first step does not vary significantly among the three groups, although there is less likelihood that Group B individuals would start by directly entering data into the computer. Even for this group, however, it is more often the practice than any of the other options offered.

5.2. Use of software features

One of the distinguishing factors between Groups A and B is the use of Excel features (e.g., functions and tools). A more extensive working knowledge of a variety of features available through Excel can enhance the sophistication and creativity of the designer and user. Those who are involved in larger, more complex and critical spreadsheets require a larger toolkit to create their models and fulfill their models' requirements.

The survey sought information on the relative frequency with which respondents use each of 14 Excel-related features. Respondents were asked to indicate the frequency of use in terms of the following options: rare use, infrequent use, occasional use, frequent use and daily use. Assigning weights of one to five, respectively, for each of these options, we created a weighted average frequency for each feature.

Fig. 3 presents the result of this analysis, ordered from most to least frequent use by all respondents. All of these Excel features are used by Group B more frequently than they are by Group A. All but two of the features (data table and goal seek tools) are used by Group B individuals, on average, at least at the "occasional use" level, represented by 3.0 on the scale in Fig. 3. The data sort tool is the only feature used by Group A respondents, on average, at that level. Overall, the relative use of these special features provides a clear measure of the difference in the practices of the two groups.

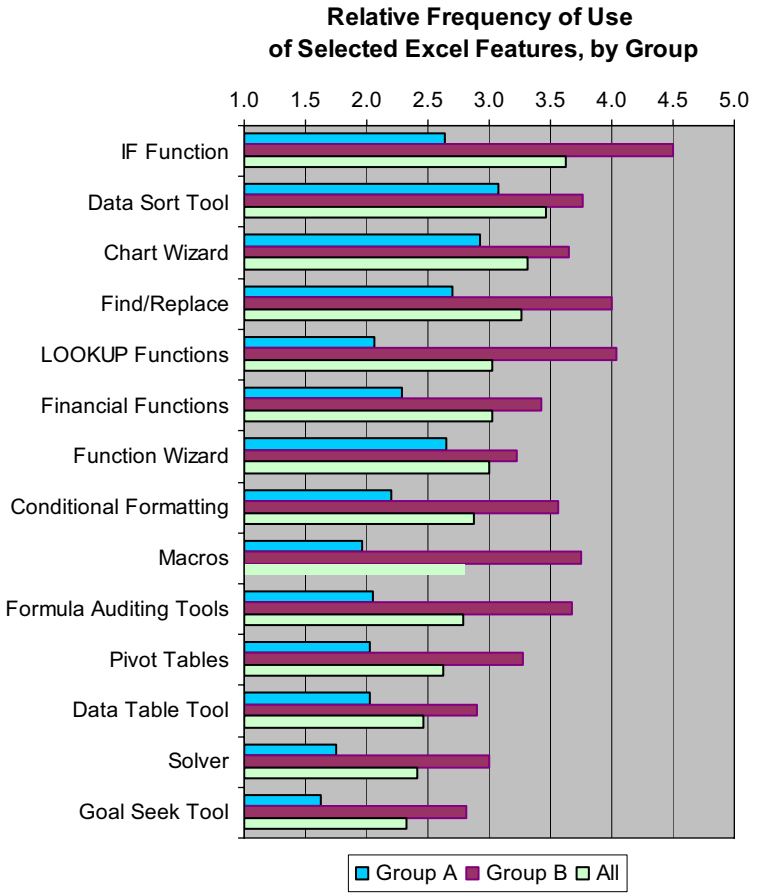

Fig. 3. Relative frequency of use of selected Excel features, by group. *Note*: 1.0, "rarely used"; 2.0, "infrequent use; 3.0, "occasional use"; 4.0, "frequent use"; 5.0, daily use.

Table 10
Frequency of testing models created or used, by group

	Group A (%)	Group B (%)	All (%)
Always	8.0	53.9	24.2
Usually	18.3	25.5	26.7
Sometimes	33.1	18.8	31.9
Never	40.6	1.8	17.1

5.3. Evaluating spreadsheets

Table 10 presents considerable differences in the manner by which Groups A and B test their spreadsheets. Over 50% of Group B respondents always test their models compared to only 8% of Group A respondents. This difference undoubtedly reflects the relative size, complexity, and importance of Group B's models as well as the experience of the spreadsheet creator.

