Carnegie Mellon University
Research Showcase

Heinz Research | Heinz College

1-1-2008

Do Spreadsheet Errors Lead to Bad Decisions: Perspectives of Executives and Senior Managers

Jonathan P. Caulkins
Carnegie Mellon University

Erika Layne Morrison
IBM Global Services

Timothy Weidemann
Fairweather Consulting

Follow this and additional works at: http://repository.cmu.edu/heinzworks

Recommended Citation
Caulkins, Jonathan P.; Morrison, Erika Layne; and Weidemann, Timothy, "Do Spreadsheet Errors Lead to Bad Decisions: Perspectives of Executives and Senior Managers" (2008). *Heinz Research.* Paper 2.
http://repository.cmu.edu/heinzworks/2

This Article is brought to you for free and open access by the Heinz College at Research Showcase. It has been accepted for inclusion in Heinz Research by an authorized administrator of Research Showcase. For more information, please contact research-showcase@andrew.cmu.edu.

Do Spreadsheet Errors Lead to Bad Decisions: Perspectives of Executives and Senior Managers

Jonathan P. Caulkins, Carnegie Mellon University, Qatar Campus and Heinz School
Erica Layne Morrison, IBM Global Services
Timothy Weidemann, Fairweather Consulting

Abstract

Spreadsheets are commonly used and commonly flawed, but it is not clear how often spreadsheet errors lead to bad decisions. We interviewed 45 executives and senior managers / analysts in the private, public, and non-profit sectors about their experiences with spreadsheet quality control and with errors affecting decision making. Almost all said spreadsheet errors are common. Quality control was usually informal and applied to the analysis and/or decision, not just the spreadsheet per se. Most respondents could cite instances of errors directly leading to bad decisions, but opinions differ as to whether the consequences of spreadsheet errors are severe. Some thought any big errors would be so obvious as to be caught by even informal review. Others suggest that spreadsheets inform but do not make decisions, so errors do not necessarily lead one for one to bad decisions. Still, many respondents believed spreadsheet errors were a significant problem and that more formal spreadsheet quality control could be beneficial.

Keywords: spreadsheets, errors, error control, managerial decision making, decision support systems, end user computing

INTRODUCTION

Spreadsheets are used in diverse domains by decision makers at all levels (Gerson, Chien, & Raval, 1992; Chan and Storey, 1996; Seal, Przasnyski, & Leon, 2000; Croll, 2005); the entire July-August, 2008 issue of the journal *Interfaces* is devoted to spreadsheet modeling success stories. However, laboratory studies and field audits consistently find that a large proportion of spreadsheets contain errors. In a dozen studies reviewed by Kruck, Maher, & Barkhi (2003), the average proportion of spreadsheets with errors was 46%. Panko's (2000a, 2005) synthesis of spreadsheet audits published since 1995 suggested a rate of 94%. Powell, Baker, & Lawson (2007b, c) critique past work and greatly advance methods of defining and measuring spreadsheet errors but at the end of the day reach the same overall error rate of 94%.

Hence, one might expect that (1) spreadsheet errors frequently lead to poor decisions and (2) organizations would invest heavily in quality control procedures governing spreadsheet creation and use. We investigated both hypotheses through 45 semi-structured interviews with executives and senior managers / analysts in the public, nonprofit, and private sectors. Field interviews raise fewer concerns about external validity than do laboratory studies, and they focus on the overall decision making process, not just the spreadsheet artifact as in audit studies. However, our approach has two important limitations. First, the respondents are a convenience sample. Second, self-report can be flawed, whether through imperfect memories, self-serving bias, conscious deception, and/or limited self-awareness. Given these limitations, we focus on broad qualitative conclusions. In brief, we found that most respondents could describe instances in which spreadsheet errors contributed to poor decisions, some with substantial consequences, yet few reported that their organization employs quality control procedures specific to spreadsheet analysis.

The literature on spreadsheet errors in general is large (see Panko, 2000b and Powell, Baker, & Lawson, 2007a for reviews), but much less has been written on these specific questions. Regarding the frequency with which spreadsheet errors lead to bad decisions, the European Spreadsheet Research Interest Group (EUSPIG) maintains a webpage of news stories reporting the consequences of

spreadsheet errors (http://www.eusprig.org/stories.htm). However, spreadsheets are used by so many organizations that even if only a small proportion were hurt badly by spreadsheet errors, there could still be scores of examples. We started with a population of individuals and organizations for which we had no a priori reason to think spreadsheet errors were a particular problem. This approach has been taken by others (e.g., Cragg and King, 1993) to explore the prevalence of defective spreadsheets, but like Powell, Baker, & Lawson (2007d), we shift the focus to assessing the impact of those spreadsheet errors.

A considerable corpus on controlling spreadsheet errors concerns what organizations *should* do. Classic recommendations lean toward application of good software engineering principles (Mather, 1999; Janvrin and Morrison, 2000; Rajalingham, Chadwick, & Knight, 2000; Grossman and Özlük, 2004) or formal theories (Isakowitz, Schocken, & Lucas, 1995). Kruck and Sheetz (2001) combed practitioner literature for practical axioms validated by empirical results, supporting aspects of the spreadsheet lifecycle theory (e.g., include planning / design and testing/debugging stages) and recommendations to decrease formula complexity.

There is also some literature describing what organizations actually do. Notably, Finlay and Wilson (2000) surveyed 10 academics and 10 practitioners on the factors influencing spreadsheet validation. Those most commonly mentioned were (a) aspects of the decision and (b) aspects of the spreadsheet underlying the decision context. However, Grossman (2002) argues that it would be valuable to have greater knowledge of what error control methods are currently used.

We seek to fill that gap. The next section describes data and methods. The third discusses results pertaining to types of errors, error control procedures and policies, and reported effects on decisions. The paper closes by discussing the decision processes within which spreadsheets were embedded and implications for practice and future research.

DATA AND METHODS

Sampling

Data collection methods were similar to Nardi and Miller (1991). Interview subjects were identified primarily by referral through personal and institutional contacts. Only one person approached through these contacts declined to be interviewed. In contrast, earlier attempts at cold-calling often led to refusals or perfunctory interviews. Even when given every assurance of anonymity, respondents seemed more wary of admitting bad decisions to a complete stranger than to someone referred by a mutual acquaintance.

