An Investigation of End User Development Success

Tanya Jane McGill BSc, Grad Dip Comp, MBA

This thesis is presented for the degree of Doctor of Philosophy of Murdoch University

2002
I declare that this thesis is my account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

(Tanya McGill)
Abstract

User development of applications provides end users with an alternative to the traditional process of systems development by allowing them to solve job related problems by developing their own software applications. User developed applications (UDAs) support decision making and organisational processes in the majority of organisations, and the ability to develop small applications forms part of the job requirements for many positions. Despite its pervasiveness, there are many risks associated with user development of applications. These risks result primarily from decreases in application quality that arise when end users have had little training and do not follow system development methodologies.

The primary aim of the research described in this thesis is to gain a better understanding of UDA success. In particular, the thesis considers the role of system quality in UDA success and the ability of end user developers to judge whether the applications they develop will have a positive impact on their performance of tasks. The research also investigates factors that might impact upon this ability.

The research objectives were addressed through two empirical studies. Two possible models of UDA success provided the starting point for Study 1. The first model is DeLone and McLean’s (1992) model of IS success, and the second model is a version of this model that was modified to address concerns about the DeLone and McLean model and to reflect current research about UDA success. The models were tested using data from a field study involving business people participating in a business policy simulation, where they developed spreadsheet applications to assist in decision making. Structural equation modelling was used to test the models. Neither of the models was well supported by the data. However, the analysis provided strong support for relationships between perceived system quality and user satisfaction, information quality and user satisfaction, user satisfaction and intended use, and user satisfaction and individual impact. It is notable that the model paths that were supported in Study 1 were primarily those that reflect user perceptions rather than objective measures. This study highlighted that user perceptions of information systems success play a significant
role in the UDA domain. The results did, however, suggest that there might be a direct relationship between system quality and individual impact.

Study 2 was a laboratory experiment and the participants were end users from a range of organisations. A revised research model was developed based on the findings of Study 1, and structural equation modelling was again used to test the model. The model paths that were supported suggest that for small to moderate applications, increases in spreadsheet development knowledge lead to increases in system quality and consequently the development of better quality spreadsheets. They also suggest that for these kinds of applications, end users have realistic perceptions of system quality and hence that user satisfaction may be an appropriate measure of UDA success. The results of Study 2 also provided insight into the role of user involvement in end user development, clarifying the process by which benefits are obtained. The study also provided insight into the importance of spreadsheet development knowledge for successful use (as well as development) of a spreadsheet application.

The results described in this thesis have practical implications for the management of user development of applications. They highlight the need either to increase end user levels of development knowledge via training so that end users can cope with applications of greater complexity, or to provide other forms of support for development. The role of organisational standards and guidelines is also be considered in the thesis and it is suggested that there is a particular need for guidelines on what kinds of applications are suitable for end user development.
Table of Contents

Chapter 1 Introduction ............................................................................................................. 1
  1.1 Background ..................................................................................................................... 1
  1.2 Research problem .......................................................................................................... 2
  1.3 Purpose of the research ................................................................................................. 4
  1.4 Importance of the research ........................................................................................... 5
  1.5 Research approach ........................................................................................................ 7
  1.6 Organisation of the thesis ............................................................................................. 9
  1.7 Definition of key terms ............................................................................................... 11

Chapter 2 Literature Review .................................................................................................. 15
  2.1 Introduction .................................................................................................................... 15
  2.2 What is end user development of applications? .......................................................... 16
    2.2.1 End user development of applications defined ..................................................... 16
    2.2.2 Benefits and risks of user development of applications ....................................... 18
  2.3 Characteristics of user developed applications ........................................................... 21
    2.3.1 Types of tasks that users develop applications for ............................................... 21
    2.3.2 Application scope and size .................................................................................... 22
    2.3.3 Sources of data and interdependence of applications ........................................... 24
    2.3.4 Frequency of use ................................................................................................. 24
  2.4 Tools used for end user development .......................................................................... 24
  2.5 End user developer characteristics ............................................................................. 26
    2.5.1 Age and gender ...................................................................................................... 26
    2.5.2 Experience and training ....................................................................................... 27
    2.5.3 Involvement .......................................................................................................... 30
  2.6 Processes used by end users in development of applications ....................................... 31
  2.7 Management of user development of applications ..................................................... 32
    2.7.1 Information centre approach ................................................................................ 34
    2.7.2 Managed free economy approach ....................................................................... 35
    2.7.3 Alavi, Nelson and Weiss’s EUC management framework .................................... 36
    2.7.4 Expansion/control approach ................................................................................ 36
    2.7.5 Brown and Bostrom’s model of EUC management effectiveness ......................... 38
    2.7.6 Other research on EUC management ................................................................... 39
  2.8 IS and UDA success ...................................................................................................... 41
    2.8.1 IS success ............................................................................................................. 41
    2.8.2 UDA success ....................................................................................................... 46
  2.9 Overview ...................................................................................................................... 50

Chapter 3 Study 1 Research Questions and Conceptual Models ........................................ 53
  3.1 Introduction .................................................................................................................... 53
  3.2 Research questions ....................................................................................................... 54
  3.3 Framework for the study ............................................................................................. 55
    3.3.1 System quality ..................................................................................................... 56
    3.3.2 Information quality ............................................................................................. 57
    3.3.3 User satisfaction ................................................................................................. 58
5.3.3 Comparison of the DeLone and McLean model and the partial alternate model .................................................................................. 129
5.3.4 Full alternate model ......................................................................................................................................................... 130
5.4 Overview ............................................................................................................................................................................. 134

Chapter 6 Study 1 Discussion ................................................................................................................................. 135
6.1 Introduction ..................................................................................................................................................................... 135
6.2 Discussion of the hypotheses ............................................................................................................................................ 135
6.2.1 Experience does not influence system quality ............................................................................................................. 136
6.2.2 Training does not influence system quality .................................................................................................................. 137
6.2.3 Experience does not influence information quality ..................................................................................................... 138
6.2.4 Training does not influence information quality .......................................................................................................... 139
6.2.5 Experience influences perceived system quality ........................................................................................................ 140
6.2.6 Training does not influence perceived system quality ................................................................................................ 141
6.2.7 System quality does not influence perceived system quality ....................................................................................... 141
6.2.8 User satisfaction reflects perceived system quality and information quality ........................................................................................................... 143
6.2.9 Information quality and perceived system quality do not directly influence use ....................................................... 144
6.2.10 User satisfaction influences use .................................................................................................................................. 146
6.2.11 Use does not influence individual impact .................................................................................................................... 146
6.2.12 User satisfaction influences individual impact ............................................................................................................. 147
6.2.13 Individual impact does not influence organisational impact ....................................................................................... 148
6.2.14 System quality influences individual impact ................................................................................................................. 149
6.2.15 Information quality does not influence individual impact .......................................................................................... 150
6.3 Discussion of the models .................................................................................................................................................. 150
6.4 Can the research questions be answered? ..................................................................................................................... 155
6.5 Further research suggested by Study 1 .......................................................................................................................... 157
6.5.1 The role of use in UDA success .................................................................................................................................. 158
6.5.2 The role of information quality in UDA success ........................................................................................................... 159
6.5.3 The roles of experience and training in UDA success .................................................................................................. 160
6.5.4 Does system quality directly influence individual impact? ............................................................................................ 161
6.5.5 The role of individual impact on organisational impact ............................................................................................... 161
6.6 Overview ........................................................................................................................................................................... 162

Chapter 7 Study 2 Research Questions and Conceptual Model .................................................................................. 165
7.1 Introduction ......................................................................................................................................................................... 165
7.2 Study 2 research questions ............................................................................................................................................... 166
7.3 Framework for the study ....................................................................................................................................................... 168
7.3.1 Spreadsheet development knowledge .......................................................................................................................... 169
7.3.2 Involvement ..................................................................................................................................................................... 171
7.4 Hypotheses ........................................................................................................................................................................... 172
7.4.1 How do experience and training impact upon system quality? ............................................................................................ 172
7.4.2 What factors influence perceived system quality? ........................................................................................................... 174
7.4.3 How do system quality and perceived system quality influence user satisfaction, perceived individual impact and individual impact? ........................................................................................................... 176
7.4.4 Does the spreadsheet development knowledge of the user play a role in their ability to successfully use a UDA? ............................................................................................................................... 179
Chapter 10 Study 2 Discussion ................................................................. 247
10.1 Introduction .......................................................................................... 247
10.2 Comparison between user developers and users who were not the developer 248
10.3 Discussion of hypotheses ........................................................................ 251
10.3.1 Experience influences spreadsheet development knowledge ............. 252
10.3.2 Training influences spreadsheet development knowledge ................. 253
10.3.3 Spreadsheet development knowledge influences system quality ...... 254
10.3.4 System quality influences perceived system quality ....................... 257
10.3.5 Spreadsheet development knowledge does not influence perceived  
    system quality .................................................................................... 259
10.3.6 Involvement influences perceived system quality ............................ 260
10.3.7 System quality influences individual impact if the user is also the  
    developer .................................................................................................. 261
10.3.8 Perceived system quality influences user satisfaction ...................... 262
10.3.9 User satisfaction influences perceived individual impact ............... 263
10.3.10 Individual impact does not influence perceived individual impact ... 263
10.3.11 User satisfaction influences individual impact ............................... 265
10.3.12 Spreadsheet development knowledge influences individual impact .. 266
10.4 Support for the Study 2 model of UDA success .................................... 266
10.4.1 Additional constructs that might be relevant .................................... 271
10.5 Contribution of Study 2 to answering the research questions .............. 274
10.6 Limitations of the research described in this thesis ............................. 278
10.7 Overview ................................................................................................ 280

Chapter 11 Conclusions .............................................................................. 283
11.1 Summary of the research and its contributions ................................... 283
11.2 Implications of these studies for UDA success research ...................... 288
11.3 Practical implications of the research .................................................. 290

Appendices .................................................................................................. 295
Appendix A Study 1 questionnaire ............................................................ 297
Appendix B Items retained after Study 1 measurement model development .... 305
Appendix C Advertisement used to recruit participants for Study 2 ............ 309
Appendix D Training session provided for participants in Study 2 ............... 313
Appendix E Study 2 participant consent form .......................................... 325
Appendix F Study 2 questionnaire 1 ............................................................ 329
Appendix G Study 2 decision questions ...................................................... 337
Appendix H Study 2 questionnaire 2 ............................................................ 343
Appendix I Confirmation of dimensionality of involvement ....................... 347
Appendix J Possible directions of the relationship between user satisfaction and  
    (perceived) individual impact – model fit information ............................ 351

References ................................................................................................... 355

Publications arising from this research ...................................................... 369
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Munro, Huff and Moore’s (1987-88) expansion/control approach</td>
<td>37</td>
</tr>
<tr>
<td>2.2</td>
<td>Brown and Bostrom’s (1989) model of EUC management effectiveness</td>
<td>38</td>
</tr>
<tr>
<td>2.3</td>
<td>DeLone and McLean’s (1992) model of IS success</td>
<td>43</td>
</tr>
<tr>
<td>2.4</td>
<td>The technology-to-performance chain (Goodhue &amp; Thompson, 1995)</td>
<td>44</td>
</tr>
<tr>
<td>2.5</td>
<td>The technology acceptance model (TAM) (Davis, 1986)</td>
<td>46</td>
</tr>
<tr>
<td>3.1</td>
<td>DeLone and McLean’s (1992) model of IS success</td>
<td>55</td>
</tr>
<tr>
<td>3.2</td>
<td>A modified and testable representation of the DeLone and McLean (1992) model of IS success factors showing the hypothesised relationships</td>
<td>64</td>
</tr>
<tr>
<td>3.3</td>
<td>The alternate model of UDA success tested in Study 1</td>
<td>69</td>
</tr>
<tr>
<td>4.1</td>
<td>Possible decision support system (spreadsheet) configurations for teams in the BPG</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Representation of the DeLone &amp; McLean (1992) model of IS success tested in Study 1</td>
<td>120</td>
</tr>
<tr>
<td>5.2</td>
<td>Partial alternate model tested in Study 1</td>
<td>121</td>
</tr>
<tr>
<td>5.3</td>
<td>Full alternate model tested in Study 1</td>
<td>121</td>
</tr>
<tr>
<td>5.4</td>
<td>Structural equation model showing the standardised path coefficients for each hypothesised path, and the $R^2$ for each dependent variable in the DeLone and McLean model</td>
<td>124</td>
</tr>
<tr>
<td>5.5</td>
<td>Structural equation model showing the standardised path coefficient for each hypothesised path and the $R^2$ for each dependent variable in the partial alternate model</td>
<td>127</td>
</tr>
<tr>
<td>5.6</td>
<td>Structural equation model showing the standardised path coefficient for each hypothesised path and the $R^2$ for each dependent variable in the full alternate model</td>
<td>132</td>
</tr>
<tr>
<td>6.1</td>
<td>Relationships between UDA success factors supported by Study 1</td>
<td>152</td>
</tr>
<tr>
<td>7.1</td>
<td>The model adopted for Study 2</td>
<td>169</td>
</tr>
<tr>
<td>7.2</td>
<td>The hypothesised relationships between experience, training, spreadsheet development knowledge and system quality</td>
<td>172</td>
</tr>
<tr>
<td>7.3</td>
<td>The hypothesised relationships between system quality, spreadsheet development knowledge, involvement and perceived system quality</td>
<td>175</td>
</tr>
<tr>
<td>7.4</td>
<td>The relationships proposed to explain how system quality and perceived system quality influence user satisfaction, perceived individual impact and individual impact</td>
<td>177</td>
</tr>
<tr>
<td>7.5</td>
<td>The hypothesised relationship between spreadsheet development knowledge and individual impact</td>
<td>179</td>
</tr>
</tbody>
</table>
Figure 8.1: Problem statement given to participants in Part 2 of the experimental session

Figure 9.1: The relationship between time to make decisions and number of decisions correct

Figure 9.2: Hypotheses tested for the scenario where an end user develops and then uses his/her own UDA

Figure 9.3: Hypotheses tested for the scenario where an end user uses an application developed by another end user

Figure 9.4: Structural equation model for the user developer data showing the standardised path coefficient for each hypothesised path and the $R^2$ for each dependent variable

Figure 9.5: Structural equation model for the non user developer data showing the standardised path coefficient for each hypothesised path and the $R^2$ for each dependent variable

Figure 10.1: Relationships supported in Study 2
List of Tables

Table 2.1: Potential benefits of user development of applications ............................................... 19
Table 2.2: Potential risks of user development of applications ...................................................... 20
Table 3.1: Summary of research on the relationships depicted in DeLone and McLean’s model .......................................................... 61
Table 4.1: Existing instruments used as a basis for measurement of the constructs included in the questionnaire for Study 1 .................................................. 82
Table 4.2: The item pool used in the measurement of perceived system quality and system quality ........................................................................................................ 84
Table 4.3: The item pool used in the measurement of information quality ................................. 87
Table 4.4: The item pool used in the measurement of user satisfaction ..................................... 88
Table 4.5: The item pool used in the measurement of individual impact ..................................... 89
Table 4.6: Organisational performance variables included in the Z-score measure of performance ........................................................................................................ 91
Table 4.7: Age and computing experience of the participants in Study 1 .............................. 102
Table 4.8: Frequency of computer use by the participants in Study 1 ................................. 103
Table 4.9: Spreadsheet experience and perceived skill of the participants in Study 1 103
Table 4.10: Previous spreadsheet training of the participants in Study 1 ........................... 104
Table 5.1: System quality summary statistics .......................................................................... 108
Table 5.2: One factor congeneric model for system quality – parameter estimates, goodness of fit measures and reliability measures .............................................. 109
Table 5.3: Perceived system quality summary statistics ........................................................ 110
Table 5.4: Goodness of fit measures and reliability measures for the one factor model for perceived system quality ................................................................. 111
Table 5.5: Information quality summary statistics .................................................................. 112
Table 5.6: One factor congeneric model for information quality - parameter estimates, goodness of fit measures and reliability measures .............................................. 112
Table 5.7: User satisfaction summary statistics ........................................................................ 113
Table 5.8: One factor congeneric model for user satisfaction – parameter estimates, goodness of fit measures and reliability measures .............................................. 114
Table 5.9: Use summary statistics ......................................................................................... 115
Table 5.10: Individual impact summary statistics .................................................................... 115
Table 5.11: Parameter estimates and reliability measures for individual impact ............... 116
Table 5.12: Summary of the information used to specify measurement model parameters in the structural models ................................................................. 118
Table 5.13: Implied correlations between the latent variables in the models ....................... 119
Table 9.11: Parameter estimates and reliability measures for perceived individual impact
Table 9.12: Individual impact summary statistics
Table 9.13: Summary of the information used to specify measurement model parameters in the structural model used to test the hypothesised relationships with the user developer data
Table 9.14: Summary of the information used to specify measurement model parameters in the structural model used to test the hypothesised relationships with the non user developer data
Table 9.15: A comparison of end user developer perceptions and performance when using their own or another application
Table 9.16: Implied correlations between the latent variables considered in the model to test the hypothesised relationships using the user developer data
Table 9.17: Implied correlations between the latent variables considered in the model to test the hypothesised relationships using the non user developer data
Table 9.18: Model coefficients, standard errors, t values and goodness of fit measures for the model tested with user developer data
Table 9.19: Standardised total effects on dependent variables estimated for the model tested with user developer data
Table 9.20: Model coefficients, standard errors, t values and goodness of fit measures for the model tested with the non user developer data
Table 9.21: Standardised total effects on dependent variables estimated for the model tested with non user developer data
Table 10.1: Summary of the relationships investigated in Study 2
Acknowledgements

I would like to thank the many people who have provided ideas, support and encouragement during this research. In particular, I would like to thank my supervisors Val Hobbs and Jane Klobas; I benefited greatly from their expertise and support. Jane’s enthusiasm, guidance, and insight made a huge difference to the progress of this thesis. My time with her in Milan was particularly valuable and most appreciated. Val’s encouragement, friendship and confidence in my ability also all helped significantly.

Chris Klisc and Helen Rowe provided substantial practical assistance at various points in conducting the two studies. This help was invaluable. Several other colleagues also provided useful advice and support: Lanny Entrekin kindly suggested the research environment used in the first study; Sam Downes and Eve Voysey provided advice on several measurement instruments; Diana Adorno helped me grapple with thesis formatting; Tom Gedeon provided practical support and encouragement; and May Lanchbury, Glenys Hoar and Dianne McShane all helped make the experimental sessions run smoothly. I would also like to thank the many participants in my research. Their interest and cooperation made these studies possible.

Finally, I would like to thank Mike Dixon who provided much encouragement and moral support, and showed me by example that it could be done, and Debbie Street, who provided practical assistance in many phases of the project and whose friendship and moral support helped sustain me over the entire period.
Chapter 1
Introduction

1.1 Background

End user computing (EUC) can be defined broadly as 'the use and/or development of information systems (IS) by the principal users of the systems’ output' (Wetherbe & Leitheiser, 1985). EUC now dominates organisational use of information technology worldwide. Its growth has been driven by increasingly inexpensive hardware, increasingly powerful and easy to use software, and user demand for control of information resources (McLean, Kappelman, & Thompson, 1993; Shayo, Guthrie, & Igbaria, 1999). Other reasons proposed for the enthusiasm with which users have adopted EUC include backlogs in application development by IS departments and general dissatisfaction with the IS environment within organisations.

End user development of applications forms a significant part of EUC. An end user developer is someone who develops applications systems to support his or her work or the work of other end users. A user developed application (UDA)\(^1\) is therefore defined as an application developed by an end user. The development of a UDA is a direct response to a particular organisational task or duty, undertaken by staff involved in that task or duty. User developers may have had little formal training for this role (Chan & Storey, 1996) and their technical abilities vary considerably (McGill & Dixon, 2001), but they are basically required to analyse, design and implement computer applications (Sipior & Sanders, 1989). Increasingly, the ability to develop small applications forms part of the job requirements for many positions (Jawahar & Elango, 2001).
1.2 Research problem

User development of applications provides users with an alternative to the traditional process of systems development; it allows end users to solve job related problems by developing their own software applications. It has been suggested that end user development offers organisations better and more timely access to information, improved quality of information, improved decision making, improved IS department/user relationships, and lower system development costs (Brancheau & Brown, 1993; Shayo et al., 1999). Many of the proposed benefits of user development of applications are seen to flow from a belief that the user has a superior understanding of the problem to be solved by the application and will therefore develop a ‘better’ application.

Notwithstanding the potential benefits of end user development of applications, there is a recognition that there are many risks associated with it that may lead to dysfunctional consequences for an organisation’s activities. These risks result from a possible decrease in quality and control as individuals with little or no formal IS training take responsibility for developing and implementing systems of their own making (Cale, 1994). The risks include ineffective use of monetary resources, threats to data security and integrity, solving the wrong problem (Alavi & Weiss, 1985-1986), unreliable systems, incompatible systems, and use of private systems when organisational systems would be more appropriate (Brancheau & Brown, 1993). With the increase in end user development of World Wide Web applications that are immediately accessible by unlimited numbers of people anywhere in the world, these risks, and their potential impact, are of increasing importance (Nelson & Todd, 1999).

1 The plural form UDAs is used throughout the thesis for user developed applications.
Despite these risks, organisations generally undertake little formal evaluation of the quality and impact of applications developed by end users (Panko & Halverson, 1996). In a 1990 study of EUC policies, Bergeron and Berube found that only 39% of organisations surveyed required user developers to have applications approved by the support group or IS department (Bergeron & Berube, 1990). Similarly, Cale (1994) found that few firms had formulated policies requiring or supporting formal testing and documentation of end user developed software. In a recent study of the effects of Web technology on EUC and its management (Nelson & Todd, 1999), participants perceived levels of control of end user development to still be significantly lower than they believed was necessary. This suggests that in many organisations, the user developer's own judgement is still the only measure of an application's suitability for use, yet it has been found that user developers' own evaluations of UDA effectiveness are not necessarily accurate (Edberg & Bowman, 1996).

There has been little research on the outcomes of user development of applications. As long ago as 1993, Brancheau and Brown stated that they found it somewhat surprising that more was not known but suggested that the rapidly moving nature of user development of applications and the lack of an overall framework for the research area contributed to the lack of cohesive research. Six years later, Shayo et al. (1999) noted that measurement of UDA success was still problematic and again called for the development of a more comprehensive and integrated model of EUC success. This call has been echoed even more recently by Powell and Moore (2002). The fact that vital organisational decision making currently relies on the individual end user’s assessment of application effectiveness suggests that more insight is needed into the ability of end user developers to judge whether the applications they develop will have a positive
impact on their effectiveness and productivity, and on the productivity of the organisation.

1.3 Purpose of the research

Information systems success models indicate that the characteristics of systems impact upon the performance of the individuals and organisations that use them (e.g. DeLone & McLean, 1992; Goodhue & Thompson, 1995). The risks associated with user development of applications (discussed in Section 1.2 and in Chapter 2) highlight the need for concern about the impact of poor quality UDAs on the work performance of individuals and on the success of organisations in general. Applications developed by end users are generally of lower quality than other organisational systems (Edberg & Bowman, 1996), yet there are fewer safeguards in place to protect against the risks (Nelson & Todd, 1999; Panko & Halverson, 1996). Organisations rely very heavily upon the judgement of individual end user developers, yet little research has been done to justify this reliance.

The objective of the research described in this thesis is to gain a better understanding of UDA success. In particular, the research considers the role of system quality in UDA success and the ability of end user developers to judge whether the systems they develop will have a positive impact on their performance of tasks. The research also investigates factors that might impact upon this ability. The target population for this research is end user developers in all organisations.

The research questions addressed in this thesis are:

1. How does UDA quality contribute to user performance on tasks?
2. Do end user developers have any misconceptions about the quality of their applications? If so, how do these misconceptions impact upon their ability to judge whether the applications they develop will have a positive impact on their performance of tasks?

3. What characteristics of end user developers influence their ability to judge whether the applications they develop will have a positive impact on their performance of the tasks the UDA is designed to support?

1.4 Importance of the research

When EUC first emerged its management ranked highly in surveys of IS management issues (e.g. Brancheau & Wetherbe, 1987; Dickson, Leitheiser, Wetherbe, & Nechis, 1984). However, the ranking of EUC as an IS management issue has dropped since the late 1980s. Facilitating and managing EUC was ranked 11th out of 31 issues in Pervan’s (1997) study of chief information officers; 16th out of 20 issues in Brancheau, Janz and Wetherbe’s (1996) study; and only 23rd out of 24 issues in Gottschalk’s (1999) study.

Munro, Huff, Marcin and Compeau (1997) have questioned this apparent lack of interest in EUC management by IS managers. One possible explanation for this perceived lack of interest is the fact that EUC has become so all pervasive it is no longer seen an issue for IS management, rather the concern of functional managers (Kreie, 1998). Munro et al. also argued that the drop in ranking is because managers are confident that they are managing the area well, assuming that the mere availability of end user tools encourages productive activity. There is, however, evidence that a substantial proportion of the potential return on the investment in end user information technology is never realised (Guimaraes, 1996; Palvia, 1991), and examples of
problems arising from user development of applications abound (e.g. Jenne, 1996; Nelson & Todd, 1999). This suggests that organisations should not be complacent with respect to the management of EUC.

Despite the drop in ranking (and unlike disaster recovery and CASE technology), EUC management has nevertheless remained on the lists of important IS management issues over a 20 year period. It also has implications for other highly ranked IS management issues such as managing IS effectiveness and productivity, and assuring software quality as well as improving data integrity and quality assurance. Issues relating to user development of applications do therefore warrant further research. In particular, because of the current reliance on end user perceptions of application success (Shayo et al., 1999), a better understanding is needed of the ability of end user developers to judge whether the applications they develop will have a positive impact on their performance, and of the factors that might impact upon this ability.

The research described in this thesis should enable organisations to gauge whether the levels of control they currently have over user development of applications and the support they provide for it are appropriate, and hence facilitate organisational change leading to improvements in EUC management. Those involved in the education and training of end user developers will be able to use the results of this research to help end users acquire the skills necessary to develop quality applications and to improve their ability to judge the impact of their applications. This in turn will have a positive influence on individual performance and hence organisational performance (DeLone & McLean, 1992).
1.5 Research approach

In order to achieve its objective this research looks first at the relationship between system quality and performance in the UDA domain. Two possible models of this relationship were initially identified. The first model is DeLone and McLean’s (1992) model of IS success, a widely discussed model (Walstrom & Leonard, 2000) of organisational IS success that has received some empirical support (Seddon & Kiew, 1996), but which has also been subject to critique (Seddon, 1997b). The second model is a modified version of DeLone and McLean’s model that was extended to include the roles of user training and experience in UDA success.

This research consisted of two studies. Study 1 was a test of the two initial research models described above. The study involved business people who were undertaking a Masters of Business Administration (MBA) and participating in a business policy simulation game which involved developing their own spreadsheet applications to assist in decision making. This approach allowed testing of the models in full, including a number of organisational outcome measures. In addition, the controlled situation allowed some control over the impact of extraneous variables. Structural equation modelling was used to test the models.

The second study was designed to specifically address several issues arising from the first study. A revised research model was developed based on the findings of Study 1. Study 2 was a controlled laboratory experiment with end users from a range of public and private organisations as participants. Structural equation modelling was again used to test the model. This study provided an opportunity to compare more closely end user perceptions of IS success measures with independent measures. It also provided the opportunity to ensure that the findings would be applicable to the wider end user
developer population, given that the Study 1 participants were solely MBA students. However, because of the constrained nature of Study 2, the applications developed were smaller and simpler than those investigated in the first study.

The three research questions presented in Section 1.3 are addressed in each of the studies as follows. Both studies investigate how UDA quality contributes to user performance on tasks (research question 1). Study 2 extends the work in Study 1 to include the role of application development tool knowledge in contributing to both system quality and user performance. Study 2 also investigates the relationship between perceived individual impact and independently measured impact on task performance.

The second research question focuses on whether end user developers have any misconceptions that may influence their ability to judge the impact of their systems on their performance of tasks. Study 1 addresses this question by looking at the impact of end user misconceptions about system quality on UDA success measures in general, and Study 2 explicitly looks at end user developers’ ability to judge whether the applications they develop will have a positive impact on their performance of tasks.

The third research question asks what characteristics of end user developers influence their ability to judge whether the applications they develop will have a positive impact. Study 1 provides an initial investigation of the roles of experience and training in UDA success. Study 2 extends this to also include the roles of application development tool knowledge, and user involvement with applications.
1.6 Organisation of the thesis

This thesis is presented in eleven chapters. Chapter 1 provides a brief overview of user development of applications and its importance, and presents the overall research questions addressed in the research.

Chapter 2 reviews the relevant literature on IS success, EUC, and end user development of applications, providing a summary of the current state of research on UDA success. The issues raised in this chapter highlight the need for further research on the ability of end user developers to judge the impact of their systems.

Chapters 3 to 6 discuss the first study undertaken. Chapter 3 provides the research questions and hypotheses for Study 1. It also presents the research models addressed in Study 1, and provides the theoretical justification for them.

Chapter 4 provides a detailed description of the research carried out in Study 1. It describes the research methodology including the sample, study design, instrument development and the data collection, and describes the data analysis technique used (structural equation modelling). It also details the demographic data about the participants and describes the applications that were developed.

The results of Study 1 are presented in Chapter 5. The measurement model development is firstly explained and described. The chapter then provides the results of the structural model evaluation including the results of each of the hypothesis tests.
Chapter 6 discusses the findings of Study 1. The results of this study are discussed in the light of their practical and research implications and issues requiring further research are highlighted for investigation in Study 2.

Chapters 7 to 10 relate to the second study. Chapter 7 presents the research questions developed in the light of the results of Study 1. A new research model is also proposed that takes into account both the results of Study 1 and the relevant literature.

A detailed description of the research carried out in Study 2 is provided in Chapter 8. The research methodology is described including the sample, study design, instrument development and data collection and data analysis. Details of the demographic data about the participants are given and the applications that were developed are described.

Chapter 9 presents the results of Study 2. The measurement model development is firstly explained and described. A comparison of end user developers using their own applications with end users using applications developed by other end users, based on a number of key constructs, is then provided. The chapter finishes with the results of the structural model evaluations including the results of the hypothesis tests.

Chapter 10 discusses the results of Study 2 with reference also to the results of Study 1. The findings from both studies are compared and conclusions drawn. The limitations of the research are also discussed and suggestions for future research are made.

The final chapter draws on the findings of the two studies to discuss the theoretical implications of the findings in relation to the research problem. The implications for organisations that rely on user development of applications are also discussed. This
thesis describes a major step towards the development of a more comprehensive and integrated model of UDA success.

1.7 Definition of key terms

To ensure that the terminology used in this thesis is clear, this section includes definitions and descriptions of the key variables and terms that are used throughout the thesis.

Application system - a system developed with the purpose of solving a problem or making decisions.

End user - a person who uses computer resources to perform job related tasks.

End user computing (EUC) - is defined as, 'the use and/or development of information systems (IS) by the principal users of the systems output' (Wetherbe & Leitheiser, 1985).

End user developer – is an end user who develops applications to support his or her work or the work of other end users.

Experience –is defined as the duration of previous use of a software development tool.

Individual impact – individual impact refers to the effect of an information system on the behaviour of the user.
**Information quality** – relates to the characteristics of the information that an information system produces. It includes issues such as timeliness, accuracy, relevance and format.

**Involvement** - is defined as ‘a subjective psychological state, reflecting the importance and personal relevance of a system to the user’ (Barki & Hartwick, 1989 p.53)

**Organisational impact** – refers to the effect of an information system on organisational performance.

**Perceived individual impact** – is defined as an end user’s perception of the impact of an information system on his or her own behaviour (see individual impact above).

**Perceived system quality** – is defined as an end user’s perception of the system quality of an information system (see system quality below).

**Spreadsheet development knowledge** – is defined as the knowledge of spreadsheet software features and spreadsheet development practices that end user developers draw upon when developing spreadsheet applications.

**Structural equation modelling (SEM)** – a multivariate data analysis technique used to estimate a series of interrelated dependence relationships simultaneously.

**System quality** – relates to the quality of an information system (as opposed to the quality of the information it produces). It is concerned with issues such as reliability, maintainability, ease of use etc.
**Training** – is defined as the level of previous training with a software development tool.

**Use** – refers to how much an information system is used.

**User developed application (UDA)** - any application system developed by an end user.

**User developed application quality** – refers to the system quality of a user developed application.

**User developed application success** – consistent with DeLone and McLean’s (1992) definition of information systems success, UDA success is considered a multidimensional construct encompassing the outcomes of user development of applications.

**User development of applications** – the use of information technology by personnel outside the IS department to develop software applications in support of organisational tasks

**User satisfaction** – relates to the attitude or response of an end user towards an information system. It is defined as ‘the affective attitude towards a particular computer application by an end user who interacts with the application directly’ (Doll & Torkzadeh, 1988).
Chapter 2
Literature Review

2.1 Introduction

Chapter 1 provided a background to the research described in this thesis and outlined the specific research problem to be addressed. The purpose of this chapter is to review the current state of research in the area of EUC focussing on the specialised area of end user development. There is a reasonably large body of literature that addresses different aspects of EUC and end user development, however much of the early research was exploratory and prescriptive in nature (Brancheau & Brown, 1993), and published research on EUC has declined since the mid-1990s (Powell & Moore, 2002). Shayo et al. (1999) noted that there was still little literature on the measurement of end user development success in 1999, and the literature review for this thesis has confirmed this.

This chapter concentrates on the theoretical and empirical literature pertinent to the research questions posed in this thesis, drawing from the organisational IS literature as well as the EUC literature. The first section of this chapter identifies the literature relevant to the definition of user development of applications and provides the definition that is used in this thesis. This section also provides a review of the benefits and risks of user development of applications, highlighting the opportunities that can be realised from encouraging it, but noting the problems that can arise if end user development is not managed well. The literature on the characteristics of end user developers, the tasks for which end users develop applications, the tools used for
application development and the processes used for application development is then reviewed. This provides the background necessary for understanding the participants in end user development and the environment in which it takes place, and provides a foundation for planning the two studies reported on in this thesis.

The next section reviews the literature on the management of EUC, discussing both the approaches to EUC that have been espoused and the empirical research that has attempted to validate these approaches, as well as other studies characterising management of EUC in organisations.

The final section reviews the literature on the measurement of IS and UDA success. It considers the various models of IS success that have been proposed and highlights the lack of previous research on UDA success. It thus draws attention to the need for further research on UDA success.

2.2 What is end user development of applications?

2.2.1 End user development of applications defined

Although there is substantial consensus in the literature about the factors that led to the growth of EUC, there has been less consensus about its definition. Definition of the phenomenon has changed over time as a result of the use of new technologies and the changing needs of end users. Some definitions of EUC are very broad, encompassing a wide range of use and development activities. For example, Wetherbe and Leitheiser (1985) define EUC as ‘the use and/or development of information systems by the principal users of the systems' output or their staff’, and Rainer and Harrison (1993)
define it as ‘individual use of computers encompassing all the computer-related activities required or necessary to accomplish one’s job’. These definitions would include activities such as the use of word processing software to produce a business letter, and looking up of customer details in a transaction processing system as well as the development of a spreadsheet to support decision making.

On the other hand, some definitions are very narrow in terms of the activities they describe. For example, Brancheau and Brown (1993 p. 439) define EUC as ‘the adoption and use of information technology by personnel outside the IS department to develop software applications in support of organisational tasks’, acknowledging only the development aspects of EUC. Of the examples of possible end user activities provided above, only the development of a spreadsheet to support decision making would be considered EUC under this narrower definition.

Rivard and Huff (1984) were among the first researchers to differentiate between EUC in general and the more specialised area of user development of applications. They wrote that ‘User development of applications should be clearly distinguished from the much broader set of activities termed end user computing (EUC). The latter includes many types of computing activities outside the scope of development and use of computer applications by end users’ (Rivard & Huff, 1985 p.89). Thus, they considered end user development of applications to be a subset of EUC.