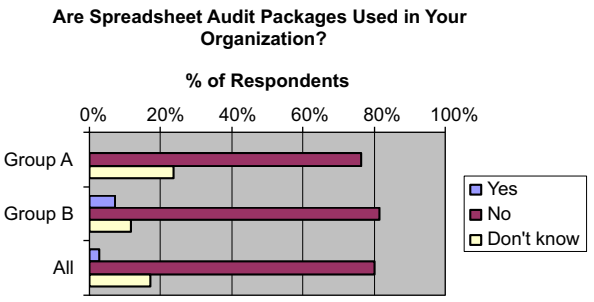

Fig. 4. Use of spreadsheet audit packages in organizations, by group.

A second way to test spreadsheets involves the use of commercial auditing software, which is increasingly available in the marketplace. Somewhat enlightening is the fact that no one in Group A is in an organization that utilizes audit software, compared to 7.1% of those in Group B (see Fig. 4). A third way to test models, and

Table 11
Types of model evaluation used by respondents, by group

	Group A (%)	Group B (%)	All (%)
Use common sense	45.7	80.6	67.4
Test extreme case	23.4	67.9	45.9
Examine formulas individually	34.3	65.5	45.6
Test performance for plausibility	24.6	64.2	43.4
Formula auditing toolbar	9.7	51.5	28.0
Use a calculator to check selected cells	29.7	46.7	38.4
Display all formulas	18.3	21.8	18.2
Use go to—special	0.6	17.0	6.3
Error checking option	4.6	16.4	10.2
Other tools	4.0	18.2	7.6

Note: Respondents could indicate use of more than one type of evaluation.

Table 12
Number of days of training offered to you each year

	Group A (%)	Group B (%)	All (%)
None	59.2	52.3	52.1
1 or 2 days	20.4	22.1	25.8
3–5 days	11.2	10.7	11.2
More than 5 days	9.2	14.8	11.0

the more traditional approach, is to use a range of techniques presented in Table 11. Group B takes advantage of *all* these techniques more frequently than those in Group A. Moreover, the average Group B respondent also uses four of these approaches while the average Group A respondent uses only two. Again, this finding reflects the size, complexity, and relative importance of spreadsheet models created and used by Group B respondents.

5.4. Training

Although there were some differences between the two groups, training was in general a soft spot for both groups and for all respondents and their organizations. Training programs are an exception rather than the rule, with no more than a few days of training each year generally offered in most organizations, as illustrated in Table 12. Training days are offered to employees in less than 50% of the organizations for both groups. The most often repeated reason for lack of training was "lack of time".

Table 13
Types of training reported by respondents, by group

	Group A (%)	Group B (%)	All (%)
Books and manuals	44.6	73.3	53.6
Demonstrations from colleagues	52.0	58.2	52.3
Formal classroom instruction	41.7	40.6	37.7
Occasional informal training sessions	29.1	34.5	29.2
None	21.1	12.7	17.6

Table 14
Types of training in spreadsheets made available by your organization

	Group A (%)	Group B (%)	All (%)
None	56.0	37.6	41.3
In-house training	24.6	40.0	38.6
Training by external party	17.1	26.7	20.3
One basic session is available	4.0	4.2	4.3
Several sessions, incl. advanced topics	6.9	18.8	14.2
Specialist to assists designers/users	3.4	7.9	5.1
Other	3.4	6.7	5.1

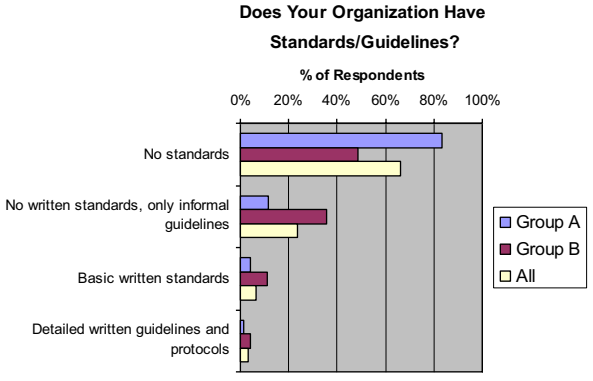

Fig. 5. Spreadsheet quality control standards and guidelines, by group.