Fifty-five people were interviewed, but ten were excluded from the analysis below. Seven were excluded because they were duplicate interviews within the same work group; we retained only the respondent with the most sophisticated perspectives concerning spreadsheet use within that work group. Two were eliminated because they had worked for more than one organization, and it became ambiguous which of their comments concerned which organizations. One was excluded because s/he used spreadsheets only for list-tracking and database functions.

Sample Characteristics

All but one interview was done in person, so most interviewees (73%) represented organizations located in one region. The one phone interview stemmed from meeting in person with the CEO of a manufacturing firm who suggested that his CFO, in another state, had more insights into the topic.

We interviewed subjects from three sectors (for-profit, non-profit, and government) and two organizational levels (executives vs. senior managers/analysts). Since we found few pronounced differences across groups, results are primarily reported in aggregate. Overall, 96% of the executives were male, as were 45% of the manager/analysts. All executives and 90% of manager/analysts were non-Hispanic White. The executives came from small to medium size organizations (ranging from several dozen to several thousand employees). Private sector managers' organizations, in contrast,

spanned the full range of sizes, up to organizations with tens of thousands of employees. Some respondents in the non-profit sector worked for large health care or educational institutions, but most came from smaller organizations. Educational attainment ranged from several PhDs to a local government manager with a high school degree. No respondent remembered any recent formal training in Excel. Most reported learning spreadsheet techniques informally, in the office environment or with the help of practical manuals.

Table 1: Description of Sample

		Gender			Location	
Group	**N**	**Women**	**Men**		**Local**	**Other**
Public Executives	6	0%	100%		100%	0%
NonProfit Executives	13	8%	92%		85%	15%
Private Executives	6	0%	100%		50%	50%
Public Managers	7	57%	43%		57%	43%
NonProfit Managers	6	83%	17%		100%	0%
Private Managers	7	29%	71%		43%	57%
Total	45	27%	73%		73%	27%

We classified interviewees by their highest reported spreadsheet use as basic calculation (n=6), simple modeling (n=20), or advanced modeling (n=19). This distinction was based on the highest mathematical complexity of formulas and models, as well as spreadsheet size (in number of cells and/or file size), links, functionality and features, such as optimization with Solver. Respondents reported using an average of 2.7 other types of software to support decision making, with database and accounting software mentioned most frequently.

Interview Protocol

The interview protocol was finalized after conducting ten exploratory interviews with subjects who are not part of this analysis. Those interviews revealed that open-ended questions elicited richer responses than did tightly-scripted or multiple choice questions. Indeed, senior executives resisted

highly structured questions. Furthermore, the most effective sequence in which to cover topics varied from interview to interview. Hence, we adopted a semi-structured interview protocol to ensure that a specific set of topics with associated probes were covered in every interview, while allowing the conversation to move nonlinearly through the topics.

Interviewees were sent a description of the research project and the interview protocol. The protocol addressed individual and organizational experience with spreadsheets, spreadsheet errors, error control, and effects on decision making. (See Appendix A.) Variables were coded by the primary interviewer based on audio recordings and detailed notes into categorical variables representing groupings of common responses. To test inter-rater reliability, a subset of the interviews were coded by both primary interviewers. Seventy-six percent of those items were coded consistently, and most discrepancies were instances in which one interviewer drew a conclusion and the other thought there was insufficient information to decide, not from the interviewers drawing contradictory conclusions. The most discordant variables pertained to advantages and disadvantages of using spreadsheets for decision making and were excluded from the analysis.

RESULTS

Types of Errors Reported

All but one respondent reported encountering errors in spreadsheets. The most commonly mentioned types were *inaccurate data* (76%), *errors inherited from reuse* of spreadsheets (49%), *model errors* (33%), and *errors in the use of functions* (also 33%). We recorded the presence or absence of a type of error, but have no information on the frequency of occurrence.

These error types emerged from respondents' statements, not a literature based classification as in Rajalingham et al. (2000), Purser and Chadwick (2006), Panko (2007), or Powell et al. (2007c) because we merely asked respondents for their opinion of the "cause" or "source" of the error, without more specific prompts. Powell et al. (2007d) found that many errors in operational spreadsheets have little or no quantitative impact. For instance, an erroneous formula could generate the correct answer

if the error pertained to a condition that did not occur in the given data. Although we did not ask respondents explicitly to distinguish between errors that did and did not have an impact, since the context of the interviews was effects on decision making, we believe most respondents discussed only errors that did have an impact or would have had an impact if they had not been detected and corrected.

Table 2: Commonly mentioned errors, by sophistication of highest spreadsheet use and by sector

Error Type	All	By Sophistication of SS Use			By Sector		
		Advanced	Simple	Basic	Public	Nonprofit	Private
Inaccurate data	76%	74%	80%	67%	77%	95%	46%
Errors inherited from reusing spreadsheets	49%	63%	40%	33%	31%	63%	46%
Model error	33%	42%	25%	33%	23%	16%	69%
Error in use of functions	33%	21%	45%	33%	46%	26%	31%
Misinterpretation of output/report	27%	47%	15%	0%	15%	32%	31%
Link broken/failed to update	22%	37%	15%	0%	8%	26%	31%
Copy/Paste	22%	21%	30%	0%	23%	26%	15%
Other	11%	11%	15%	0%	8%	5%	23%
Lost file/saved over file	7%	5%	10%	0%	0%	5%	15%
No errors	2%	5%	0%	0%	8%	0%	0%
N	45	19	20	6	13	19	13

Inaccurate Data

We expected that inaccurate data would come primarily from miskeyed data and other "mechanical errors" to use Panko's (1999, 2007) term. Such "typos" were frequently mentioned, but there were other sources of data value errors.

One was bad data piped into a spreadsheet automatically, for example, from a database query or web-based reporting service. A health insurance firm reported that a change in prescription pill units triggered automatic reimbursements that were too high because of a units inconsistency between the revised database and the spreadsheet. Such systems integration problems illustrate an issue raised

by multiple respondents and the literature (Panko, 2005). Figures in spreadsheets can sometimes be attributed with an aura of inerrancy, lulling users into not reviewing them as critically as they might have in another context.