In this thesis EUC is defined as the use and/or development of IS by the principal users of the system’s output. This definition explicitly acknowledges that development and use are two distinct activities within EUC. It was adopted so that when literature
referring to EUC was being discussed, it would be clear that a broad range of end user activities might be included.

The primary focus of this thesis is, however, user development of applications not EUC. In this thesis user development of applications is defined as the use of information technology by personnel outside the IS department to develop software applications in support of organisational tasks, and an end user developer is defined as someone who develops applications to support his or her work or the work of other end users. Thus, an end user who develops a computer application is an end user developer. An end user who uses an application to undertake job related tasks is participating in EUC but is not considered an end user developer. So, while the technical abilities of user developers may vary considerably, their basic tasks are to analyse, design and implement computer applications (Sipior & Sanders, 1989).

A UDA is thus defined in this thesis as any application system developed by an end user. The development of a UDA is a direct response to a particular organisational task or duty, undertaken by staff involved in that task or duty. UDAs can support a very wide range of organisational tasks and vary greatly in size and complexity (McLean, Kappelman, & Thompson, 1993; Shayo et al., 1999).

2.2.2 Benefits and risks of user development of applications

As mentioned in Chapter 1, a number of possible benefits of user development of applications to individual end users and to organisations have been proposed. Table 2.1 lists the potential benefits from the perspective of each of the possible beneficiaries. This table is adapted from tables in Amoroso and Cheney (1992) and Kreie (1998).
Many of these potential benefits are of course interrelated, for example, user development of applications should allow the IS staff to focus more on the remaining, presumably larger, requests and hence to reduce the application development backlog. This, in turn, improves relationships between IS staff and end users.

Table 2.1: Potential benefits of user development of applications

<table>
<thead>
<tr>
<th>Benefits to end users</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Improved decision making effectiveness (Amoroso &amp; Cheney, 1992)</td>
</tr>
<tr>
<td>• Improved productivity (Amoroso &amp; Cheney, 1992; Davis, 1988)</td>
</tr>
<tr>
<td>• Improved user computer literacy (Amoroso &amp; Cheney, 1992)</td>
</tr>
<tr>
<td>• Increased satisfaction (Amoroso &amp; Cheney, 1992)</td>
</tr>
<tr>
<td>• Faster response to information requests (Davis, 1988)</td>
</tr>
<tr>
<td>• Improved relationships with IS staff (Benson, 1983)</td>
</tr>
<tr>
<td>• Encourages experimentation and innovation (Davis, 1988)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Benefits to IS staff and IS department</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A reduction in the backlog of IS development projects (Rivard &amp; Huff, 1984)</td>
</tr>
<tr>
<td>• A decreased proportion of IS resources spent on application maintenance and programming (Rivard &amp; Huff, 1984)</td>
</tr>
<tr>
<td>• Improved programmer job satisfaction (Benson, 1983)</td>
</tr>
<tr>
<td>• Better use of limited resources (Benson, 1983)</td>
</tr>
<tr>
<td>• Improved relationships with users (Benson, 1983)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Benefits to management</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Fewer user/IS conflicts (Davis, 1988; Rivard &amp; Huff, 1985)</td>
</tr>
<tr>
<td>• More satisfied end users and IS staff (Amoroso &amp; Cheney, 1992)</td>
</tr>
<tr>
<td>• Direct control over departmental information and applications (Davis, 1988)</td>
</tr>
<tr>
<td>• Increased end user productivity (Davis, 1988)</td>
</tr>
</tbody>
</table>

The benefits in Table 2.1 are however only proposed benefits. Section 2.7.2 addresses the attempts to measure the success of UDAs and discusses research that provides support for some of these claims.
As mentioned in Chapter 1, user development of applications also brings with it many risks to the organisation. Table 2.2 below presents some of the risks that have been identified in the literature.

Table 2.2: Potential risks of user development of applications

- Incompatible end user tools preventing sharing of applications and information (Alavi & Weiss, 1985-1986)
- Solving the wrong problem (Alavi & Weiss, 1985-1986)
- Inability to identify correct and complete information requirements (Davis, 1988)
- Mismatch between tools and applications (Alavi & Weiss, 1985-1986; Davis, 1988; O'Donnell & March, 1987)
- Redundant development effort (Alavi & Weiss, 1985-1986)
- Lack of documentation for applications (Benson, 1983)
- Use of private systems when organisational systems would be more appropriate (Brancheau & Brown, 1993)
- Inefficient use of personnel time (Alavi & Weiss, 1985-1986; Davis & Srinivasan, 1988; O'Donnell & March, 1987)
- Ineffective use of monetary resources (Alavi & Weiss, 1985-1986; Davis & Srinivasan, 1988)
- Unreliable systems (Brancheau & Brown, 1993).
- Failure to backup data (Benson, 1983)
- Lack of data security (Benson, 1983)

In her review of research on end user development of applications, Lally (1995) found reports that many UDAs are never developed to completion (Klepper & Sumner, 1990), are error-prone (Edberg & Bowman, 1996), and are never used for the purpose for which they are developed. Even UDAs that are successfully adopted by their developers may be duplications of other systems already available in the organisation and therefore can be seen as a waste of resources, or may have been developed by highly paid employees when they could have been developed by IS professionals faster and at cheaper hourly rates.
The risks to the success of user development of applications can occur at various stages throughout the application lifecycle. For example, lack of application documentation may cause particular risks when maintenance is attempted, and inefficient expenditure of staff time may be a particular problem during development but not necessarily once the application is in use. However, some problems such as tool/application mismatch will cause problems during all stages of the application lifecycle.

In summary, the proposed benefits of user development of applications have been seen to flow from a belief that the user has a superior understanding of the problem to be solved by the application, and the proposed risks from the belief that users have less understanding of the process of system development than do IS professionals.

2.3 Characteristics of user developed applications

2.3.1 Types of tasks that users develop applications for

A task is a particular activity that needs to be done or a problem that needs to be solved. In user development of applications, an application is developed to help the end user accomplish a task. Users develop applications to cover a wide range of tasks. In a survey to determine the types of applications developed by end users, Rittenberg and Senn (1990) identified over 130 different types of applications. The majority of these (56%) were accounting related but marketing, operations and human resources applications were also heavily represented. The accounting applications most frequently implemented were budget analysis, financial analysis, financial performance forecasting, and “what-if” analysis.
Early studies tended to find that UDAs were typically used for queries, reports and simple analyses (e.g. Rivard & Huff, 1985; Rockart & Flannery, 1983). However, the range of tasks has expanded as the sophistication of both software development tools and user developers has increased. This has led to a degree of convergence with corporate computing, so that the tasks for which UDAs are developed are less distinguishable from tasks for corporate computing applications (McLean et al., 1993). In addition to the traditional tasks that UDAs have been developed to support, Web applications are becoming increasingly common (Nelson & Todd, 1999; Ouellette, 1999).

As well as the types of tasks applications are used for, dimensions also considered in studies attempting to characterise UDAs include the scope and size of applications, the source of data used in applications, and the frequency of use.

### 2.3.2 Application scope and size

Application scope refers to whether an application’s major use is by individuals, single departments or by multiple departments. Several early studies found many UDAs to be broad in scope. In a 1983 study more than half the applications surveyed were relevant to the operations of an entire department and 17 percent were multi-department applications (Rockart & Flannery, 1983). Similarly Sumner and Klepper (1987) found the majority of the systems in their study to be departmental in nature. However, Bergeron and Berube (1988) found that although 44% of their sample had developed applications that were used by more than two people, only 14.6% considered their applications to be useful department wide (i.e. departmental applications). In a more
recent study of spreadsheet development, Hall (1996) found that only 17% of the spreadsheets contributed by participants were solely for the developer's own use, and the output of the remainder was distributed to others. Twenty nine percent of the total sample were distributed beyond the developer's own organisation.

Little has been written about the size of UDAs and comparison between types of applications is difficult, as there is no agreed basis for comparison. In her study of 106 user developed spreadsheet applications from a wide range of organisations, Hall (1996) found most applications were ‘large’ (ranging from 800 bytes to 5.3Mb); 45% included macros, and 36% were linked to other spreadsheets. However, Taylor, Moynihan and Wood-Harper (1998) undertook case studies in 34 UK organisations and concluded that UDAs were not usually ‘large’, rarely taking more than 1 person-week development time.

The results of these studies on scope and size may not, however, reflect the wider population of UDAs as either convenience samples were used or participants were asked to contribute a UDA of their choice. In fact, it is likely that when researchers sought UDAs they were shown the larger, high profile applications. As end users often develop applications without organisational knowledge (Shayo et al., 1999), those brought to the attention of researchers by departmental managers would most likely be highly visible applications. Hall (1996) also recognised that there had been developer bias in the selection of spreadsheets chosen to respond to her survey, in which user developers were contacted directly.
2.3.3 **Sources of data and interdependence of applications**

Klepper and Sumner (1990) found that many UDAs required rekeying of data, but that 31% took input from other systems and 16% created output for other systems, thus highlighting the interdependence among UDAs. Chan and Storey (1996) reported that 65% of their spreadsheet users said that manual input was their most common source of data, 21% most commonly downloaded data from a server, and 10% usually copied from a floppy disk, suggesting that the data was obtained from other UDAs. In Hall’s (1996) study of spreadsheets the majority of applications involved corporate rather than purely local data.

2.3.4 **Frequency of use**

Many end users believe that the applications they develop will be ‘one off’ (i.e. used for only one finite task) and hence they may not put much effort into the way they develop and how they document their application. However even these ‘one-off’ applications tend to be used more than just once (Cragg & King, 1993; Kroenke, 1992). Hall (1996) found that most of the spreadsheets in her sample (67%) were used on a regular basis (daily, weekly, monthly or frequently), and a much smaller portion (17%) were used once or only a few times.

2.4 **Tools used for end user development**

Application development by end users did not really become practical until 4th generation tools became available (Lally, 1995). Fourth generation tools or languages (4GL) are designed to allow users to focus on what they want done, rather than on how it should be done. This means that end user developers do not require such detailed
knowledge of software development or of how hardware functions. It also means that development can be done much faster than was previously possible.

Although a wide range of tools is available for use by end user developers, the most commonly used software tools have been spreadsheets followed by databases and 4th generation query languages (Rittenberg et al., 1990). Amoroso and Cheney (1991) found that 50% to 70% of the 1100 applications developed by participants in their study were spreadsheets while Raymond and Bergeron (1992) found the figure to be about 78%. The majority of the 34 organisations participating in Taylor, Moynihan and Wood-Harper’s (1998) study (88%) also used spreadsheets for end user development whereas only 35% used query languages and 12% used databases.

The figures for spreadsheet use are probably higher than they should be. In their study of 256 business analysts Chan and Storey (1996) found that once users have mastered a tool they can be reluctant to adopt other software packages that may be more suitable for their applications. Seeley and Targett (1997) attribute this to the large initial investment end users make in learning to use spreadsheets. They are therefore reluctant to invest more time in learning to use another tool if it is possible to do what they want with a spreadsheet regardless of whether or not it is the most efficient way to accomplish the task.

With the increasing popularity of Web application development by end users, Website development tools are being used more often by end user developers (O’Brien, 2002; Ouellette, 1999). However, to date the extent of their use by end user developers has not been reported in the literature.
2.5 End user developer characteristics

There have been many studies identifying and evaluating end user characteristics, although research relating specifically to end user developers is more limited. The end user developer characteristics that have been studied include age, gender, computer experience and training. However, most studies of developer characteristics simply correlate individual characteristics with computer attitudes or user satisfaction and don't address the impact of characteristics on work outcomes.

2.5.1 Age and gender

The research on age and gender has suggested that older end users and women are less confident and less satisfied with their EUC experiences, and possibly have less skill with a computer. For example, Harrison and Rainer (1992) surveyed 776 knowledge workers from a university and found that male gender and younger age were associated with higher computer skill. Similarly, in a study of 104 end users in six large companies Igbaria and Nachman (1990) found that user age was negatively related to end user satisfaction. While studying the impacts of different spreadsheet design approaches, Janvrin and Morrison (2000) found that females were less confident that their applications were error free and felt that they had less expertise. However, this lack of confidence was not reflected in the number of errors they made, and thus may only reflect their self-confidence rather than their skill as application developers.

Any differences due to age and gender may however be declining, as the use of information technology becomes an integral part of working life. Munro et al. (1997)
found no relationship between age and EUC competence, and obtained conflicting results with respect to gender. As in the Janvrin and Morrison study mentioned above, women reported lower levels of EUC knowledge than men, however there was no difference in their scores in software knowledge tests.

2.5.2 Experience and training

Previous experience and previous training have been hypothesised to play a role in determining the impact of user development of applications on individual productivity. A number of studies have shown positive associations between experience and outcomes, and between training and outcomes. In the Harrison and Rainer (1992) study mentioned above, greater experience with computers was associated with higher computer skill. Similarly, Kasper and Cerveny (1985) found that greater end user experience increased the likelihood of successful decision making using UDAs. Nelson and Cheney (1987) also found that more training was associated with greater computer-related ability, and the preliminary results of Babbitt, Galletta and Lopes’s (1998) study of spreadsheet development by novice users suggested that end users whose training emphasises planning and testing of spreadsheets develop better quality spreadsheets.

However, several studies have had conflicting results. For example, Al-Shawaf (1993) found no relationship between development experience and user satisfaction, although Crawford (1986) found that greater user developer experience was associated with higher levels of satisfaction. In Crawford’s study, despite the positive association between experience and satisfaction, higher levels of education and training were generally associated with lower levels of satisfaction. In addition, Chan and Storey (1996) found no relationship between computer training and spreadsheet proficiency.
Many of these early studies used subjective outcome measures. A possible explanation for the conflicting results reported in the literature is that the more end users learn, the more they may recognise the flaws in their applications. This might lead to less satisfaction with applications, counteracting the anticipated increases in satisfaction due to improvements in quality. This would be consistent with Yaverbaum and Nosek’s (1992) speculation that computer training increases one’s expectations of IS, and hence may actually cause negative perceptions. It also highlights the risk that end user developers with little experience or training may be unable to recognise the flaws in their applications and may be less able to judge whether the applications they develop will have a positive impact on their performance of tasks.

Research on experience and training is also complicated by the variety of means used to measure these variables. Experience has been measured variously as frequency of use (Blili, Raymond, & Rivard, 1996) or years of use (Harrison & Rainer, 1992; Yaverbaum, 1988), with either computers in general (Harrison & Rainer, 1992; Yaverbaum, 1988) or a particular application package (Agboola, 1998; Panko, 1998a). Rarely have these measures been tied more closely to the actual application development experience. However, Amoroso (1986) developed a self-report measure that attempted to be more inclusive by asking end user developers to rate their experience with each of a number of areas of system development.

Measurement of previous training has also been problematic. Bowman (1988) recognised that the systems development expertise of an end user developer is a function of two dimensions: firstly, expertise in application development methods, and
secondly, expertise with the development tool or language used to develop the
application. However very few of the studies that have measured training as an
independent variable have distinguished between training leading to skills with the
development tool and training leading to skills in analysis and design. This is despite
the fact that a number of authors have commented on end user developers’ lack of
analysis and design knowledge (Benham, Delaney, & Luzi, 1993; Rivard & Huff,
1988).

It appears that end user developers generally receive little training (Taylor et al., 1998)
and that the major means of training is self-study (Benham et al., 1993; Chan & Storey,
1996). When organisations do provide training it rarely covers application development
methodologies (Jenson, 1993). It has also been suggested that when end users are self-
taught the emphasis is predominantly on how to use the software rather than broader
analysis and design considerations (Benham et al., 1993). There are many books that
teach introductory spreadsheet or database skills typically giving a detailed, step-by-
step coverage of examples that illustrate the main product features. However, the very
proliferation of these features in recent software versions means that, increasingly, the
fundamentals of ‘what’ the end user is attempting to do are being obscured by the
multiplicity of ways ‘how’ to achieve it. Examples are presented as solutions to
problems without the design stages being made explicit. Thus, end users may have a
narrow knowledge focused on software development tool features but lacking in
techniques for developing applications that are user-friendly, reliable, and maintainable.

Given the uncertainty about how experience and training influence the success of user
development of applications further research is needed in this area.
2.5.3 Involvement

Another characteristic of end user developers that may affect the success of their applications is the level of involvement they have with the development process and the resulting applications. Barki and Hartwick (1989 p.53) defined user involvement as ‘a subjective psychological state, reflecting the importance and personal relevance of a system to the user’. They noted that users may also attach personal importance and relevance to the activities of systems development, thereby becoming involved not just with a system but with a process. The literature on user involvement with organisational systems indicates that increased involvement is associated with increased user satisfaction (Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Doll & Torkzadeh, 1988; Lawrence & Low, 1993). The emphasis on achieving involvement through user participation is hence an integral part of participative design methods (Cavaye, 1995).

In the UDA domain, Doll and Torkzadeh (1989) found that end user developers had much higher involvement levels than users who were involved in the development process but where the application was primarily developed by a systems analyst or another end user. End user developers were also found to be more satisfied with applications they had developed themselves than with applications developed by another end user (McGill, Hobbs, Chan, & Khoo, 1998), or with applications developed by a systems analyst (despite involvement in the systems development process) (Doll & Torkzadeh, 1989). However, the relationship between involvement and the impact of UDAs on end user performance has not been clarified.
2.6 Processes used by end users in development of applications

Although many methodologies have been advocated for end user developers to follow, little follow up research has been done identifying what development processes end user developers actually follow and whether any of the methodologies proposed in the literature are adopted and used successfully.

The system development methodologies that have been proposed tend to be based on either the traditional system development lifecycle (SDLC) or a prototyping approach. For example, Salchenberger (1993) proposed a comprehensive structured design methodology intended for use with a wide range of applications and tools, and Ronen, Palley and Lucas (1989) proposed a structured approach for the development of spreadsheets that was based upon modifying data flow diagram symbols to create Spreadsheet Flow Diagrams. Pliskin and Shoval (1987) advocated end user prototyping for sophisticated end user developers and Kroenke (1992) provided a very detailed prototyping methodology for end user development.

None of the above methodologies has been empirically tested, although Salchenberger (1993), and Pliskin and Shoval (1987) provided several case studies of successful use. Alavi, Phillips, and Freedman (1990) and Janvrin and Morrison (2000) conducted experiments using students as surrogate end users to test the success of structured development methodologies. In both studies, use of a structured methodology had positive outcomes in terms of the quality of applications developed.
Studies that have not been tied to specific methodologies seem to indicate that development activities for end user developed applications tend to be very unstructured. For example, in an experimental study of spreadsheet creation, Brown and Gould (1987) observed that their participants spent little time planning before launching into coding. Sumner and Klepper (1987) also found that formal systems development practices such as data validation, documentation, and data security were not being followed. Similarly, Cragg and King (1993) found that spreadsheets were usually built in an informal, iterative manner. This still seems to be the case, as in a study of the effects of systems analysis and design training on spreadsheet quality Kreie (1998) found that only 11% of her subjects undertook any design before actually starting to build the spreadsheet. This lack of formal analysis and design also appears to be condoned at an organisational level. In case study research involving 34 UK organisations, Taylor, Moynihan, and Wood-Harper (1998) found that none of the organisations studied had adopted a formal system development methodology for end user developed systems.

### 2.7 Management of user development of applications

Brancheau and Brown (1993 p.439) defined the management of user development of applications as ‘planning, organising, staffing, directing, controlling, supporting, and coordinating the adoption and use of information technology by end users to develop software applications in support of organisational tasks’. The potential benefits and risks of end user development have prompted a number of researchers to examine management approaches to it. Successful management of end user development should reduce the risks and help realise the benefits that can arise from it.
Organisations manage user development of applications as part of their overall management of EUC (Martin, Dehayes, Hoffer, & Perkins, 1991), therefore the discussion below includes both management of user development of applications and more general management of EUC. Although EUC overlaps with organisational computing it can be distinguished from it by reporting relationships within the organisation (McLean et al., 1993). EUC is generally outside the direct control of the IS department and is of concern to functional managers as well as IS management.

A number of approaches have been suggested in the literature for how management of EUC should be undertaken. However, there has been little consensus among academics or practitioners on what constitutes the effective management of EUC (Brancheau & Brown, 1993). The following discussion gives an overview of some of the main strategies that have been proposed and a review of any research that has been done to validate them. The approaches discussed are: the information centre approach; the managed free economy approach (Gerrity & Rockart, 1986); Alavi, Nelson and Weiss’s (1987-1988) EUC management framework; the expansion/control approach (Munro, Huff, & Moore, 1987-88); and Brown and Bostrom’s (1989) model of EUC management effectiveness. Other research on EUC management that is not tied to a specific model is then briefly discussed.

### 2.7.1 Information centre approach

Information centres (ICs) are organisational units designed to facilitate and coordinate end user development activities. Typical services offered by an IC include: training, consulting, technical and operations support, hotline assistance, management of data,
software evaluation, debugging assistance, and documentation support for user applications.

The IC approach was very popular in the 1980s (Brancheau & Brown, 1993). However, although ICs were an improvement on the way EUC was previously being managed in many organisations, the traditional implementation of ICs had a number of shortcomings (Brancheau & Brown, 1993). These included: lack of localised support; lack of functional expertise within the IC; and lack of provision of strategic initiatives.

Much of the research on ICs has focused on listing and categorising the services offered. However, in an empirical investigation of 16 organisations, Saaksjarvi, Heikkila and Saarinen (1988) found no clear connection between the existence of a formal IC and the success of EUC. They therefore concluded with the recommendation that IS managers should critically re-evaluate their EUC strategies as an IC alone does not guarantee success.

Guimaraes and Igbaria (1994) surveyed 252 internal auditing directors to investigate the relationship between IC effectiveness and impact on the organisation. The study provided strong evidence that there is a direct relationship between IC effectiveness or performance level, and two major dependent variables: (1) the benefits that the organisation has derived from EUC activities and (2) the organisation's business performance. However, they found that organisations on average have only had 'a little improvement' from EUC.
A study by Vijayaraman and Ramarkrishna (1990) compared successful and unsuccessful ICs and found that as the end users who were being supported became more advanced they tended to be less satisfied with the support they received. The researchers also found that successful ICs had a higher budget per end user and fewer end users per organisational staff member than other, less successful ICs.

2.7.2 Managed free economy approach

Gerrity and Rockart (1986) identified four EUC management strategies: monopolist, laissez-faire, information centre and managed free economy. They recommended the managed free economy approach which advocates the development of a partnership between the IS department and users in exploiting information technology for the benefit of the organisation. In this approach, standards and policies are established to control certain aspects of EUC that have an organisational impact. If end users work within these standards and policies, they are free to do what they want.

Whilst Gerrity and Rockart offered comprehensive guidance for managers of EUC, they did not indicate how the approach was developed or what the basis for it was, and provided no case studies of successful use. The approach also does not appear to have been subsequently subjected to any empirical testing.

2.7.3 Alavi, Nelson and Weiss’s EUC management framework

Alavi, Nelson and Weiss (1987-1988) identified five EUC management strategies: the laissez-faire strategy, the monopolist strategy, the acceleration strategy, the marketing strategy and the operations-based strategy. These strategies were identified based on
prior literature and interviews conducted with five companies they believed to be successful in managing EUC. The proposed strategies overlap substantially with those described by Gerrity and Rockart (1986) (see above). Alavi et al. viewed EUC management as evolutionary, with organisations using different strategies over time to facilitate movement along an S-shaped EUC growth curve. Marketing and operations-based strategies were perceived as being more mature approaches. The validity of this model was investigated in a case study by Brown and Wynne (1990), and received some support. However, the results of the study were exploratory as only one case study was used, and the measurement instruments were not validated.

2.7.4 Expansion/control approach

Munro, Huff and Moore (1987-88) also provided an evolutionary model. They mapped EUC strategies on to a two-dimensional grid (see Figure 2.1 below). Their model suggests that EUC can be managed by varying the rate of EUC expansion and the degree of control exercised over EUC, and implies that effective management of EUC occurs when organisations match their implementation of control and expansion tactics with their strategic choice for the expansion.

Munro, Huff and Moore believed that all firms start with both low expansion and low control (laissez-faire approach). From this starting point they believed that organisations tend to move either clockwise or counter-clockwise, either becoming expansionist with little control or following a containment strategy with strong control and little acceleration. They also identified 19 expansion and 5 control tactics being used by their sample of 37 organisations to achieve a desired degree of EUC growth. The firms in the sample exhibited significant variation in their implementation tactics.
Saarinen, Heikkila and Saaksjarvi (1988) undertook case studies of four organisations to determine if the four strategies identified by Munro et al. (1987-88) were actually alternative growth strategies or only different stages of EUC evolution. Two of the organisations employed an acceleration strategy and the other two a control strategy. In the two using the control approach, EUC had grown slowly and needed more resources allocated to it. The other firms had a much higher level of EUC growth. In an empirical investigation of 16 organisations Saaksjarvi, Heikkila and Saarinen (1988) also found that accelerating strategies resulted in significantly better success than controlling strategies.

### 2.7.5 Brown and Bostrom’s model of EUC management effectiveness
Brown and Bostrom’s model of EUC management effectiveness (1989) proposes that the most effective management structure for EUC is contingent on the organisation’s current EUC implementation phase and its current EUC growth objective. The model uses three dimensions to define appropriate management structures: centralisation, formalisation and complexity. The model defines two management structures that are appropriate for different EUC conditions (see Figure 2.2). Organic structures are characterised as having low centralisation, low formalisation and high complexity. Mechanistic structures have high centralisation, high formalisation and low complexity. The appropriate infrastructure is determined by the organisation’s objective for EUC growth and the position of the organisation on the EUC learning curve.

Figure 2.2: Brown and Bostrom’s (1989) model of EUC management effectiveness

The validity of this model was investigated in a case study by Brown and Wynne (1990). They found that while this model and both the expansion/control approach
(Munro et al., 1987-88) and Alavi, Nelson and Weiss’s (1987-1988) EUC management framework were useful in explaining the ratings of EUC management effectiveness obtained in their case study, the two contingent factors of the Brown and Bostrom model (EUC implementation phase and EUC growth objective) seemed to enhance the predictions of the models. However, as mentioned above, the results of this study were only exploratory.

2.7.6 Other research on EUC management

The previously discussed studies all related to specific models of EUC management. There have also been a number of studies that were not tied to specific models. Bergeron and Berube (1988) conducted an empirical study to investigate the adequacy of EUC management and support practices in 31 corporations. They found no evidence that having an EUC management plan improved user satisfaction. In fact, there was a negative correlation between the number of policies and user satisfaction. They suggested that although establishment of policies is a good way to assure efficient operations, it is wise to be careful about the number of policies formulated, as users tend to become dissatisfied if their freedom is restricted.

Bowman (1988) found that UDAs developed under situations of greater control have better documentation, testing and backup but, like Bergeron and Berube (1988), he found that users are less satisfied with the development process. In contrast, in a study of 97 management accountants Jenson (1993) found no support for the hypothesis that more highly controlled environments lead to better quality applications.
Crawford (1986) found that organisational support policies and procedures were more influential in determining user development success than were control policies and procedures. User development training and consulting support were most influential, along with the ability to access quality data.

Bowman (1988) found low compliance with standards and procedures and concluded that it is inappropriate to insist on a single, inflexible development process control environment. In a more recent study, Taylor et al. (1998) also argued that the bureaucracy associated with formal methodologies forms too great an overhead, and made several suggestions that would enable end users to adopt ‘cut down’ versions of system development methodologies that are appropriate for the scale of particular projects.

Nelson and Todd (1999) investigated what strategies organisations were using to manage EUC on the Web. They noted that most firms in their study appeared to be relying on a monopolistic control strategy (as described by Gerrity and Rockart (1986) and Alavi et al. (1987-1988)). They then concluded that while such a strategy may be the best approach given the relative infancy of Web technology, it could prove to be an unstable strategy in the future.

As can be seen from the discussion above, the majority of the published research on EUC management is relatively old, and more recent research such as that of Nelson and Todd (1999) draws heavily upon early models of EUC management that have received little validation. Research into EUC management has not yet provided answers to concerns about the risks associated with user development of applications, thus
organisations continue to rely heavily on the judgements of individual end user
developers as to the suitability of applications for use. The ability of end users to judge
the success of applications they develop clearly requires further research.

2.8 IS and UDA success

2.8.1 IS success

The literature on organisational IS effectiveness has proposed numerous measures of
success. These include user satisfaction (e.g. Nelson & Cheney, 1987), decision making
performance (e.g. Fuerst & Cheney, 1982), frequency of use (e.g. Srinivasan, 1985) and
cost-benefit analysis (e.g. Dickson, Leitheiser, Wetherbe, & Nechis, 1984). DeLone and
McLean (1992) provide examples of many more measures.

The most commonly used measure of IS success is user satisfaction (Gelderman, 1998;
Melone, 1990). However, user satisfaction has been defined in various ways. In a
review of the research on user satisfaction, Kim (1989) recognised that this construct
has been considered from three different perspectives: (1) user satisfaction in terms of
attitudes towards IS related ‘objects’ (including IS staff, organisational support and
information attributes) (e.g. Bailey & Pearson, 1983; Ives, Olson, & Baroudi, 1983), (2)
user satisfaction in terms of information quality (e.g. Larcker & Lessig, 1980) and (3)
user satisfaction in terms of IS effectiveness (Hamilton & Chervany, 1981). Doll and
Torkzadeh (1988) noted the change in usage from traditional to EUC environments and
developed an instrument to measure the satisfaction of users who directly interact with
the computer for a specific application. End user computing satisfaction (EUCS) was
defined as ‘the affective attitude towards a particular computer application by an end user who interacts with the application directly’ (Doll & Torkzadeh, 1988 p.261).

One of the reasons that user satisfaction is commonly used to assess success is that it is a more convenient measure than performance related measures. However, there is also an implicit assumption that user satisfaction with an IS results in some positive change in user behaviour resulting in increased effectiveness (Davis & Srinivasan, 1988). This, however, has not been conclusively demonstrated.

The use of such a variety of IS success measures has been problematic. Firstly, the value of some measures is doubtful (Melone, 1990; Trice & Treacy, 1988) and secondly, comparison between studies has been difficult due to variations in measures used (Amoroso, 1991). There have been several attempts to provide a conceptual framework to support research on IS success. DeLone and McLean (1992) conducted an extensive review of the IS success literature and concluded that IS success is a multidimensional construct and that it should be measured as such. From their literature review, they identified the following constructs: system quality, information quality, use, user satisfaction, individual impact and organisational impact. They proposed the following model of IS success (Figure 2.3) which shows the interdependencies of the categories of success measures and indicates the serial, temporal dimension of information flow and impact.
DeLone and McLean’s work makes two important contributions to the understanding of IS success. First, it provides a scheme for categorising the multitude of IS success measures that have been used in the literature. Second, it suggests a model of temporal and causal interdependencies between the categories. The model has received much attention amongst IS researchers (Walstrom & Hardgrave, 1996; Walstrom & Leonard, 2000), and attempts to validate parts of it with respect to organisational IS have had some success (e.g. Roldán & Millán, 2000; Seddon & Kiew, 1996).

Seddon and Kiew (1994) suggested that user involvement is a fundamental factor which should be present in a model of IS success, and Seddon then later went on to propose a respecification and extension to DeLone and McLean’s model. Seddon (1997b) argued that, rather than a single sequence of relationships, there were two linked IS success sub-systems: one that explained use, and another that explained impact. He argued that use is not an indicator of IS success, but that user satisfaction is because it is associated with impact. There are as yet no published empirical tests of Seddon’s proposed respecification of the model.
The DeLone and McLean model was also analysed critically by Ballantine et al. (1998) who, like Seddon, proposed but did not test an alternative. The Ballantine model suggests that a three dimensional model of success may be more appropriate. Myers, Kappelman and Prybutok (1998) also presented a respecification of DeLone and McLean’s model. Their model incorporates the roles of both service quality and workgroup impact in addition to the original success constructs.

A different approach has been followed by Goodhue and his colleagues (Goodhue, 1988; Goodhue, 1995; Goodhue, Klein, & March, 2000; Goodhue & Thompson, 1995). Drawing on the job satisfaction literature, they proposed that an explanation of IS success needs to recognise both the task for which the technology is used and the fit between the task and the technology. They proposed the technology-to-performance chain shown in Figure 2.4 below.

![Figure 2.4: The technology-to-performance chain (Goodhue & Thompson, 1995)](image-url)
This model is consistent with DeLone and McLean’s model in that both use and user attitudes about the technology lead to individual performance impacts. However, Goodhue et al. claim that their model goes beyond DeLone and McLean’s model because it highlights the importance of task-technology fit in explaining how technology leads to performance impacts.

Behavioural intention models may also be useful in understanding UDA success. The most popular use model in recent IS literature, the technology acceptance model (Davis, 1986) has been used consistently to demonstrate that perceived usefulness of a system is associated with its use (Adams, Nelson, & Todd, 1992; Davis, 1989, 1993; Taylor & Todd, 1995) (see Figure 2.5 below). It makes intuitive sense to propose that perceived usefulness is associated with actual usefulness and therefore with the impact of an IS. Several richer use models have been developed from Ajzen and Fishbein's work on the social psychology of human behaviour (the theory of reasoned action (Fishbein & Ajzen, 1975) and the theory of planned behavior (Ajzen, 1991)). These models characterise use as a human behaviour influenced by beliefs about, and attitudes to, the outcomes of use, and usefulness as one of the desired outcomes associated with use. One such model, the planned behaviour in context (PBiC) model (Klobas & Clyde, 2000; Klobas & Morrison, 1999) has been used to demonstrate that users' attitudes to a range of individual impacts (outcomes), including but not limited to usefulness, influence their intention to use Internet-based ISs. Provided there is a relationship between the outcomes of use that are valued by individual users and the impact of systems on individuals and organisations, use models based on Ajzen and Fishbein's work may contribute to more satisfactory explanations of IS success.
2.8.2 UDA success

Edberg and Bowman (1996) noted that while the EUC literature had provided many prescriptions for managing EUC (as described in Section 2.6), there had been little direct empirical examination of the effectiveness of end users as application developers. Similarly, Shayo et al. (1999 p.8) noted that ‘while research continues to grow on IS success, it remains scanty on measurement of end user success’. They attributed this lack of research to the fact that UDAs are rarely tracked formally by organisations and that users may be reluctant to allow measurement of the efficiency or effectiveness of their applications. They suggested that benign measures of success, such as user satisfaction, are less threatening and easier to obtain than independent measures. However, this is problematic because users are asked to place a value on something about which they may be far from objective.

In their early review of the EUC literature, Brancheau and Brown (1993) found only seven empirical studies of outcomes of user development of applications. Examples of early studies include an exploratory study of UDA success by Rivard and Huff (1985) who found user development of applications to be associated with improved
productivity (assessed subjectively) and higher satisfaction. Similarly, Doll and Torkzadeh (1989) reported significantly higher satisfaction for users who did their own development compared with users who depended on others for development, and Cronan and Douglas (1990) reported that savings of around seven hours per week were claimed by end user developers. Alavi, Phillips and Freedman (1990) examined the impact of structured development methodologies and found that the use of a structured methodology had positive outcomes in terms of application quality, but negative outcomes in terms of user satisfaction with the development process.

More recently, Edberg and Bowman (1996) compared application development by end users and surrogate IS professionals and, not surprisingly, found that the surrogate IS professionals were much more productive and produced higher quality applications than did the end users. Rivard et al. (1997) also contributed to the research on the success of UDAs by developing an instrument to enable end users to assess the system quality of their applications. They also demonstrated a significant relationship between UDA quality and user satisfaction. McGill, Hobbs, Chan and Khoo (1998) compared the user satisfaction ratings of applications that were evaluated by their end user developers with ratings of the same application by other end users. End users were significantly more satisfied with applications they had developed themselves. The authors suggested that measures of user satisfaction commonly used to evaluate organisational systems might be inappropriate when end users assess applications they have developed themselves. However, they did not investigate the relationship between user satisfaction and performance.
Other recent research on user development of applications has focussed largely on the implications of design methodologies for system quality (e.g. Janvrin & Morrison, 2000; Kreie, Cronan, Pendley, & Renwick, 2000; Panko, 1998b; Taylor et al., 1998), rather than on impact at the individual or organisational level.