When asked what types of training respondents have had in their careers, there were some subtle differences between Groups A and B (as well as for all respondents). As Table 13 illustrates, Group B individuals have received more "occasional informal training", more "demonstrations from colleagues," and substantially more training from "books and manuals" than Group A. This informal and self-taught learning system seems to characterize the spreadsheet training of Group B. This pattern also shows up when we compare the types of training offered by their *current* organization (see Table 14), where all types are more frequently

Table 15
Summary of spreadsheet user characteristics and practices, by group

	Group A—least experienced	Group B—most experienced
Characteristics		
Time spent on spreadsheets	> 92% spend less than 25% of their time	> 88% spend more than 25% of their time
Spreadsheets used per week	> 85% use 5 or fewer	> 78% use more than 5 (> 61% use more than 10)
Spreadsheet users who determine trends, make projections or evaluate alternatives (%)	< 36	> 65
Users who create spreadsheets that become "permanent assets" for their organizations (%)	< 3	> 47
Users whose spreadsheets present medium or high risks to their organization (%)	< 31	74
Users in financial function in organization (%)	< 11	> 46
Received training through books and manuals (in addition to other methods) (%)	< 45	> 73
Age of user	57% are older than 40 years	56% are 40 years old or younger
Male/female (%)	70/30	91/9
Practices		
Always divide spreadsheets into integrated modules (%)	< 3	> 51
Always separate data inputs from formulas in spreadsheets (%)	< 9	> 43
Enter data and formulas directly into computer as a first step in creating a spreadsheet (%)	> 54	< 38
First sketch spreadsheet on paper or write fundamental relationships in algebra (%)	< 18	> 29
Excel features used more frequently than occasionally (of 14 select features)	1 Excel feature	8 Excel features
Always test their spreadsheet models (%)	8	> 53
Average number of model evaluation approaches used (of 9 select approaches)	< 2.0	> 4.0

offered to Group B respondents. Moreover, the topics covered in this training are generally more advanced in Group B's organizations.

5.5. Organizational quality control

With the prevalence of errors in spreadsheets, it is disappointing to note that most organizations represented in our survey do not have standards related to spreadsheet quality. As Fig. 5 portrays, even for those respondents in Group B, 80% or more of their organizations have no written standards and, at most, only informal guidelines (35%).

With this result in mind, it is not surprising to discover that audit packages that can help spreadsheet creators check their designs are used by only a small percentage of Group B respondents and not at all by any of the Group A respondents' organizations (although there were some respondents who "do not know").

6. Conclusions and implications of the research

The results of this analysis underscore the fact that spreadsheet practices vary substantially from person to person and organization to organization. Some of these differences relate to the relative importance of the spreadsheet and the level of expertise of the user as well as the size and complexity of the spreadsheet itself. Other differences may reflect the context within which the spreadsheet creator works. As summarized in Table 15, this paper has shown that there are measurable differences in the practices of respondents from the two groups studied here. These include some differences among corporate cultures as reflected in policies and guidelines and training offered, the type of training undertaken, the work styles, specific design and creation practices, the types of tools used, and the methods used (and frequency of that use) to test spreadsheets.

It appears that while some organizations in our survey provide advanced training, this is true of fewer than half. Furthermore, organizations most often leave it to the spreadsheet designer to pursue appropriate training, to determine and employ good design practices, and to test and evaluate their spreadsheets. The question of what constitutes *best* corporate or individual spreadsheet practice may be a function of the spreadsheet's use, its size and complexity, the degree of sharing, and its importance. There may be a few hard and fast rules or practices, but the context of the use of the spreadsheet is significant.