Respondents mentioned human bias as another source of inaccurate data, akin to Panko's (2007) "blameful acts", varying in culpability from "wishful thinking" to outright fraud. Wishful thinking included choosing "base case" parameter values that "made sense" to the analyst because they gave the (perhaps unconsciously) preferred conclusion. The more extreme version was willful and self-serving manipulation of input values. Bias can contaminate any analysis, but bias buried in a spreadsheet can be hard to detect.

Different types of quality control are required for these different sources of inaccurate data. Asking analysts to check parameter values a second time might help catch typos, but it would do little to address a fraudulent self-serving bias.

Errors Inherited from Reusing Spreadsheets

Almost all respondents said they reused their spreadsheets, and almost half described errors from reuse of their own or colleagues' spreadsheets. One respondent inherited a model containing a vestigial 'assumptions page' that did not link to the model. Several described spreadsheet errors that endured for an extended period of time and noted that small errors replicated many times led to substantial losses over time. One nonprofit had relied on a faulty ROI worksheet for several years, affecting contracts worth ~$10 million. Reports of reuse errors increased with the complexity of spreadsheet applications.

Opinions differed on the value of reuse. For many respondents, the ability to reuse templates was a key advantage of spreadsheets. Some echoed the observation by Nardi and Miller (1991) that templates enable sharing of domain expertise as well as spreadsheet skills. However, a sizable minority felt reused spreadsheets were difficult to control, since updating worksheets can introduce more errors.

Errors in the Use of Functions

Thirty-three percent of respondents described errors in functions, ranging from inappropriate use of built-in functions to mistaken operators and cell addressing problems. Careful inspection of formulas was reported to be rare unless motivated by an observed discrepancy. In the words of one respondent, "Formulas are only examined in depth if there's a reason." The most frequently mentioned explanation was the difficulty of review. One senior analyst observed: "Spreadsheets are not easy to debug or audit…It's a very tedious process to check someone else's cells, especially two or three levels [of cell references] down."

Model Error

One-third of respondents reported model errors, including errors in assumptions, overall structure, errors of omission, and other major distortions of the modeled situation (as identified by the respondent). Model errors are not programming mistakes, but rather erroneous judgments about how to model real world situations, so they cannot be detected solely by reviewing the spreadsheet (Grossman, 2003). The same can be said for the fifth most commonly cited type of error, namely misinterpretation of spreadsheet output, since a correct spreadsheet may be misinterpreted by a person who has flawed understanding of the spreadsheet model's assumptions or limitations.

Spreadsheet Quality Control Methods

Most respondents thought about quality control in terms of detecting errors ("inspecting in quality") rather than in terms of preventing them ("building in quality") from the outset. The twelve quality control methods mentioned repeatedly by respondents can usefully be divided into three categories: (1) informal quality control, (2) organizational methods such as peer review, and (3) technical tools that are specific to spreadsheets. (See Table 3.)

The academic literature focuses on extending formal software quality control practices to the end user environment (e.g., Rajalingham, Chadwick, Knight, and Edwards, 2000), but the respondents did not report following common design and development recommendations (Teo and Tan, 1999, Clermont, Hanin, & Mittermier, 2002), a software lifecycle approach, or formal software tools (Morrison, Morrison, Melrose, & Wilson, 2002). Instead, the methods described might be characterized as stemming primarily from the application of general managerial acumen. They do not differ in obvious ways from quality control steps that would be employed to review any other form of analysis. What seems to distinguish the conscientious from the lackadaisical was not necessarily technical sophistication. Rather, it was the formality with which general purpose quality control procedures such as peer review were employed.

Table 3: Proportions reporting use of various quality control methods, of those for whom sufficient information was gathered to ascertain whether the method was used (Method Type 1 = Informal methods; Type 2 = Organizational methods; Type 3 = Technical tools)

				Spreadsheet Sophistication (Highest Level)		
Quality Control Method	Method Type	N	Avg of All Groups	Advanced Modeling	Simple Modeling	Basic Calculation
Gut check against the bottom line	I	45	96%	100%	90%	100%
Review by developer	2	44	86%	84%	85%	100%
Review by someone other than developer	2	44	73%	67%	70%	100%
Crossfooting	3	45	60%	63%	55%	67%
Review by multiple reviewers other than developer	2	44	45%	56%	45%	17%
Documentation	2	45	42%	53%	40%	17%
Keep it simple	1	43	33%	33%	30%	40%
Input controls	3	45	22%	37%	15%	0%
Prevent deletion of calculation cells (protection)	3	45	20%	21%	25%	0%
Other		45	16%	11%	15%	33%
Test cases	3	45	13%	26%	5%	0%
Separate page for audit/change tracking	3	45	13%	21%	10%	0%
Audit tools	3	45	7%	16%	0%	0%

Informal Quality Control: The 'Sniff Test' and 'Keeping It Simple'

The most frequently cited quality control procedure was the 'gut check' or 'sniff test,' a cursory examination of bottom line figures for reasonableness. Many respondents acknowledged the limitations of 'sniff tests'. One analyst reported finding a major modeling error worth several million dollars by "dumb, blind luck" the night of the presentation to the board. Pointing to a blank cell accidentally excised a significant portion of the analysis, but even that obvious error passed many sniff tests. However, a minority held the contrary view that elementary review would detect all errors of consequence before they misinformed a decision.

The other informal method mentioned was to 'keep it simple.' This included simplifying the analysis as well as the spreadsheet itself. Advanced developers reported using advanced quality control methods in addition to, not instead of, these informal methods, although it is possible that what advanced developers call a gut check or keeping it simple might be richer than what less experienced mean when they use the same term.