As the examples given above indicate, UDA success research has focussed primarily on user satisfaction as the outcome measure of success. Various researchers have justified this because of the extensive use of user satisfaction in the organisational IS literature (e.g. Rivard & Huff, 1988). However, it has been pointed out that problems exist with the use of satisfaction as a measure of success (e.g. Etezadi-Amoli & Farhoomand, 1996; Galletta & Lederer, 1989; Melone, 1990; Thong & Chee-Sing, 1996). As discussed earlier, Kim (1989) noted that the user satisfaction construct has been used to refer to a user’s perception of the system in some instances and an assessment of its output goals in others. Etezadi-Amoli and Farhoomand (1996) also noted that user satisfaction instruments rely on the affective/cognitive dimension of satisfaction without accounting for the performance-related dimensions. There has been little research attempting to link user satisfaction with UDAs directly with any measures of user behaviour such as improved productivity, fewer errors or better decision making.

Melone (1990) provides a critique of the theoretical issues involved in the user satisfaction construct, noting that many theoretical and practical issues remain to be resolved. User satisfaction is an attitude, and attitudes that users hold may play a role in establishing and maintaining self-esteem (Pratkanis & Greenwald, 1989). Melone (1990 p.85) cautions that ‘to the extent that attitudes are held to establish or maintain a positive image of the self, they are less likely to serve the evaluative function implied in
information system research’. Similarly, Hufnagel (1990) suggests that an individual’s affective response to a given performance outcome is often highly subjective and in some cases, highly ego-defensive.

The appropriateness of user satisfaction as a measure of system effectiveness may be even more questionable in the UDA domain. Users who assess their own computer applications may be less able to be objective than users who assess applications developed by others (McGill et al., 1998). The actual development of an application, which may involve a significant investment of time and creative energy, may be satisfying other needs beyond the immediate task. User satisfaction with a UDA could therefore reflect satisfaction with the (highly personal) development process as much as with the application itself.

As with organisational IS there is a need to bring together the disparate measures of UDA success and to look at the relationships between them. Caudle, Gorr, and Newcomer (1991) called for the development of valid and reliable measures to help in management of EUC. Bergeron, Rivard, and Raymond (1993) attempted to determine the actual importance, from an organisational perspective, of 30 EUC success criteria identified in the literature. However, they did not attempt to address the relationships between these potential success measures and their work has not formed a basis for further research.

Very little has been written about UDA success since the mid-1990s and Shayo et al. (1999) stressed the need for a more comprehensive and integrated model of UDA success. Models of organisational IS success such as DeLone and McLean’s (1992)
provide a starting point for studies on how UDAs can impact upon organisational performance and on the ability of end user developers to judge whether the applications they develop will have a positive impact on their productivity and the productivity of the organisation. The research undertaken for this thesis progresses this work.

2.9 Overview

This chapter reviewed the literature relating to the research questions posed in this thesis. The literature covers the development of applications by end users, the management of EUC, IS success and UDA success. The potential benefits of user development of applications were described, but the lack of empirical examination of the attainment of these benefits was noted. The potential risks of end user development were also highlighted and attempts to mitigate these via management of user development described. This review of the EUC management literature highlights the fact that there appears to be a consensus that increased organisational control over user development of applications results in dissatisfied end users which may result in a loss of potential benefits to the organisation. However, it suggested that management strategies that encourage end user development also increase risk.

Research characterising end users and the applications they develop was also described in the chapter. End users develop applications to support a wide range of organisational tasks. In general, they appear to receive little training from their organisations and are mostly self-taught. The literature on the roles of training and experience in application success was reviewed and it was noted that there has been some uncertainty about how
previous experience and training influence the success of user development of applications.

The review of models of IS success and UDA success highlighted the lack of recent research on UDA success and thus draws attention to the need for further work in this area. In particular, the issues raised in this literature review highlight the need for both research into the ability of end users to judge the likely impact of their applications, and further research on the role of UDA quality in individual and organisational outcomes. Chapter 3 follows from this review of the literature, and introduces and justifies the research questions asked in Study 1.
Chapter 3
Study 1 Research Questions and Conceptual Models

3.1 Introduction

The objective of the research described in this thesis is to gain a better understanding of UDA success, by focussing on the role of system quality in UDA success and the ability of end user developers to judge whether the applications they develop will have a positive impact on their performance of tasks. This research objective was addressed via two complementary studies. This chapter is the first of four chapters that describe Study 1 and it presents the research questions and conceptual models that form the basis for the study.

This chapter first presents the research questions for Study 1. Answering these research questions should provide insight into the ability of end user developers to judge application success. The DeLone and McLean (1992) model of IS success is then presented as a means of identifying and categorising relevant concepts. As there is some previous evidence to support the usefulness of the DeLone and McLean model, it is taken as the starting point for this study. The variables in the model are each defined and some attempts at measuring them described. An alternate model that takes into account research in the UDA domain is also presented. The research hypotheses derived from both models are also presented.
3.2 Research questions

The objective of the research described in this thesis is to gain a better understanding of UDA success. In particular, the research considers the role of system quality in UDA success and the ability of end user developers to judge whether the systems they develop will have a positive impact on their performance of tasks. Study 1 was conducted in order to address the following research questions.

1. How does UDA quality contribute to user performance on tasks?
   
   This research question seeks to elucidate the process by which IS success is mediated for UDAs.

2. Do end user developers have any misconceptions about system quality? If so, how do they impact upon UDA success?

   This research question attempts to explicitly address concerns in the literature about end user developer perceptions of system quality (e.g. Edberg & Bowman, 1996; Kreie et al., 2000; Shayo et al., 1999). If end user developers are found to have problems with gauging the quality of applications they develop, the potential of this to impact upon the UDA success process warrants further investigation. This question is followed up in Study 2 with a more specific question addressing the impact of any misconceptions on the ability of end users to judge the impact of their applications.

3. Do experience and training influence the ability of end user developers to judge whether the applications they develop will have a positive impact on the performance of the tasks the UDA is designed to support?
End user developers differ from IS professionals in terms of their system development training and experience. This research question addresses the role that experience and training play in the UDA success process.

### 3.3 Framework for the study

A framework is needed to form a basis for undertaking this study. In view of the scarcity of literature on UDA success (Shayo et al., 1999) models of organisational IS success can provide a starting point. DeLone and McLean’s (1992) model of IS success (see Figure 3.1) was selected because, as well as providing a means of categorising relevant concepts that have been considered in the IS success literature, it suggests a model of causal interdependencies between the categories.

![Figure 3.1: DeLone and McLean’s (1992) model of IS success](image)

In DeLone and McLean’s model, system quality and information quality singularly and jointly influence both user satisfaction and use, which also reciprocally influence one another. User satisfaction and use jointly influence individual impact, which in turn influences organisational impact. The concepts included in DeLone and McLean’s model were all considered relevant to this study. They are described below.
3.3.1 System quality

System quality relates to the quality of the IS itself and is concerned with matters such as whether or not there are ‘bugs’ in the system, the consistency of the user interface, ease of use, response rates in interactive systems, documentation, and sometimes, quality and maintainability of the program code. A wide range of measures of system quality has been used in empirical studies. Attempts at measurement of system quality range from direct measures of system quality from a software engineering perspective, such as defects per 1000 lines of code (e.g. Chow, 1985), to perceptual measures such as Amoroso and Cheney’s (1992) system quality measure which incorporates aspects of end user information satisfaction and application utilisation.

Existing direct measures tend to focus on coding and are less applicable to applications developed in 4GLs or form-based environments, but perceptual measures are also of concern because of the potential for lack of objectivity (Edberg & Bowman, 1996; Igbaria, 1990; Rivard, Poirier, Raymond, & Bergeron, 1997). In addition, measures such as Amoroso and Cheney’s (1992) confound system quality with other IS success measures. Amoroso and Cheney’s measure considers information quality and use as dimensions of system quality, however, these have been identified as independent first order factors (Seddon, 1996).

Rivard et al. (1997) have developed a system quality instrument that was designed to be suitable for end user developers to complete, yet to be sufficiently deep to capture their detailed perceptions of components of quality. It addresses the dimensions of system quality discussed in the software engineering literature and is consistent with the quality criteria proposed for UDAs by Salchenberger (1993). Its inclusion of items that directly
address technical quality criteria encourages objective reflection on system quality, nevertheless, as it uses a self-report instrument with Likert scales it only measures perceived system quality.

3.3.2 Information quality

Information quality relates to the characteristics of the information that the IS produces (DeLone & McLean, 1992). It is concerned with issues such as the timeliness, accuracy, relevance and format of information generated by an IS. Instruments that have attempted to measure aspects of information quality include those of O’Reilly (1982), Swanson (1987) and Wang and Strong (1996).

While shown as a separate concept in DeLone and McLean’s model, information quality has often been included in measures of user satisfaction. For example, both the Bailey and Pearson (1983) and the Doll and Torkzedah (1988) user satisfaction instruments include a number of items that measure information quality. It has been argued that these user satisfaction instruments measure independent variables that are likely to cause satisfaction rather than user satisfaction itself (Seddon & Kiew, 1996). An empirical study by Seddon and Yip (1992) found that factors such as information quality, system quality and user knowledge explained over 70% of the variance in an instrument designed to measure user satisfaction directly. Seddon and Kiew (1996) considered this as strong support for information quality being a major determinant of user satisfaction rather than a component of it.
3.3.3 User satisfaction

User satisfaction relates to the attitude or response of an end user towards an IS. It has been defined as ‘the affective attitude towards a particular computer application by an end user who interacts with the application directly’ (Doll & Torkzadeh, 1988). Several instruments have been developed to measure user satisfaction; amongst the most widely used are those of Bailey and Pearson (1983), Baroudi and Orlikowski (1988) and Doll and Torkzadeh (1988).

As discussed in Chapter 2, several researchers have pointed out that problems exist with the conceptualisation of user satisfaction (e.g. Kim, 1989; Melone, 1990; Thong & Chee-Sing, 1996). For example, Melone (1990) commented that user satisfaction has been used to refer to both a user’s perception of the system and an assessment of its output goals. In a review of user satisfaction research, Kim (1989) concluded that this construct has been conceptualised from three different perspectives: (1) user satisfaction in terms of attitudes towards IS, (2) user satisfaction in terms of information quality, and (3) user satisfaction in terms of IS effectiveness. Kim argued that this has resulted in confusion leading to mis-specification of research models and an inability to generate cumulative evidence across studies. Doll and Torkzadeh (1988) noted that general measures of user satisfaction developed for traditional IS environments are less relevant in end user environments and developed a measure for end users who directly interact with a specific application. This measure, EUCS, has been the most used surrogate measure of EUC success (Shayo et al., 1999).
3.3.4 Use

Use refers to how much an IS is used. Many researchers have adopted use as a measure of IS success, assuming that the more a system is used the more successful or effective it is. Examples of measures of IS usage include self-reported hours of use (Seddon & Kiew, 1996), self-reported intended frequency of use (Amoroso & Cheney, 1992), number of system inquiries (DeSanctis, 1983) and connect time (Srinivasan, 1985). However the role of use in IS success is a controversial one. There are many examples of systems that are used infrequently, yet when used provide crucial information. Furthermore, when usage is mandatory the value of use as an indicator of IS success becomes very questionable (DeLone & McLean, 1992).

3.3.5 Individual impact

Individual impact relates to the effect of the IS on the behaviour of the user. DeLone and McLean (1992) stated that individual impact is the most difficult category to define in unambiguous terms. Most commonly the behaviours considered in IS success research relate to management performance and decision making, but individual impact can also be interpreted in a number of other ways (DeLone & McLean, 1992). For example, impacts such as whether the IS has provided the user with a better understanding of the decision making process or whether it has produced a change in user activity might be considered. Many of the studies investigating individual impact have used laboratory experiments and measured variables such as time to complete a task (Benbasat, Dexter, & Todd, 1986), confidence in the decision made (Goslar, Green, & Hughes, 1986), and ability to forecast firm performance (Kasper, 1985).
3.3.6 Organisational impact

Organisational impact refers to the effect of the IS on organisational performance. While of central importance to IS practitioners, many IS research studies have avoided the use of organisational impact measures (except in laboratory studies) because of the difficulty of isolating the effect of an IS from the many other effects that influence organisational performance (DeLone & McLean, 1992).

Quantitative measures that have been used in laboratory simulation studies include return on assets, market share and stock price (Kasper & Cerveny, 1985) and profit (Benbasat et al., 1986). Attempts to use quantitative measures such as these in field studies have, however, been disappointing (Gelderman, 1998). The emergence of business performance frameworks such as the balanced scorecard approach (Kaplan & Norton, 1992) and the business value framework (Rubin, 1991) has provided a broader approach to the assessment of the impact of systems on organisations.

3.4 Support for DeLone and McLean’s model of IS success

Until recently there had been no complete empirical test of the relationships implied by the DeLone and McLean model. Roldán and Millán (2000) tested the entire model for executive information systems and found support for some of the relationships. Studies of parts of the model or individual relationships implied by it also provide empirical support for a number of the relationships. The key supporting research is summarised in Table 3.1.
<table>
<thead>
<tr>
<th>Relationship</th>
<th>Study</th>
</tr>
</thead>
</table>
| System quality → user satisfaction | Seddon and Kiew (1996)  
Roldán and Millán (2000)  
Rivard et al. (1997) |
| Information quality → user satisfaction | Seddon and Kiew (1996)  
Roldán and Millán (2000) |
| User satisfaction → use | Baroudi, Olson and Ives (1986)  
Igbaria and Tan (1997)  
Fraser and Salter (1995) |
| Use → individual impact | Snitkin and King (1986)  
Igbaria and Tan (1997) |
| User satisfaction → individual impact | Gatian (1994)  
Gelderman (1998)  
Igbaria and Tan (1997)  
Etezadi-Amoli and Farhoomand (1996)  
Roldán and Millán (2000) |
| Individual impact → organisational impact | Millman and Hartwick (1987)  
Kasper and Cerveny (1985)  
Roldán and Millán (2000) |

*Seddon and Kiew (1996) undertook the first attempt to explicitly test part of the model. They tested the ‘upstream’ portion of the model and their results provided substantial support for the proposed relationships among system quality, information quality and user satisfaction. Roldán and Millán (2000) also found support for these relationships. In addition, their study also considered the relationships between system quality and use, and information quality and use, but failed to find a relationship.*

*A study by Baroudi, Olson and Ives (1986) showed that, although user satisfaction influenced use, use did not significantly influence user satisfaction. Igbaria and Tan...*
(1997) and Fraser and Salter (1995) also found support for the influence of user satisfaction on system usage.

The results of a study of decision support system use by Snitkin and King (1986) provided support for the proposed relationship between use and individual impact. However, neither Gelderman (1998) nor Roldán and Millán (2000) found any evidence of this relationship. The relationship between user satisfaction and individual impact received support in Gatian’s (1994) study, in which significant positive relationships were found between user satisfaction and both direct and subjective measures of individual impact. Gelderman’s (1998) survey of Dutch managers also confirmed the relationship between satisfaction and both subjective and direct individual impact measures. Etezadi-Amoli and Farhoomand (1996) and Roldán and Millán (2000) used only perceptual measures of individual impact but their results were consistent with the previously mentioned studies on this relationship. Igbaria and Tan (1997) found that user satisfaction has the strongest direct effect on individual impact, but identified a significant role for system usage in mediating the relationship between user satisfaction and individual impact.

Empirical support for the relationship between individual impact and organisational impact has been provided by Millman and Hartwick (1987) in their study of middle managers’ perceptions of the impact of systems, and by Roldán and Millán (2000).

Despite the number of studies that provide a degree of support for DeLone and McLean’s model of IS success it is difficult to compare and interpret their results due to differences in measurement approaches.
Only two of the proposed relationships shown in Table 3.1 have also been specifically investigated for UDAs. The proposed relationship between system quality and satisfaction is supported by Rivard et al. (1997) who found a significant positive correlation between perceived system quality and user satisfaction for UDAs. Kasper and Cerveny’s (1985) study provided evidence for the link between individual impact and organisational impact, with the improved performance of the end user developers flowing through to their firm's stock price, market share, and return on assets.

3.5 Respecification of DeLone and McLean’s model

As DeLone and McLean’s model has served as a focus for discussion of IS success (Walstrom & Hardgrave, 1996; Walstrom & Leonard, 2000), its potential ability to model UDA success was of interest. DeLone and McLean’s model is taken as the starting point for Study 1. Two modifications were made to the model to recognise earlier research results. DeLone and McLean had included both direct and subjective measures of system quality in their single system quality category. However, because of concerns about the ability of end user developers to make judgments about system quality (Edberg & Bowman, 1996; Kreie et al., 2000; Shayo et al., 1999), system quality as assessed by end users and system quality as assessed by experts were specified as separate constructs in the model to be tested. In order to differentiate the two constructs, system quality as assessed by end users is referred to as perceived system quality.

In addition, because prior research suggests that user satisfaction causes system usage rather than vice versa (Baroudi et al., 1986) the causal path between satisfaction and use was specified in this direction. This direction for the hypothesised relationship is
consistent with Fishbein and Ajzen’s (1975) model of the relationship between attitudes and behaviour. The first model tested in the study is therefore the model presented in Figure 3.2 below.

![Figure 3.2: A modified and testable representation of the DeLone and McLean (1992) model of IS success factors showing the hypothesised relationships](image)

The hypotheses that follow directly from this model are:

DM1: Perceived system quality reflects system quality
DM2: User satisfaction reflects information quality
DM3: User satisfaction reflects perceived system quality
DM4: Use reflects information quality
DM5: Use reflects perceived system quality
DM6: Use reflects user satisfaction
DM7: The impact of a UDA on an individual’s work performance increases as use increases
DM8: The impact of a UDA on an individual’s work performance increases as user satisfaction increases
DM9: The organisational impact of a UDA increases as the impact on an individual’s work performance increases.
3.6 Alternate model proposed for Study 1

Despite the literature support for DeLone and McLean’s (1992) model in the organisational domain (see Section 3.4), little is known about its applicability in the UDA domain. As Table 3.1 shows, most support for elements of the model has come from research in the organisational domain with only two of the relationships proposed in the model having been specifically investigated for UDAs. Further investigation of the organisational IS literature and the UDA literature indicate potential problems with use of this model in the UDA domain. In line with DeLone and McLean’s comment that ‘The success model clearly needs further development and validation before it could serve as a basis for the selection of appropriate IS measures’ (DeLone & McLean, 1992 p.88), an alternate model of UDA success was proposed for Study 1. The discussion below introduces this model.

The grouping together of subjective and direct measures of system quality in DeLone and McLean’s categorisation of IS success variables, which led to their model, relies upon the ability of end user developers to accurately judge the quality of applications they develop. Edberg and Bowman (1996) pointed out that users may not only lack the skills to develop quality applications but may also lack the knowledge to make reasonable judgements of the quality of applications that they develop. A user developer may be pleased with the quality of their ‘creation’ when in fact the application includes serious errors such as incorrect formulae. Several empirical studies provide evidence that supports this. Hobbs, McGill and Rowe (1998) found significant differences between the system quality assessments of student developers and independent experts with respect to applications developed by the students. A study by McGill et al. (2000) also provides indirect evidence that questions the proposed relationship, as the
relationship between system quality (measured by independent assessors) and user satisfaction hypothesised by Delone and McLean was not found when the user was also the developer.

Edberg and Bowman (1996) called for future researchers to consider using direct examination of applications as well as subjective measures of satisfaction, and both Igbaria (1990) and Al-Shawaf (1993) called for the development of more direct and objective measures of EUC and UDA effectiveness. Because of these concerns, in the alternate model tested in Study 1, it was hypothesised that when end user developers use UDAs, their perceptions of system quality may not reflect system quality as assessed by experts.

If the relationship between system quality and perceived system quality is not as is implied by DeLone and McLean (1992) then the influence of system quality on individual impact may be more direct than that suggested by DeLone and McLean’s model. Therefore, in the alternate model a direct influence of system quality on individual impact is shown (see Figure 3.3). Hubona and Cheney (1994) investigated the possibility of a direct relationship between system characteristics and decision making performance in their study of the impact of user interface design on decision making. They found partial support for a direct relationship in addition to one mediated by user satisfaction.

Much has been written discussing concerns about use as an indicator of IS success. Szajna (1993) noted that while the utilisation of an IS is widely regarded as an indicator of its success, past research has found inconsistent associations between usage and other
measures of system success. While all would agree that a system that is not used at all
cannot be judged successful, system quality and information quality need not be
positively related to use. In fact, the opposite may be true. A user may need to spend a
longer time using a system to obtain the required information if the system quality or
information quality is poor. Seddon (1997a) did not find a relationship between
information quality and IS use in his investigation of individual user, single application
IS effectiveness, and Roldán and Millán (2000) found no relationship between either
system quality or information quality and use in their test of DeLone and McLean’s
model in the executive information system domain. Thong and Chee-Sing (1996) also
noted that ineffective systems may be used extensively because of subjective
motivations such as political motivations or self-protection for justifying ‘poor’
decisions. Therefore, relationships between perceived system quality and use, and
information quality and use were not included in the alternate model.

Similarly, it could be argued that the relationship between use and individual impact
proposed by DeLone and McLean may not exist. Longer use of a system may in fact
result in decreased productivity. As mentioned previously, neither Gelderman (1998)
nor Roldán and Millán (2000) found a relationship between usage and individual
impact. Potential problems with use influencing individual impact (beyond a binary
situation where it must be used to have any impact) may be magnified in the UDA
domain, as time spent using the system may be confounded with time spent on iterative
enhancement of the system. In their 18 month study of 51 UDAs, Klepper and Sumner
(1990) found that evolutionary change occurred in nearly all the UDAs.
Given the direct relationship between system quality and individual impact proposed in this model, an equivalent direct relationship between information quality and individual impact seems justified. Kim (1989) reviewed prior research and found some empirical support for there being a direct relationship between information quality and decision making performance.

The potential for a user developer’s perceptions of system quality to be coloured by ignorance indicates the need to investigate what characteristics of end user developers influence their ability to tell whether the applications they develop will have a positive impact on their performance of tasks the UDA is designed to support. As discussed earlier (see Section 2.5.2), in previous studies that have related previous experience and training to UDA success (e.g., Al-Shawaf, 1993; Chan & Storey, 1996; Janvrin & Morrison, 2000; Raymond & Bergeron, 1992) the dependent variable used has mainly been user satisfaction and the results have not been conclusive. Hence in Study 1 the research model was extended to enable investigation of the influence of years of spreadsheet experience and level of previous spreadsheet training on system quality, perceived system quality and information quality.

Yaverbaum and Nosek (1992) speculated that computer training increases one’s expectations of systems, and therefore may actually cause negative perceptions of systems. This may be the case for both training and experience in the UDA domain and may go some way to explaining the lack of conclusive results in the literature. Therefore, in this model it was proposed that while experience and training should be positively related to system quality and information quality, they would be negatively related to perceived system quality.
Based on these considerations the alternate model shown below (see Figure 3.3) was developed for comparison with the DeLone and McLean model.

Figure 3.3: The alternate model of UDA success tested in Study 1

The hypotheses associated with this model are:

H1: Perceived system quality does not reflect system quality (note: this hypothesis tests the same relationship as DM1)

H2: Individual impact reflects system quality

H3: User satisfaction reflects perceived system quality (note: this hypothesis is the same as DM3)

H4: User satisfaction reflects information quality (note: this hypothesis is the same as DM2)

H5: Individual impact reflects information quality
H6: Level of use reflects user satisfaction (note: this hypothesis is the same as DM6)
H7: The impact of a UDA on an individual’s work performance increases as user satisfaction increases (note: this hypothesis is the same as DM8)
H8: The organisational impact of a UDA increases as the impact on an individual’s work performance increases (note: this hypothesis is the same as DM9)
H9: System quality reflects the end user developer’s level of experience in the use of the software development tool
H10: System quality reflects the end user developer’s level of training in the use of the software development tool
H11: Information quality reflects the end user developer’s level of experience in the use of the software development tool
H12: Information quality reflects the end user developer’s level of training in the use of the software development tool
H13: Perceived system quality decreases as the end user developer’s level of experience in the use of the software development tool increases
H14: Perceived system quality decreases as the end user developer’s level of training in the use of the software development tool increases.

3.7 Overview
This chapter is the first of four chapters describing the first study of two undertaken during the research described in this thesis. This chapter presented the research questions and conceptual models for Study 1. The objective of the research described in this thesis is to gain a better understanding of UDA success. In particular, the research was designed to investigate the role of system quality in UDA success and to investigate the ability of end user developers to judge whether the applications they develop will
have a positive impact on their performance of tasks. Study 1 was designed to address three research questions relating to these objectives. The research questions are:

1. How does UDA quality contribute to user performance on tasks?

2. Do end user developers have any misconceptions about system quality? If so, how do they impact upon UDA success?

3. Do experience and training influence the ability of end user developers to judge whether the applications they develop will have a positive impact on the performance of the tasks the UDA is designed to support?

Two possible models were identified and used to provide a framework to support the investigation of these research questions. The first model was DeLone and McLean’s (1992) model of IS success. This model has received some empirical support (Seddon & Kiew, 1996), but has also been subject to critique (Seddon, 1997b). The chapter reviewed the relevant literature. The second model was a version of DeLone and McLean’s model that was modified for this study to reflect current research about UDA success (e.g. Al-Shawaf, 1993; Edberg & Bowman, 1996; Hobbs et al., 1998; Igbaria, 1990; Janvrin & Morrison, 2000), and extended to include the roles of user training and experience in UDA success. Testable versions of the two models and their associated hypotheses were presented.
Chapter 4
Method for Study 1

4.1 Introduction
This chapter describes the methodology used to achieve the objectives of Study 1. The methodology is first described in general and reasons for the choice of methodology given. There is then a detailed description of the methodology including the participants, the research environment, the procedure for data collection and the development of the measurement instruments. The chapter continues with a discussion of the statistical techniques chosen to evaluate the models and test the hypotheses, and concludes with descriptive information about the participants in the study and the UDAs that they developed.

4.2 Methodology
This study was a field study in an environment where UDAs are used to support business decision making. The UDAs studied were spreadsheet applications and the decision making took place in a simulated business environment. The participants were postgraduate business students with substantial previous work experience who were participating in a course on strategic management. They developed and used spreadsheet applications to support decision making in a business policy simulation ‘game’. System quality was assessed independently and organisational impact was determined via outputs from the game. End user perceptions of quality and impact were obtained via questionnaire.
This research environment was chosen for the study because it provided an opportunity to explore the nature of end user development of applications, the impact of UDAs on organisational outcomes and the ability of end user developers to make judgments about the quality and success of the applications they develop, in a controlled setting. The major advantages of the approach chosen were firstly that the participants were ‘real’ end user developers, developing applications to support them in their ‘work’, in this case for a fictitious organisation, but ultimately to achieve a good performance in this unit of academic study. This situation was less artificial than an experiment.

The second advantage was that because the participants were involved in a business simulation it was possible to obtain organisational performance measures that should have been directly linked to the performance of the individuals involved. Goodhue and Thompson (1995) stressed the need to go beyond perceived performance impacts and make objective measurements of performance. However, it has proved to be difficult to measure the organisational impact of individual applications (DeLone & McLean, 1992) and in particular UDAs (Shayo et al., 1999), so this situation provided a unique opportunity to explore the full series of relationships represented in DeLone and McLean’s (1992) model of IS success. The opportunity to undertake the study in a partially controlled environment where the possible impact of a number of factors on organisational outcomes could be investigated, with minimum confounding by extraneous variables, was considered worth trading off against the greater generalisability that could have been obtained from a study of end user development in actual organisations.
A further reason for the choice of research environment was the fact that spreadsheets were the tool recommended for participants to develop their applications. Spreadsheets are the most commonly used tool for end user development of applications (Taylor, Moynihan, & Wood-Harper, 1998) and by studying their use maximum generalisability of results would be possible.

4.3 The game

The Business Policy Game (BPG) (Cotter & Fritzche, 1995) simulates the operations of a number of manufacturing companies. Teams compete with one another as members of the management of these companies, producing and selling a consumer durable good. Individual participants assume the roles of managers, and make decisions in the areas of marketing, production, financing and strategic planning. Typical decisions to be made include:

- product pricing
- advertising expenditure
- hiring and firing of sales staff
- product upgrades
- obtaining finance
- research and development expenditure
- production scheduling.

As the simulation model is interactive, decisions made by one company influence the performance of other companies as well as their own.
In this study, the decisions required for the operation of each company were made by teams with four or five members. Each team was free to determine its management structure, but in general the groups adopted a functional structure, with each member responsible for a different area of decision making. Formal group decision making sessions of about one hour were held before each set of decisions was recorded, and these were preceded by substantial preparation. Decisions were recorded twice a week and the simulation run immediately afterwards so that results were available for teams to begin work on the decisions for the next period.

The simulation was run over 13 weeks as part of a capstone course in strategic management. It simulated five years of business performance with each bi-weekly decision period equating to one financial quarter. The simulation accounted for 50% of the participants’ overall course grade, so successful performance was very important to them. Participants drew upon both their previous business knowledge, and that acquired during all of the previous units of their MBA. Successful decision making required applications of equivalent complexity to those used in ‘real’ businesses.

4.4 Participants

The participants in this study were end user developers, developing applications to support decision making as part of their ‘work’, in this case for a fictitious manufacturing company as part of the BPG, but ultimately to have an impact on their performance in their unit of academic study. They were all Masters of Business Administration (MBA) students who had at least two years of previous professional employment experience, as this was a condition of entry to the MBA. The fact that they
were also students in an MBA is secondary to their role as end users, and so this study can be considered a field study of end users. The university simply provided an environment in which end user development and subsequent use of UDAs for decision making occurred.

The applicability of research findings derived from student samples has been raised as an issue of concern (Cunningham, Anderson, & Murphy, 1974). However, in this study participants developed spreadsheets because they recognized the potential value of a UDA for decision support rather than because of any compulsion resulting from the research study. They were not undertaking a contrived task purely for experimental purposes. In addition, Briggs et al. (1996) found MBA students to be good surrogates for executives in studies relating to the use and evaluation of technology, suggesting that the participants in this study can be considered as typical of professionals who would be involved in user development of applications in organisations. The target population for both studies discussed in this thesis is that of all end user developers, so it is probable that this group of participants were more highly educated than end user developers in general.

Three successive cohorts of MBA students participated in the study but they did not differ significantly with respect to any of the background variables described in Section 4.7. Section 4.13.1 provides a descriptive analysis of the participants and their backgrounds with respect to computing and spreadsheet use in particular.
4.5 User developed applications

To help in their decision making, the teams developed their own decision support systems using spreadsheets. These decision support systems could consist of either a workbook containing a number of linked worksheets, or a number of standalone workbooks, or a combination of standalone and integrated worksheets and workbooks. Where several members of a team worked on one workbook, each was responsible for one worksheet, that relating to their area of responsibility. Figure 4.1 provides an example of the possible decision support configurations for the teams. In each case a single individual was responsible for the development of an identifiable application: either a whole workbook or one or more worksheets within a team workbook. Hence, the unit of the analysis in the study was an individual’s application, whether workbook or worksheet(s)\(^1\).

If they wished, the participants were able to use simple templates available with the BPG as a starting point for their applications, but they were not constrained with respect to what they developed, how they developed it, or the hardware and software tools they used. The majority of applications were developed in Microsoft Excel\(^\circledast\) but some participants also used Lotus 1-2-3\(^\circledast\) and Claris Works\(^\circledast\). The spreadsheets themselves were not part of the course assessment, so there were no formal requirements beyond the students’ own needs for the game. The fact that development of applications was optional and unrelated to the purposes of this study reflects the situation in industry where the ability to develop small applications is a necessary part of many jobs (Jawahar & Elango, 2001), yet few spreadsheet developers have spreadsheet development in their job descriptions (Panko, 2000).
Figure 4.1: Possible decision support system (spreadsheet) configurations for teams in the BPG

1 The term spreadsheet is used to refer to individual applications (both worksheets and workbooks) from now on.
4.6 Procedure for data collection

The data for Study 1 was collected from three sources: questionnaires, copies of the UDAs supplied on floppy disk, and output reports from the BPG. As a major weakness of data collection via questionnaire is non-response (Emory & Cooper, 1991), a concerted effort was made to ensure a good response rate. Participants were told about the project during the first week of the BPG and introduced to the two researchers who would be collecting the questionnaires and spreadsheets. They were provided with an opportunity to ask questions about the study and were reassured that information collected was not related in any way to course assessment.

Each participant was asked to complete a written questionnaire and provide a copy of their spreadsheet on disk after eight ‘quarterly’ decisions had been made (four weeks after the start of the simulation). This point was chosen to allow sufficient time for the development and testing of the applications. The majority of completed questionnaires and spreadsheets were collected in person during the time when participants were submitting their decisions, but where this wasn’t possible participants were sent a follow up letter with a reply paid envelope.

Ninety one questionnaires were distributed in total (39 in year 1, 33 in year 2 and 19 in year 3) and 79 useable responses (36 in year 1, 28 in year 2 and 15 in year 3) were received, giving a response rate of 86.8%. Unfortunately, the number of potential participants was less than had been anticipated due to a decline in enrolment in the MBA. This small sample size had implications for the method of data analysis chosen (see Section 4.10).
Output reports from the BPG were obtained from the technician supporting the game immediately after the simulation had been run for the eighth ‘quarterly’ decision.

\section*{4.7 Development of the questionnaire}

The development of the questionnaire for this study involved a review of many existing survey instruments. To ensure the reliability and validity of the measures used, previously validated measurement scales were adopted wherever possible. The questionnaire consisted of two sections. The first section asked questions about the participant and the second section asked questions about the spreadsheet they had developed, and its potential impact on their team’s performance. The complete questionnaire is included in Appendix A.

In the first section, the following background information was collected for each participant:

- Age in years
- Gender
- Number of years of computing experience
- Frequency of use of computers, measured using a 6 category scale from Igbaria (1990)
- Perceived spreadsheet skill relative to the others in the study, measured using a scale where (1) was labelled ‘little or no skill’ and (7) was labelled ‘very skilful’.
This information was collected to enable description of the participants and to confirm that the three cohorts of participants in Study 1 were similar. It was also collected to enable comparison of participants between Study 1 and Study 2.

The following sections discuss the development of each of the measurement scales relating to the constructs included in the two models. See Table 4.1 for a summary of the constructs and the existing instruments used as a basis for their measurement.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived system quality</td>
<td>Rivard et al. (1997)</td>
</tr>
<tr>
<td>Information quality</td>
<td>Fraser and Salter (1995)</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Seddon and Yip (1992)</td>
</tr>
<tr>
<td>Use</td>
<td>Amoroso and Cheney (1992)</td>
</tr>
<tr>
<td>Individual impact</td>
<td>Goodhue and Thompson (1995)</td>
</tr>
<tr>
<td>Training</td>
<td>Igbaria (1990)</td>
</tr>
</tbody>
</table>

### 4.7.1 Perceived system quality

A discussed in Chapter 3, system quality relates to the quality of the IS itself and is concerned with matters such as reliability, ease of use, maintainability and documentation. Perceived system quality relates to the system user’s perceptions of quality, and in the EUC domain the user is often also the developer. In this study, the goal was to develop equivalent instruments to measure both perceived system quality and system quality (see Section 4.9 for a description of the measurement of system quality).
The use of an instrument that focused on UDAs created using 4GLs was desirable, as it would enable applications to be assessed relatively quickly by answering a series of simple questions. The perceived system quality construct was operationalised based upon the instrument developed by Rivard et al. to assess specifically the quality of UDAs (Rivard et al., 1997).