We conclude that some practices are more often undertaken by the most advanced spreadsheet designers and users, and that these practices improve the quality of their spreadsheets. These practices help them meet the requirements necessitated by the critical nature of their spreadsheets. As with any more highly skilled artist or technician, the practices employed by Group B respondents are adopted for good reason: they work! Short of being a recipe for best spreadsheet practices, their significance is undeniable. These include being better trained, working more closely with colleagues or in a team, planning the design of spreadsheets before entering data into a computer, separating spreadsheets into integrated modules, separating data from formulas, utilizing version and document control, protecting the work, and employing more testing methods. While all of these are relevant for individuals, it seems that some institutional guidelines, if not standards, could help assure organizations that quality control is extended to spreadsheets in their decision-support systems. Some of our results and conclusions, particularly those related to corporate practices, are consistent with the results of prior research as summarized earlier in our literature review. Other results provide a detailed list of characteristics that we would associate with advanced users. Most importantly, for our purposes in this paper, we have empirically identified a number of "best practices" for the design and use of spreadsheets.

References

[1] Rasmussen R. QAP—not so hard in spreadsheets. Omega 2007;35:541–52.
[2] Ioannou G, Mavri M. Performance-net: a decision support system for reconfiguring a bank's branch network. Omega 2007;35:190–201.
[3] Rasmussen R. On time series data and optimal parameters. Omega 2004;32:111–20.
[4] Faizul H, Cutright K, Martin C. Employee scheduling and makespan minimization in a flow shop with multi-processor work stations: a case study. Omega 2004;32:121–9.
[5] Carter AE, Ragsdale C. Scheduling pre-printed newspaper advertising inserts using genetic algorithms. Omega 2002;30:415–21.
[6] Cragg P, King M. Spreadsheet modelling abuse: an opportunity for OR?. Journal of the Operational Research Society 1993;44:743–52.
[7] Schultheis R, Sumner M. The relationship of application risks to application controls: a study of micro-computer-based spreadsheet applications. Journal of End User Computing 1994;6:11–8.
[8] Floyd B, Walls J, Marr K. Managing spreadsheet model development. Journal of Systems Management 1995;46:38–43.
[9] Chan YE, Storey VC. The use of spreadsheets in organizations: determinants and consequences. Information and Management 1996;31:119–34.
[10] Hall M, Johnstone J. A risk and control-oriented study of the practices of spreadsheet application developers. In: Proceedings of the 29th annual Hawaii international conference on systems sciences, 1996.
[11] Pemberton JD, Robson AJ. Spreadsheets in business. Industrial Management and Data Systems 2000;100:379–88.
[12] Caulkins JP, Morrison EL, Weidemann, T. Spreadsheet errors and decision making: evidence from field interviews. Journal of Organizational and End User Computing 2007;19:1–23.
[13] Conway DG, Ragsdale CT. Modeling optimization problems in the unstructured world of spreadsheets. Omega 1997;25: 313–22.
[14] Edwards J, Finlay P, Wilson J. The role of OR specialists in do-it-yourself spreadsheet development. European Journal of Operational Research 2000;127:14–27.
[15] Kreie J, Cronan TP, Pendley J, Renwick JS. Applications development by end-users: can quality be improved. Decision Support Systems 2000;29:143–52.
[16] Lawrence RJ, Lee J. Financial modelling of project financing transactions. In: Institute of actuaries of Australia financial services forum, 2004.
[17] McGill T, Klobas JE. The role of spreadsheet knowledge in user-developed application success. Decision Support Systems 2004;39:355–69.
[18] Croll G. The importance and criticality of spreadsheets in the City of London. In: Proceedings of the EuSpRIG conference, London, 2005.
[19] Grossman TG, Mehrotra V, Özlük Ö. Spreadsheet information systems are essential to business. University of San Francisco working paper, 2005.
[20] Read N, Batson J. Spreadsheet modelling best practice. Institute of Chartered Accountants for England and Wales, 1999.
[21] Raffensperger J. New guidelines for spreadsheets. In: Proceedings of the EuSpRIG Conference, Amsterdam, 2001.
[22] BPM Analytical Empowerment. Best practice—spreadsheet modelling standards version 3.1, 2004.