Formal Quality Control Methods that Are Organizational in Nature

Respondents mentioned two methods that were organizational in nature: review and documentation. Most (73%) respondents reported that spreadsheets were reviewed by someone besides the developer. For the forty-five percent using multiple outside reviewers, this ranged from two colleagues in the office to review by several executives in teleconference. One organization mentioned inadvertent outside review. A public manager confessed that their annual budget, which invariably contained errors, would be scrutinized by unions and other groups for their own interests, helping to correct the document. Although review seems common, one-quarter of respondents did not mention any kind of review by others. Furthermore, almost no respondents reported spending even the minimum time on validation that is suggested by Olphert and Wilson (2004). Documentation concerns focused on structural assumptions and parameter values, rather than formulas and was reported more frequently than in Baker, Powell, Lawson, & Foster-Johns (2006a).

Technical Tools

We define technical tools as quality control methods that are specific to spreadsheets, such as protecting portions of the spreadsheet as a form of input control. Advanced modelers were the most likely to report using these technical tools, particularly those beyond crossfooting (redundant calculations). Test cases and separate pages for audit/change tracking were rarely cited. As in Baker et al. (2006b), least frequently mentioned were automatic audit tools, such as add-ins or Excel's own (limited) built-in audit feature, used by just three advanced modelers.

Other Methods

Several other error control methods were mentioned by only one respondent, but are still interesting since they represent "outside the box" thinking relative to methods typically discussed in the literature. Two were personnel-related: firing people who produced flawed spreadsheets and, on the positive side, hiring talented people in the first place, where talent referred to general analytical competence not spreadsheet skills. These actions might be seen as consistent with Panko's (2007) point that spreadsheets are not the cause of spreadsheet errors, and are usefully seen as a special case of human error more generally. Another method cited was to avoid using spreadsheets altogether, for example, by converting the organization to some enterprise resource planning (ERP) system or by encouraging use of other analytical software.

Spreadsheet Quality Control Policies

Consistent with Cragg and King (1993) and Baker et al. (2006a, 2006b), few respondents reported that their organizations had formal policies intended to ensure spreadsheet quality. Financial services sector organizations were the exception. They often reported using standardized spreadsheet models, created by IT personnel at corporate offices and distributed to branches.

Even without formal policies, several workgroups described high quality environments where review and documentation were routine. While there has been little success identifying correlates to

an individual's error rate (Howe and Simkin, 2006), we felt it was likely that some backgrounds facilitate more awareness within organizations. Of our respondents who stressed quality control, one was a computer scientist familiar with the risks of programming errors and two others were trained in accounting controls and taught spreadsheet modeling to graduate students.

For most of our respondents, the absences of policies were not the result of any thoughtful decision balancing the benefits of improved quality against the overhead associated with rigorous quality control (cf., Grossman, 2002). Indeed, many responded by saying, in essence, "Never thought about it, but it sounds like a great idea."

Others could cite reasons why a more structured approach might not work, including lack of time. One government analyst reported that she often had to create or modify spreadsheets in as little as half an hour. Sometimes errors were later discovered in those spreadsheets, but timeliness was paramount.

Several managers perceived that checking spreadsheets more carefully would require hiring another employee, and one public manager in particular had resisted an auditor's recommendations to hire additional staff to check spreadsheet accuracy, citing budget constraints. These views may help explain the absence of formal policies even if they are inconsistent with the conventional wisdom in software development that time invested in quality control is more than recouped in reduced rework.

The literature suggests that formal policies often encounter resistance because a principal advantages of spreadsheets is empowering end users to complete analysis independently (Cragg and King, 1993 and Kruck and Sheetz, 2001). Some respondents expressed a related yet distinct concern. They worked in small, non-corporate environments and suggested that guidance must be informal and/or implicit in these close-knit workplaces. Their arguments centered around unintended consequences of formality on workplace culture, such as signaling lack of trust in staff competence, rather than effects on spreadsheet productivity per se.

Most organizations likewise had no formal policies governing how to respond when errors were detected, and respondents described a range of responses. A common response was to fix the

specific error but do nothing else. Forty percent claimed to go further by investigating the associated processes to detect other, related errors. Other responses were less obviously effective. One individual threw out the computer on which a perpetually buggy spreadsheet was being run, in the belief that the hardware was somehow at fault.

Thirty percent of respondents mentioned sometimes rebuilding a spreadsheet from scratch when errors were detected. This could be eminently sensible if the original design was not well conceived. Often a cumbersome exploratory spreadsheet can be replaced by one whose design is better engineered and less error-prone. On the other hand, this may also reflect undue confidence that rebuilding the spreadsheet will not introduce new, more serious errors. Some respondents seemed to view errors as aberrations that can be eliminated if the work is redone carefully, rather than a predictable outcome, as in the software engineering perspective anticipating a certain number of errors per thousand lines of code (cf., Panko, 1999).

Other Factors Mediating the Application of Spreadsheet Quality Control

Distinct from the question of what methods are used is how *consistently* they are are applied. Our protocol did not address this directly, so we do not have data from all 45 respondents, but it came up spontaneously in many interviews.

One ideal espoused in the literature is *not* to apply all methods to all spreadsheets. Rather, the risk analysis philosophy suggests investing more heavily in quality control for certain high-risk and/or high-stakes spreadsheets (Whittaker, 1999; Finlay and Wilson, 2000; Grossman, 2002; Madahar, Cleary, & Ball, 2007). However, at least eight respondents mentioned situations in which the level of review for important spreadsheets was less, not more, rigorous:

- Highly independent executives often completed one-off, ad hoc and first-time analysis for an important decision without the benefit of review.
- When the spreadsheet was highly confidential, few people in the organization had access to it, making effective review difficult.

- Important decisions were often associated with time pressures that precluded formal review.

Several respondents noted that details of the decision context matter, not just the stakes, including whether the decision was internal as opposed to being part of a public or adversarial proceeding such as labor negotiations or law suits. One example stemmed from a highly partisan political budgeting battle. The respondent noted that the opposition's budget literally did not add up. The sum of itemized subcategories did not quite match the category total. It was a minor discrepancy both in absolute dollars and as a percentage of the total budget. However, the respondent was able to exploit that small but incontrovertible error to cast doubt on the credibility of all of the other party's analysis, leading fairly directly to a dramatic political victory before the legislative body.