Rivard et al.’s instrument was designed to be suitable for end user developers to complete, yet to be sufficiently deep to capture their perceptions of components of quality. It assesses eight dimensions of quality: reliability, effectiveness, portability, economy, user-friendliness, understandability, verifiability, and maintainability. Each dimension is measured via a number of criterion variables that, in turn, are measured through a series of questionnaire items. Rivard et al. reported that their instrument had a Cronbach alpha of 0.70.

For this study, items that were not appropriate for the applications under consideration (e.g. those specific to database applications) were excluded. Minor adaptations to wording were also made to reflect the environment in which application development and use occurred (i.e. decision making for the BPG).

The resulting item set used as the initial pool for measuring perceived system quality consisted of 53 items (see Appendix A: questions 40 – 92), each scored on a 7 point Likert scale where (1) was labelled ‘strongly agree’ and (7) was labelled ‘strongly disagree’. However, 13 of these items proved not to be amenable to independent assessment (e.g. required access to the hardware configurations on which the spreadsheets were originally used) so were not used in the measurement of system
quality (see Section 4.9). As it was important to be able to directly compare end user assessments of system quality (perceived system quality) with independent expert assessments (system quality), these 13 items were excluded from further analysis. The resulting 40 items are shown below grouped by Rivard et al.’s quality dimensions (Table 4.2).

Table 4.2: The item pool used in the measurement of perceived system quality and system quality

<table>
<thead>
<tr>
<th>Economy</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>The system increased my data processing capacity</td>
<td>The system provides all the information it should</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maintainability</th>
<th>Portability</th>
</tr>
</thead>
<tbody>
<tr>
<td>This system provides the capability to import data from other applications</td>
<td>The system can be run on computers other than the one presently used</td>
</tr>
<tr>
<td>It is possible to copy parts of the system (outputs or data) into other systems or to link with other systems</td>
<td>The system could be used in other similar organisational environments, without any major modification</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unauthorised users could easily access all the data or a part of it</td>
<td>Should an error arise, the system provides the capability to perform some checking in order to locate the source of error</td>
</tr>
<tr>
<td>Each user owns a unique password</td>
<td>This system (rather than the spreadsheet package) always issues an error message when it detects an error</td>
</tr>
<tr>
<td>Unauthorised access is controlled in several parts of the system</td>
<td>All outputs provided by this system are required</td>
</tr>
<tr>
<td>Errors in the system are easy to identify</td>
<td>The data entry sections provide the capability to easily make corrections to data</td>
</tr>
<tr>
<td>Each password limits the access to specific parts of the system</td>
<td>Outputs provided by this system are comprehensive</td>
</tr>
<tr>
<td>This system (rather than the spreadsheet package) automatically corrects certain types of errors, at data-entry time</td>
<td>The system contains all the information required to produce comprehensive outputs</td>
</tr>
<tr>
<td>Should an error arise, the system provides the capability to perform some checking in order to locate the source of error</td>
<td>The system does not destroy any information without asking for a confirmation and getting a positive response</td>
</tr>
</tbody>
</table>
• The system provides default values in the data-entry section
• The system performs an automatic backup of the data
• Data is labelled so that it can be easily matched with other parts of the system
• The system never modifies a cell without asking for a confirmation and getting a positive response
• Corrections to errors in the system are easy to make

**Understandability**
• The same terminology is used throughout the system
• Data entry sections are organised in such a way that the data elements are logically grouped together
• The data entry areas clearly show the spaces reserved to record the data
• The format of a given piece of information is always the same, where ever it is used in the system
• Headings provide information related to the nature of data in the system (e.g. emp-no = employee number)
• The system is broken up into separate and independent sections
• Each section has a unique function
• Each section includes enough information to help you understand its functioning
• The documentation provides all the information required to use the system
• Message presentation is always the same (position, terminology, style…)
• The documentation explains the functioning of the system

**Userfriendliness**
• Using the system is easy, even after a long period of non-utilisation
• The system is easy to learn by new users
• The outputs are easy to understand
• The terms used in data-entry sections are familiar to users
• Queries are easy to make

The items used by Seddon and Kiew (1994) to measure the perceived system quality of organisational applications in their test of DeLone and McLean’s (1992) model of IS success were also included in the questionnaire in case they were required for any comparative analysis. However, they were not appropriate measures of perceived system quality for this study as they only measure ease of use. All the participants in Seddon and Kiew’s study were using the same system and it was thus assumed that the technical dimensions of quality were the same for all participants and thus did not need to be measured. These items were not ultimately included in the analysis for Study 1.
4.7.2 Information quality

Information quality relates to the characteristics of the information that the IS produces (DeLone & McLean, 1992). It is concerned with issues such as the timeliness, accuracy, relevance and format of information generated by an IS. While shown as a separate concept in DeLone and McLean’s (1992) model it has often been included in measures of user satisfaction (e.g. Doll & Torkzadeh, 1988). In Study 1 it was operationalised as a separate construct.

The item pool used to measure perceived information quality consisted of Fraser and Salter’s (1995) 14 item, 7 point scale instrument where (1) is labelled ‘never’ and (7) is labelled ‘always’ (see Table 4.3 below). This instrument contains the items from Seddon and Kiew’s (1994) study, which were based on Doll and Torkzadeh’s (1988) questions relating to information quality. Fraser and Salter also included four extra items from Gelinas, Oram and Wiggins’s (1993) definition of information quality to capture additional dimensions of information quality. They found that their instrument had a Cronbach alpha of 0.95.

All items in this established scale can be interpreted in relation to UDAs. It was considered that end user developer perceptions would be suitable measures of information quality as end users have been considered as ‘experts’ about the knowledge they require from systems to support them in their work (Amoroso, 1988).

Table 4.3: The item pool used in the measurement of information quality

- Do you think the output from your system is presented in a useful format?
• Are you satisfied with the accuracy of your system?
• Is the information provided by your system clear?
• Is the system accurate?
• Does the system provide sufficient information?
• Does the system provide up to date information?
• Do you get the information you need in time?
• Does the system provide output that seems to be just about exactly what you need?
• Does the system provide the precise information you need?
• Does the system’s information content meet your needs?
• Is the information provided by your system understandable?
• Is the information produced by your system valid?
• Is the information provided by your system verifiable?
• Is the information provided by your system complete?

4.7.3 User satisfaction

User satisfaction relates to the attitude or response of an end user towards an IS.

DeLone and McLean (1992) noted the inclusion of information quality measurements in a number of user satisfaction instruments. System quality measurements have also featured in some user satisfaction instruments. It has been argued that these user satisfaction instruments measure independent variables that are likely to cause satisfaction, rather than user satisfaction itself (Seddon & Kiew, 1996). Furthermore, perceived system quality and information quality were already being measured separately in Study 1. Hence, an instrument to measure only user satisfaction was sought. Seddon and Yip (1992) constructed a 4 item, 7 point semantic differential that attempts to measure user satisfaction directly, and Seddon and Kiew (1996) used this user satisfaction instrument in their partial test of DeLone and McLean’s (1992) model of IS success. They found that their instrument had a Cronbach alpha of 0.95. In Study 1, user satisfaction was therefore operationalised using Seddon and Yip’s (1992) instrument (see Table 4.4 below).
Table 4.4: The item pool used in the measurement of user satisfaction

- How adequately do you feel the system meets the information processing needs of your area of responsibility in the BPG?
- How efficient is the system used for your area of responsibility?
- How effective is the system?
- Overall, are you satisfied with the system?

4.7.4 Use

Use refers to the amount an IS is used. Many researchers have measured use as a surrogate for IS success, assuming that the more a system is used the more successful or effective it is. However, the role of use in IS success is a controversial one (Szajna, 1993). There are many examples of systems that are used infrequently, yet when used provide crucial information. Furthermore, usage must be voluntary to be relevant as a measure of IS success. Development and use of decision support systems was optional in the BPG, so use is a pertinent measure of success in this study (DeLone & McLean, 1992).

Two items were included to measure use in Study 1. Firstly, participants were asked how many hours a week they used their system on average. They were also asked to indicate their intended use of the system over the next four quarterly decisions in the BPG. This item was based on Amoroso and Cheney’s (1992) item to measure use and is measured on a 5-point scale ranging from (1) ‘rarely’ to (5) ‘often’. The inclusion of this item is consistent with testing of the relationship between user satisfaction and use in the direction proposed (see Section 3.5) as it suggests that current user satisfaction influences future use. Intended use has also been shown to be a satisfactory surrogate for actual use (Ajzen, 1988; Klobas, 1995).
4.7.5 Individual impact

Individual impact relates to the effect of the IS on the behaviour of the user. Most commonly the behaviours considered in IS success research relate to management performance and decision making (DeLone & McLean, 1992).

In Study 1, individual impact was measured by perceived individual performance impact since direct measures of individual impact were not available from the BPG. The two items used by Goodhue and Thompson (1995) in their study on task-technology fit and individual performance were adopted for this study (see Table 4.5 below). These items are measured on a 7 point Likert scale ranging from (1) ‘agree’ to (7) ‘disagree’. Goodhue and Thompson found that the questions had a Cronbach alpha of 0.61, which they considered marginally acceptable.

Table 4.5: The item pool used in the measurement of individual impact

- The system has a large, positive impact on my effectiveness and productivity in my role in the BPG
- The system is an important and valuable aid to me in the performance of my role in the BPG

4.7.6 Experience and training

As the roles of previous experience and training were included in the alternate model, instruments to measure them were sought. Previous experience with spreadsheets was measured with the single item ‘How long have you been using spreadsheets?’. This is consistent with previous studies that measured experience with a particular software development tool (e.g. Agboola, 1998; Panko, 1998).
Previous spreadsheet training was measured using a 4 item, 5 point scale from Igbaria (1990) where (1) was labelled ‘none’ and (5) was labelled ‘very intensive’. The items asked for level of training received in each of four types of training (college or university; vendor; in-company; self-study).

4.8 Measurement of organisational impact

Organisational impact refers to the effect of the IS on organisational performance. DeLone and McLean (1992) provide an extensive list of individual measures which have been used to assess the organisational impact of systems. Goodhue and Thompson (1995) stressed the need to go beyond perceived performance impacts and to test models using objective measures of performance. In Study 1 it was considered important, given the reliance on perceived measures of individual impact, to obtain direct measures of organisational performance.

A number of objective measures of organisational outcomes were available from the results of the BPG. The Z-score measure of organisational performance is a weighted sum of Z-scores on 17 performance variables (see Table 4.6 below for a list of the performance variables included). The Z-score for each performance variable represents the number of standard deviations that a company’s average value is from the mean for all companies over the period of the simulation (in Study 1 performance was measured over eight quarterly decisions). The weights used in summing the individual Z-scores reflect the relative importance of each of the 17 variables included. Cotter and Fritzche (1995) consider that the Z-score measure closely matches both the subjective
assessments of the writers of the BPG and those of business people who have judged intercollegiate competitions of the game. It was thus chosen as a single composite measure of organisational impact.

Table 4.6: Organisational performance variables included in the Z-score measure of performance

- Net income
- Sales/assets
- Sales in dollars
- Net income/assets
- Sales (percent of market)
- Net income/sales
- Total equity
- Net income/equity
- Total assets
- Unit production cost
- Plant and equipment
- Investor’s ROI
- Stock price
- Interest coverage
- Earnings per share
- Bonds/equity
- Dividends per share

4.9 Independent assessment of system quality

The system quality of each UDA supplied by the participants was assessed by two independent expert assessors using the item pool described in Section 4.7.1 and Table 4.2 above. Both assessors were IS academics with substantial experience teaching spreadsheet design and development.

Before assessing the study sample, the assessors spent a substantial amount of time familiarising themselves with the BPG and then completed four pilot evaluations of
applications not included in the study sample. Differences were discussed and adjustments made to ensure consistency between the assessors. Assessments of the actual UDAs were then undertaken. The consistency of the system quality ratings for the two independent expert assessors was compared by calculating a composite measure of system quality for each assessor for each participant using the results of the measurement model development described in Section 5.2.1. The system quality ratings of the two independent expert assessors were highly correlated ($r = 0.73$, $p < 0.001$).

4.10 Data analysis technique

The major data analysis technique chosen for Study 1 was structural equation modelling (SEM). The term structural equation modelling encompasses a variety of second generation multivariate data analysis techniques, but all SEM techniques are distinguished by two characteristics. These characteristics are (1) simultaneous estimation of multiple and interrelated dependence relationships, and (2) the ability to represent unobserved concepts in these relationships and to account for measurement error in the estimation process (Hair, Anderson, Tatham, & Black, 1998). SEM techniques provide researchers with a comprehensive means for assessing and modifying theoretical models and have become increasingly popular in IS research as they offer great potential for furthering theory development (Gefen, Straub, & Boudreau, 2000).

SEM involves consideration of two types of model: structural and measurement. A structural model consists either wholly or primarily of unobservable constructs (latent variables) and the theoretical relationships (paths) among them. Estimated path
coefficients indicate the strength and sign of the theoretical relationships. By taking measurement error into account SEM can provide more accurate estimates of the causal relationships of interest. Measurement error is taken into account via a measurement model. Each construct in a structural model has a corresponding measurement model that specifies which variables are indicators of that construct. The measurement model can be used to assess the contribution of each scale item as well as to incorporate how well a scale measures a concept (i.e. its reliability) into the estimation of relationships between latent variables.

Whilst both structural and measurement models can be estimated simultaneously many authors have advocated a two step process in which the measurement model is first estimated and then fixed in the second stage when the structural model is estimated (e.g. Anderson & Gerbing, 1988; James, Mulaik, & Brett, 1982; Mulaik et al., 1989). These authors have recommended a two-step approach because testing of the structural model (i.e. the theory being researched) may be meaningless unless it can be established that the indicators for a construct do actually measure that construct reliably and validly (Joreskog & Sorbom, 1993). Use of a two-step approach is of particular importance when measures are less reliable or when the theory being tested is only tentative (Hair et al., 1998), thus a two-step approach to SEM was adopted for this study.

The sample size for Study 1 constrained the analytical techniques available for data analysis. One strategy for dealing with this problem is to develop one factor congeneric measurement models as a means of data reduction in order to obtain a manageable number of composite variables which can be used in subsequent structural models (Holmes-Smith & Rowe, 1994). Whilst using congeneric modelling the validity and
reliability of the composite scales can also be improved. As fitting a congeneric model allows for differences in the degree to which each measure contributes to the overall composite scale, it can provide a more realistic representation of the data (Fleishman & Benson, 1987).

Once measurement model development has been completed, it is possible to build structural equation models that examine relationships amongst the latent variables underlying these composite scales (McDonald, 1996). Because the reliabilities of the composite variables have already been calculated, it is possible to build this information into the models and account for the known amount of error associated with the measurement of each latent variable. Munck (1979) showed that it is possible to fix both the regression coefficients, which reflect the loading of each composite variable on its latent variable, and the measurement error variances associated with each composite variable. The following section describes the process undertaken to develop measurement scales for the latent variables in Study 1.

4.11 Measurement model development

The aim of the measurement model development was to establish a set of items that permitted measurement of each of the latent constructs in the structural models.

Responses to questionnaire items act as indicators that permit measurement of the latent variables. Although where possible the items used came from previously validated instruments, changes in wording and in the domain in which the research was conducted necessitated a rigorous assessment of the measurement models. Parsimonious scales
(sets of indicators) that validly and reliably measured each latent variable were sought from the pool of items used for each construct (see Table 4.2 to Table 4.5).

Correlations between potential indicator items were used both to confirm that the items might validly be combined in a single scale and as a basis for selection of a parsimonious set of indicators for each latent variable. Any items with low inter-item correlation were omitted. Inter-item correlations were also examined to identify variables that were so highly correlated that they might be considered colinear and their contribution to the scale represented by one item. No threshold value for colinearity was set, rather high correlations were examined in relation to theory (Hayduck, 1987).

The evaluation of the measurement models was undertaken using Amos 3.6. Maximum likelihood estimates were used. Missing data was handled by mean substitution. A maximum of 3 out of 79 (3.8%) responses to any item required mean substitution. Mean substitution gave comparable results to listwise deletion and was the approach used by Seddon and Kiew (1996 see Appendix 1) in their partial test of DeLone and McLean’s (1992) model of IS success.

The first measurement property to be assessed was the unidimensionality of each construct. The unidimensionality of a construct is demonstrated when the indicators of the construct have acceptable fit on a single factor model (Hair et al., 1998). Reliability measures such as Cronbach alpha do not ensure unidimensionality, rather they assume it exists. There is little consensus on how to measure model fit, except for agreement that several measures of fit be used (Hair et al., 1998). None of the available measures (except the chi-square statistic) has an associated statistical test. Although various
guidelines have been published no absolute test is available, so the researcher must 
make a judgment of acceptability. Goodness of fit for the single factor congeneric 
measurement models in this study was measured by the likelihood ratio chi-square ($\chi^2$), 
the ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df), the goodness of fit index (GFI), the 
adjusted goodness of fit index (AGFI), the root mean square error of approximation 
(RMSEA) and the Tucker-Lewis index (TLI). After considering the values of these 
goodness of fit indexes using the guidelines discussed below, potential items that did 
not allow acceptable model fit, and hence unidimensionality, were excluded.

For the $\chi^2$ goodness of fit measure a non-significant $\chi^2$ ($p > 0.05$) indicates good fit. 
However, nonsignificance does not guarantee that the ‘correct’ model has been 
identified and a significant $\chi^2$ does not necessarily indicate a poor fit (Gefen et al., 
2000). It has been suggested that consideration of $\chi^2$/df is also useful and Chin and 

The GFI was developed to measure goodness of fit in a way that is less dependent on 
sample size than $\chi^2$ (Joreskog & Sorbom, 1993). A guideline for good fit as measured 
by the GFI is GFI > 0.9 (Schumacker & Lomax, 1996). The AGFI is an extension of the 
GFI, adjusted by the ratio of degrees of freedom for the proposed model to the degrees 
of freedom for the null model. In their guidelines for SEM use in IS research, Gefen, 
Straub and Boudreau (2000) urged IS researchers to report AGFI. AGFI > 0.80 
indicates a good model fit (Segars & Grover, 1993).

The RMSEA measures the mean discrepancy (per degree of freedom) between 
population estimates based on the model and observed sample values. RMSEA < 0.05
indicates a good model fit (Schumacker & Lomax, 1996), but values from 0.05 to 0.08 have been deemed acceptable (Hair et al., 1998). The $\chi^2$, GFI and RMSEA are all measures of absolute fit; no distinction is made as to whether the model fit is better or worse than other possibilities. The TLI is an incremental fit measure that compares the proposed model to the null model. A recommended value of TLI is 0.90 or greater (Hair et al., 1998).

After overall model fit was considered, the significance of each estimated coefficient was examined to determine the probability that the item was a valid indicator of the construct. The ratio of coefficient to standard error, or $t$ value, was used to test each coefficient for significance. Ratios greater than 1.96 are acceptable, but the higher the ratio the more likely the item is a valid indicator of the construct (Hayduck, 1987). Any items with non-significant coefficients were removed. In this way convergent validity was demonstrated. This process also contributes to the establishment of discriminant validity.

Once the item set was reduced to a valid, parsimonious, unidimensional scale, three estimates of reliability were calculated for each latent variable: Cronbach alpha coefficient, composite reliability, and average variance extracted. In the past, reliability has commonly been assessed using Cronbach alpha coefficient. However, Cronbach alpha is calculated with the restrictive assumption that all items weight equally on their corresponding construct. It was calculated in this study along with the other reliability measures for comparative purposes. For unidimensional scales, values for Cronbach alpha of 0.7 or higher indicate acceptable internal consistency (Nunnally, 1978). Cronbach alpha was calculated using SPSS for Windows 7.0.
Composite reliability is a more general measure of reliability as it uses the item loadings estimated within the model. It was calculated as:

\[
\text{Composite reliability} = \frac{(\Sigma \text{standardised loadings})^2}{(\Sigma \text{standardised loadings})^2 + \Sigma \varepsilon_j}
\]

where standardised loadings are the standardised loadings of the indicator on the latent variable, and \( \varepsilon_j \) is the measurement error for each indicator. Standardised loadings and measurement errors are reported by Amos 3.6. Although composite reliability is more general than Cronbach alpha, the interpretation of values is similar. A commonly used threshold value for composite reliability is 0.7, although values below 0.7 have been considered acceptable for exploratory research (Hair et al., 1998).

The variance extracted reliability measure reflects the overall amount of variance in the indicators accounted for by the latent construct. The variance accounted for by the indicators was calculated as:

\[
\text{Variance extracted} = \frac{\Sigma \text{standardised loadings}^2}{\Sigma \text{standardised loadings}^2 + \Sigma \varepsilon_j}
\]

A variance extracted value greater than 0.5 indicates acceptable reliability on this measure (Hair et al., 1998).

For those measurement models where a composite variable was created, the loading of the composite variable on its associated latent variable and the error term needed to be specified for subsequent use. The loading of the composite variable on its associated latent variable was calculated as described in Hair et al. (1998):
The error term was specified as:

$$\theta = \sigma^2(\text{composite})(1 - r)$$

where:

- $\lambda$ is the loading of the composite variable on the associated latent variable
- $\sigma$ is the standard deviation of the composite variable
- $r$ is the composite reliability of the measurement scale
- $\theta$ is the error term.

In the absence of information to estimate reliability, for single indicator variables the loading of the indicator on its associated latent variable was specified at one and the error term was specified as zero. The results of the measurement model development are reported in Section 5.2.

### 4.12 Structural model evaluation

Following the assessment of the measurement models and the creation of composite variables, AMOS 3.6 was used to evaluate the structural models. The evaluation of the structural models was conducted with the entire sample. Maximum likelihood estimates were used. A competing models strategy was used, but as the models were not nested models they could not be directly compared using a difference $\chi^2$ test (Hair et al., 1998).

Although only two models were presented in Chapter 3, three structural models were actually tested. The DeLone and McLean model was tested and compared with a subset
of the alternate model that includes the same constructs. This was done to enable a more
direct comparison on the evaluation criteria. The full alternate model was also tested.

Each structural model was evaluated on three criteria. The first criterion was overall
goodness of fit between the model and the sample data. This was measured using the
same goodness of fit measures used to assess the measurement models: the likelihood
ratio chi-square (χ^2), the ratio of χ^2 to degrees of freedom (χ^2/df), the goodness of fit
index (GFI), the adjusted goodness of fit index (AGFI), the root mean square error of
approximation (RMSEA) and the Tucker-Lewis index (TLI). Section 4.11 above
provides a description of these measures and the ‘rules of thumb’ for interpreting them.

The second criterion considered was the ability of the models to explain the variance in
the dependent variables. The dependent variables of most interest in the DeLone and
McLean model and the partial alternate model are individual impact and organisational
impact. As the full alternate model was proposed to provide further insight into the roles
of experience and training in influencing individual impact and organisational impact
through system quality, perceived system quality and information quality, the variance
explained for these variables was also of interest. An estimate of variance explained is
provided by the squared multiple correlations (R^2) of the structural equations for these
variables (Hair et al., 1998). AMOS 3.6 reports these values. Although no test of
statistical significance can be performed, R^2 provides a relative measure of fit for each
structural equation in the model.

The third criterion was the significance of estimated model coefficients. Structural
models represent propositions about relationships between constructs. If a model is a
valid representation of UDA success, all proposed relationships should be significant. As well as estimated coefficients, SEM techniques provide standard errors and calculate \( t \) values for each coefficient. All of the hypotheses except one specify a direction for the proposed relationship so a one-tailed \( t \) value of 1.645 indicates significance at the \( p < 0.05 \) level (Hair et al., 1998). For the one non-directional hypothesis a two-tailed \( t \) value of 1.96 indicates significance at the \( p < 0.05 \) level.

In addition to the statistical significance of the estimated model coefficients, the strength of the relationships they represent was also of interest. There is no substantial consensus about definitions of correlation strength (Bryman & Cramer, 1999), but for this research correlations of less than 0.20 have been defined as weak, correlations between 0.20 and 0.50 have been defined as moderate, and correlations of 0.50 and over have been defined as strong. This categorisation is consistent with Cohen’s (1988) conventions.

The models were compared on each of the criteria described above. It was considered that an acceptable model should explain a moderate to high proportion of the variance in the dependent variables of interest, would contain only valid paths, and would meet the criteria for acceptable fit.

### 4.13 Descriptive analysis

This section presents background information about the participants in Study 1 and about the applications developed during the study.
4.13.1 The participants

Of the 79 participants in Study 1, 78.5% were male and 21.5% female (62 males, 17 females). Their ages ranged from 21 to 49 with an average age of 31.8 (see Table 4.7). They had an average of 9.5 years experience using computers (with a range from 2 to 24 years, see Table 4.7) and the majority of participants (56 or 70.6%) used a computer more than once a day (in addition to activities relating to the BPG, see Table 4.8).

Table 4.7: Age and computing experience of the participants in Study 1

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>31.8</td>
<td>21</td>
<td>49</td>
<td>7.49</td>
</tr>
<tr>
<td>Computing experience (years)</td>
<td>9.5</td>
<td>2</td>
<td>24</td>
<td>4.46</td>
</tr>
</tbody>
</table>
Table 4.8: Frequency of computer use by the participants in Study 1

<table>
<thead>
<tr>
<th>Frequency of Computer Use (apart from BPG)</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than once a month</td>
<td>1</td>
<td>1.3</td>
</tr>
<tr>
<td>Once a month</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A few times a month</td>
<td>3</td>
<td>3.8</td>
</tr>
<tr>
<td>A few times a week</td>
<td>13</td>
<td>16.5</td>
</tr>
<tr>
<td>About once a day</td>
<td>6</td>
<td>7.6</td>
</tr>
<tr>
<td>Several times a day</td>
<td>56</td>
<td>70.6</td>
</tr>
</tbody>
</table>

Table 4.9 and Table 4.10 summarise the participants’ spreadsheet experience and training and their perceptions of their own skill. Participants reported an average of 5.9 years experience using spreadsheets (with a range from 0 to 15 years) and perceived themselves as moderately skilful (average 3.5, with only 9 or 11.4% perceiving themselves as lacking in skill – i.e. 1 or 2 on the scale).

Table 4.9: Spreadsheet experience and perceived skill of the participants in Study 1

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet experience (years)</td>
<td>5.9</td>
<td>0</td>
<td>15</td>
<td>3.81</td>
</tr>
<tr>
<td>Perceived spreadsheet skill</td>
<td>3.5</td>
<td>1</td>
<td>5</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 4.10 indicates that the participants had received relatively little spreadsheet training. More than 50% of the participants had received no in-company, vendor or university / technical college training. Self-study was the predominant means by which participants had acquired their knowledge of spreadsheets. This is consistent with other studies of spreadsheet development, as Chan and Storey (1996) and Hall (1996) also found self-study to be the most common form of training for end user developers.
Table 4.10: Previous spreadsheet training of the participants in Study 1

<table>
<thead>
<tr>
<th>Training Source</th>
<th>Mean</th>
<th>Number in each category</th>
<th>Number in each category</th>
<th>Number in each category</th>
<th>Number in each category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1) None</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>N %</td>
<td>N %</td>
<td>N %</td>
<td>N %</td>
</tr>
<tr>
<td>University or college</td>
<td>2.0</td>
<td>46</td>
<td>58.2</td>
<td>8</td>
<td>10.1</td>
</tr>
<tr>
<td>Vendor</td>
<td>1.5</td>
<td>62</td>
<td>78.5</td>
<td>3</td>
<td>3.8</td>
</tr>
<tr>
<td>In-company</td>
<td>1.7</td>
<td>52</td>
<td>65.8</td>
<td>6</td>
<td>7.6</td>
</tr>
<tr>
<td>Self-study</td>
<td>3.3</td>
<td>8</td>
<td>10.1</td>
<td>8</td>
<td>10.1</td>
</tr>
</tbody>
</table>

As three cohorts of MBA students were involved in Study 1, a comparison across cohorts was made. No significant differences were found on any of the characteristics described in this section.

4.13.2 The user developed applications

As discussed in Section 4.5 the majority of applications were developed in Microsoft Excel©, but there were also some developed in Lotus 1-2-3© and Claris Works©. The average file size of the workbooks in Study 1 was 182K with a minimum of 23K and a maximum of 814K. This average size is similar to the average file size found in Hall’s (1996) field study of end user developed spreadsheets (218K) and therefore suggests that these applications were typical of spreadsheet applications in the workplace in this respect. The UDAs also varied greatly in terms of sophistication, with the simplest making little use of functions, charts or macros.

Each participant in the study belonged to one of 19 groups in the BPG. As previously described, groups were free to decide on the degree of interconnectedness that they
wanted for the spreadsheets used in decision making. In six of the groups each of the team members developed a standalone application. These applications were not linked to the other related applications from the group. Eight groups chose to develop a workbook in which all the worksheets of the individual team members were linked. The remaining five groups made use of a mixture of standalone applications and linked worksheets.

**4.14 Overview**

This chapter is the first of two chapters that describe the research approaches used in answering the research questions for this thesis. This chapter described the design of Study 1 (see Chapter 8 for the research approach used in Study 2). Study 1 was a field study in an environment where UDAs are used to support business decision making. The UDAs studied were spreadsheet applications and the decision making took place in a simulated business environment. The participants were MBA students who were participating in a course on strategic management. They developed and used spreadsheet applications to support decision making in a business policy simulation ‘game’. UDA quality was assessed independently and organisational impact was determined via outputs from the game. End user perceptions of quality and impact were obtained via questionnaire.

The chapter included a general discussion of SEM, the data analysis technique chosen for this study, and provided details of how the measurement model development and structural model evaluation were carried out. This is followed up in Chapter 5, which presents the results of both the measurement model development and structural model
testing. The chapter concluded with a descriptive analysis of the participants and the applications they developed. The participants had an average of about six years experience with spreadsheets and perceived themselves as being moderately skilful; they had, however, received relatively little spreadsheet training, with self-study being the predominant means by which participants had acquired their knowledge of spreadsheet use and development. The UDAs were mainly developed in Microsoft Excel® and varied widely in size and sophistication, but the average file size was similar to the average file size found in Hall’s (1996) field study of end user developed spreadsheets, suggesting that these applications were typical of spreadsheet applications in the work place (e.g. Hall, 1996).
Chapter 5
Study 1 Results

5.1 Introduction
This chapter reports the results of the Study 1 data collection and statistical analyses carried out as described in Chapter 4. It first describes the results of the measurement model development for each of the latent variables in both the DeLone and McLean (1992) model and the alternate model presented in Chapter 3. It then presents the results of the structural model evaluations. Three criteria were used to evaluate the models: overall goodness of fit of each model; the amount of variance in key constructs explained by the model; and the significance of model coefficients. The chapter describes how each of the models compares on these criteria and provides the results of the hypothesis testing.

5.2 Measurement models
This section presents the measurement model for each latent variable in Study 1 and provides a summary of the measurement model information used in the evaluation of the structural models.

5.2.1 System quality
As previously discussed, the Rivard et al. (1997) system quality instrument measures a number of dimensions of quality, each of which is composed of a set of criteria. The assessment of the measurement model for system quality was performed bottom up as
described by Rivard et al. (1997), starting with the averaged item scores for the two independent assessors. One factor congeneric models for each criterion were first assessed. Once a unidimensional item set for each criterion was obtained, the criterion scores for each dimension were averaged. Of the original 40 items considered (see Table 4.2) 19 items relating to system quality were retained in the final measurement model (see Appendix B for a list of the items retained). This process ensured the convergent validity of the system quality construct (Kline, 1998).

The overall measurement model for system quality was then assessed using the dimension scores. Although Rivard et al. averaged the dimension scores to obtain overall system quality, better model fit was obtained in this study with a congeneric model, thus composite system quality scores were calculated using the factor score weightings from the overall system quality measurement model. Table 5.1 below shows summary statistics for the overall measurement model.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Mean</th>
<th>SD</th>
<th>Economy</th>
<th>Portability</th>
<th>Reliability</th>
<th>Understandability</th>
<th>Userfriendliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>4.27</td>
<td>0.71</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portability</td>
<td>4.51</td>
<td>0.68</td>
<td>0.439***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>2.08</td>
<td>0.69</td>
<td>0.408***</td>
<td>0.523***</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understandability</td>
<td>3.14</td>
<td>0.75</td>
<td>0.455***</td>
<td>0.385**</td>
<td>0.516***</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Userfriendliness</td>
<td>3.13</td>
<td>0.80</td>
<td>0.411***</td>
<td>0.520***</td>
<td>0.780***</td>
<td>0.648***</td>
<td>1.00</td>
</tr>
</tbody>
</table>

** p < 0.01  
*** p < 0.001  

all n = 78

Table 5.2 shows the parameter estimates, factor score weights and fit statistics for the system quality measurement model. All items had t values substantially greater than 1.96 and thus can be considered valid indicators of the construct. The model fits the data

108
relatively well, with all goodness of fit measures except RMSEA meeting the
guidelines. The measures of reliability for this scale were all acceptable: Cronbach
alpha was 0.84, composite reliability was 0.84 and variance extracted was 0.52. The
scale was therefore considered satisfactory for SEM.

Table 5.2: One factor congeneric model for system quality – parameter estimates,
goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>System quality</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economy</td>
<td>0.349</td>
<td>0.079</td>
<td>4.426</td>
<td>0.059</td>
</tr>
<tr>
<td>Portability</td>
<td>0.398</td>
<td>0.073</td>
<td>5.469</td>
<td>0.087</td>
</tr>
<tr>
<td>Reliability</td>
<td>0.573</td>
<td>0.066</td>
<td>8.662</td>
<td>0.266</td>
</tr>
<tr>
<td>Understandability</td>
<td>0.509</td>
<td>0.077</td>
<td>6.572</td>
<td>0.111</td>
</tr>
<tr>
<td>Userfriendliness</td>
<td>0.732</td>
<td>0.074</td>
<td>9.919</td>
<td>0.476</td>
</tr>
<tr>
<td><strong>Goodness of fit measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>10.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.055</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>2.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.849</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.927</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reliability measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronbach alpha</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average extracted variance</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2.2 Perceived system quality

Perceived system quality was measured using the responses of the participants to the
same item set as were retained in the final measurement model for system quality. This
was done to ensure compatibility between the two constructs. The criterion scores for
each dimension were averaged to give dimension scores. Table 5.3 below provides summary statistics for the individual dimensions of perceived system quality.

Table 5.3: Perceived system quality summary statistics

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Mean</th>
<th>SD</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>3.85</td>
<td>1.74</td>
<td>1.00</td>
</tr>
<tr>
<td>Portability</td>
<td>3.89</td>
<td>1.32</td>
<td>0.465***</td>
</tr>
<tr>
<td>Reliability</td>
<td>3.28</td>
<td>1.02</td>
<td>0.179</td>
</tr>
<tr>
<td>Understandability</td>
<td>3.79</td>
<td>0.84</td>
<td>0.218</td>
</tr>
<tr>
<td>Userfriendliness</td>
<td>3.63</td>
<td>0.96</td>
<td>0.441***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Economy</th>
<th>Portability</th>
<th>Reliability</th>
<th>Understandability</th>
<th>Userfriendliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>1.00</td>
<td>0.347**</td>
<td>1.00</td>
<td>0.504***</td>
<td></td>
</tr>
<tr>
<td>Portability</td>
<td>0.465***</td>
<td>1.00</td>
<td>0.209</td>
<td>0.573***</td>
<td>1.00</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>0.179</td>
<td>0.347**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Understandability</td>
<td></td>
<td>0.218</td>
<td>0.504***</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Userfriendliness</td>
<td></td>
<td>0.441***</td>
<td>0.573***</td>
<td>0.471***</td>
<td>0.558***</td>
</tr>
</tbody>
</table>

** p < 0.01
*** p < 0.001

* Maximum of 3 out of 79 (3.8%) responses to any item required mean substitution

The overall measurement model for perceived system quality was then assessed using the dimension scores with regression weights fixed as in the final system quality measurement model. The factor weightings used to create the composite system quality measure (see Table 5.2 above) were then also used to create the composite perceived system quality measure. This was done to enable direct comparison between the system quality and perceived system quality measures.

