

# Chapter 5

## Expert and Novice End–User Spreadsheet Debugging: A Comparative Study of Performance and Behaviour

**Brian Bishop**

*Dundalk Institute of Technology, Ireland*

**Kevin McDaid**

*Dundalk Institute of Technology, Ireland*

### ABSTRACT

*The reliability of end-user developed spreadsheets is poor. Research studies find that 94% of ‘real-world’ spreadsheets contain errors. Although some research has been conducted in the area of spreadsheet testing, little is known about the behaviour or processes of individuals during the debugging task. In this paper, the authors investigate the performance and behaviour of expert and novice end-users in the debugging of an experimental spreadsheet. To achieve this aim, a spreadsheet debugging experiment was conducted, with professional and student participants requested to debug a spreadsheet seeded with errors. The work utilises a novel approach for acquiring experimental data through the unobtrusive recording of participants’ actions using a custom built VBA tool. Based on findings from the experiment, a debugging tool is developed, and its effects on debugging performance are investigated.*

### INTRODUCTION

#### Spreadsheet Prevalence

Commercial off-the-shelf spreadsheet programs are one of the most popular end-user programming environments in use today (Burnett, Cook, & Rothermel, 2004), and have become ubiquitous,

and in many cases indispensable software tools throughout industry and within all levels of the business world. Some of the reasons spreadsheets are so popular in end-user computing is that they are easy to develop and modify, they are highly scalable, and they support numeric data analysis without the specific need to use a ‘behind the scenes’ programming environment. Spreadsheet

DOI: 10.4018/978-1-4666-2059-9.ch005

programs are also extremely flexible in terms of size and complexity: the 2007 version of Microsoft Excel can support over 1 million rows and 16 thousand columns per worksheet, with a maximum formula length of 8 thousand characters per cell.

It has been estimated that by 2012 there will be 90 million end-users in the U.S. alone and of these 13 million will be end-user programmers; which is significantly higher than the estimated 3 million professional programmers in the U.S. by the same year (Scaffidi et al., 2005). Purser and Chadwick (2006) stated that “*commonly known – though unpublished and intentionally unconfirmed by Microsoft – Excel has an install base of 90% on end-user desktops*” (p. 185).

The reported usage of spreadsheet programs spans a wide variety of job functions, purposes and industries. In a survey of nearly 1600 respondents, Baker et al. (2006) found that spreadsheets were used by end-users in various job functions including finance, engineering, manufacturing, marketing, sales, HR etc., and for many different purposes, such as maintaining lists, analysing data, tracking data, determining trends etc. Maybe more so than of any other industry, spreadsheets are of critical importance to the finance sector (Croll, 2005). In a study on the use of spreadsheets in organisations in the City of London, Croll (2005) found that with regard to the financial markets “*Excel is utterly pervasive. Nothing large (good or bad) happens without it passing at some time though Excel*” (p. 3).

## Spreadsheet Reliability

Worryingly, the reliability of end-user developed spreadsheets has been shown to be poor following empirical and anecdotal evidence collected from studies investigating operational (real world) spreadsheet error rates, most of which are detailed in previous studies (Panko, 1998; Panko, 2000; Powell, Baker, & Lawson, 2008). The most recent study that investigated spreadsheet error rates was conducted by Powell, Baker and Lawson (2007),

who reported that of the 50 real-world operational spreadsheets they audited, 94% contained errors. Due to the nature of spreadsheet use, it is the case that when failures do occur the results can be quite significant. Many real world examples of the impact of errors in spreadsheets are published on the European Spreadsheet Risks Interest Group website ([www.eusprig.org](http://www.eusprig.org)). The following is just a small selection of these reported errors:

- A government housing authority had to pay \$216,352 to cover expenses incurred as a result of a spreadsheet data-entry error that overpaid landlords.
- A spreadsheet error caused the Nevada City 2006 budget to show a deficit of \$5 million for a water fund.
- In 2005, a Kodak employee was mistakenly accrued an \$11 million severance. A Kodak spokesman said the severance error was traced to a faulty spreadsheet, and the money was not paid out.
- Stock prices for an online retailer tumbled by 25% in one day, and the Chief Financial Officer resigned due to a single number mis-recorded in one cell of a spreadsheet.
- Numbers entered as text lost a school £30,000 from its budget.
- A finance company was undercharged \$25,652 for rent in 1999 due to an error in a spreadsheet formula.
- A Legislative Auditor found a government department’s key performance indicator value overstated by \$58 million (45%). The miscalculation was due to a spreadsheet formula error.

The reported and published spreadsheet errors are just the tip of the iceberg. Very few organisations report embarrassing spreadsheet errors and their impacts, and details of these errors only enter the public domain when it becomes unavoidable. From correspondence with professionals that cannot be published, there appears to be many

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/expert-novice-end-user-spreadsheet/69613](http://www.igi-global.com/chapter/expert-novice-end-user-spreadsheet/69613)

## Related Content

---

### Inhibitors of Two Illegal Behaviors: Hacking and Shoplifting

Lixuan Zhang, Randall Young and Victor Prybutok (2008). *End-User Computing: Concepts, Methodologies, Tools, and Applications* (pp. 1652-1666).

[www.irma-international.org/chapter/inhibitors-two-illegal-behaviors/18277](http://www.irma-international.org/chapter/inhibitors-two-illegal-behaviors/18277)

### A Meta-Analysis Approach toward the Development of an Integrative Framework for Online Consumer Behavior Research

Khawaja A. Saeed, Yujung Hwang and Mun Y. Yi (2004). *Advanced Topics in End User Computing, Volume 3* (pp. 69-101).

[www.irma-international.org/chapter/meta-analysis-approach-toward-development/4458](http://www.irma-international.org/chapter/meta-analysis-approach-toward-development/4458)

### Friend or Foe?: Leveraging – and Mitigating – the Access and Immediacy of Social Media and Information Technologies

Linda L. Brennan (2011). *Organizational and End-User Interactions: New Explorations* (pp. 376-388).

[www.irma-international.org/chapter/friend-foe-leveraging-mitigating-access/53100](http://www.irma-international.org/chapter/friend-foe-leveraging-mitigating-access/53100)

### Twitter Predicting the 2012 US Presidential Election?: Lessons Learned from an Unconscious Value Co-Creation Platform

Miguel Maldonado and Vicenta Sierra (2016). *Journal of Organizational and End User Computing* (pp. 10-30).

[www.irma-international.org/article/twitter-predicting-the-2012-us-presidential-election/154000](http://www.irma-international.org/article/twitter-predicting-the-2012-us-presidential-election/154000)

### User Behaving Badly: Phenomena and Paradoxes from an Investigation into Information System Misfit

Panagiotis Kanellis and Ray J. Paul (2008). *End-User Computing: Concepts, Methodologies, Tools, and Applications* (pp. 637-657).

[www.irma-international.org/chapter/user-behaving-badly/18213](http://www.irma-international.org/chapter/user-behaving-badly/18213)