Spreadsheets Role in Decision Making

This project began with a vision of spreadsheets and decision making that was shaped by experience teaching decision modeling courses. It might be caricatured as follows. "Leaders sometimes analyze decisions with spreadsheets that estimate bottom-line consequences of different courses of action. The spreadsheet includes cells whose values can be selected to achieve managerial goals. Decision makers combine that analysis with expert judgment and factors outside the model to select a course of action."

This view suggests an almost one-to-one connection between spreadsheet errors and decision errors. Instead, we observed a continuum in terms of how spreadsheet output interacted with human judgment that is much broader than our original "academic" view of organizational decision making. We discretize this continuum into a five-part typology based on how tightly coupled the spreadsheet analysis is to the decisions:

Automatic Data Processing: When spreadsheets are used for automatic data processing, errors do directly translate into adverse consequences. A typical administrative horror story was a mis-sort, such as a mailing list matching incorrect names to addresses. As a result, many organizations

avoided using spreadsheets for processing data automatically. As one respondent put it, “It’s not like we’re cutting a check off a spreadsheet.” Yet others described instances in which output from spreadsheets was formalized without any review, including a spreadsheet-based program to automatically generate invoices and this could be problematic.

Making Recommendations That Are Subject to Human Review: Some applications matched our academic view in the sense of computing a bottom-line performance metric that pointed directly to a recommended course of action, but the results were still subject to human review. A typical example would be a spreadsheet used to project the ROI from a contemplated investment. In principle, the decision recommendation is simple; if the ROI is favorable, then take the contemplated action. However, borderline cases triggered more intense scrutiny. In effect, the spreadsheet plus minimal human oversight made the easy calls, but for tough decisions spreadsheet analysis was just the first step. Furthermore, the user was ultimately responsible, and when bad decisions were made, our respondents did not scapegoat the spreadsheets. As one executive in the financial sector said, quantitative analysis is only “the first 75% of a decision. … this [spreadsheet] is a decision making tool, not an arbiter.” Another senior financial manager emphasized that it was “his job” to interpret the spreadsheet analysis in light of other critical qualitative factors. Because human decision making based on experience and domain knowledge was critical, these interviewees were less concerned about the impact of spreadsheet errors.

Projecting Decision Consequences: Many spreadsheets supported “what if” analysis but not of bottom-line performance metrics. Their output becomes just one of several inputs to a higher-level decision process. A typical example was a senior financial manager at a school district modeling the budget impacts of alternative property tax rate scenarios. The spreadsheet output was directly relevant to a specific decision before the school board, yet the board members could have been at least as interested in effects on educational attainment, relations with the teachers’ union, and/or voter anger and their reelection prospects, considerations that were entirely outside of the spreadsheet model.

Understanding A System's Relationships: Managers sometimes used spreadsheets to understand interrelationships among a system's variables even if they did not model a decision's consequences directly. Respondents in this category might "use the spreadsheet to think with," as one respondent put it. Another respondent noted, "The spreadsheet exists as part of the analytical framework in any strategic or operational decision, but it's not the spreadsheet alone." In these situations, the spreadsheet is used more for insight than for computing a specific number. If spreadsheet errors distorted the key relationships, those errors could harm understanding and, hence, decision making. However, the effect would be mediated through something else, namely a person's understanding. Furthermore, the understanding that was the proximate source of any decision errors was subject to independent quality control not tied to the spreadsheet itself.

Providing Background, Descriptive Information: Respondents described many instances where the spreadsheet was used as an information management system to synthesize, organize, and process facts that were relevant to the decision maker, but the spreadsheet did not project the consequences of any specific decision per se. Decision makers might use the spreadsheet to compare current year-to-date costs to those of the previous year, to check progress relative to plan, or to estimate budgets for the coming year. Such basic information could inform myriad decisions, but there is no sense in which the spreadsheet is modeling the consequences of a particular decision. Rather, such spreadsheets estimated parameters that fed into some other decision making system, typically human judgment.

Severity of the Impact of Spreadsheet Errors on Decision Making

A slim majority of subjects whose responses could be coded (25 of 44) expressed strong concern about the consequences of spreadsheet errors in their organization. Nine of the nineteen who were not strongly concerned said they simply did not use spreadsheets in ways that were integral to high-stakes decision making. Almost by definition, spreadsheet errors could not cause those organizations grave harm. Looking solely at organizations that reported using spreadsheets to inform

high-stakes decisions, the slim majority of concerned subjects (25 of 44 or 57%) becomes a substantial majority (25 of 35 or 71%). Still, ten organizations reported both using spreadsheets to inform important decisions and experiencing spreadsheet errors, yet they still had no major concern about adverse impact on the decisions made.

Opinions about whether spreadsheet errors led to bad decisions could also be categorized into three groups in a slightly different way: (1) a small minority who thought spreadsheet errors were always caught before decisions were made, (2) a larger group who acknowledged that not all errors are detected but who thought any errors of consequence would be detected before they misinformed a decision, and (3) the plurality who thought spreadsheet errors could have a significant adverse impact on decisions.

These responses support two conclusions. First, spreadsheet errors sometimes lead to major losses and/or bad decisions in practice. Indeed, we heard about managers losing their jobs because of inadequate spreadsheet quality control. Second, many decision makers whose organizations produce erroneous spreadsheets do not report serious losses or bad decisions stemming from those flawed spreadsheets. Those reassuring self-reports could simply be false; our respondents may have been over-confident about their organizations ability to withstand spreadsheet errors even if all but one admitted they had flawed spreadsheets. However, the interviews suggested another possibility, namely that organizational factors can help prevent errors in spreadsheets that inform decisions from automatically or inevitably leading to bad decisions. Investigating that possibility was not part of our interview protocol and we do not have systematic data concerning this hypothesis, so we elaborate it in the next section, as part of the discussion.

DISCUSSION

We found that spreadsheet errors do lead directly to bad decisions, but the more common scenario may be spreadsheet errors contributing indirectly to suboptimal decisions. That is, our respondents could offer their share of classic spreadsheet errors (e.g., incorrect formulae) causing

someone to make the wrong decision. Often, however, it may be more useful to think of spreadsheet errors as misinforming deliberations rather than as recommending the wrong course of action. Spreadsheets are used for monitoring, reporting, and a host of other managerial activities besides deliberative decision making, and even when a well-defined decision was being made, spreadsheet analysis often provided merely descriptive or contextual information. Spreadsheets were often more of a management information system, than a decision support system. Furthermore, when spreadsheets are used in decision support mode, there is still, by definition, a human in the loop.