Table 5.4 shows the fit statistics and reliability measures for the perceived system quality measurement model. As the regression estimates were fixed to match those of the system quality measurement model they were not estimated by AMOS. All of the measures of reliability were acceptable with a Cronbach alpha of 0.73, composite reliability of 0.84 and variance extracted of 0.54. However, none of the goodness of fit measures was satisfactory. This is not surprising given that congeneric models generally provide better fit than do models with fixed parameters such as parallel models (Fleishman & Benson, 1987) and suggests that end user developers may perceive
system quality differently than experts do. However, despite the poor model fit the composite was retained in this form to enable direct comparison with system quality.

Table 5.4: Goodness of fit measures and reliability measures for the one factor model for perceived system quality

<table>
<thead>
<tr>
<th>Goodness of fit measures</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>48.04</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>10</td>
</tr>
<tr>
<td>Probability (p)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>4.80</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.798</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.698</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.221</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.641</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability measures</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach alpha</td>
<td>0.73</td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.84</td>
</tr>
<tr>
<td>Average extracted variance</td>
<td>0.54</td>
</tr>
</tbody>
</table>

5.2.3 Information quality

After congeneric modelling, six items were retained as representative unidimensional measures of information quality (see Table 4.3 for the complete set of items considered and Appendix B for the items retained). This process ensured the convergent validity of the information quality construct (Kline, 1998). Table 5.5 below contains summary statistics for these items.

Table 5.6 shows the parameter estimates, factor score weights and fit statistics for the information quality measurement model. All items had $t$ values substantially greater than 1.96 and thus can be considered valid indicators of the construct. The model fits
the data relatively well. Despite a significant $\chi^2$ and a higher than desirable RMSEA, the $\chi^2$/df, GFI, AGFI and TLI all indicate good fit.

Table 5.5: Information quality summary statistics

<table>
<thead>
<tr>
<th>Variable $^a$</th>
<th>Mean</th>
<th>SD</th>
<th>Q18</th>
<th>Q19</th>
<th>Q20</th>
<th>Q21</th>
<th>Q22</th>
<th>Q25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q18</td>
<td>5.54</td>
<td>1.24</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q19</td>
<td>5.16</td>
<td>1.29</td>
<td>0.63 ***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q20</td>
<td>5.06</td>
<td>1.30</td>
<td>0.58 ***</td>
<td>0.84 ***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q21</td>
<td>5.24</td>
<td>1.21</td>
<td>0.72 ***</td>
<td>0.78 ***</td>
<td>0.83 ***</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q22</td>
<td>5.51</td>
<td>1.05</td>
<td>0.68 ***</td>
<td>0.61 ***</td>
<td>0.63 ***</td>
<td>0.71 ***</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Q25</td>
<td>4.95</td>
<td>1.24</td>
<td>0.68 ***</td>
<td>0.74 ***</td>
<td>0.75 ***</td>
<td>0.74 ***</td>
<td>0.69 ***</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*** p < 0.001

$^a$ all $n = 79$

Table 5.6: One factor congeneric model for information quality - parameter estimates, goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>Information quality</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q18</td>
<td>0.951</td>
<td>0.119</td>
<td>7.971</td>
<td>0.140</td>
</tr>
<tr>
<td>Q19</td>
<td>1.074</td>
<td>0.119</td>
<td>9.017</td>
<td>0.091</td>
</tr>
<tr>
<td>Q20</td>
<td>1.118</td>
<td>0.119</td>
<td>9.408</td>
<td>0.115</td>
</tr>
<tr>
<td>Q21</td>
<td>1.105</td>
<td>0.106</td>
<td>10.442</td>
<td>0.327</td>
</tr>
<tr>
<td>Q22</td>
<td>0.815</td>
<td>0.101</td>
<td>8.107</td>
<td>0.170</td>
</tr>
<tr>
<td>Q25</td>
<td>1.042</td>
<td>0.114</td>
<td>9.140</td>
<td>0.156</td>
</tr>
</tbody>
</table>

| **Goodness of fit measures** | | | | |
| Chi-square ($\chi^2$) | 17.72 |
| Degrees of freedom (df) | 8 |
| Probability (p) | 0.023 |
| Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df) | 2.22 |
| Goodness of fit index (GFI) | 0.932 |
| Adjusted goodness of fit index (AGFI) | 0.823 |
| Root mean square error of approximation (RMSEA) | 0.125 |
| Tucker-Lewis index (TLI) | 0.953 |

| **Reliability measures** | |
| Cronbach alpha | 0.93 |
| Composite reliability | 0.94 |
| Average extracted variance | 0.72 |

112
The measures of reliability for this scale were all acceptable: Cronbach alpha was 0.93, composite reliability was 0.94 and average extracted variance was 0.72. The scale was therefore considered satisfactory for SEM, and a composite variable for information quality was created using the factor score weights shown in Table 5.6.

### 5.2.4 User satisfaction

Three of the four potential items were retained as representative unidimensional measures of user satisfaction, thus ensuring the convergent validity of the construct (Kline, 1998). Question 26 was excluded because it had low correlation with Question 29 (r = 0.149, p = 0.192) (see Appendix B for a list of the items retained). Table 5.7 contains summary statistics for the retained items.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
<th>Q27</th>
<th>Q28</th>
<th>Q29</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q27</td>
<td>4.89</td>
<td>1.34</td>
<td>79</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q28</td>
<td>4.86</td>
<td>1.28</td>
<td>79</td>
<td>0.53***</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Q29</td>
<td>4.80</td>
<td>1.42</td>
<td>78</td>
<td>0.60***</td>
<td>0.38**</td>
<td>1.00</td>
</tr>
</tbody>
</table>

** p < 0.01  
*** p < 0.001

Table 5.8 shows the parameter estimates, factor score weights and fit statistics for the user satisfaction measurement model. All items had t values substantially greater than 1.96 and can therefore be considered valid indicators of the construct. All goodness of fit measures met the recommended guidelines. The model can thus be considered to fit the data well.
Table 5.8: One factor congeneric model for user satisfaction – parameter estimates, goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>User Satisfaction</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q27</td>
<td>1.213</td>
<td>0.166</td>
<td>7.302</td>
<td>0.726</td>
</tr>
<tr>
<td>Q28</td>
<td>0.735</td>
<td>0.152</td>
<td>4.840</td>
<td>0.120</td>
</tr>
<tr>
<td>Q29</td>
<td>0.933</td>
<td>0.165</td>
<td>5.660</td>
<td>0.153</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Goodness of fit measures</strong></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>0.034</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.853</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.998</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>1.050</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Reliability measures</strong></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach alpha</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average extracted variance</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The measures of reliability for this scale were all acceptable: Cronbach alpha was 0.75, composite reliability was 0.77, and average extracted variance was 0.53. The scale was therefore considered satisfactory for SEM, and a composite variable for user satisfaction was created using the factor score weights shown in Table 5.8 above.

### 5.2.5 Use

As described in Chapter 4, two potential measures of system use were included in the questionnaire: current use and intended use. These measures were not significantly correlated (see Table 5.9 below) and thus could not be validly combined in a single
scale. It appeared likely that some responses given for current use included time invested in ongoing iterative development of the system. Thus, the lack of correlation between the two measures could be due to development time being differentially included. Current use, as estimated by the participants, was thus considered to be a less reliable indicator of use and a decision was made to include only intended use in the structural models. This is consistent with the direction of the relationship between user satisfaction and use proposed in Section 3.5. As only one measure of use was retained, no measures of reliability could be calculated.

Table 5.9: Use summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Use (hours)</td>
<td>78</td>
<td>4.04</td>
<td>2.72</td>
<td>-0.024</td>
</tr>
<tr>
<td>Intended Use (scale 1 to 5)</td>
<td>79</td>
<td>3.62</td>
<td>1.29</td>
<td></td>
</tr>
</tbody>
</table>

5.2.6 Individual impact

Individual impact was measured using two items. Summary statistics for the items are shown in Table 5.10.

Table 5.10: Individual impact summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q30</td>
<td>78</td>
<td>4.99</td>
<td>1.42</td>
<td>0.858***</td>
</tr>
<tr>
<td>Q31</td>
<td>79</td>
<td>5.14</td>
<td>1.40</td>
<td></td>
</tr>
</tbody>
</table>

*** p < 0.001
At least three indicator variables must be available to undertake one factor congeneric modelling. As only two items to measure individual impact were included in the questionnaire it was not possible to determine measurement model fit in the initial modelling phase, hence no goodness of fit statistics are provided in Table 5.11. However, the measures of reliability for this scale were all very good: Cronbach alpha was 0.92, composite reliability was 0.92, and average extracted variance was 0.86. The scale was therefore considered satisfactory for SEM.

Table 5.11: Parameter estimates and reliability measures for individual impact

<table>
<thead>
<tr>
<th>Individual Impact</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter estimates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q30</td>
<td>1.307</td>
<td>0.122</td>
<td>10.732</td>
</tr>
<tr>
<td>Q31</td>
<td>1.290</td>
<td>0.121</td>
<td>10.695</td>
</tr>
<tr>
<td>Reliability measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronbach alpha</td>
<td>0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average extracted variance</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2.7 Organisational impact

As discussed in Section 4.8, the Z-Score measure of organisational performance is a weighted sum of Z-scores on 17 performance variables. The creators of the BPG, Cotter and Fritzche (1995), consider that it closely matches both the subjective assessments of the writers of the game and those of business people who have judged intercollegiate competitions of the BPG. It was thus chosen as a single composite measure of organisational impact. The mean for Z-score was 0.046 and the standard deviation was 0.61.
5.2.8 Experience
As described in Section 4.7.6, experience was measured using the single item: years of spreadsheet experience (mean = 5.92, SD = 3.81). This was treated as a single item indicator in the structural models.

5.2.9 Training
A summary measure of training was created for each participant by summing the responses for each of the four types of training (after adjusting the scale to start at zero). The average level of training on this scale was 4.60 and the standard deviation was 3.01. This summary variable was treated as a single item indicator in the structural models.

5.2.10 Summary of the information used to specify the structural models
The sections above detail the development of the individual measurement models required prior to testing of the structural models. Table 5.12 provides a summary of information used to specify parameters in the structural models. It includes the composite variables derived from the measurement models, and the single indicator variables.
Table 5.12: Summary of the information used to specify measurement model parameters in the structural models

<table>
<thead>
<tr>
<th>Construct</th>
<th>Mean</th>
<th>SD</th>
<th>Loading</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Composites created</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>3.03</td>
<td>0.64</td>
<td>0.594</td>
<td>0.067</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>3.60</td>
<td>0.80</td>
<td>0.686</td>
<td>0.174</td>
</tr>
<tr>
<td>Information quality</td>
<td>5.25</td>
<td>1.06</td>
<td>1.030</td>
<td>0.070</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>4.86</td>
<td>1.21</td>
<td>1.057</td>
<td>0.336</td>
</tr>
<tr>
<td><strong>Single item indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>3.62</td>
<td>1.29</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Organisational impact</td>
<td>0.046</td>
<td>0.61</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Experience</td>
<td>5.92</td>
<td>3.81</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Training</td>
<td>4.60</td>
<td>3.01</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

5.3 **Analysis of the structural models for Study 1**

As described in Section 4.12, following the assessment of the measurement models and the creation of composite variables, the structural models were evaluated. This section describes the results of these evaluations.

Table 5.13 displays the implied correlations between all of the latent variables considered in the models. As can be seen the estimated correlations between the latent variables were not excessively high, indicating that discriminant validity was achieved (Kline, 1998).
Table 5.13: Implied correlations between the latent variables in the models

<table>
<thead>
<tr>
<th></th>
<th>Experience</th>
<th>Training</th>
<th>System quality</th>
<th>Perceived system quality</th>
<th>Information quality</th>
<th>User satisfaction</th>
<th>Use</th>
<th>Individual impact</th>
<th>Organisational impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.452</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>0.002</td>
<td>-0.015</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>-0.213</td>
<td>0.007</td>
<td>-0.175</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>0.088</td>
<td>0.036</td>
<td>0.000</td>
<td>-0.019</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>-0.012</td>
<td>0.026</td>
<td>-0.058</td>
<td>0.318</td>
<td>0.661</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>-0.005</td>
<td>0.012</td>
<td>-0.027</td>
<td>0.149</td>
<td>0.309</td>
<td>0.467</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.013</td>
<td>0.020</td>
<td>0.126</td>
<td>0.146</td>
<td>0.594</td>
<td>0.698</td>
<td>0.326</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Organisational impact</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.006</td>
<td>-0.007</td>
<td>-0.027</td>
<td>-0.032</td>
<td>-0.015</td>
<td>-0.045</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Three structural models were tested. The DeLone and McLean model was tested and compared with the subset of the alternate model that includes the same constructs. This was done to enable a direct comparison on the evaluation criteria. The full alternate model was also tested. The three models tested are shown in Figure 5.1, Figure 5.2 and Figure 5.3 below. The models were evaluated and compared on three criteria: overall goodness of fit, ability to explain the variance in the dependent variables, and the significance of estimated model coefficients (see Section 4.12 for a detailed explanation of the criteria chosen). It was considered that an acceptable model should explain a moderate to high proportion of the variance in the dependent variables of interest, would contain only valid paths, and would meet the criteria for acceptable fit.

Figure 5.1: Representation of the DeLone & McLean (1992) model of IS success tested in Study 1
Figure 5.2: Partial alternate model tested in Study 1

Figure 5.3: Full alternate model tested in Study 1
5.3.1 DeLone and McLean’s model of IS success

The goodness of fit measures, model coefficients, standard errors and $t$ values for the version of the DeLone and McLean model of IS success tested in Study 1 are reported in Table 5.14. They are shown along side the results for the partial alternate model to facilitate comparison. Figure 5.4 shows the standardised coefficients for each hypothesised path in the DeLone and McLean model and the $R^2$ for each dependent variable.

The goodness of fit measures for the DeLone and McLean model provided conflicting information. Model $\chi^2$ was 27.74, with 16 degrees of freedom and $\chi^2$ was significant at 0.034. RMSEA was also above the recommended level at 0.097. However, the $\chi^2$/df, GFI, AGFI and TLI all indicated good fit.

The second criterion of good fit considered was the proportion of variance in individual impact and organisational impact explained by the model. The $R^2$ for individual impact was 0.577 (i.e. 57.7% of the variance was explained). Thus, the model explained the variance in individual impact moderately well. However, the $R^2$ for organisational impact was only 0.002, indicating that almost none of the variance in organisational impact was explained by the model.

As can be seen from the $t$ values in Table 5.14, although some of the hypothesised paths were significant a number were not significant. They also varied in strength (see Figure 5.4).
Table 5.14: Model coefficients, standard errors, $t$ values and goodness of fit measures for the DeLone and McLean model and the partial alternate model

<table>
<thead>
<tr>
<th>Path</th>
<th>From</th>
<th>To</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
<th>Partial alternate model</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>System quality</td>
<td>Perceived system quality</td>
<td>System quality</td>
<td>-0.179</td>
<td>0.144</td>
<td>-1.240</td>
<td>-0.186</td>
<td>0.144</td>
<td>-1.252</td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>User satisfaction</td>
<td>Information quality</td>
<td>0.643</td>
<td>0.095</td>
<td>6.798***</td>
<td>0.617</td>
<td>0.100</td>
<td>6.139***</td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>User satisfaction</td>
<td>Perceived system quality</td>
<td>0.310</td>
<td>0.105</td>
<td>2.955**</td>
<td>0.324</td>
<td>0.113</td>
<td>2.876**</td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>Use</td>
<td>Information quality</td>
<td>-0.113</td>
<td>0.258</td>
<td>-0.439</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>Use</td>
<td>Perceived system quality</td>
<td>-0.111</td>
<td>0.195</td>
<td>-0.568</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Use</td>
<td>User satisfaction</td>
<td>0.843</td>
<td>0.336</td>
<td>2.513**</td>
<td>0.637</td>
<td>0.158</td>
<td>4.031***</td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>Individual impact</td>
<td>User satisfaction</td>
<td>-0.183</td>
<td>0.118</td>
<td>-1.547</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Individual impact</td>
<td>User satisfaction</td>
<td>1.131</td>
<td>0.197</td>
<td>5.735***</td>
<td>0.746</td>
<td>0.210</td>
<td>3.543***</td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>Organisational impact</td>
<td>Individual impact</td>
<td>-0.022</td>
<td>0.058</td>
<td>-0.376</td>
<td>-0.022</td>
<td>0.058</td>
<td>-0.383</td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>Individual impact</td>
<td>Individual impact</td>
<td>0.202</td>
<td>0.122</td>
<td>1.655*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>Individual impact</td>
<td>Information quality</td>
<td>0.277</td>
<td>0.174</td>
<td>1.587</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

<table>
<thead>
<tr>
<th></th>
<th>DeLone and McLean model</th>
<th>Partial alternate model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>27.74</td>
<td>25.52</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.034</td>
<td>0.084</td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>1.73</td>
<td>1.50</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.924</td>
<td>0.930</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.828</td>
<td>0.852</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.097</td>
<td>0.080</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.904</td>
<td>0.934</td>
</tr>
</tbody>
</table>

* $p < 0.05$ (one-tailed test)
** $p < 0.01$ (one-tailed test)
*** $p < 0.001$ (one-tailed test)
The results for each of the hypotheses are stated below.

DM1: System quality did not demonstrate a significant influence on perceived system quality, so this hypothesis was not supported.

DM2: Information quality had a strong significant positive influence on user satisfaction, so this hypothesis was supported.

DM3: Perceived system quality displayed a moderate significant positive influence on user satisfaction, therefore this hypothesis was supported.

DM4: Information quality did not demonstrate a significant influence on use, so this hypothesis was not supported.

DM5: Perceived system quality did not demonstrate a significant influence on use. Therefore, this hypothesis was not supported.
DM6: User satisfaction displayed a strong positive significant influence on use, so this hypothesis was supported.

DM7: Use did not demonstrate a significant influence on individual impact. This hypothesis, therefore, was not supported.

DM8: User satisfaction displayed a strong significant influence on individual impact, so this hypothesis was strongly supported.

DM9: Individual impact did not demonstrate a significant influence on organisational impact. This hypothesis, therefore, was not supported.

In addition to the direct relationships reported on in Table 5.14 and Figure 5.4, relationships may be indirect, such that the relationship between two variables in a model is mediated by one or more intervening variables. Table 5.15 below reports the standardised total effects (direct plus indirect) estimated for the DeLone and McLean model.

The total effects reported in Table 5.15 indicate that information quality had a significant indirect effect on both use and individual impact. Perceived system quality also had a significant indirect effect on individual impact, but the effect was only of approximately half the strength of that of information quality. Perceived system quality did not have a significant indirect effect on use.
Table 5.15: Standardised total effects on dependent variables estimated for the DeLone and McLean model

<table>
<thead>
<tr>
<th>System quality</th>
<th>Information quality</th>
<th>Perceived system quality</th>
<th>User satisfaction</th>
<th>Use</th>
<th>Individual impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived system quality</td>
<td>-0.177</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>-0.059</td>
<td>0.697**</td>
<td>0.335*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>-0.021</td>
<td>0.335*</td>
<td>0.117</td>
<td>0.608*</td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>-0.046</td>
<td>0.522**</td>
<td>0.259*</td>
<td>0.725**</td>
<td>-0.188</td>
</tr>
<tr>
<td>Organisational impact</td>
<td>0.002</td>
<td>-0.023</td>
<td>-0.011</td>
<td>-0.032</td>
<td>0.008</td>
</tr>
</tbody>
</table>

* p < 0.05 (two-tailed)
** p < 0.01 (two-tailed)

Note: Table shows total effects of each variable listed across the top of the table on each dependent variable listed in the left-hand column.

5.3.2 Partial alternate model

The goodness of fit measures, model coefficients, standard errors and $t$ values for the partial alternate model are reported in Table 5.14. Figure 5.5 shows the standardised coefficients for each hypothesised path and the $R^2$ for each dependent variable in the partial alternate model. The goodness of fit measures for the alternate model suggest that this model provided a good fit for the data. Model $\chi^2$ was 25.52, with 17 degrees of freedom and was nonsignificant at 0.084, and $\chi^2$/df (1.50) also indicated good fit. GFI, AGFI and TLI all indicated good fit and RMSEA was just within the range that has been considered acceptable (Hair et al., 1998).

The second criterion of good fit considered was the proportion of variance in individual impact and organisational impact explained by the model. The $R^2$ for individual impact was 0.539 (i.e. 53.9% of the variance was explained). Thus, the model explains the variance in individual impact moderately well. However, the $R^2$ for organisational
impact was only 0.002, indicating that almost none of the variance in organisational impact was explained by the model.

Figure 5.5: Structural equation model showing the standardised path coefficient for each hypothesised path and the $R^2$ for each dependent variable in the partial alternate model

The results for each of the hypotheses associated with the alternate model are stated below.

**H1:** System quality did not demonstrate a significant influence on perceived system quality, so this hypothesis was supported.

**H2:** System quality had a weak positive and just significant influence on individual impact. Therefore, this hypothesis was cautiously supported.
H3: Perceived system quality displayed a moderate positive significant influence on user satisfaction, therefore this hypothesis was supported. (Note: this hypothesis is the same as DM3).

H4: Information quality had a strong positive influence on user satisfaction, so this hypothesis was supported. (Note: this hypothesis is the same as DM2).

H5: Information quality did not demonstrate a significant influence on individual impact. Therefore, the hypothesis is not supported.

H6: User satisfaction displayed a moderate positive influence on use, so this hypothesis was supported. (Note: this hypothesis is the same as DM6).

H7: User satisfaction displayed a large significant influence on individual impact, so this hypothesis was supported (Note: this hypothesis is the same as DM8).

H8: Individual impact did not demonstrate a significant influence on organisational impact. This hypothesis, therefore, was not supported. (Note: this hypothesis is the same as DM9).

Table 5.16 below reports the standardised total effects estimated for the partial alternate model. These results confirm the strong indirect effect of information quality on both use and individual impact detected with the Delone and McLean model. Perceived system quality also had a significant indirect effect on use and individual impact, but the effects were much weaker than those of information quality.
Table 5.16: Standardised total effects on dependent variables estimated for the partial alternate model

<table>
<thead>
<tr>
<th></th>
<th>System quality</th>
<th>Information quality</th>
<th>Perceived system quality</th>
<th>User satisfaction</th>
<th>Individual impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived system quality</td>
<td>-0.185</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>-0.063</td>
<td>0.656**</td>
<td>0.344*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>-0.030</td>
<td>0.308**</td>
<td>0.161*</td>
<td>0.469*</td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.127</td>
<td>0.599*</td>
<td>0.196*</td>
<td>0.570**</td>
<td></td>
</tr>
<tr>
<td>Organisational impact</td>
<td>-0.006</td>
<td>-0.027</td>
<td>-0.009</td>
<td>-0.025</td>
<td>-0.045</td>
</tr>
</tbody>
</table>

* p < 0.05 (two-tailed)
** p < 0.01 (two-tailed)

Note: Table shows total effects of each variable listed across the top of the table on each dependent variable listed in the left-hand column.

5.3.3 Comparison of the DeLone and McLean model and the partial alternate model

The two models can be compared on the basis of goodness of fit measures, the proportion of variance explained and the significance of the individual path coefficients. The partial alternate model provided a slightly better overall fit for the data than did the DeLone and McLean model. Both models explained a moderate proportion of the variance in individual impact but neither model was useful in explaining the variance in organisational impact. Both models had some insignificant path coefficients. Five of the nine relationships hypothesised by the DeLone and McLean model were found to be nonsignificant. Of the eight relationships in the partial alternate model, two hypothesised relationships were found to be nonsignificant (note: it was also hypothesised that there would be no relationship between system quality and perceived system quality, and this was supported). Therefore, the alternate model could be considered superior on this criterion.
Despite the superiority of the partial alternate model on two of the three criteria considered it still cannot be considered a good model of UDA success because it does not explain the variance in organisational impact.

5.3.4 Full alternate model

Table 5.17 provides the model coefficients, standard errors, \( t \) values and goodness of fit measures for the full alternate model. Figure 5.6 shows the standardised path coefficient for each hypothesised path and the \( R^2 \) for each dependent variable in the model. The model did not provide a good fit for the data as only \( \chi^2/df \) (1.80) indicated good fit.

As the full model was proposed to provide further insight into the roles of experience and training in influencing individual impact and organisational impact through system quality, perceived system quality and information quality, the variance explained for all these variables is of importance. Figure 5.6 shows the \( R^2 \) for each of these variables. Apart from user satisfaction and individual impact, the model demonstrated little capacity to explain the variance in the variables of interest.
Table 5.17: Model coefficients, standard errors, $t$ values and goodness of fit measures for the full alternate model

<table>
<thead>
<tr>
<th>Path</th>
<th>From</th>
<th>To</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>System quality</td>
<td>0.003</td>
<td>0.036</td>
<td>0.077</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>System quality</td>
<td>-0.007</td>
<td>0.046</td>
<td>-0.143</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>Perceived system quality</td>
<td>-0.071</td>
<td>0.038</td>
<td>-1.881*</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>Perceived system quality</td>
<td>0.042</td>
<td>0.048</td>
<td>0.879</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>Information quality</td>
<td>0.024</td>
<td>0.034</td>
<td>0.693</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>Information quality</td>
<td>-0.002</td>
<td>0.044</td>
<td>-0.040</td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>Perceived system quality</td>
<td>-0.175</td>
<td>0.141</td>
<td>-1.237</td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>User satisfaction</td>
<td>0.625</td>
<td>0.101</td>
<td>6.201***</td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>User satisfaction</td>
<td>0.310</td>
<td>0.112</td>
<td>2.766**</td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Use</td>
<td>0.636</td>
<td>0.159</td>
<td>4.007***</td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Individual impact</td>
<td>0.789</td>
<td>0.217</td>
<td>3.629***</td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>Organisational impact</td>
<td>-0.021</td>
<td>0.054</td>
<td>-0.388</td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>Individual impact</td>
<td>0.211</td>
<td>0.128</td>
<td>1.645*</td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>Individual impact</td>
<td>0.294</td>
<td>0.185</td>
<td>1.594</td>
<td></td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

- Chi-square ($\chi^2$): 50.42
- Degrees of freedom (df): 28
- Probability (p): 0.006
- Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df): 1.80
- Goodness of fit index (GFI): 0.895
- Adjusted goodness of fit index (AGFI): 0.794
- Root mean square error of approximation (RMSEA): 0.101
- Tucker-Lewis index (TLI): 0.852

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$
Figure 5.6: Structural equation model showing the standardised path coefficient for each hypothesised path and the $R^2$ for each dependent variable in the full alternate model

The results for each of the additional hypotheses (beyond those presented with the partial alternate model) are stated below.

H9: Experience did not demonstrate a significant influence on system quality. Therefore, this hypothesis was not supported.

H10: Training did not demonstrate a significant influence on system quality, so this hypothesis was not supported.

H11: Experience did not demonstrate a significant influence on information quality, so this hypothesis was not supported.

H12: Training did not demonstrate a significant influence on information quality. Therefore, this hypothesis was not supported.
H13: Experience had a moderate significant negative influence on perceived system quality. Therefore, this hypothesis was supported.

H14: Training did not demonstrate a significant influence on perceived system quality, so this hypothesis was not supported.

Hence, of the additional six relationships tested in the full alternate model only one was supported. Table 5.18 below reports the standardised total effects estimated for the two additional independent variables included in the alternate model: training and experience (see Table 5.16 above for the total effects of the other variables). These results support the results of the hypotheses testing the direct effects of these variables; neither had a significant indirect effect on any of the dependent variables modelled. The inclusion of experience and training in this model was not very successful in providing additional insight into the UDA success process.

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>System quality</td>
<td>-0.020</td>
<td>0.011</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>0.129</td>
<td>-0.271</td>
</tr>
<tr>
<td>Information quality</td>
<td>-0.005</td>
<td>0.090</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>0.039</td>
<td>-0.029</td>
</tr>
<tr>
<td>Use</td>
<td>0.018</td>
<td>-0.014</td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.018</td>
<td>0.005</td>
</tr>
<tr>
<td>Organisational impact</td>
<td>-0.001</td>
<td>0.000</td>
</tr>
</tbody>
</table>

* p < 0.05 (two-tailed)
** p < 0.01 (two-tailed)
5.4 Overview

This chapter reported the results of the data collection and analyses undertaken during Study 1. Measurement models were presented for the variables considered in the structural models. These measurement models were considered satisfactory, and hence were used to create composite variables for structural model evaluation. The results of the structural model evaluation were also presented in detail. Three structural models were tested. The DeLone and McLean model was tested and compared with the subset of the alternate model that includes the same constructs. This was done to enable a direct comparison on the evaluation criteria. The full alternate model was also tested.

The partial alternate model provided a better overall fit for the data than did the DeLone and McLean model. Both models explained a moderate proportion of the variance in individual impact, but neither model was useful in explaining the variance in organisational impact. Both models had some insignificant path coefficients. Five of the nine relationships hypothesised by the DeLone and McLean model were found to be nonsignificant. Of the eight relationships in the alternate model, two hypothesised relationships were found to be nonsignificant. Therefore, the partial alternate model could be considered superior on the criterion of significance of hypothesised relationships. The full alternate model did not however fit the data well, and only one of the additional six relationships tested in the full alternate model was supported. This model was therefore not very successful in modelling the roles of experience and training in UDA success. The results reported in this chapter are discussed in detail in Chapter 6, and are also considered in the light of the results of Study 2 in Chapters 10 and 11.
Chapter 6
Study 1 Discussion

6.1 Introduction
Two models of UDA success were tested in Study 1 in order to investigate the UDA success process and gain insight into the ability of end user developers to judge application success. The results of the evaluation of these models were presented in Chapter 5 and this chapter presents a discussion of the results. It first considers the results of each of the hypotheses tested and seeks explanations for the results obtained. This is then followed by a discussion of the contributions of the two models. The chapter then assesses what progress has been made towards answering the research questions posed for this study and identifies further research required.

6.2 Discussion of the hypotheses
This section discusses the results of each of the relationships tested in Study 1 (see Section 5.3). Table 6.1 below summarises these relationships, indicating which were supported and which were not. For those relationships supported by the results of Study 1 the relative strength of the relationship is indicated in brackets.
Table 6.1: Summary of the relationships investigated in Study 1

**Relationships supported**
- Experience $\rightarrow$ perceived system quality (moderate)
- Perceived system quality $\rightarrow$ user satisfaction (moderate)
- Information quality $\rightarrow$ user satisfaction (strong)
- User satisfaction $\rightarrow$ use (moderate to strong)
- User satisfaction $\rightarrow$ individual impact (strong)
- System quality $\rightarrow$ individual impact (weak)

**Relationships not supported**
- Experience $\rightarrow$ system quality
- Training $\rightarrow$ system quality
- Experience $\rightarrow$ information quality
- Training $\rightarrow$ information quality
- Training $\rightarrow$ perceived system quality
- System quality $\rightarrow$ perceived system quality
- Information quality $\rightarrow$ use
- System quality $\rightarrow$ use
- Use $\rightarrow$ individual impact
- Individual impact $\rightarrow$ organisational impact
- Information quality $\rightarrow$ individual impact

### 6.2.1 Experience does not influence system quality

It was hypothesised in the alternate model that system quality would reflect experience (hypothesis H9); however, this hypothesis was not supported in Study 1. Despite being more aware of the limitations of their applications (see the discussion of the effect of experience on perceived system quality (H13) in Section 6.2.5 below), those with more experience may not have aimed to develop high quality applications. This would be consistent with the results of Kreie’s (1998) study where she found no relationship between spreadsheet knowledge and spreadsheet quality and concluded that those with knowledge do not necessarily apply it. Panko and Sprague (1999) also found very little difference in spreadsheet development error rates across undergraduates, MBAs with little spreadsheet experience and MBAs with substantial spreadsheet development experience.
If those with more experience do not aim to develop high quality applications it could suggest a lack of awareness of the consequences of using applications of low quality (Ronen, Palley, & Lucas, 1989). A lack of concern for consequences might have been exacerbated by two factors in this study. Firstly, the applications did not form part of the formal assessment for the course, and secondly, the participants were aware that the applications would only be required for a limited period of time (the duration of the simulation). However, these circumstances are often mirrored in the workplace, with no external controls being placed on development and with end users developing applications that they believe will only be used once but then using them repeatedly (Kroenke, 1992). It might also be the case that participants knew that they were developing applications of lower than ideal quality, but believed that they could compensate for any quality ‘short cuts’ in the way they used the application.

6.2.2 Training does not influence system quality

Training was also found not to influence system quality in Study 1 (hypothesis H10). The points mentioned above with respect to experience could also apply to this finding. Thus, even end users with relatively large amounts of training may not have considered it important to develop applications of high quality. In Kreie’s (1998) study of the effect of training on system quality, only half of those who had received specific systems analysis and design training showed improvements in the quality of their spreadsheet designs despite having scored 100% on the training test. The participants in her study knew a process for good spreadsheet design, but chose not to apply it.
The lack of relationship between training and system quality may also relate to the amounts and types of training received by the participants in this study. The participants had received relatively little formal training, and the predominant means of training was self-study (see Table 4.10). It has been suggested that when end users are self-taught the emphasis is predominantly on how to use the software rather than broader analysis and design considerations (Benham, Delaney, & Luzi, 1993). Thus, the participants in Study 1 may not have received training of a type conducive to improvements in system quality. As self-training has been shown to be the major form of training in a number of studies (e.g. Amoroso & Cheney, 1991; Benham, 1993; Chan & Storey, 1996) the results of Study 1 may highlight potential problems in a wide range of organisations.

It is also possible that the measurement of previous training in this study was not sensitive enough to detect the hypothesised relationship. The items used relied on end users’ perceptions of the intensity of their previous training, and it is possible that some end users had trouble gauging how their previous training compared to that of other people.

### 6.2.3 Experience does not influence information quality

The hypothesis that experience would have a positive influence on information quality was not supported (hypothesis H11). This result mirrors the lack of relationship found between experience and system quality, and is consistent with the results of Kreie’s (1998) study, as her system quality measure also addressed information quality.

One of the major benefits proposed for user development of applications has been improved quality of information. Increased quality of information is expected because
end users should have a better understanding of the data (Amoroso, 1988; Brancheau & Brown, 1993). Could this mean that even end user developers with little experience develop systems of high information quality? The mean score for information quality in this study was 5.25 out of 7 (75%), indicating that the participants perceived their systems to have relatively high information quality.

There was no independent measure of information quality in Study 1. Given the lack of relationship found between system quality and perceived system quality, and the lack of correlation between system quality and information quality, it would be valuable to further explore the assumption that all end user developers are ‘experts’ about the quality of information they use. This might provide further insight into the relationship between experience and information quality.

6.2.4 Training does not influence information quality

The hypothesis that training would have a positive influence on information quality was also not supported (hypothesis H12). Whilst this could be due to the nature of the training received, the high perceived levels of information quality of the applications suggest that this may not be the case. It is possible that their expertise with respect to the information required from the applications enables even novice end user developers to develop applications of relatively high perceived information quality (despite a lack of corresponding system quality). For example, applications might appear to produce the required information when first developed but might not be maintainable or be able to handle errors in data input. However, as discussed above, further research is needed to explore the objectivity of end user perceptions of information quality.
6.2.5 Experience influences perceived system quality

It was hypothesised that perceived system quality would be negatively associated with level of experience in the use of the software tools (hypothesis H13). This hypothesis was supported. This suggests that lack of experience might impede the ability of user developers to assess the quality of their applications. This result is consistent with a study by Harmon (1995) on differences in the ability of novice and experienced multimedia users to evaluate hypermedia information resources. The study found that novices were less capable evaluators than experienced users and suggested that this might be because novices did not have adequate conceptual understanding of the dimensions of the evaluation scale. This may also have occurred in Study 1.