Couched in terms of Simon's (1960) intelligence, design, choice decision making framework, spreadsheet errors may lead to errors in the intelligence phase, but their ramifications are buffered by human judgment at the choice phase. Furthermore, it has long been observed that managers' activities do not fit well into the classic vision of planning, organizing, and coordinating (Mintzberg, 1975) and that in practice decision making in organizations departs fundamentally from the classical decision making paradigm (March and Simon, 1958; Cyert and March, 1963; Lipshitz, Klein, Orasanu, & Salas, 2001).

The information context underpinning an organization's decision making process process is almost always murky, ill structured, and incomplete. Since poor information is the norm, decision making processes in successful organizations have evolved to cope with that reality, and a piece of bad information is not like a monkey wrench lodged in the gears of a finely tuned machine. Respondents' reports were consistent with Willemain, Wallace, Fleischmann, Waisel, & Ganaway's (2003) argument that there is a "robust human ability to overcome flawed decision support." Indeed, a flawed spreadsheet might even help dispel some of the fog of uncertainty surrounding a decision, just less effectively than it would have if it had not contained an error (cf., Hodges, 1991; Grossman, 2003). This metaphor of spreadsheet errors clouding an already hazy decision situation is relevant whether the poor information deriving from the erroneous spreadsheet pertains to historical facts, future projections/random quantities, relationships among variables, and/or basic structural assumptions about how to think about the problem.

These observations do not minimize the problem of spreadsheet errors or question the premise that reducing spreadsheet errors can improve organizational performance. However, they may help explain why, if spreadsheets are as riddled with errors as Panko (2005) and Powell et al. (2007b,c) suggest, organizations continue to use them to support key decisions and why a substantial minority of respondents seemed relatively unconcerned about the ramifications of spreadsheet errors.

Implications for Research

To the extent that results from this sample generalize, they have implications for research on spreadsheet errors and decision making. An obvious point is simply that field interviews contribute interesting insights on this important topic. There is a small but important organizational literature that studies spreadsheet use in situ (Nardi and Miller, 1991; Nardi, 1993; Hendry and Green, 1994; Chan and Storey, 1996). Extending its focus slightly from spreadsheet errors to spreadsheet errors' effect on decision making appears useful.

A second point is that the spreadsheet itself is not the only productive unit of analysis. This research took the organization as the unit of analysis. Further research should also consider the decision and/or decision process as the unit of analysis. The impact of spreadsheet errors depends on various factors beyond the spreadsheet itself or even the decision. Our research suggests also paying attention to (1) how the spreadsheet informs the decision and (2) the larger organizational decision making context, such as whether the decision in question is made by an individual or group, whether it is a final decision or a recommendation, and whether the spreadsheet will remain private or be subject to (potentially hostile) external review.

In particular, if our five-part typology of spreadsheet's roles in decision making generalizes to other samples, it would be interesting to assess how concern about errors' effects on decisionmaking varies across those five types. Likewise, it would be useful to move beyond self-report to obtain objective measures of the consequences of flawed spreadsheets misinforming decisions in each of the ways outlined in the typology.

Implications for Practice

Discussion of implications for practice must be prefaced by a large caveat. We did not ask about interventions to improve spreadsheet quality, let alone collect evaluation data demonstrating that an intervention improved organizational performance. All we can do is offer some opinions.

Our first suggestion is that organizations such as those in this sample ought to consider investing more in spreadsheet quality control. There was a yawning chasm between what the research literature suggests and what respondents' described as typical practice. There was also a great gulf between the most and least quality conscious organizations encountered. The gap between literature and practice was most apparent in the types of quality control tools used (formal and technical vs. informal and organizational); the gap among respondents was apparent in the varying intensity with which informal and organizational methods were pursued. Hence, most organizations can ramp up quality control efforts even if they lack the technical sophistication to use high-end methods described in the literature.

A related suggestion is that an organization should not eschew appointing a "chief spreadsheet quality control officer" just because it does not employ software engineers. Certain other educational backgrounds appear to prepare people to appreciate readily the concepts of spreadsheet quality control, notably industrial engineering, systems analysis, and accounting.

Third, the range of methods relevant for preventing bad decisions is broader than is the range of methods relevant for preventing spreadsheet errors. There are two ways an organization can prevent spreadsheet errors from leading to bad decisions: (1) preventing spreadsheet errors and (2) preventing any spreadsheet errors that do occur from translating into bad decisions. For our respondents, the latter was as important as the former. Even when it comes to preventing spreadsheet errors, some actions are organizational not technical, such as insisting that spreadsheet training include error control methods not just functionality, explicitly budgeting time for spreadsheet quality

assurance testing, and having reviewers publicly sign-off on a spreadsheet before it is used, the way a professional engineer must certify the quality of building plans before construction begins.

The final suggestion is for executives simply to raise awareness in their organizations about the idea of establishing spreadsheet quality control standards and procedures. Many managers seemed not to have thought about the possibility of being proactive in spreadsheet quality management.

SUMMARY

Our interviewees affirmed two common findings: (1) Spreadsheets are frequently used to inform decisions and (2) spreadsheets frequently have errors. Given this, one might expect these respondents to be able to recount many instances of spreadsheet errors leading to bad decisions. Indeed, the majority could cite such instances and viewed them as a serious problem. However, a significant minority did not view them as a serious problem and even among those who did, the sky was not falling. No respondent suggested that the proportion of flawed decisions in any way approached the proportion of spreadsheets the literature finds to be flawed.

Disaster was not being avoided because of systematic application of formal, spreadsheet-specific quality control policies and procedures. Indeed, few organizations outside the financial sector had such policies, and the actual practices seem to reflect common concern for the quality of analysis generally more than they did technical or spreadsheet-specific tools or procedures.

Three alternative but not mutually exclusive explanations emerged as to why spreadsheet errors lead to some, perhaps even many, but still not an overwhelming number of flawed decisions. The first view, espoused by a significant minority of respondents, was that informal quality control methods work for precisely those errors that could be most problematic. When the spreadsheet analysis is wildly off, experienced decision makers can sniff that out. Small errors might not be noticed, but small errors were believed to have minor consequences.

The second explanation is that for some organizations (nine of forty-five in our sample), spreadsheets are not used in ways that are tied to specific high-stakes decisions. The spreadsheets

might be used for various types of information processing, ranging from database-like functions to synthesizing and graphing data drawn from another system, but the spreadsheets are not being used in ways that guide specific decisions.

The third explanation is that even if large errors might go undetected in spreadsheets that inform specific, strategic decisions, the spreadsheet analysis is merely informing not driving the decisions. The image one should have is not that an analyst enters all relevant considerations into a spreadsheet, analyzes that spreadsheet, and the organization implements whatever course of action the spreadsheet suggests. Instead, there is some organizational decision process often involving multiple people. Those people bring to the table a great deal of judgment and wisdom, as well as a range of data, mental models, forecasts, and so forth. Spreadsheets may have been used to inform or even generate some of those data, mental models, and forecasts, but other sources of information are also drawn upon. At the end of the day, it is humans exercising human judgment that make the decision.

Usually that judgment is exercised in the face of terribly incomplete and imperfect information. A good spreadsheet analysis might fill in some but not all of that incomplete information. A bad spreadsheet analysis might increase the amount of imperfect information. The murkier the information, the greater the risk of bad decisions, so spreadsheet errors contribute to bad decisions. Ultimately, however, organizational decision processes do not necessarily break down in the face of some bad information, whether it comes from a spreadsheet error or some other source.

References

Baker, K.R., Powell, S.G., Lawson, B. & Foster-Johns, L. (2006a). A Survey of Spreadsheet Users (in submission).

Baker, K.R., Powell, S.G., Lawson, B. & Foster-Johns, L. (2006b) Comparison of Characteristics and Practices Among Spreadsheet Users with Different Levels of Experience. Presented at the European Spreadsheet Risks Interest Group 6th Annual Symposium, Cambridge.

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

Cragg, P.B., & King, M. (1993). Spreadsheet Modeling Abuse: An Opportunity for OR. J, 44(8), 743-752.

Clermont, M., Hanin, C., & Mittermier, R. (2002). A Spreadsheet Auditing Tool Evaluated in an Industrial Context. https://143.205.180.128/Publications/pubfiles/pdffiles/2002-0125-MCCH.pdf.

Croll, G.J. (2005) The Importance and Criticality of Spreadsheets in the City of London. Paper presented to the EuSpRIG Conference and available at http://www.eusprig.org/tiacositcol4.pdf.

Cyert, R. & March, J. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice Hall.

Finlay, P.N., & Wilson, J.M. (2000). A Survey of Contingency Factors Affecting the Validation of End-use Spreadsheet-based Decision Support Systems. JORS, 51, 949-958.

Gerson, M., Chien, I.S., & Raval, V. (1992). Computer Assisted Decision Support Systems: Their Use in Strategic Decision Making. Proceedings of the 1992 ACM SIGCPR Conference on Computer Personnel Research held in Cincinnati, Ohio. New York: ACM Press 152-160.

Grossman, T.A. (2002). Spreadsheet Engineering: A Research Framework. European Spreadsheet Risks Interest Group 3rd Annual Symposium, Cardiff . Available at: http://www.usfca.edu/sobam/faculty/grossman_t.html

Grossman, T.A. (2003). Accuracy in Spreadsheet Modeling Systems. European Spreadsheet Risks Interest Group 4th Annual Symposium, Dublin.

Grossman, T.A., & Özlük, O. (2004). A Paradigm for Spreadsheet Engineering Methodologies. European Spreadsheet Risks Interest Group 5th Annual Symposium, Klagenfurt, Austria. Available at: http://www.usfca.edu/sobam/publications/AParadigmforSpreadsheetEngineeringMethodologies2004.pdf.

Hendry, D.G., & 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, 1033-1065.

Hodges, J. (1991). Six (or so) Things You Can Do With a Bad Model. Operations Research, 39, 355-365.

Howe, H. & Simkin, M.G. (2006). Factors Affecting the Ability to Detect Spreadsheet Errors. Decision Sciences Journal of Innovative Education. 4 (1), 101-122.

Isakowitz, T., Schocken, S., & Lucas, Jr., H.C. (1995). Toward a Logical/Physical Theory of Spreadsheet Modeling. ACM Transactions on Information Systems, 13(1), 1-37.

Janvrin, D., & Morrison, J. (2000). Using a Structured Design Approach to Reduce Risks in End User Spreadsheet Development. Information and Management, 37, 1-12.

Kruck, S.E., Maher, J.J., & Barkhi, R. (2003). Framework for Cognitive Skill Acquisition and Spreadsheet Training. Journal of End User Computing, 15(1), 20-37.

Kruck, S.E., & Sheetz, S.D. (2001). Spreadsheet Accuracy Theory. Journal of Information Systems Education, 12(2), 93-107.

Lipshitz, R., Klein, G., Orasanu, J. & Salas, E. (2001) Taking Stock of Naturalistic Decision Making. Journal of Behavioral Decision Making, 14, 331-352.

Madahar, M., Cleary, P. & Ball, D. (2007) Categorisation of Spreadsheet Use within Organisations Incorporating Risk: A Progress Report. Proceedings of the European Spreadsheet Risks Interest Group 8th Annual Conference, University of Greenwich, London, pp.37-45.

March, J.G. & Simon, H.A. (1958) *Organizations*. New York: Wiley.

Mather, D. (1999). A Framework for Building Spreadsheet Based Decision Models. Journal of the Operational Research Society, 50, 70-74.

Mintzberg, H. (1975) The Manager's Job: Folklore and Fact. Harvard Business Review, June – August, 49-61.

Morrison, M., Morrison, J., Melrose, J., & Wilson E.V. (2002). A Visual Code Inspection Approach to Reduce Spreadsheet Linking Errors. Journal of End User Computing, 14(3), 51.