Although perceived system quality was not specifically analysed at the dimension of quality level in this study, a subsequent study has considered this (McGill, in press). End user developers with high levels of experience rated their applications lower on the more technical dimensions of maintainability, reliability and understandability than did the low experience end user developers. It seems that despite Rivard et al.’s (1997) concerns about end user awareness of the technical dimensions of quality, experience brings some increase in awareness.

The fact that the relationship between experience and perceived system quality was not very strong, and that no relationship was found between system quality and perceived system quality, suggests that other factors are influencing perceived system quality. This is explored further in Section 6.2.7 and in Chapter 7.
6.2.6  **Training does not influence perceived system quality**

It was hypothesised that perceived system quality would be negatively associated with level of spreadsheet training as training was expected to make end user developers more critical of their applications (hypothesis H14). However, this hypothesis was not supported. In this study, level of spreadsheet training did not appear to play a role in determining either system quality itself or perceived system quality. As has already been discussed for system quality, both the amount of training and types of training that the participants had received could explain the results. The participants may not have received training of a type conducive to reflection on system quality.

6.2.7  **System quality does not influence perceived system quality**

The relationship between system quality and perceived system quality (hypotheses DM1/H1) was of interest because of concerns expressed in the literature about the ability of end users to make realistic judgements of system quality (Edberg & Bowman, 1996). As the system quality construct described in DeLone and McLean’s model had been operationalised in the literature they reviewed as a variety of both subjective and direct measures, it was modelled as two separate constructs for Study 1. As follows from DeLone and McLean’s (1992) discussion, the version of the DeLone and McLean model tested hypothesised a positive relationship between the constructs. However, for the alternate model it was hypothesised that there would be no relationship between system quality and perceived system quality. No significant relationship between system quality and perceived system quality was found in Study 1, so the hypothesis of a positive relationship in the DeLone and McLean model (DM1) was rejected and the hypothesis of no relationship (H1) was accepted for the alternate model. This lack of
relationship between system quality and perceived system quality provides justification for the concerns expressed in the literature about the ability of end users to make realistic judgements of system quality. It should be noted, however, that the measurement model for perceived system quality was the least well fitting of all the measurement models (despite showing good reliability), and it is possible that the lack of relationship is due to measurement problems.

End user developers’ perceptions of system quality might be compromised through lack of knowledge. This was explored in hypotheses H13 and H14 (see Section 6.2.5 and Section 6.2.6) and received some support. Whilst not explored in this study, the nature of differences between end user developer perceptions of quality and independent expert assessments has been explored at the quality dimension level in work subsequent to this study (McGill, in press). End user developers were found to rate the reliability, understandability and userfriendliness of their applications significantly higher than did the independent assessors. These differences were consistent with the findings of Nelson (1991), who identified the major skill deficiencies of end users as being in technical areas, and with those of Edberg and Bowman (1996) who found major data integrity problems with the end user applications examined in their study. Rivard et al. (1997) noted that they would not be surprised to find user attitudes quite impervious to the technical dimensions of application quality as the more ‘technical’ dimensions of quality would be expected to preoccupy information technology professionals but probably not end users unless they had been trained to focus on them.

Huitfeldt and Middleton (2001) argued that the standard system quality criteria are oriented towards information technology maintenance staff rather than end users and
that 'it is still difficult for an end user, or software development client, to evaluate the quality of the delivered product' (p. 3). Although the instrument used to measure perceived system quality in Study 1 was designed specifically for end users (Rivard et al., 1997), informal feedback from participants suggests they found quality assessment a difficult task.

The judgement of end user developers might be further clouded by their close involvement with both the application and with the application development process itself. Cheney, Mann and Amoroso (1986) argued that end user development can be considered as the ultimate user involvement. End user developers are not only the major participants in the development process but also often the primary users of their applications. Applications can come to be viewed as much more than merely problem solving tools. The literature on user involvement indicates that increased involvement is associated with increased user satisfaction (Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Blili, Raymond, & Rivard, 1998; Doll & Torkzadeh, 1988; Lawrence & Low, 1993), and Seddon and Kiew (1994) recognised this by including involvement in the version of DeLone and McLean’s model that they tested. The possibility that high levels of involvement distort the role of user satisfaction in the UDA success process requires investigation and this is discussed further in Chapter 7.

6.2.8 User satisfaction reflects perceived system quality and information quality

As predicted by the DeLone and McLean model (and the alternate model), information quality and perceived system quality were both positively associated with user satisfaction (hypotheses DM3/H3 and DM2/H4). This is consistent with the findings of
Seddon and Kiew (1996) for organisational systems. In discussing their findings, Seddon and Kiew stated that user satisfaction may be interpreted as a response to three types of user aspirations for a system: information quality, system quality and usefulness. Perceptions of information quality and system quality should then explain a large proportion of variance in user satisfaction. The relationship between information quality and user satisfaction was much stronger than the relationship between perceived system quality and user satisfaction. This suggests that user responses to perceptions of the quality of information produced by a UDA are more important than user responses to perceptions of the system itself in determining user satisfaction.

### 6.2.9 Information quality and perceived system quality do not directly influence use

Neither information quality nor perceived system quality was found to influence system use directly, thus the positive relationships hypothesised by the DeLone and McLean model were rejected (hypotheses DM4 and DM5). These relationships were not included in the alternate model. These findings support the arguments made in Chapter 3 that while a system that is not used at all cannot be judged successful, system quality and information quality need not be positively related to frequency of use. A user may need to spend a longer time using a system if the system quality or information quality is poor. The lack of relationship between information quality and use is consistent with Seddon’s (1997a) study investigating measuring individual user, single application IS effectiveness.
The lack of association between quality and use may be particularly marked for UDAs as the high involvement end user developers have with both the development process and the completed system may result in them making more use of the system than is strictly necessary. The average number of hours of use reported by the participants in this study was four hours per week with a maximum of 12 hours per week. This appears to be high given the relative importance of the BPG in their overall programs of study. Amoroso and Cheney (1991) found that motivation to develop applications explained about 50% of the variance in application use. This recognises that use of a UDA is driven by more than just specific information needs.

Analysis of total effects showed that information quality had a moderate significant indirect effect on use via user satisfaction in both models. The indirect effect of perceived system quality on use via user satisfaction was not significant in the DeLone and McLean model but significant in the alternate model and was a much weaker effect than that of information quality. The indirect influence of perceived information quality on intended use has been demonstrated in research on other types of systems (Klobas & Clyde, 2000; Klobas & Morrison, 1999). The lack of evidence for any influence (either direct or indirect) of perceived system quality on intended frequency of use may point to a different influence function. Users may need to use a poor quality system more frequently to meet their needs. Alternatively, they may choose to use a high quality system more frequently because it meets their needs well. Further research is needed to understand reasons for differences in intended frequency of use.
6.2.10 User satisfaction influences use

User satisfaction had a moderate to strong positive influence on use, as predicted by both models (hypotheses DM6/H6). Thus, the more satisfied with an application an end user was, the more they intended to use the application in future. This is consistent with Baroudi, Olson and Ives’s (1986) findings in the organisational domain.

The issue of a two way relationship between use and satisfaction, as shown in DeLone and McLean’s original model, whilst not formally explored in Study 1, was addressed in post hoc analysis. Results indicated that the model was probably unidentified and hence could not be accepted. This analysis does not preclude a more complex relationship, which should be tested in future research: user satisfaction may explain intended use, while actual use may influence subsequent user satisfaction.

6.2.11 Use does not influence individual impact

No significant relationship was found between use and individual impact, thus the positive relationship hypothesised by the DeLone and McLean model was rejected (hypothesis DM7). This relationship was not included in the alternate model. These findings support the argument made in Chapter 3 that the relationship between use and individual impact proposed by DeLone and McLean may not exist, and are consistent with the results of Gelderman (1998), Seddon (1997a) and Roldán and Millán (2000). Seddon (1997a) noted that further work is required to explain why, for different users, levels of use can vary so much for similar perceived impact.
One possible reason for the lack of relationship between use and individual impact was identified earlier in this discussion: higher frequency of use may reflect an inefficient system and therefore low productivity, rather than frequent use in order to obtain substantive benefits. In the UDA domain, an additional issue is that time spent using the system may be confounded with time spent on iterative enhancement of the system, as evolutionary change has been shown to occur in nearly all UDAs (Cragg & King, 1993; Klepper & Sumner, 1990). Frequency of use may be a less valuable indicator of system success in the UDA domain than in the organisational domain, unless researchers are able to differentiate time spent on further development and unproductive work, from time spent using the system to obtain information or to assist directly with decision making.

6.2.12 User satisfaction influences individual impact

User satisfaction displayed a strong influence on individual impact as was hypothesised in both models (hypotheses DM8/H7). This finding is consistent with the results of studies conducted with organisational systems (e.g. Gatian, 1994; Gelderman, 1998; Roldán & Millán, 2000) and is encouraging as it suggests that the reliance of organisations on end user developers’ satisfaction with the applications they develop may not be misplaced.

The major concern with this result is that individual impact was only measured using self-report items. While Gelderman’s (1998) study found user satisfaction to be positively associated with both subjective and direct measures of individual impact, it would be useful to have this finding confirmed in the UDA domain using a direct measure of individual impact. This would ensure that any differences attributable to the
user also being the developer were highlighted. This issue is explored further in Study 2 (see Chapter 7).

6.2.13 Individual impact does not influence organisational impact

Both models hypothesised that increased individual impact would lead to increased organisational impact (hypotheses DM9/H8); however individual impact did not have a significant influence on organisational impact in Study 1. The participants in the study evidently felt their UDAs were contributing to their individual performance, yet this was not reflected in the BPG outcomes. The relationship between individual impact and organisational impact is acknowledged to be complex. Organisational impact is a broad concept, and there has been a lack of consensus about what organisational effectiveness is and how it should be measured (Thong & Chee-Sing, 1996). Roldán and Millán (2000) used four measures of individual impact and four measures of organisational impact in their investigation of the applicability of DeLone and McLean’s model in the executive IS domain. They tested relationships between each possible pair of individual impact and organisational impact measures and obtained inconsistent results.

Whilst changes in the quantitative indicators of organisational effectiveness chosen for this study would provide a clear signal of organisational impact, more subtle impacts may be involved. DeLone and McLean (1992 p. 74) recognised that difficulties are involved in 'isolating the effect of the I/S effort from the other effects which influence organisational performance'. Again, this issue is likely to be magnified in the UDA domain, where system use may be very local in scope. Any changes in organisational impact for a particular organisation would be the result of the combined individual
effects of the UDAs in the organisation, which may well be of varying quality. Individual UDAs could potentially have mutually conflicting effects making it difficult to detect a systematic effect.

In the study in which they reported a relationship between individual impact and organisational impact, Kasper and Cerveny (1985) used direct measures for both constructs. It is possible that perceived individual impact is not a good indicator of direct individual impact in terms of decision making performance, but rather is biased because of factors not included in this model, distorting its relationship with organisational impact. This would suggest that user developers are not only poor judges of the quality of their systems, but could also be poor judges of the impact of their systems on their own performance.

6.2.14 System quality influences individual impact

It was hypothesised in the alternate model that system quality would have a direct positive influence on individual impact (hypothesis H2). As a weak positive and just significant effect was found, this hypothesis was cautiously supported. Hubona and Cheney (1994) tested a model in which user performance using a decision support system was influenced by both user satisfaction and system characteristics (as well as user characteristics). They found a strong relationship between user satisfaction and two direct measures of user performance, but also a significant direct relationship between system characteristics (in this case user interface) and user performance. The direct relationship found between system quality and individual impact in Study 1 is consistent with the results of Hubona and Cheney’s study. If this finding were confirmed with a larger sample and with other types of UDAs it would suggest that at
least part of the effect of a UDA on performance is a direct one, or mediated via constructs not included in these models. However, the weakness of the relationship suggests that the effect is less important than the effect of user satisfaction.

### 6.2.15 Information quality does not influence individual impact

It was also hypothesised in the alternate model that information quality would have a direct positive influence on individual impact (hypothesis H5). This hypothesis was not supported, and the relationship between information quality and individual impact appeared to be mediated through user satisfaction. However, as there was near statistical significance, and the sample size for this study was small, the relationship warrants further study.

### 6.3 Discussion of the models

In order to try to answer the research questions for this study two models of UDA success were tested. Although the alternate model appeared to provide a better fit for the subset of constructs included in the DeLone and McLean model, neither model meets all the criteria established in Chapter 4, and thus neither model can be considered a good model of UDA success. The inclusion of experience and training in the alternate model did not enhance model fit. In fact, the full alternate model provided a less satisfactory fit for the data than did the DeLone and McLean model, and little of the variability in system quality, perceived system quality and information quality was explained. Further research is required to clarify the roles of experience and training in UDA success. It may be that the model is missing other important concepts or that the operationalisations of experience and training were poor.
Nevertheless, the alternate model can be considered to have provided a contribution to research on UDA success. The results suggest that end user developers may not aim to develop good quality applications despite experience and training. It can only be hoped that despite the fact that their applications were not of significantly better quality, the additional insight into the quality of their applications (suggested by the negative relationship between experience and perceived system quality) would lead experienced end users to treat their results with more caution.

The hypothesised model paths that were supported in Study 1 (see Figure 6.1) suggest that the perceived individual impact of a UDA is largely mediated via user satisfaction. Increases in perceived system quality and information quality result in increased user satisfaction, which, in turn, is associated with increased use and increased perceived individual impact. The relative strengths of the influences of perceived system quality and information quality appear to differ, however, with information quality having approximately twice the effect on user satisfaction and individual impact. System quality also appeared to have a weak direct effect on individual impact separate from the effect of perceived system quality mediated via user satisfaction.
A major benefit claimed for user development of applications is improved quality of information because end users should have a better understanding of the information they require. If end users are ‘experts’ with respect to their information, then the strong positive relationship between perceived information quality and user satisfaction is a valuable one. It should reassure organisations that rely on user satisfaction with UDAs as the sole measure of application success that the satisfaction of end users will not be disproportionate to the quality of information provided by the applications, and that end user developers can recognise when use of an application might require caution or be inadvisable. This conclusion, however, rests on the assumption that end user developers are ‘experts’ with respect to the quality of information they use. Given the lack of relationship between system quality and perceived system quality in this study, this assumption should be explored in future research.

The lack of relationship between system quality and perceived system quality suggests another reason for caution on the part of organisations. Most organisations place a heavy reliance on the individual end user’s perceptions of the value of applications they...
develop. If the satisfaction of the user developer is the sole measure of application success, and satisfaction does not reflect system quality, then the benefits anticipated from end user development of applications may be compromised, and organisations may be put at risk.

It appears that Melone’s (1990) caution that the evaluative function of user satisfaction can be compromised by the role of attitude in maintaining self-esteem is particularly relevant in the UDA domain. The literature on user involvement indicates that increased involvement is associated with increased user satisfaction (Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Doll & Torkzadeh, 1988; Lawrence & Low, 1993), and that this might be mediated via increased perceived quality, but if perceived system quality does not reflect conventional notions of system quality, other benefits of higher involvement must be demonstrated.

On the other hand, the observed influence of user satisfaction on perceived individual impact is encouraging. It suggests that organisational reliance on end user developers’ satisfaction with the applications they develop may not be misplaced. However, it would be useful to have this finding confirmed using a direct measure of individual impact, particularly given the lack of a relationship between individual impact (as perceived by the participants) and organisational impact in Study 1. Differences attributable to the user also being the developer could be identified, and an explanation of the relationship between perceived and direct individual impact and organisational impact sought.
The weak direct relationship between system quality and individual impact predicted by the alternate model suggests the possibility that there are two separate subsystems involved in UDA success: one mediated via user satisfaction and one that involves a direct influence of system characteristics on performance. This would mean that problems with system quality that end users are unaware of may be influencing their work performance, and suggests that it may not be wise for organisations to rely entirely upon user satisfaction with UDAs. However, as the relationship is weak and the sample size small, further research is required to establish the existence of the direct relationship.

Seddon (1997b), identifying some problems with DeLone and McLean’s model as a model of IS success, suggested that, rather than a single sequence of relationships, there were two linked sub-systems: one that explained use, and another that explained impact. He argued that use is not an indicator of IS success, but that user satisfaction is because it is associated with impact. There are no published empirical tests of this proposed model, but Study 1 provides support for Seddon's proposal to separate impact measures from one another and from use: there was no evidence of correlation between use, individual impact, or organisational impact. This study, however, does not support Seddon's proposal for two separate sub-systems; rather, it suggests that user satisfaction is a key indicator of subsequent outcomes, including use and individual impact. A single model that explains user satisfaction therefore appears more appropriate in the UDA domain than Seddon's proposed dual system model.

As discussed in Section 2.8.1, Goodhue and his colleagues (Goodhue, 1988; Goodhue, 1995; Goodhue, Klein, & March, 2000; Goodhue & Thompson, 1995) adopted a
different approach to modelling IS success. They proposed that an explanation of IS success needs to recognise the task for which the technology is used, and the fit between the task and the technology, and proposed the technology-to-performance chain (see Figure 2.4). Reflection on Goodhue's concept of task-technology fit suggests that the lack of observed relationship between use and impact in this study may be explained by the need to use the system for more tasks (learning and development) than the functional tasks on which impact (performance) measures were based. Nonetheless, Goodhue's model does not resolve the questions of relationship between use and user attitudes raised by both the results reported here and the criticisms of the DeLone and McLean model offered by Seddon (1997b) and by Ballantine and his colleagues (1998).

6.4 Can the research questions be answered?

Three main research questions were posed for this study addressing parts of the main research questions for the thesis. This section discusses the progress made in answering the research questions.

The first research question was:

*How does UDA quality contribute to user performance on tasks?*

This research question seeks to elucidate the process by which IS success is mediated for UDAs. Neither the DeLone and McLean model nor the alternate model provided a complete explanation of the process. However, the results of Study 1 did contribute new insight into the process. The results suggest that both information quality and perceived system quality influence user satisfaction, which in turn influences individual impact. The results of the study also suggest that system quality (and possibly information
quality) may have a direct influence on individual impact as well as the influence that is mediated by user satisfaction.

The alternate model provided additional clarification of the process beyond that of the DeLone and McLean model. The alternate model clarified the role of use in the UDA success. Whilst increased satisfaction with a UDA was associated with greater use of the application, this greater use was not found to be associated with increased individual impact. Neither perceived system quality nor information quality appeared to play a direct role in the level of usage.

The second research question was:

*Do end user developers have any misconceptions about system quality? If so, how do they impact upon UDA success?*

This research question attempts to address explicitly concerns in the literature about end user developer perceptions of system quality (e.g. Edberg & Bowman, 1996; Kreie et al., 2000; Shayo et al., 1999). End user assessments of system quality were not significantly related to independent expert assessments of system quality in this study, suggesting that end user developers do have misconceptions. Despite this, perceived system quality was significantly related to user satisfaction and via that to individual impact. However, the lack of relationship between system quality and perceived system quality suggests that user satisfaction with applications may be based upon erroneous perceptions. As individual impact was measured using subjective measures only, the answer to this research question could not be fully explored. Further research is required to determine if there are mismatches between perceived individual impact and independent measures of individual impact. If end user developers do not have reliable
perceptions of the quality of applications they develop this may compromise the effectiveness of end users as application developers and could have major consequences when the systems developed are used to support decision making in organisations.

The third research question asked was:

*Do experience and training influence the ability of end user developers to judge whether the applications they develop will have a positive impact on the performance of the tasks the UDA is designed to support?*

End user developers differ from IS professionals in terms of their system development training and experience. This research question addresses the role that experience and training play in the UDA success process. The results of Study 1 suggest that lack of experience may contribute to inaccuracy of end user perceptions of system quality, which in turn influences user satisfaction. As user satisfaction is the sole indicator of application success in many organisations, this could have serious consequences.

Level of training was not found to have a significant influence on perceived system quality. This could have been due to the nature and amount of training end users had received and requires further research.

### 6.5 Further research suggested by Study 1

Whilst the results of Study 1 have provided partial answers to the research questions, they have also raised a number of new questions and suggested areas for further research. This section identifies areas that require future research. Some of these areas are followed up in Study 2, but some are beyond the scope of this thesis. In particular, the small sample size used for the study is acknowledged, and given the sample sizes
generally recommended for SEM (Hair et al., 1998), it would be valuable to have all results confirmed with a larger sample.

6.5.1 The role of use in UDA success

Several issues relating to the role of use in UDA success were raised by the results of Study 1. Intended use and use of systems are major concepts in the main models of IS success (e.g. DeLone & McLean, 1992; Goodhue & Thompson, 1995) and in behavioural intention models such as the technology acceptance model (Davis, 1986). However, despite a strong relationship between user satisfaction and use, use did not appear to influence individual impact. This is inconsistent with some models of user behaviour and IS success but consistent with some previous research (e.g. Gelderman, 1998; Roldán & Millán, 2000). The role of use in UDA success is not investigated further in this thesis, but some of the issues relating to use that may require further research are noted below:

- Study 1 investigated the influence of user satisfaction on use but did not investigate the opposite relationship or the possibility of a reciprocal relationship as specified by DeLone and McLean (1992). This should be tested in future research.

- Users may need to use a poor quality system more frequently to meet their needs. Alternatively, they may choose to use a high quality system more frequently because it meets their needs well. Further research is needed to understand reasons for differences in frequency of use.
Further work is required to understand the lack of relationship between use and perceived individual impact.

6.5.2 The role of information quality in UDA success

The assumption was made in Study 1 that end user developers are ‘experts’ with respect to the quality of information they use. This assumption runs through the literature encouraging user development of applications (Amoroso, 1988; Brancheau & Brown, 1993), but should be explored in future research. However, given the difficulty of obtaining independent assessments of information quality this issue is not explored further in this thesis.

There is also a second reason for omitting information quality from Study 2. An aim of this thesis is the investigation of the role of system quality in UDA success, and omitting information quality should facilitate examination of the relative roles of system quality and perceived system quality. However, the issues relating to information quality that may require further research are noted below:

- Given the lack of relationship found between system quality and perceived system quality, and the lack of correlation between system quality and information quality, it would be valuable to further explore the nature of end users’ perceptions of information quality.

- It was hypothesised in the alternate model that information quality would have a direct positive influence on individual impact. As the effect was not significant, the hypothesis could not be supported. However, further research is warranted given the
small sample size and the $t$ value of 1.587, which was just under the level set for significance of 1.645 ($p = 0.05$).

### 6.5.3 The roles of experience and training in UDA success

Previous experience and training have been hypothesised to play a positive role in determining the impact of user development of applications on individual productivity, and there is some evidence to support this (Babbitt, Galletta, & Lopes, 1998; Kasper & Cerveny, 1985; Nelson & Cheney, 1987). However, the results of Study 1 were inconclusive: neither experience nor training led to applications of improved quality, and only experience appeared to be related to perceived system quality. Further research is required to clarify the roles of experience and training in UDA success and Study 2 addresses this. Issues that require further investigation include:

- The lack of relationship between spreadsheet experience and system quality.
- The lack of relationship between spreadsheet training and system quality.
- The lack of relationship between system quality and perceived system quality.
- The relationship between experience and perceived system quality.
- The relationship between training and perceived system quality.

The possibility that level of involvement might influence the above relationships also needs to be investigated and is explored in Study 2 (see Chapter 7).
6.5.4 Does system quality directly influence individual impact?

Whilst DeLone and McLean’s (1992) model of IS success shows individual impact being mediated via user satisfaction and use, Goodhue’s (1995) model shows technology characteristics having a more direct influence on performance via task-technology fit. Further research is required to clarify the relationship between system quality and individual impact. This is undertaken in Study 2. Issues relating to this relationship that may require further research are noted below:

- Study 1 suggested that there is a direct relationship between system quality and individual impact. However, as the relationship was weak and the sample size small, further research is required to establish the existence of the direct relationship.

- Further research is required to determine if there are mismatches between perceived individual impact and independent measures of individual impact.

6.5.5 The role of individual impact on organisational impact

Neither model provided insight into the relationship between individual impact and organisational impact in the UDA domain. This requires further research, but is beyond the scope of this thesis.
6.6 Overview

This chapter presents a discussion of the results of Study 1. This study has contributed to the literature on UDA success by providing the first empirical test of the entire DeLone and McLean model in the UDA domain. Overall, the model was not supported by the data. Of the nine hypothesised relationships tested by SEM, four were found to be significant while the remainder were not significant. The analysis did, however, provide strong support for relationships between perceived system quality and user satisfaction, information quality and user satisfaction, user satisfaction and intended use, and user satisfaction and individual impact. An alternate model of UDA success was also tested. Of the 13 relationships hypothesised, only six were found to be significant.

It is notable that the model paths that were supported in Study 1 are primarily those that reflect user perceptions. User satisfaction reflects a user's perceptions of both quality of the system itself and the quality of the information that can be obtained from it. Intended ongoing use of the UDA reflects user satisfaction, and the impact that an individual feels a UDA has on their work reflects their satisfaction with it. The results of this study indicate that user perceptions of IS success play a significant role in the UDA domain.

In the DeLone and McLean model no significant paths were found involving the independently measured constructs (system quality and organisational impact). System quality did not influence perceived system quality, and individual impact as perceived by the participants did not influence organisational impact. The results of the test of the alternate model did, however, suggest that there may be a direct relationship between system quality and individual impact. Further research is required to understand the
relationship between user perceptions of the elements of IS success and independent measures of success, and to provide a model of IS success appropriate to end user development. Several other issues requiring further research were also raised but were considered beyond the scope of this thesis. These included the roles of use and information quality in UDA success. The remaining chapters of this thesis describe further research designed and conducted to answer some of the questions raised by Study 1.
Chapter 7
Study 2 Research Questions and Conceptual Model

7.1 Introduction
This thesis is concerned with understanding the UDA success process, specifically addressing three research questions relating to the ability of end user developers to judge application success. In Study 1, two models of IS success were used as frameworks for the investigation. The results of the study provided answers to some of the questions asked, but also raised new questions and highlighted areas where further research was necessary. Most of the model paths that were supported were those that reflect user perceptions, indicating that user perceptions of IS success play a significant role in the UDA domain. The results of the test of the alternate model did, however, suggest that there may be a direct relationship between system quality and individual impact. Study 2 was designed to pursue some of the issues raised in Study 1 and hence provide further insight into the research questions for the thesis.

This chapter is the first of four chapters that describe Study 2 and presents the research questions, conceptual framework and hypotheses for the study. The chapter first highlights the specific aspects of the overall research questions for the thesis that are addressed in this study. It then presents the research model used as a framework for the study. This model includes two new concepts that were proposed in the discussion of Study 1 as having a possible role in explaining the findings: spreadsheet development knowledge and involvement. The research hypotheses derived from the model are then presented along with previous research to support them. An additional set of hypotheses
is also proposed to enable comparison of end user developers using their own applications with end users using applications developed by other end users.

### 7.2 Study 2 research questions

The objective of the research described in this thesis is to gain a better understanding of UDA success. In particular, the research considers the role of system quality in UDA success and the ability of end user developers to judge whether the systems they develop will have a positive impact on their performance of tasks. Chapter 1 states the research questions for the thesis, and Chapter 6 discusses the contribution Study 1 made towards answering these questions. This section identifies the specific aspects of the research questions to be investigated in Study 2.

The first research question for the thesis is:

*How does UDA quality contribute to user performance on tasks?*

Neither of the models tested in Study 1 provided a completely satisfactory explanation of how UDA quality contributes to user performance on tasks. Although perceived system quality, user satisfaction and individual impact as perceived by the participants were all strongly associated, no relationship was found between system quality and perceived system quality, nor between individual impact and organisational impact. If perceived system quality does not reflect system quality, it raises the issue of how UDA success is mediated. A limitation of Study 1 is that individual impact was measured using self-report items only. It is possible that perceived individual impact is not consistent with independently assessed individual impact, but rather is biased because of factors not included in the models evaluated in Study 1, distorting its relationship with organisational impact. This could suggest that user developers may not only be
poor judges of the quality of their systems, but could also be poor judges of the impact of their systems on their own performance. Study 2 extends the work undertaken in Study 1 to clarify the roles of both perceived individual impact and independently measured individual impact in UDA success, allowing this research question to be more fully addressed.

The second research question for this thesis is:

\[ \text{Do end user developers have any misconceptions about the quality of their applications? If so, how do these misconceptions impact upon their ability to judge whether the applications they develop will have a positive impact on their performance of tasks?} \]

In Study 1, no relationship was found between perceived system quality and system quality. This finding is of major concern, and Study 2 provides an opportunity to compare more closely end user perceptions of both system quality and individual impact with independent measures. It was considered important both to investigate possible reasons for the lack of relationship between perceived system quality and system quality and to explore the impact of this lack of relationship on the ability of end user developers to judge whether the applications they develop will have a positive impact on their performance of tasks.

The third research question asked in this thesis is:

\[ \text{What characteristics of end user developers influence their ability to judge whether the applications they develop will have a positive impact on their performance of the tasks the UDA is designed to support?} \]
Study 1 provided an initial investigation of the roles of experience and training in UDA success; however, the results were inconclusive. Neither amount of previous spreadsheet experience nor amount of previous spreadsheet training influenced system quality. Several possible reasons for this were suggested in Chapter 6; however, these findings demand further investigation. Study 2 was designed to extend the investigation undertaken in Study 1 to try to determine more clearly the roles played by experience and training. It introduces two additional concepts, spreadsheet development knowledge and involvement, to help investigate where the hypothesised relationships between experience and both perceived and independently measured system quality, and training and both perceived and independently measured system quality might be breaking down. The study also explicitly explores whether involvement plays a role in influencing perceived system quality.

7.3 Framework for the study
A revised research model was developed as a framework for Study 2. It addresses a subset of the constructs considered in Study 1, but introduces two new constructs not considered in Study 1: spreadsheet development knowledge and involvement. It also explicitly acknowledges that end users’ perceptions of the impact of a UDA on their performance may differ from independent measurements of impact by modelling these as two separate constructs. Figure 7.1 shows the model developed for this study. The next two sections describe the additional constructs included in the model. The hypotheses associated with the model are then presented and justified in Section 7.4.
7.3.1 **Spreadsheet development knowledge**

Many EUC studies have measured previous experience (Chan & Storey, 1996; Harrison & Rainer, 1992; Kasper & Cerveny, 1985; Rivard & Huff, 1988) and training (Chan & Storey, 1996; Nelson & Cheney, 1987) as individual characteristics of importance (see Chapter 2 for a detailed discussion of this research). The implicit assumption in this research is that experience and training lead to greater levels of end user knowledge and skill.

Cheney and Nelson (1988) made an early attempt to measure end user computing abilities, developing an instrument that measured three dimensions: technical abilities, modelling abilities, and applications abilities. A similar construct, user competence, has also been addressed in several studies (Marcolin, Compeau, Munro, & Huff, 2000; Munro et al., 1997). Marcolin et al. (2000 p.38) define user competence as ‘the user's
potential to apply technology to its fullest possible extent so as to maximize
performance of specific job tasks’. The construct they propose is a general one relating
to the range of an end user’s skills and has three dimensions: breadth, depth and finesse.
Other terms have also been used to describe similar or related constructs. For example,
computer proficiency (Nelson, 1991), computer literacy (Rainer & Harrison, 1993) and
user sophistication (Blili, Raymond, & Rivard, 1996; Marcolin, Huff, & Munro, 1992;
Zinatelli, Cragg, & Cayaye, 1996) have also been used.

Very little research has looked explicitly at spreadsheet development knowledge or
skill. Both Kreie (1998) and Marcolin et al. (2000) measured spreadsheet knowledge
but used only multiple choice questions that focused specifically on knowledge of
spreadsheet features in a particular package. Kreie did not report the reliability of her
measure and Marcolin et al. found theirs to have low reliability. Hall (1996) called for
the development of a spreadsheet expertise metric, noting that expertise is a difficult
feature to self-assess, particularly for end users who may have little understanding of the
variation within the broader developer population.

Bowman (1988) noted that the systems development expertise of an end user developer
is a function of two dimensions: expertise in application development methods, and
expertise with the development tool or language used to develop the application.
Spreadsheet development knowledge should thus be considered to include both
dimensions. Therefore, it is defined in this study as: the knowledge of spreadsheet
software features and spreadsheet development practices that end user developers draw
upon when developing spreadsheet applications. Thus spreadsheet development
knowledge determines what end user developers are capable of doing rather than what they actually do.

7.3.2 Involvement

As discussed in Chapter 2, Barki and Hartwick (1989 p.53) defined user involvement as ‘a subjective psychological state, reflecting the importance and personal relevance of a system to the user’. They also noted that users may also attach personal importance and relevance to the activities of systems development, thereby becoming involved not just with a system but with a process. Their definition of user involvement has been adopted for this thesis. Prior to Barki and Hartwick’s work, the terms user involvement and user participation had been used interchangeably. Barki and Hartwick drew attention to the difference between the concepts and defined user participation as ‘a set of behaviors and activities performed by users in the systems development process’ (Barki & Hartwick, 1989 p.53). Participation is presented as a precursor to involvement in the subsequent user involvement literature (McKeen, Guimaraes, & Wetherbe, 1994), although Lin and Shao (2000) did not find that user participation was significantly associated with user involvement.

The early instruments designed to measure involvement (e.g. Baroudi, Olson, & Ives, 1986; Doll & Torkzadeh, 1988; Franz & Robey, 1986) in fact measured user participation. Barki and Hartwick (1991) developed an involvement scale for IS based on the involvement scale for products proposed by Zaichkowsky (1985). This scale is the basis for the measure used in Study 2.
7.4 Hypotheses

This section presents and justifies the hypotheses tested in Study 2. Figure 7.1 above shows all of the hypothesised relationships and Figure 7.2 to Figure 7.5 illustrate different parts of the model.

7.4.1 How do experience and training impact upon system quality?

As discussed in Chapter 2, whilst intuitively increased experience and training would be expected to lead to improved quality of UDAs, previous research has been inconclusive (e.g., Amoroso, 1986; Janvrin & Morrison, 2000). In Study 1 neither training nor experience significantly influenced system quality. In Study 2 the role of spreadsheet development knowledge as an intervening variable was considered important as there has been very little research that looked explicitly at spreadsheet development knowledge. Figure 7.2 summarises the relationships considered in the first three hypotheses.

Figure 7.2: The hypothesised relationships between experience, training, spreadsheet development knowledge and system quality
Harrison and Rainer (1992) found that end users with more computer experience reported higher levels of computer skill, however these results were not specific to spreadsheet knowledge and were self-reports. Regular use of a spreadsheet over time should provide repeated opportunities to encounter advanced features of the software. The user sees toolbar icons for, as yet, unused features and when attempting to solve a new problem has recourse to the help menu. However, regular use of a spreadsheet will not necessarily ensure that the end user learns more about system design processes or about quality assurance. Nevertheless, on balance, more experienced end user developers would be expected to have higher levels of knowledge, hence it was hypothesised that:

H1: Spreadsheet development knowledge reflects previous spreadsheet experience.

Despite the conventional wisdom that training leads to an increase in knowledge, in a survey of 254 end users Chan and Storey (1996) found no relationship between training and self-reported spreadsheet proficiency. This lack of association between training and spreadsheet knowledge may, however, be an artefact of the use of self-report measures. As Yaverbaum and Nosek (1992) concluded, training can increase one’s expectations and hence actually cause negative perceptions. For Study 2, in which it was intended that spreadsheet development knowledge be measured independently, it was hypothesised that:

H2: Spreadsheet development knowledge reflects previous spreadsheet training.
Kreie (1998) found no relationship between spreadsheet knowledge and spreadsheet quality. However, she speculated that this was because all of the subjects in her study had high levels of spreadsheet knowledge. Her spreadsheet knowledge instrument also focused specifically on spreadsheet features rather than spreadsheet design knowledge. End users may have a narrow knowledge focused on spreadsheet features but be lacking in techniques for developing spreadsheets that are userfriendly, reliable, and maintainable.