Nardi, B.A. (1993). *A Small Matter of Programming: Perspectives on End User Computing*. MIT Press, Cambridge, MA.

Nardi, BA. &. Miller, J.R. (1991). Twinkling Lights and Nested Loops: Distributed Problem Solving and Spreadsheet Development. International Journal of Man-Machine Studies, 34, 161-184.

Olphert, C.W., & Wilson, J.M. (2004). Validation of Decision-Aiding Spreadsheets: The Influence of Contingency Factors. Journal of the Operational Research Society, 55, 12-22.

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

Panko, R. (2000a). Two Corpuses of Spreadsheet Error. IEEE: Proceedings of the 33rd Hawaii International Conference on Systems Sciences. Available at: http://panko.cba.hawaii.edu/ssr/Mypapers/HICSS33-Panko-Corpuses.pdf.

Panko, R.R. (2000b) Spreadsheet Errors: What We Know. What We Think We Can Do. Proceedings of the European Spreadsheet Risks Interest Group Conference at the University of Greenwich, London, 7-17, available at www.arxiv.org.

Panko, R.R. (2005). What We Know About Spreadsheet Errors. Available at: http://panko.cba.hawaii.edu/ssr/Mypapers/whatknow.htm, updated 2005. Previously published in Journal of End User Computing 10(2).

Panko, R.R. (2007) Thinking is Bad: Implications of Human Error Research for Spreadsheet Research and Practice. Proceedings of the European Spreadsheet Risks Interest Group 8th Annual Conference, pp. 69-80, available at www.arxiv.org.

Powell, S.G., Lawson, B.,Baker, K.R. (2007a) A Critical Review of the Literature on Spreadsheet Errors. Spreadsheet Engineering Research Project working paper.

Powell, S.G., Lawson, B.,Baker, K.R. (2007b) An Auditing Protocol for Spreadsheet Models. Spreadsheet Engineering Research Project working paper.

Powell, S.G., Lawson, B.,Baker, K.R. (2007c) Errors in Operational Spreadsheets. Spreadsheet Engineering Research Project working paper.

Powell, S.G., Lawson, B.,Baker, K.R. (2007d) Impact of Errors in Operational Spreadsheets. Proc. European Spreadsheet Risks Int. Grp., 57-68.

Purser, M. & Chadwick, D. (2006). "Does An Awareness of Differing Types of Spreadsheet Errors Aid End-Users in Identifying Spreadsheet Errors?" Proceedings of the European Spreadsheet Risk Interest Group Annual Conference, Cambridge, UK, pp. 185–204.

Rajalingham, K., Chadwick, D., & Knight, B. (2000). Classification of Spreadsheet Errors. British Computer Society (BCS) Computer Audit Specialist Group (CASG) Journal, 10(4), 5-10.

Rajalingham, K., Chadwick, D., Knight, B., Edwards, D. (2000). Quality Control in Spreadsheets: A Software Engineering-Based Approach to Spreadsheet Development. Proceedings of the Thirty-Third Hawaii International Conference on System Sciences, Maui, Hawaii, January 2000.

Seal, K.; Przasnyski, Z., and Leon, L. A Literature Survey of Spreadsheet Based MR/OR Applications: 1985-1999. OR Insight. 2000 Oct-2000 Dec 31; 13(4):21-31.

Simon, H.A. (1960) *The New Science of Management Decision*. New York: Harper & Row.

Teo, T.S. H., & Tan, M. (1999). Spreadsheet Development and What if Analysis: Quantitative Versus Qualitative Errors. Accounting, Management and Information Technology, 9, 141-160.

Whittaker, D. (1999). Spreadsheet Errors and Techniques for Finding Them. Management Accounting, 77(9), 50-51.

Willemain, T.R., Wallace, W.A., Fleischmann, K.R., Waisel, L.B., & Ganaway, S.N. (2003). Bad Numbers: Coping with Flawed Decision Support. Journal of the Operational Research Society, 54, 949-957.

Appendix A – Interview Protocol (SS = spreadsheet)

1. Introduction of researchers and project topic
   - Managers frequently use SS to analyze and inform decisions; research has shown that many of these SS contain errors.
   - This project will investigate how these errors affect the quality of decision making, and propose recommendations on the best ways of reducing these errors.

2. How often do you build SS decision making tools?
   - Do you personally create SS to support decisions?
   - How complex are these SS (in terms of the calculations performed in them, or the amount of data contained in them)?

3. How often do you use SS for making decisions?
   - Does your staff present you with SS or SS-based analysis on a regular basis?
   - How complicated are the SS you encounter (in terms of the calculations performed in them, or the amount of data contained in them)?
   - What decisions are these SS being used to support?
   - What makes SS useful for this decision making?
   - Do you use other quantitative tools for decision making?

4. What is your level of expertise with SS modeling and/or other development environments?
   - Have you had formal training in Excel or software programming/development?
   - Have you ever held a position that required daily creation and manipulation of SS?
   - Which features of Excel are familiar to you?

5. Please describe your experiences with SS that were known to contain errors
   - Do SS errors have the potential to cause significant impact?
   - Was the source of the error(s) ever determined?
   - Were the errors caught before damage was done? If not, what was the extent of damage?
   - Describe the errors and what you think caused them.
   - How were the errors fixed?

6. What are the advantages and disadvantages of using SS for decision making?
   - Are there particular features or tools that you have found most useful in your SS?
   - What are the limitations of SS?
   - Is the quality and reliability of your SS a concern for you?
   - Is there anything that might reduce your concerns?

7. Please describe any processes or tools you have used to ensure the integrity of SS
   - Have you or your staff used Excel's built in tools for error-checking?
   - Have you or your staff used Add-ins provided by another vendor?
   - Does your organization follow a particular development process for creating SS models?
   - What other methods are used to detect errors?

8. Are there any other issues related to the topic that you would like to talk about?
   - Do you have advice for other decision-makers?
   - Any stories/anecdotes about particularly helpful solutions to SS problems, or horror-stories about the impact of errors?
   - Recommended readings, web sites, other resources?