Taylor, Moynihan and Wood-Harper (1998) found that few, if any, design principles are applied in end user development. However, in a study on end user development of databases, Agboola (1998) found a strong relationship between the data modelling knowledge of the developer and database quality. It would seem to follow that development of good quality spreadsheet applications requires not only knowledge of spreadsheet features but knowledge of design and development techniques. Given the more inclusive definition of spreadsheet development knowledge adopted in this thesis, it was hypothesised that:

H3: The system quality of a UDA reflects the end user developer’s level of spreadsheet development knowledge.

### 7.4.2 What factors influence perceived system quality?

Whilst there has been an implicit assumption in the organisational IS literature that user perceptions of system quality reflect ‘actual’ system quality (DeLone & McLean, 1992) there have been concerns expressed in the literature about the ability of end users to make realistic judgements of system quality (e.g. Edberg & Bowman, 1996; Kreie et al.,
2000; Shayo et al., 1999). The results of Study 1 showed that this concern is justified, and several possible reasons for the lack of relationship found between perceived system quality and system quality were proposed in Chapter 6. For Study 2 it was hypothesised that system quality influences perceived system quality, but that this relationship may be perturbed by the influences of spreadsheet development knowledge and involvement. Figure 7.3 shows these hypotheses.

![Diagram](image)

Figure 7.3: The hypothesised relationships between system quality, spreadsheet development knowledge, involvement and perceived system quality

It was hypothesised that:

H4: Perceived system quality reflects system quality.

However, end users’ perceptions of system quality might be compromised because of lack of knowledge. In particular, end users with low levels of knowledge may not recognise system quality problems and hence may have inflated perceptions of quality, whereas end users with high levels of knowledge may become more critical. As
discussed earlier, Yaverbaum and Nosek (1992) raised this possibility with respect to the role of training. Therefore, for this study it was hypothesised that:

H5: Perceived system quality **decreases** as the end user’s level of spreadsheet development knowledge increases.

The role of involvement in end user perceptions of system quality is also of interest. User developers’ judgement might also be clouded by their close involvement with both the application and with the application development process itself. Cheney, Mann and Amoroso (1986) argued that end user development can be considered as the ultimate user involvement. End user developers are not only the major participants in the development process but also often the primary users of their applications. Applications can come to be viewed as much more than merely problem solving tools. The literature on user involvement indicates that increased involvement is associated with increased user satisfaction (Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Doll & Torkzadeh, 1988; Lawrence & Low, 1993). In Study 2 it was proposed that this increased satisfaction is mediated through increased levels of perceived system quality. Hence it was hypothesised that:

H6: Perceived system quality reflects involvement.

**7.4.3 How do system quality and perceived system quality influence user satisfaction, perceived individual impact and individual impact?**

The possibility that perceived individual impact is not consistent with direct measures of individual impact was raised in the discussion of Study 1 (see Chapter 6). Further
investigation of the relationships of both system quality and perceived system quality, and user satisfaction, perceived individual impact and individual impact is required. The hypothesised relationships between these constructs are summarised in Figure 7.4.

Figure 7.4: The relationships proposed to explain how system quality and perceived system quality influence user satisfaction, perceived individual impact and individual impact

The small direct influence of system quality on individual impact as perceived by the participants in Study 1 suggests that at least part of the effect of a UDA on performance is a direct one, not mediated by user satisfaction. This is consistent with a study by Hubona and Cheney (1994) in which they found a strong relationship between user satisfaction and two objective measures of user performance, but also a significant direct relationship between system characteristics (in this case user interface) and user performance. Therefore it was hypothesised that:

H7: Individual impact reflects system quality.
In addition to the results from Study 1, there is substantial previous evidence to support the existence of positive relationships between perceived system quality and user satisfaction (Rivard et al., 1997; Seddon & Kiew, 1996) and between user satisfaction and perceived individual impact (Gatian, 1994; Gelderman, 1998). Hence it was again hypothesised that:

H8: User satisfaction reflects perceived system quality.

H9: Perceived individual impact reflects user satisfaction.

Measures such as perceived individual impact are commonly used as surrogates for ‘actual’ impact on individual performance (DeLone & McLean, 1992). However, user evaluations have been criticised as lacking a clearly articulated theoretical basis for linking them to systems effectiveness (Goodhue, Klein, & March, 2000). Goodhue, Klein and March (2000) noted that there has been little research that explicitly tests the link between user evaluations and objectively measured performance. They used an experiment to test the relationship between user evaluations of task-technology fit and performance and found significant support for one measure of performance but not for a second. In Study 2, ‘actual’ individual impact is conceptualised as an antecedent to perceived individual impact such that an end user’s perception of impact is formed in part by objective evidence of the system’s impact. This evidence might include reduced time to make decisions or to undertake a task, or increased accuracy of decision making. It was therefore hypothesised that:

H10: Perceived individual impact reflects individual impact.
As previously mentioned, Hubona and Cheney (1994) found a direct relationship between system characteristics and objective measures of user performance, as well as a relationship mediated by user satisfaction. Therefore it is necessary to determine the relationship of user satisfaction with individual impact as well as with perceived individual impact. Consistent with the literature, it was hypothesised that:

H11: Individual impact reflects user satisfaction.

7.4.4 Does the spreadsheet development knowledge of the user play a role in their ability to successfully use a UDA?

The previous hypotheses are consistent with what might be expected for the use of organisational IS. However, UDAs are generally of lower quality than organisational IS (Cale, 1994; Edberg & Bowman, 1996; Palvia, 1991) and successful use may require more from the user than would successful use of an organisational system. In particular, use of user developed spreadsheet applications may require substantial prior knowledge because of the lack of separation of data and processing that is commonly found (Hall, 1996; Rajalingham, Chadwick, Knight, & Edwards, 2000; Ronen, Palley, & Lucas, 1989). For example, if data is embedded within formulas, the user must be able to edit formulas in order to update data.

![Diagram](image)

Figure 7.5: The hypothesised relationship between spreadsheet development knowledge and individual impact
Hence it was hypothesised that:

H12: Individual impact reflects the level of spreadsheet development knowledge of the application user.

### 7.4.5 How do end user developers differ from other end users?

UDAs are commonly used by other end users as well as by the developer (Hall, 1996). Whilst a main aim of this thesis is to investigate the ability of end user developers to judge whether the systems they develop will have a positive impact on their performance of tasks, the UDA success process for end user developers using applications developed by other end users is also of interest. The model proposed for this study (Figure 7.1) is intended to apply to both end users using applications they have developed themselves, and to end users using applications developed by other end user developers. If the relationships observed are consistent for both groups, then this will help establish the concurrent validity of the model. However, the literature on end user development suggests that end user developers who use their own applications may differ from other end users on a number of criteria. In a further analysis of the data from Study 1, end user developers perceived their applications to be of higher quality on some dimensions than did independent assessors (McGill, in press). Doll and Torkzadeh (1989) found much higher involvement scores for end user developers than for users who were involved in the development process but where the application was primarily developed by a systems analyst or another end user. End user developers were also found to be more satisfied with applications they had developed themselves than with
applications developed by another end user (McGill et al., 1998) or with applications developed by a systems analyst (despite involvement in the systems development process) (Doll & Torkzadeh, 1989). In a laboratory experiment, Kasper (1985) identified improvements in both decision accuracy and speed by end user developers using applications they had developed when compared to end users using existing decision support tools.

The involvement construct was included in the model for Study 2 to account for a prime difference claimed about end user developers: their high levels of involvement. However, in order to ensure that any specific differences due to the user being the developer could be identified the set of hypotheses below were also developed and tested.

H13: End user developers perceive their own applications to be of higher system quality than applications developed by another end user with a similar level of spreadsheet knowledge.

H14: End user developers have higher levels of involvement with their own applications than with applications developed by another end user with a similar level of spreadsheet knowledge.

H15: End user developers have higher levels of user satisfaction when using their own applications than when using applications developed by another end user with a similar level of spreadsheet knowledge.
H16: End user developers have higher levels of perceived individual impact when using their own applications than when using applications developed by another end user with a similar level of spreadsheet knowledge.

H17: End user developers make more accurate decisions when using their own applications than when using applications developed by another end user with a similar level of spreadsheet knowledge.

H18: End user developers make faster decisions when using their own applications than when using applications developed by another end user with a similar level of spreadsheet knowledge.

7.5 Overview

This chapter is the first of four chapters describing the second study undertaken during the research described in this thesis. This chapter presents the research questions and conceptual model for Study 2. The objective of the research described in this thesis is to better understand UDA success by gaining a better understanding of both the role of system quality in UDA success and of the ability of end user developers to judge whether the applications they develop will have a positive impact on their performance of tasks. Study 2 was designed to pursue some of the issues raised in Study 1 and hence provide further insight into the research questions for the thesis. In particular, it was intended to:

- Clarify the roles of both perceived individual impact and independently measured individual impact in UDA success
• Investigate possible reasons for the lack of relationship demonstrated between perceived system quality and system quality in Study 1

• Investigate where the hypothesised relationships between experience and both perceived system quality and system quality, and training and both perceived system quality and system quality, might be breaking down.

A revised research model was proposed to support the research in Study 2. The model draws upon both the results of Study 1, and the UDA and IS success literature, and includes two concepts that were not included in the Study 1 models: spreadsheet development knowledge and involvement. The model provides an opportunity to compare more closely end user perceptions of IS success measures with independent measures and to explore the mechanisms by which UDA success is achieved. The hypotheses associated with this model were presented and justified.

An additional set of hypotheses was also proposed to enable comparison of end user developers using their own applications with end users using applications developed by other end users. UDAs are commonly used by other end users as well as by the developer (Hall, 1996), so any differences in outcomes between these two types of users should be of interest to organisations trying to maximise the efficiency and effectiveness of UDA use.
Chapter 8
Method for Study 2

8.1 Introduction
This chapter describes the methodology used to carry out Study 2. The methodology is first described in general terms and reasons for the choice explained. There is then a detailed description of the methodology including the participants, the experimental procedure, and the development of the measurement instruments. The chapter also provides a discussion of the statistical techniques used to answer the research questions. The chapter concludes with descriptive information about the participants and about the applications they developed during the study.

8.2 Methodology
Study 2 was a laboratory experiment and the UDAs considered were spreadsheet applications. The participants were end users recruited from a variety of organisations. They developed spreadsheet applications to assist in a hypothetical problem scenario and then used them to support decision making. Both UDA system quality and individual impact were assessed independently and end user perceptions of quality and impact were obtained via survey instruments.

Brancheau and Brown (1993) called for research questions on end user development to be better defined and narrower in scope, and encouraged the use of experimental designs to supplement previous field studies with data collection by interview and questionnaire. As Study 2 was designed to build upon the results of the field study
undertaken during Study 1 and to clarify and extend its results, a laboratory situation was appropriate. Laboratory research offers the advantage of controlled manipulation of independent variables, minimisation of the impact of extraneous variables, and a high level of internal validity (Emory & Cooper, 1991). The specific benefits of a laboratory situation for this study were: the ability to specify the task to be undertaken so that all participants and UDAs could be compared directly; the ability to measure individual impact directly; and the ability to compare end user developers using their own applications with end users using applications developed by another end user.

This study used a within-subjects research design as this has been shown to provide superior control for individual subject differences (Maxwell & Delaney, 1990). This kind of design also offers more data points than a between-subjects design, and this was considered valuable given the sample size requirements for SEM (Hair et al., 1998). The participants in Study 1 were end users who were postgraduate business students. A broader range of end users was sought for Study 2 to ensure that the results would be generalisable to the population of all end user developers in organisations.

Spreadsheets were again selected as the development tool. As previously discussed, spreadsheets are the most commonly used tool for end user development of applications (Taylor, Moynihan, & Wood-Harper, 1998), and by studying their use maximum generalisability of results was hoped for. The use of spreadsheets also ensured that the results of Study 2 would be directly comparable with the results of Study 1.
8.3 Participants

The target population for this study was all end users who develop their own applications using spreadsheets. As discussed above, given that Study 1 only included MBA students it was important for the external validity of the results of the thesis that Study 2 included a more representative sample of end user developers. Participants were recruited for the study by advertisements in the local press (see Appendix C), e-mails to three large organisations, and word of mouth. The criterion for inclusion in the study was previous experience using Microsoft Excel©. As an incentive to participate, end users who completed the experimental session were offered a one hour training course entitled ‘Developing Spreadsheet Applications’. This session focused on spreadsheet planning, design and testing (see Appendix D for copy of training material provided to participants). They were also given $20 to compensate them for petrol and parking costs, and inconvenience.

The minimum sample size recommended for SEM is 150 (Hair et al., 1998); therefore at least 150 participants were needed. One hundred and fifty nine end users participated in Study 2.

8.4 Experimental Session

Fourteen separate experimental sessions of approximately four hours were held over a period of five months. Each session involved between seven and 17 participants (depending on availability) and was held in one of two 15 PC computer laboratories. Each session consisted of four parts. The sequencing and timing of the sessions was piloted with four end users who were not involved in the study. Table 8.1 summarises
the sequence of the steps followed, and a full description of the activities undertaken in
each experimental session is given below.

Table 8.1: Experimental session outline

<table>
<thead>
<tr>
<th>Part</th>
<th>Activities</th>
<th>Approx. Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduce study and obtain participant consent (see Appendix E) Collect background information and assess spreadsheet knowledge (see Appendix F)</td>
<td>30 min.</td>
</tr>
<tr>
<td>2</td>
<td>Develop spreadsheets (see Figure 8.1 for the problem statement)</td>
<td>1.5 hour</td>
</tr>
<tr>
<td>3</td>
<td>Use spreadsheets to answer decision questions (see Appendix G) and complete perceived system quality, involvement, user satisfaction and perceived individual impact questions (see Appendix H)</td>
<td>1 hour</td>
</tr>
<tr>
<td>4</td>
<td>Training session (see Appendix D)</td>
<td>1 hour</td>
</tr>
</tbody>
</table>

In Part 1 participants were asked to complete a written questionnaire to provide demographic information about themselves, and information about their background with computers and spreadsheets. This questionnaire also tested their knowledge of spreadsheets (see Appendix F). Section 8.6.4 describes the development and composition of the spreadsheet knowledge instrument.

Participants were given a problem statement in Part 2 of the session and asked to develop a spreadsheet to solve it using Microsoft Excel 97© (the development of this task and the associated problem statement is described in Section 8.5). The problem related to making choices between car rental companies (see Figure 8.1 for the complete problem statement). Participants were provided with blank paper to use for planning if
they desired, but otherwise were left to develop the application as they wished. They were encouraged to treat the development exercise as they would a task at work, rather than as a test. Participants could use on-line help or ask for help from the two researchers present in the laboratory during each session.

Once all participants in the session had completed their spreadsheet (and had a coffee break) they undertook Part 3 of the session. In order to ensure that the model proposed in Chapter 7 applies to both the situation in which an end user developer uses their own application, and the situation in which an end user uses an application developed by another end user, each participant was asked to participate in both situations. Each participant was given a floppy disk containing both the spreadsheet they had developed and a spreadsheet from another participant in the session. The spreadsheet they received was from the person with the closest spreadsheet development knowledge score to their own. This matching was done in the expectation that participants with a similar level of spreadsheet knowledge would develop spreadsheets of similar sophistication.

This within-subjects design also provided a way to ensure that any specific differences due to the user being the developer could be identified (hypotheses 13 to 18). Use of a within-subjects design creates a need for stringent learning effect controls (Cook & Campbell, 1979). To control for presentation order effects, each participant was randomly assigned to use either their own or the other spreadsheet first. They then used this spreadsheet to answer 10 questions relating to making choices about car rental hire (See Appendix G for the complete set of questions and Section 8.6.9 for a description of the development of these questions). The time taken to answer these questions was
recorded. Participants were told that they could make corrections and adjustments to the
spreadsheets, but not to attempt to redesign them, as the intention of this part of the
session was to use the spreadsheets, not redevelop them.

After completing the questions, participants then completed a questionnaire containing
items to measure: perceived system quality, involvement, user satisfaction and
perceived individual impact (see Appendix H for the complete set of questions and
Section 8.6.5 to Section 8.6.8 for descriptions of the development of these items). Once
the questionnaire and their answers to the car rental decision questions were collected
each participant then repeated the process with the other spreadsheet on their floppy
disk. A different but equivalent set of car rental decision questions was used (see
Appendix G).

The final part of the session was a complimentary optional training session that focused
on learning to use a structured process to develop spreadsheets (see Appendix D for the
training notes provided to participants). Participants were also given twenty dollars. The
informal feedback from participants about the session was very positive. Many
commented how on how much they had learned from the experience in general, and in
particular they valued the opportunity to see how another end user had tackled the same
problem. Overall at least 80% of the participants remained for the training session.
8.5 Development of the experimental task

The following requirements were identified for the task to be used in Study 2 and for the resulting UDAs.

- The goals of the task needed to be at least partially achievable by novices.
- No specific domain knowledge should be required.
- The task should be of interest to a wide range of potential participants.
- The UDAs developed needed to be complex enough to allow users to answer a range of questions, so that they could gain enough experience with their applications to be able to complete the questionnaire measuring perceived system quality, involvement, user satisfaction and perceived individual impact.
- The UDAs developed needed to be sufficiently complex so that usage would have a benefit (over not using) and would provide a range of possible individual impacts.
- It had to be possible to develop a UDA to solve the problem in less than one and a half hours in order to keep the time demands on participants feasible.

Kreie (1998) used two spreadsheet problems in her experiment that investigated an approach to improving application quality. Both were similar and fairly simple (each could be completed in less than half an hour). She found no effect of systems analysis and design training on the completeness or accuracy of the spreadsheets (just on design quality) and suggested that the task may have been too simple. Neither of her problem statements would result in applications supporting the level of decision making required for Study 2. So whilst being mindful of the need to enable novice end user developers to at least develop a partial solution, a task was chosen that allowed for a wide range of sophistication.
The initial problem statement for the chosen task was reviewed by two IS academics with substantial experience in teaching spreadsheet development, and feedback was provided on the task’s ability to meet the requirements. Several modifications were made in response to this feedback. The task was also piloted with four spreadsheet end users following which several additional minor wording changes were made. The final problem statement is shown in Figure 8.1 below.

**CAR RENTAL PROBLEM**

Deciding which car rental company to choose when planning a holiday can be quite difficult. A local consumer group has asked you to set up a spreadsheet to help people make decisions about car rental options. The spreadsheet will enable users to determine which company provides the cheapest option for them, given how long they need to hire a car and how much driving they intend to do.

After investigating the charges of the major companies you have the following information about the options for hiring a compact size car in Australia.

- Advantage Car Rentals charge $35 per day for up to 100 kilometres per day. Extra driving beyond 100 kilometres per day is charged a $0.25/km excess.
- OnRoad Rentals charge $41 per day. This rate includes 200 free kilometres per day. Extra kilometres beyond that are charged at the rate of $0.30/km.
- Prestige Rent-A-Car charge $64 per day for unlimited kilometres.

Your task is to create a spreadsheet that will allow you or someone else using it to type in the number of days they will need the car and the number of kilometres they expect to drive over the time of the rental. The spreadsheet should then display the rental cost for each of the above three companies.

Figure 8.1: Problem statement given to participants in Part 2 of the experimental session

### 8.6 Development of the measurement instruments

The development of the research instruments for Study 2 involved a review of many existing survey instruments. To ensure the reliability and validity of the measures used, previously validated measurement scales were adopted wherever possible. As noted
below, several of the instruments were based on those used in Study 1 and several were
exactly as used in Study 1.

8.6.1 Background information

In order to enable description of the participants and to enable comparison between
studies, the following background information was collected for each participant:

• Age in years
• Gender
• Number of years of computing experience
• Frequency of use of computers, measured using the six category scale from Igbaria
  (1990)
• Perceived spreadsheet skill relative to the others in the study, measured using a scale
  where (1) was labelled ‘little or no skill’ and (7) was labelled ‘very skilful’
• What purposes they used spreadsheets for. The options provided were: ‘work’,
  ‘home use’ and ‘other’.

8.6.2 Experience

As in Study 1, previous experience with spreadsheets was measured with the single item
‘How long have you been using spreadsheets?’. This was also consistent with previous
studies that measured experience with a particular software development tool (Agboola,
8.6.3 Training

As in Study 1, previous spreadsheet training was measured using a 4 item, 5 point scale from Igbaria (1990) where (1) was labelled ‘none’ and (5) was labelled ‘very intensive’. The items asked for the level of training received in each of four types of training (college or university; vendor; in-company; self-study). These four items were used to create a summary measure of training as described in Section 9.2.2.

8.6.4 Spreadsheet development knowledge

Spreadsheet development knowledge relates to the knowledge that end user developers make use of when developing UDAs. A major component of this knowledge is knowledge about the features and functionality of spreadsheet packages. Kreie’s (1998) spreadsheet knowledge instrument was used as a starting point for the development of these items. Knowledge of development processes and quality assurance are also important components of the knowledge required to develop good quality spreadsheets (Janvrin & Morrison, 2000). Items to test knowledge of spreadsheet development processes were developed specifically for the study and drew upon two published methodologies for the development of spreadsheets (Ronen, Palley, & Lucas, 1989; Salchenberger, 1993), whilst attempting to ensure that participants were not disadvantaged by a lack of specialised terminology. The items covered areas such as the need for planning and methods of testing. Items were also included to test knowledge of quality assurance of spreadsheets. These items were developed specifically for the study using Rivard et al.’s (1997) instrument to measure the quality of end user developed applications as a source of material.
The initial instrument was examined for content validity by four information technology academics who have been involved in teaching spreadsheet use and design, and a few revisions were made on the basis of their suggestions. A 32 item questionnaire was piloted with 60 predominantly mature aged students enrolled in undergraduate business degrees. This instrument was shown to be reliable with a Cronbach alpha of 0.77, however a Guttman analysis of the potential of items to discriminate between students (Guttman, 1950) showed that some items did not discriminate well. These items were removed.

The final instrument had 25 multiple choice items. Each item was presented as a multiple choice question with five options. In each case the fifth option was ‘I don’t know’ or ‘I am not familiar with this spreadsheet feature’. Appendix F contains the full set of items and possible answers. Nine of the items relate to knowledge about the features and functionality of spreadsheet packages, eight items relate to the spreadsheet development process and eight items relate to spreadsheet quality assurance.

8.6.5 Involvement

Involvement is a subjective psychological state reflecting the importance and personal relevance of a system to the user. As discussed in Section 7.3.2, many early instruments to measure involvement actually measured participation (e.g. Doll & Torkzadeh, 1988; Franz & Robey, 1986; Hawk, 1993). Barki and Hartwick (1991) developed a scale for involvement with an IS, based on the general involvement scale proposed by Zaichkowsky (1985). The resulting scale was a seven point semantic differential with 11 items (see Table 8.2 below). Barki and Hartwick found that their instrument had a
Cronbach alpha of 0.91. In Study 2, the involvement construct was operationalised using Barki and Hartwick’s (1991) instrument.

Table 8.2: The item pool used in the measurement of involvement (NB: variable names are shown in brackets)

This car rental spreadsheet is:

- unimportant ............ important (INV21)
- not needed ............ needed (INV22)
- nonessential ............ essential (INV23)
- trivial ............ fundamental (INV24)
- insignificant ............ significant (INV25)
- means nothing to me ............ means a lot to me (INV26)
- unexciting ............ exciting (INV27)
- of no concern to me ............ of concern to me (INV28)
- not of interest to me ............ of interest to me (INV29)
- irrelevant to me ............ relevant to me (INV30)
- doesn’t matter to me ............ matters to me (INV31)

8.6.6 Perceived system quality

The item set used to measure perceived system quality (and system quality) was derived from the instrument used in Study 1. The Rivard et al. (1997) instrument that was designed specifically for UDAs was the starting point for the measurement of perceived system quality in Study 1. After the measurement model analysis had been undertaken the item set was reduced to a group of 19 items used in the creation of composite variables for both constructs (see Appendix B). These items were reviewed to ensure that they were appropriate for the applications to be developed in Study 2 and four items were removed because they measured aspects of quality that would not be expected in the experimental situation (eg. portability to other organisational environments and security).
Four items from the original Rivard et al. instrument were reinstated despite having been removed during measurement model analysis for Study 1. Because of the close linkage in Study 2 between application development and feedback on performance via the decision making, it was crucial to obtain valid perceptions about accuracy and it was anticipated that the participants would be able to provide insightful answers to these items. Three of the items related to errors and accuracy and one related to formatting of the spreadsheet. Minor changes were also made to reflect the terminology of the task. An additional item relating to errors and accuracy was created specifically for Study 2.

Table 8.3: The item pool used in the measurement of perceived system quality in Study 2

- Using the spreadsheet would be easy, even after a long period of not using it
- Errors in the spreadsheet are easy to identify
- The spreadsheet increased my data processing capacity
- The spreadsheet is easy to learn by new users
- Should an error occur, the spreadsheet makes it straightforward to perform some checking in order to locate the source of error
- The data entry sections provide the capability to easily make corrections to data
- The same terminology is used throughout the spreadsheet
- This spreadsheet does not contain any errors
- The terms used in the spreadsheet are familiar to users
- Data entry sections of the spreadsheet are organised so that the different bits of data are grouped together in a logical way
- The data entry areas clearly show the spaces reserved to record the data
- The format of a given piece of information is always the same, where ever it is used in the spreadsheet
- Data is labelled so that it can be easily matched with other parts of the spreadsheet
- The spreadsheet is broken up into separate and independent sections
- Use of this spreadsheet would reduce the number of errors you make when choosing a rental car
- Each section has a unique function or purpose
- Each section includes enough information to help you understand what it is doing
- Queries are easy to make
- The spreadsheet provides all the information required to use the spreadsheet (this is called documentation)
- Corrections to errors in the spreadsheet are easy to make
The resulting system quality scale consisted of 20 items, each scored on a Likert scale of 1 to 7 where (1) was labelled ‘strongly agree’ and (7) was labelled ‘strongly disagree’ (see Table 8.3 above).

### 8.6.7 User satisfaction

As can be seen in Table 8.4, user satisfaction was measured using the same items as were used in Study 1 (with a slight modification to the first item to reflect the terminology of the task). These items are from Seddon and Yip’s (1992) study of user satisfaction.

<table>
<thead>
<tr>
<th>Table 8.4: The item pool used in the measurement of user satisfaction in Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>• How adequately do you feel the spreadsheet meets your information processing needs when answering car rental queries? (US32)</td>
</tr>
<tr>
<td>• How efficient is the spreadsheet? (US33)</td>
</tr>
<tr>
<td>• How effective is the spreadsheet? (US34)</td>
</tr>
<tr>
<td>• Overall, are you satisfied with the spreadsheet? (US35)</td>
</tr>
</tbody>
</table>

### 8.6.8 Perceived individual impact

In Study 2 it was explicitly recognised that an individual’s perception of the impact of an IS on their performance may not be consistent with other external measures of individual impact. Perceived individual impact was measured using the items used to measure individual impact in Study 1 (see Table 8.5). These items derive from Goodhue and Thompson (1995).
8.6.9 Individual impact

DeLone and McLean (1992) claimed that individual impact is the most difficult IS success category to define in unambiguous terms. For example, the individual impact of a UDA could be related to a number of measures such as impact on performance, understanding, decision making or motivation. Given that perceived individual impact was also a construct in Study 2, a decision was made to focus on ‘objective’, easily quantifiable aspects of individual impact. Individual impact was measured in two different ways: accuracy of decision making (number of questions correct) and time taken to answer a set of questions. These measures were also used by Goodhue, Klein, and March (2000) in their study on user evaluations of systems.

Two sets of 10 different but equivalent questions involving the comparison of costs of three car rental companies under a variety of scenarios were created (see Appendix G). The second set of questions was derived by altering the figures in the first set, but keeping the form of the questions the same.

The questions ranged from comparison of the three firms when no excess kilometre charges are imposed, through to questions where excesses are applied and basic parameters are assumed to have changed from those given in the original problem description. A typical question is ‘Which rental company is the cheapest if you wish to
hire a car for 6 days and drive approximately 1500 kilometres with it?’ Participants were asked to provide both the name of the cheapest firm and its cost.

The questions were piloted by four end users and slight changes made to clarify them. The equivalence of the two sets of questions in terms of difficulty and time to complete was also confirmed by measuring the time taken to answer each set using the four applications created during piloting of the task (see Section 8.5).

8.7 Measurement of system quality

As mentioned above (Section 8.6.6), the same item set was used to measure both perceived system quality and system quality (see Table 8.3 for the item pool). System quality was assessed by two independent assessors. Both assessors were IS academics with substantial experience teaching spreadsheet design and development. Before assessing the study sample, the assessors completed four pilot evaluations not included in the study sample to ensure consistency between the assessors. Differences were discussed and adjustments made to ensure consistency. Assessments of the actual UDAs were then undertaken. The consistency of the system quality ratings for the two independent assessors was compared by calculating a composite measure of system quality for each assessor for each participant using the results of the measurement model development described in Chapter 9. The two final sets of assessments were highly correlated ($r = 0.80, p < 0.001$).
8.8 Data analysis techniques

Comparisons between end user developers using their own applications and end users using applications developed by another end user were made using paired samples $t$ tests. As in Study 1, evaluation of the model was undertaken using SEM. Section 4.10 provides an introduction to the technique, and to the two-step estimation process used in Study 1. This process was also used in Study 2.

8.8.1 Measurement model development

The aim of the measurement model development was to establish a set of items that permitted measurement of each of the latent constructs in the structural models. Parsimonious scales (sets of indicators) that validly and reliably measured each latent variable were sought from the pool of items used for each construct (see Table 8.2 to Table 8.5).

As described for Study 1 in Chapter 4, correlations between potential indicator items were used both to confirm that the items might validly be combined in a single scale, and as a basis for selection of a parsimonious set of indicators for each latent variable.

The evaluation of the measurement models was again undertaken using Amos 3.6, using maximum likelihood estimation, and missing data was handled by mean substitution. One question (Q15) had 18 out of 316 responses missing (5.6%) due to the scale not being printed on the questionnaires used in the first session. The maximum number of missing responses for other items was 5 out of 316 (1.6%).
The first measurement property to be assessed was the unidimensionality of each construct. As in Study 1, goodness of fit for the single factor congeneric measurement models was measured by the likelihood ratio chi-square ($\chi^2$), the ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df), the goodness of fit index (GFI), the adjusted goodness of fit index (AGFI), the root mean square error of approximation (RMSEA) and the Tucker-Lewis index (TLI). The same guidelines for acceptability of these measures were used (see Section 4.11 for a discussion of the guidelines, and Table 8.6 below provides a summary of them). After considering the values of these goodness of fit indexes using the guidelines, potential items that did not allow acceptable model fit and hence unidimensionality were excluded.

Table 8.6: Summary of the guidelines for model fit used in Study 2

<table>
<thead>
<tr>
<th>Model fit measures</th>
<th>Guidelines for fit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goodness of fit measures</strong></td>
<td></td>
</tr>
<tr>
<td>• Chi-square ($\chi^2$)</td>
<td>non-significant $\chi^2$ (p&gt;0.05)</td>
</tr>
<tr>
<td>• Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>&lt; 3:1</td>
</tr>
<tr>
<td>• Goodness of fit index (GFI)</td>
<td>&gt; 0.90</td>
</tr>
<tr>
<td>• Adjusted goodness of fit index (AGFI)</td>
<td>&gt; 0.80</td>
</tr>
<tr>
<td>• Root mean square error of approximation (RMSEA)</td>
<td>&lt; 0.05 indicates a good model fit, but values from 0.05 to 0.08 have been considered acceptable</td>
</tr>
<tr>
<td>• Tucker-Lewis index (TLI)</td>
<td>&gt; 0.90</td>
</tr>
<tr>
<td><strong>Reliability measures</strong></td>
<td></td>
</tr>
<tr>
<td>• Cronbach alpha coefficient</td>
<td>&gt; 0.70</td>
</tr>
<tr>
<td>• Composite reliability</td>
<td>&gt; 0.70</td>
</tr>
<tr>
<td>• Average variance extracted</td>
<td>&gt; 0.50</td>
</tr>
</tbody>
</table>

After overall model fit was considered, the significance of each estimated coefficient was examined to determine the probability that the item was a valid indicator of the construct. The ratio of coefficient to standard error, or $t$ value, is used to test each.
coefficient for significance. Ratios greater than 1.96 are acceptable, but the higher the ratio the more likely it is that the item is a valid indicator of the construct (Hayduck, 1987). Any items with non-significant coefficients were removed. In this way, convergent validity was demonstrated.

As in Study 1, once the item set was reduced to a valid, parsimonious, unidimensional scale, three estimates of reliability were calculated for each latent variable. These were Cronbach alpha coefficient, composite reliability, and average variance extracted (see Section 4.11 for a discussion of these measures). The guidelines used for these reliability measures are summarised in Table 8.6 above.

For those measurement models where a composite variable was created, the loading of the composite variable on its associated latent variable and the error term needed to be specified for subsequent use (Hair et al., 1998). These were calculated as described in Section 4.11. The results of the measurement model development are reported in Section 9.2.

8.8.2 Structural model evaluation

Following the assessment of the measurement models and creation of composite variables, AMOS 3.6 was again used to evaluate the structural models. The model proposed in Chapter 7 (see Figure 7.1) was first evaluated with the subset of the data where each participant answered the decision questions using their own UDA and completed the instruments with respect to their own UDA. The model was then evaluated with the subset of data where each end user used an application developed by another end user.
As in Study 1, both the structural models were evaluated on three criteria. The first criterion was overall goodness of fit between the model and the sample data. This was measured using the same goodness of fit measures used to assess the measurement models: the likelihood ratio chi-square ($\chi^2$), the ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df), the goodness of fit index (GFI), the adjusted goodness of fit index (AGFI), the root mean square error of approximation (RMSEA) and the Tucker-Lewis index (TLI) (see Section 4.11 for a description of these measures and the rules of thumb for interpreting them). The guidelines for interpreting these measures are summarised in Table 8.6 above.

As in Study 1, the second criterion considered was the ability of the models to explain the variance in the dependent variables. The squared multiple correlations ($R^2$) of the structural equations for these variables were used as estimates of variance explained. Although no test of statistical significance can be performed, $R^2$ provides a relative measure of fit for each structural equation in the model.

The third criterion was the significance of the estimated model coefficients. If a model is a valid representation of UDA impact, all proposed relationships should be significant. As well as the estimated coefficients, SEM techniques provide standard errors and calculate $t$ values for each coefficient. All of the hypotheses specify a direction for the proposed relationship so a one-tailed $t$ value of 1.645 indicates significance at the $p < 0.05$ level (Hair et al., 1998).
Each of the models was evaluated on the above criteria. It was considered that an acceptable model should explain a moderate to high proportion of the variance in the dependent variables, would contain only valid paths, and would meet the criteria for acceptable fit.

8.9 Descriptive analysis

This section describes the participants in Study 2 and the applications that they developed. It also provides a comparison with the participants in Study 1 and with the applications developed in Study 1.

8.9.1 The participants

Table 8.7 and Table 8.8 present background information about the participants in Study 2 and about their previous use of computers. Of the 159 participants in Study 2, 32.7% were male and 67.3% female (52 males, 107 females). Their ages ranged from 14 to 77 with an average age of 42.7. They had an average of 11.2 years experience using computers (with a range from 4 months to 38 years) and the majority of participants (105 or 66%) used a computer more than once a day.

Table 8.7: Age and computing experience of the participants in Study 2

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>42.7</td>
<td>14.0</td>
<td>77</td>
<td>12.22</td>
</tr>
<tr>
<td>Computing experience (years)</td>
<td>11.2</td>
<td>0.3</td>
<td>38</td>
<td>6.42</td>
</tr>
</tbody>
</table>
Table 8.8: Frequency of computer use by the participants in Study 2

<table>
<thead>
<tr>
<th>Frequency of computer use</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than once a month</td>
<td>2</td>
<td>1.3</td>
</tr>
<tr>
<td>About once a month</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>A few times a month</td>
<td>8</td>
<td>5.0</td>
</tr>
<tr>
<td>A few times a week</td>
<td>17</td>
<td>10.7</td>
</tr>
<tr>
<td>About once a day</td>
<td>27</td>
<td>17.0</td>
</tr>
<tr>
<td>Several times a day</td>
<td>105</td>
<td>66.0</td>
</tr>
</tbody>
</table>

Table 8.9 and Table 8.10 summarise the participants’ spreadsheet experience and previous training. Participants reported an average of 4.5 years experience using spreadsheets (with a range from 0 to 21 years). One hundred and twelve (70.4%) used spreadsheets at work and 92 (57.9%) used spreadsheets for personal use. They did not perceive themselves as being particularly skilful with spreadsheets (average 2.2, with only 21 or 13.3% rating themselves above the midpoint of the scale).

Table 8.9: Spreadsheet experience and perceived skill of the participants in Study 2

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet experience (years)</td>
<td>4.5</td>
<td>0</td>
<td>21</td>
<td>4.43</td>
</tr>
<tr>
<td>Perceived spreadsheet skill</td>
<td>2.2</td>
<td>1</td>
<td>5</td>
<td>1.07</td>
</tr>
</tbody>
</table>

Table 8.10 indicates that the participants had received relatively little spreadsheet training. Over 80% had received no in-company or vendor training and nearly 70% had received no university or technical college training. As in Study 1 and other studies reported in the literature (Chan & Storey, 1996; Hall, 1996), self-study was the predominant means by which participants had acquired their knowledge of spreadsheets.
Table 8.10: Previous spreadsheet training of the participants in Study 2

<table>
<thead>
<tr>
<th>Training Source</th>
<th>Mean</th>
<th>Level of training</th>
<th>Number in each category</th>
<th>(5) Extr. Intensive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1) None</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>University or college</td>
<td>1.6</td>
<td>108</td>
<td>67.9</td>
<td>22</td>
</tr>
<tr>
<td>Vendor</td>
<td>1.1</td>
<td>151</td>
<td>95.0</td>
<td>5</td>
</tr>
<tr>
<td>In-company</td>
<td>1.2</td>
<td>129</td>
<td>81.1</td>
<td>25</td>
</tr>
<tr>
<td>Self-study</td>
<td>2.3</td>
<td>41</td>
<td>25.8</td>
<td>59</td>
</tr>
</tbody>
</table>

The participants in Study 2 were older on average than those in Study 1, but had similar levels of computer experience and use. The Study 2 participants had less spreadsheet experience and training on average than did those in Study 1, and they perceived themselves as less skilful.

8.9.2 The user developed applications

As mentioned in Section 8.4, all applications were developed in Microsoft Excel©. The average file size of the spreadsheets in Study 2 was 16K with a minimum of 14K and a maximum of 36K. This average file size is much smaller than that of the applications in Study 1 (182K), and smaller than the average file size reported in Hall’s (1996) field study of end user developed spreadsheets (218K), and therefore suggests that these applications would be at the smaller end of spreadsheet applications in the workplace in that respect.
8.10 Overview

This chapter is the second of two chapters that describe the research approaches used in answering the research questions for this thesis. This chapter describes the design of Study 2 (see Chapter 4 for the research approach used in Study 1). Study 2 was a laboratory study and the UDAs considered were spreadsheet applications. The participants were end users from a variety of organisations. During the study, they developed spreadsheet applications to assist in a hypothetical problem scenario and then used them to support decision making. Both UDA quality and individual impact were assessed independently and end user perceptions of quality and impact were obtained via survey instruments.

The specific benefits of a laboratory situation for Study 2 were: the ability to specify the task to be undertaken so that all participants and UDAs could be compared directly; the ability to measure individual impact directly; and the ability to compare end user developers using their own applications with end users using applications developed by another end user. The study was also designed to provide the opportunity to ensure that the findings would be applicable to the wider end user developer population, given that the participants in Study 1 were solely MBA students. However, because it was designed as a laboratory study, with a limited time commitment possible from participants, the UDAs developed were by necessity not as large or complex as those typically used in the workplace.

The chapter also included a general discussion of SEM, the major data analysis technique chosen for Study 2, and details of how the measurement model development and structural model evaluation were carried out. This is followed up in Chapter 9,
which presents the results of both the measurement model development and structural model testing. The chapter concluded with descriptive information about the participants and the UDAs they developed. As in Study 1, participants were found to have had little formal training. This is consistent with the findings of earlier studies. The UDAs developed in this study were smaller than those developed in Study 1, and based on reports in the literature, appear to be smaller on average than UDAs that have been studied in organisational settings.
9.1 Introduction

This chapter reports on the results of the Study 2 data collection and statistical analyses that were carried out as described in Chapter 8. It first describes the measurement models for the latent variables in the model presented in Chapter 7. The chapter then provides a comparison of end user developers using their own applications with end users using applications developed by other end users. This comparison is based on a number of key IS success variables and is intended to ensure that any differences in UDA success due to the developer also being the user are identified. The chapter concludes with the results of the structural model evaluations for the two groups. The models are again evaluated against three criteria: overall goodness of fit; the amount of variance explained for key constructs; and the significance of model coefficients.

9.2 Measurement models

This section presents the measurement model for each latent variable in Study 2 and provides a summary of the measurement model information used in the evaluation of the structural models. The measurement models were developed as described in Section 8.8.1.
9.2.1 Experience
As described in Section 8.6.2, experience was measured using the single item: years of spreadsheet experience (mean = 4.46, SD = 4.43). This was treated as a single item indicator in the structural models.

9.2.2 Training
A summary measure of training was created for each participant by summing the responses for each of the four types of training (after adjusting the scale to start at zero). The average level of training on this scale was 2.13 (out of a maximum of 16) and the standard deviation was 1.67. This summary variable was treated as a single item indicator in the structural models.

9.2.3 Spreadsheet development knowledge
Each multiple choice item in the spreadsheet development knowledge instrument was scored as 1 if correct and 0 if incorrect or if ‘I don’t know’ was chosen. The total spreadsheet development knowledge score was used as a summary measure of spreadsheet development knowledge for each participant. The average spreadsheet development knowledge of the participants was 12.27 out of 25 (minimum = 1, maximum = 23, standard deviation = 4.46). Cronbach alpha was calculated as a measure of the reliability of the spreadsheet development knowledge scale and found to be 0.78, which met the guideline for reliability. Cronbach alpha for dichotomous items calculated using SPSS is equivalent to the Kuder-Richardson (KR20) coefficient for reliability of dichotomous items (SPSS Inc., 1999).
9.2.4 Involvement

The item pool to measure involvement included the 11 items from Barki and Hartwick (1991) (see Table 8.2). Barki and Hartwick (1994) conducted further validation and identified two subscales in the instrument, and factor analysis confirmed the two subscale structure using the Study 2 data (see Appendix I). The two subscales are importance and personal relevance. Barki and Hartwick retained nine of the original items after their further validation. In Study 2, 10 of the 11 items met the criteria for retention, and Table 9.1 contains summary statistics for these items.

Table 9.1: Involvement summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance</td>
<td></td>
<td></td>
<td></td>
<td>INV21 INV22 INV23 INV24 INV25</td>
</tr>
<tr>
<td>INV21</td>
<td>4.51</td>
<td>1.85</td>
<td>316</td>
<td>1.00</td>
</tr>
<tr>
<td>INV22</td>
<td>4.38</td>
<td>1.92</td>
<td>315</td>
<td>0.84*** 1.00</td>
</tr>
<tr>
<td>INV23</td>
<td>4.06</td>
<td>1.85</td>
<td>314</td>
<td>0.78*** 0.89*** 1.00</td>
</tr>
<tr>
<td>INV24</td>
<td>4.39</td>
<td>1.65</td>
<td>313</td>
<td>0.78*** 0.80*** 0.80*** 1.00</td>
</tr>
<tr>
<td>INV25</td>
<td>4.44</td>
<td>1.71</td>
<td>315</td>
<td>0.84*** 0.79*** 0.79*** 0.88*** 1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Personal Relevance</th>
<th></th>
<th></th>
<th></th>
<th>INV26 INV28 INV29 INV30 INV31</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV26</td>
<td>4.43</td>
<td>1.75</td>
<td>315</td>
<td>1.00</td>
</tr>
<tr>
<td>INV28</td>
<td>4.33</td>
<td>1.67</td>
<td>315</td>
<td>0.74*** 1.00</td>
</tr>
<tr>
<td>INV29</td>
<td>4.66</td>
<td>1.75</td>
<td>316</td>
<td>0.76*** 0.85*** 1.00</td>
</tr>
<tr>
<td>INV30</td>
<td>4.27</td>
<td>1.82</td>
<td>316</td>
<td>0.77*** 0.74*** 0.81*** 1.00</td>
</tr>
<tr>
<td>INV31</td>
<td>4.34</td>
<td>1.80</td>
<td>316</td>
<td>0.75*** 0.82*** 0.85*** 0.84*** 1.00</td>
</tr>
</tbody>
</table>

*** p < 0.001

One factor congeneric models for each subscale were first assessed. Although Barki and Hartwick (1991) averaged all of the item scores to obtain an overall involvement measure, better model fit was obtained with a congeneric model for each subscale.
Table 9.2: One factor congeneric model for importance – parameter estimates, goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>Importance</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV21</td>
<td>1.695</td>
<td>0.081</td>
<td>21.042</td>
<td>0.334</td>
</tr>
<tr>
<td>INV22</td>
<td>1.723</td>
<td>0.085</td>
<td>20.266</td>
<td>0.197</td>
</tr>
<tr>
<td>INV23</td>
<td>1.596</td>
<td>0.084</td>
<td>19.086</td>
<td>0.115</td>
</tr>
<tr>
<td>INV24</td>
<td>1.421</td>
<td>0.074</td>
<td>19.094</td>
<td>0.127</td>
</tr>
<tr>
<td>INV25</td>
<td>1.539</td>
<td>0.076</td>
<td>20.332</td>
<td>0.226</td>
</tr>
</tbody>
</table>

Goodness of fit measures

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>36.25</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>3</td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.000</td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>12.08</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.957</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.784</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.188</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.939</td>
</tr>
</tbody>
</table>

Reliability measures

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach alpha</td>
<td>0.96</td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.96</td>
</tr>
<tr>
<td>Average extracted variance</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 9.2 and Table 9.3 show the parameter estimates, factor score weights, goodness of fit statistics and reliability measures for each of the measurement models. All items had $t$ values substantially greater than 1.96 and thus can be considered valid indicators of the subscales, thus ensuring the convergent validity of the construct (Kline, 1998). The importance subscale was less well fitting than the personal relevance subscale. Despite a significant $\chi^2$, RMSEA higher than desirable and AGFI just below the recommended level, GFI and TLI indicated good fit for the importance subscale. The personal relevance subscale had a moderately good fit. Although $\chi^2$ was significant, $\chi^2$/df and RMSEA were acceptable and all the other measures indicated good fit.
Table 9.3: One factor congeneric model for personal relevance – parameter estimates, goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>Personal Relevance</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV26</td>
<td>1.442</td>
<td>0.081</td>
<td>17.911</td>
<td>0.083</td>
</tr>
<tr>
<td>INV28</td>
<td>1.502</td>
<td>0.073</td>
<td>20.557</td>
<td>0.285</td>
</tr>
<tr>
<td>INV29</td>
<td>1.601</td>
<td>0.075</td>
<td>21.262</td>
<td>0.188</td>
</tr>
<tr>
<td>INV30</td>
<td>1.638</td>
<td>0.080</td>
<td>20.526</td>
<td>0.260</td>
</tr>
<tr>
<td>INV31</td>
<td>1.654</td>
<td>0.078</td>
<td>21.289</td>
<td>0.183</td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

- Chi-square ($\chi^2$): 10.812
- Degrees of freedom (df): 4
- Probability (p): 0.029
- Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df): 2.70
- Goodness of fit index (GFI): 0.987
- Adjusted goodness of fit index (AGFI): 0.952
- Root mean square error of approximation (RMSEA): 0.074
- Tucker-Lewis index (TLI): 0.989

**Reliability measures**

- Cronbach alpha: 0.95
- Composite reliability: 0.95
- Average extracted variance: 0.79

The measures of reliability for both subscales were all very good. For the importance subscale, Cronbach alpha was 0.96, composite reliability was 0.96 and average extracted variance was 0.88. For the personal relevance subscale, Cronbach alpha was 0.95, composite reliability was 0.95 and average extracted variance was 0.79. The scales were therefore considered satisfactory for SEM and a composite score for each subscale was created using the factor scores from the measurement models. The two subscales were significantly correlated ($r = 0.698, p < 0.001$) and it was concluded that they could be combined. An overall involvement score was then created by averaging the two subscale scores.
9.2.5 **System quality**

As described in Section 8.6.6, the item set used to measure system quality (and perceived system quality) was derived from the set used in Study 1, but differed slightly. The assessment of the measurement model for system quality was performed bottom up as described by Rivard et al. (1997). One factor congeneric models for each criterion were first assessed, using the average expert responses to the questionnaire items. Once a unidimensional item set for each criterion was obtained, the criterion scores for each dimension were averaged. The overall measurement model for system quality was then assessed using the dimension scores. Although Rivard et al. averaged the dimension scores to obtain overall system quality, better model fit was obtained in this study with a congeneric model, thus composite system quality scores were calculated using the factor score weightings from the overall system quality measurement model as was done in Study 1. Table 9.4 below shows summary statistics for the overall measurement model.

Table 9.4: System quality summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Economy</th>
<th>Effectiveness</th>
<th>Reliability</th>
<th>Understandability</th>
<th>Userfriendliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>3.89</td>
<td>1.68</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>3.88</td>
<td>1.73</td>
<td>0.973***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>3.87</td>
<td>1.22</td>
<td>0.664***</td>
<td>0.675***</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understandability</td>
<td>3.92</td>
<td>1.08</td>
<td>0.472***</td>
<td>0.470***</td>
<td>0.729***</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Userfriendliness</td>
<td>4.00</td>
<td>1.21</td>
<td>0.543***</td>
<td>0.548***</td>
<td>0.841***</td>
<td>0.867***</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*** p < 0.001  

a all n = 159
Table 9.5 shows the parameter estimates, factor score weights, goodness of fit statistics and reliability measures for the system quality measurement model. All items had $t$ values substantially greater than 1.96 and thus can be considered valid indicators of the construct. The model fit measures provide a mixed picture. GFI and TLI met the guidelines, but $\chi^2$ was significant and neither AGFI nor RMSEA met the guidelines. The measures of reliability for this scale were all acceptable: Cronbach alpha was 0.94, composite reliability was 0.89 and average extracted variance was 0.63. On balance, the scale was considered satisfactory for SEM.

Table 9.5: One factor congeneric model for system quality – parameter estimates, goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>System Quality</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economy</td>
<td>0.964</td>
<td>0.124</td>
<td>7.780</td>
<td>0.007</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>1.002</td>
<td>0.127</td>
<td>7.862</td>
<td>0.013</td>
</tr>
<tr>
<td>Reliability</td>
<td>1.050</td>
<td>0.078</td>
<td>13.441</td>
<td>0.105</td>
</tr>
<tr>
<td>Understandability</td>
<td>0.948</td>
<td>0.068</td>
<td>13.900</td>
<td>0.140</td>
</tr>
<tr>
<td>Userfriendliness</td>
<td>1.182</td>
<td>0.071</td>
<td>16.647</td>
<td>0.734</td>
</tr>
<tr>
<td><strong>Goodness of fit measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td></td>
<td>37.092</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability (p)</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td></td>
<td>9.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td></td>
<td>0.918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td></td>
<td>0.692</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td></td>
<td>0.229</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td></td>
<td>0.914</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reliability measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronbach alpha</td>
<td></td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite reliability</td>
<td></td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average extracted variance</td>
<td></td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Perceived system quality was measured using the responses of the participants to the same item set as were retained in the final measurement model for system quality. This was done to ensure compatibility between the two constructs. The criterion scores for each dimension were averaged to give dimension scores. Table 9.6 below shows summary statistics for the individual dimensions of perceived system quality.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
<th>Economy</th>
<th>Effectiveness</th>
<th>Reliability</th>
<th>Understandability</th>
<th>Userfriendliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>4.15</td>
<td>2.00</td>
<td>315</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>3.98</td>
<td>2.14</td>
<td>298</td>
<td>0.609***</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>3.83</td>
<td>1.40</td>
<td>311</td>
<td>0.585***</td>
<td>0.605***</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understandability</td>
<td>4.48</td>
<td>1.38</td>
<td>312</td>
<td>0.532***</td>
<td>0.610***</td>
<td>0.641***</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Userfriendliness</td>
<td>4.36</td>
<td>1.54</td>
<td>314</td>
<td>0.637***</td>
<td>0.650***</td>
<td>0.631***</td>
<td>0.821***</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*** p < 0.001

The overall measurement model for perceived system quality was then assessed using the dimension scores with regression weights fixed as in the final system quality measurement model. The factor weightings used to create the composite system quality measure (see Table 9.5 above) were then also used to create the composite perceived system quality measure. This was done to enable direct comparison between the system quality and perceived system quality measures.

Table 9.7 shows the goodness of fit statistics and reliability measures for the perceived system quality measurement model. As the regression estimates were fixed to match those of the system quality measurement model they were not estimated by AMOS. All of the measures of reliability were acceptable with a Cronbach alpha of 0.94, composite
reliability of 0.84 and average extracted variance of 0.51. However, the model did not fit the data well, as only AGFI was within guidelines. This is not surprising given that congeneric models generally provide better fit than do models with fixed parameters such as parallel models (Fleishman & Benson, 1987) and suggests that end user developers may perceive system quality differently to experts. However, despite the less than ideal model fit the composite was retained in this form to enable direct comparison with system quality.

Table 9.7: Goodness of fit measures and reliability measures for the one factor model for perceived system quality

<table>
<thead>
<tr>
<th>Perceived System Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goodness of fit measures</strong></td>
</tr>
<tr>
<td>Chi-square ($\chi^2$)</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
</tr>
<tr>
<td>Probability (p)</td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
</tr>
</tbody>
</table>

| **Reliability measures** | |
| Cronbach alpha | 0.94 |
| Composite reliability | 0.84 |
| Average extracted variance | 0.51 |

### 9.2.7 User satisfaction

The item pool to measure user satisfaction contained the same four items that were used in Study 1. Table 9.8 contains summary statistics for the items.
Table 9.8: User satisfaction summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>US32</th>
<th>US33</th>
<th>US34</th>
<th>US35</th>
</tr>
</thead>
<tbody>
<tr>
<td>US32</td>
<td>4.13</td>
<td>2.18</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US33</td>
<td>4.02</td>
<td>2.06</td>
<td>0.84***</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US34</td>
<td>4.15</td>
<td>2.08</td>
<td>0.88***</td>
<td>0.87***</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>US35</td>
<td>3.77</td>
<td>2.16</td>
<td>0.84***</td>
<td>0.85***</td>
<td>0.86***</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*** p < 0.001
a all n = 316

Table 9.9: One factor congeneric model for user satisfaction – parameter estimates, goodness of fit measures and reliability measures

<table>
<thead>
<tr>
<th>User Satisfaction</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
<th>Factor score weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US32</td>
<td>1.972</td>
<td>0.093</td>
<td>21.124</td>
<td>0.218</td>
</tr>
<tr>
<td>US33</td>
<td>1.970</td>
<td>0.088</td>
<td>22.454</td>
<td>0.233</td>
</tr>
<tr>
<td>US34</td>
<td>1.895</td>
<td>0.089</td>
<td>21.398</td>
<td>0.345</td>
</tr>
<tr>
<td>US35</td>
<td>2.004</td>
<td>0.094</td>
<td>21.364</td>
<td>0.204</td>
</tr>
</tbody>
</table>

Goodness of fit measures

| Chi-square ($\chi^2$) | 1.963 |
| Degrees of freedom (df) | 2 |
| Probability (p) | 0.375 |
| Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df) | 0.98 |
| Goodness of fit index (GFI) | 0.997 |
| Adjusted goodness of fit index (AGFI) | 0.984 |
| Root mean square error of approximation (RMSEA) | 0.000 |
| Tucker-Lewis index (TLI) | 1.000 |

Reliability measures

| Cronbach alpha | 0.96 |
| Composite reliability | 0.96 |
| Average extracted variance | 0.86 |

Table 9.9 above shows the parameter estimates, factor score weights, goodness of fit statistics and reliability measures for the user satisfaction measurement model. All items
had $t$ values substantially greater than 1.96 and therefore can be considered valid indicators of the construct. All goodness of fit measures met the recommended guidelines. The model can thus be considered to fit the data well. The measures of reliability for this scale were very good: Cronbach alpha was 0.96, composite reliability was 0.96, and average extracted variance was 0.86. The scale was therefore considered satisfactory for SEM, and a composite variable for user satisfaction was created using the factor score weights shown in Table 9.9 above.

### 9.2.8 Perceived individual impact

As in Study 1, perceived individual impact was measured using two items. Summary statistics for the items are shown in Table 9.10.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$n$</th>
<th>Mean</th>
<th>SD</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PII36</td>
<td>316</td>
<td>4.10</td>
<td>2.12</td>
<td>0.931***</td>
</tr>
<tr>
<td>PII37</td>
<td>316</td>
<td>4.22</td>
<td>2.20</td>
<td>*** p &lt; 0.001</td>
</tr>
</tbody>
</table>

Table 9.11 shows the parameter estimates and reliability measures for the user satisfaction measurement model. All items had $t$ values substantially greater than 1.96 and therefore can be considered valid indicators of the construct. At least three indicator variables must be available to undertake one factor congeneric modelling. As only two items to measure perceived individual impact were included in the questionnaire it was not possible to determine factor loadings and measurement model fit in the initial modelling phase, hence no goodness of fit statistics are provided in Table 9.11. However, the measures of reliability for this scale were all very good: Cronbach alpha
was 0.96, composite reliability was 0.95 and average extracted variance was 0.90. The scale was therefore considered satisfactory for SEM.

<table>
<thead>
<tr>
<th>Perceived Individual Impact</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PII36</td>
<td>1.858</td>
<td>0.116</td>
<td>16.030</td>
</tr>
<tr>
<td>PII37</td>
<td>1.969</td>
<td>0.122</td>
<td>16.197</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability measures</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach alpha</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average extracted variance</td>
<td>0.90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**9.2.9 Individual impact**

As described in Chapter 8, two potential items to measure individual impact were included in the design of this study: number of correct decisions; and time to make decisions. Summary statistics for these items are shown in Table 9.12. These measures were not significantly correlated and hence not valid indicators of the same construct.

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to make decisions</td>
<td>313</td>
<td>16.53 min</td>
<td>8.80 min</td>
<td>0.046</td>
</tr>
<tr>
<td>Decisions correct score (/10)</td>
<td>316</td>
<td>3.94</td>
<td>3.30</td>
<td></td>
</tr>
</tbody>
</table>

Examination of the data revealed that a sizeable proportion of participants appeared to have quickly discovered that they were unlikely to be able to answer the questions correctly with their spreadsheet, and may have abandoned their attempt.
Figure 9.1 below illustrates this, showing a concentration of participants in the left bottom corner with low times and low numbers of correct decisions. This suggests that time to make decisions was not an appropriate indicator of individual impact for this study. It was therefore decided to retain number of correct decisions as the only indicator for individual impact in the structural model evaluations. As only one measure of individual impact was retained, measures of reliability could not be calculated.

Figure 9.1: The relationship between time to make decisions and number of decisions correct
9.2.10 Summary of information used to specify the structural models

The sections above detail the development of the individual measurement models required prior to testing of the structural models. Table 9.13 provides a summary of the information used to specify parameters in the structural model used to test the hypothesised relationships with the user developer data. It includes the composite variables derived from the measurement models and the single indicator variables. The loadings and errors are calculated as described in Chapter 4 based on the approach described in Hair et al. (1998).

Table 9.13: Summary of the information used to specify measurement model parameters in the structural model used to test the hypothesised relationships with the user developer data

<table>
<thead>
<tr>
<th>Constructa</th>
<th>Mean</th>
<th>SD</th>
<th>Loading</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Composites created</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement</td>
<td>9.37</td>
<td>2.72</td>
<td>2.669</td>
<td>0.297</td>
</tr>
<tr>
<td>System quality</td>
<td>3.97</td>
<td>1.15</td>
<td>1.088</td>
<td>0.146</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>4.62</td>
<td>1.28</td>
<td>1.219</td>
<td>0.165</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>4.43</td>
<td>1.86</td>
<td>1.812</td>
<td>0.138</td>
</tr>
<tr>
<td><strong>Single item indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>4.46</td>
<td>4.43</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Training</td>
<td>2.13</td>
<td>1.67</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>12.27</td>
<td>4.56</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Individual impact</td>
<td>4.39</td>
<td>3.33</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

a n= 159

Table 9.14 provides a summary of the information used to specify parameters in the structural model used to test the hypothesised relationships using the data from end users using applications created by other end user developers. This data set is referred to as the non user developer data. The slight differences in means for composite variables that would be expected to be stable between the two data sets are due to the fact that two
cases were not included in the non user developer data set due to missing data. The data set for end user developers using their own applications has 159 cases and the data set for end users using UDAs developed by other end user developers has 157 cases.

Table 9.14: Summary of the information used to specify measurement model parameters in the structural model used to test the hypothesised relationships with the non user developer data

<table>
<thead>
<tr>
<th>Construct</th>
<th>Mean</th>
<th>SD</th>
<th>Loading</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Composites created</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement</td>
<td>8.18</td>
<td>3.19</td>
<td>3.130</td>
<td>0.408</td>
</tr>
<tr>
<td>System quality</td>
<td>3.96</td>
<td>1.17</td>
<td>1.088</td>
<td>0.146</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>4.00</td>
<td>1.48</td>
<td>1.409</td>
<td>0.220</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>3.65</td>
<td>2.07</td>
<td>2.030</td>
<td>0.172</td>
</tr>
<tr>
<td><strong>Single item indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>4.57</td>
<td>4.42</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Training</td>
<td>2.17</td>
<td>1.66</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>12.27</td>
<td>4.82</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Individual impact</td>
<td>3.49</td>
<td>3.22</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*a n= 157

9.3 Comparison between user developers and non user developers

As mentioned in Chapter 7, UDAs are commonly used by other end users as well as by the developer (Hall, 1996). In Study 2, a within-subjects design was used to ensure that any specific differences due to the user being the developer could be identified.

Table 9.15 below provides descriptive information about each of the constructs on which differences due to the user being the developer might be relevant. These constructs were measured as described in Section 9.2 above. Perceived individual impact was calculated as the average of the two indicator variables PII36 and PII37 as no composite variable was created for the structural model evaluation. Both the number
of decisions correct and the time taken to make decisions were of interest in this comparison so both are reported.

Each end user’s interaction with their own application was directly compared to their interaction with another UDA using paired samples $t$ tests (see Table 9.15). End users perceived their own applications to be of higher quality than applications developed by other end users. On average, there was a 15.5% difference in perceived quality when the developer was assessing his/her own application. This difference was significant ($t = 4.21$, $df = 156$, $p < 0.001$) therefore, hypothesis H13 was supported.

End user developers were also significantly more involved with their own applications ($t = 4.93$, $df = 156$, $p < 0.001$) and significantly more satisfied with them ($t = 3.43$, $df = 156$, $p = 0.001$). The average difference in involvement if the user was also the developer was 15.0% and the average difference in user satisfaction was 21.9%.

Hypotheses H14 and H15 were therefore both supported.

Table 9.15: A comparison of end user developer perceptions and performance when using their own or another application

<table>
<thead>
<tr>
<th></th>
<th>Developer + User</th>
<th>User Only</th>
<th>% diff.</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived system quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.62</td>
<td>4.00</td>
<td>15.5</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>SD</td>
<td>1.28</td>
<td>1.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>9.41</td>
<td>8.18</td>
<td>15.0</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>SD</td>
<td>2.69</td>
<td>3.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.45</td>
<td>3.65</td>
<td>21.9</td>
<td>0.001</td>
</tr>
<tr>
<td>SD</td>
<td>1.85</td>
<td>2.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived individual impact</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>9.38</td>
<td>7.29</td>
<td>28.7</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>SD</td>
<td>3.92</td>
<td>4.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of decisions correct (/10)</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.43</td>
<td>3.49</td>
<td>26.9</td>
<td>0.002</td>
</tr>
<tr>
<td>SD</td>
<td>3.33</td>
<td>3.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time to make decisions (minutes)</td>
<td>154</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>17.62</td>
<td>15.22</td>
<td>15.8</td>
<td>0.020</td>
</tr>
<tr>
<td>SD</td>
<td>10.00</td>
<td>7.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
End users perceived their own applications as having a significantly greater impact on their decision performance \((t = 4.35, \text{df} = 156, p < 0.001)\), and this was confirmed as they made a significantly larger number of correct decisions \((t = 3.08, \text{df} = 156, p = 0.002)\). The average difference in perceived individual impact of the application was 28.7% and the average difference in the number of decisions correct was 26.9%. Hypotheses H16 and H17 were therefore also supported. However, end users took significantly longer to make the decisions using their own application \((t = 2.36, \text{df} = 153, p = 0.02)\). On average, the difference in decision time was 15.8%. Hypothesis H18 was therefore not supported.

Given the significant differences reported above it was considered necessary to ensure that the model proposed in Chapter 7 applies to both the situation in which an end user developer uses their own application, as well as the situation in which an end user uses an application developed by another end user. The following sections describe the testing of the model with two sets of data: end user developers using applications they have developed themselves, and end users developers using applications developed by another end user.
9.4 Analysis of the structural models for Study 2

As described in Section 8.8.2, following the assessment of the measurement models and creation of the composite variables, the structural models were evaluated. Two slightly different structural models were used in order to test the hypotheses. All hypotheses were tested for the scenario where end users develop and then use their own UDAs (see Figure 9.2). That is, the full proposed model was evaluated with the half of the data where each participant answered the decision questions using his/her own UDA and completed the instruments with respect to his/her own UDA.

![Diagram of hypotheses](image)

Figure 9.2: Hypotheses tested for the scenario where an end user develops and then uses his/her own UDA

All hypotheses except H3 were also tested in a second structural model for the scenario where end users use applications developed by other end users (see Figure 9.3). That is, this part of the proposed model was evaluated with the half of the data where each
participant answered the decision questions using the UDA created by another participant in the same session with a similar level of spreadsheet development knowledge (non user developer data).

The two structural models used to test the model for the different scenarios differ in the following ways:

- The structural model shown in Figure 9.3 does not propose a relationship between spreadsheet development knowledge and system quality. The quality of the initial UDA could not be affected by the user’s spreadsheet development knowledge as the user was not the developer.

Figure 9.3: Hypotheses tested for the scenario where an end user uses an application developed by another end user
In Figure 9.2, the construct spreadsheet development knowledge represents the knowledge of both the developer and the user (as they are the same person). It is hypothesised to influence both system quality (via the developer’s role) and perceived system quality (via the user’s perception). Figure 9.3 models only the user role, so spreadsheet development knowledge is that of the user but not the developer.

Table 9.16 displays the implied correlation matrix for the latent variables considered in the model to test the hypothesised relationships using the user developer data, and Table 9.17 displays these correlations for the test of the model using the non user developer data. These estimated correlations can be used to examine the discriminant validity of the latent variables (Kline, 1998). Discriminant validity appeared to be satisfactory for all operationalisations except for possibly user satisfaction and perceived individual impact. The high implied correlation between them is of some concern. However, as the instruments were used in Study 1 and discriminant validity demonstrated for that study a decision was made to accept these operationalisations.
Table 9.16: Implied correlations between the latent variables considered in the model to test the hypothesised relationships using the user developer data

<table>
<thead>
<tr>
<th></th>
<th>Experience</th>
<th>Training</th>
<th>Spreadsheet development knowledge</th>
<th>System quality</th>
<th>Perceived system quality</th>
<th>Involvement</th>
<th>User satisfaction</th>
<th>Perceived individual impact</th>
<th>Individual impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.375</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>0.438</td>
<td>0.424</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>0.203</td>
<td>0.196</td>
<td>0.463</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>0.150</td>
<td>0.110</td>
<td>0.286</td>
<td>0.414</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement</td>
<td>0.070</td>
<td>-0.022</td>
<td>0.016</td>
<td>0.008</td>
<td>0.393</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>0.128</td>
<td>0.094</td>
<td>0.245</td>
<td>0.354</td>
<td>0.854</td>
<td>0.336</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived individual impact</td>
<td>0.124</td>
<td>0.092</td>
<td>0.239</td>
<td>0.339</td>
<td>0.805</td>
<td>0.315</td>
<td>0.941</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.196</td>
<td>0.182</td>
<td>0.435</td>
<td>0.378</td>
<td>0.366</td>
<td>0.092</td>
<td>0.381</td>
<td>0.381</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 9.17: Implied correlations between the latent variables considered in the model to test the hypothesised relationships using the non user developer data

<table>
<thead>
<tr>
<th></th>
<th>Experience</th>
<th>Training</th>
<th>Spreadsheet development knowledge</th>
<th>System quality</th>
<th>Perceived system quality</th>
<th>Involvement</th>
<th>User satisfaction</th>
<th>Perceived individual impact</th>
<th>Individual impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.370</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>0.335</td>
<td>0.381</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>0.083</td>
<td>0.068</td>
<td>0.134</td>
<td>0.451</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement</td>
<td>0.096</td>
<td>0.050</td>
<td>0.037</td>
<td>0.000</td>
<td>0.462</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>0.061</td>
<td>0.049</td>
<td>0.098</td>
<td>0.328</td>
<td>0.727</td>
<td>0.335</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived individual impact</td>
<td>0.060</td>
<td>0.049</td>
<td>0.098</td>
<td>0.319</td>
<td>0.705</td>
<td>0.325</td>
<td>0.970</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.121</td>
<td>0.132</td>
<td>0.335</td>
<td>0.209</td>
<td>0.331</td>
<td>0.125</td>
<td>0.401</td>
<td>0.400</td>
<td>1.000</td>
</tr>
</tbody>
</table>
The models were evaluated and compared on three criteria: overall goodness of fit, ability to explain the variance in the dependent variables, and the significance of estimated model coefficients (see Section 4.12 for a detailed explanation of the criteria chosen). It was considered that an acceptable model should explain a moderate to high proportion of the variance in the dependent variables of interest, would contain only valid paths, and would meet the criteria for acceptable fit. The next section describes testing of the first structural model.

9.4.1 Test of the model for end user developers using their own applications

The goodness of fit measures, model coefficients, standard errors and t values for the first structural model are reported in Table 9.18 below (see Figure 9.4 for the standardised coefficients). The goodness of fit measures provided conflicting information. Model $\chi^2$ was 57.37, with 28 degrees of freedom and $\chi^2$ was significant at 0.001, however $\chi^2$/df was within the guideline. RMSEA was marginal at 0.081. However, GFI (0.935), AGFI (0.873) and TLI (0.949) all indicated good fit.

The second criterion of good fit considered was the proportion of variance in dependent variables explained by the model (see Figure 9.4). Both user satisfaction ($R^2 = 0.729$, i.e. 72.9% of the variance is explained) and perceived individual impact ($R^2 = 0.886$) are explained well by the model. The model explains 27.0% of the variance in spreadsheet development knowledge, 21.5% of the variance in system quality, 33.4% of the variance in perceived system quality and 28.5% of the variance in individual impact, indicating that the role of other possible influences needs to be considered.
Table 9.18: Model coefficients, standard errors, $t$ values and goodness of fit measures for the model tested with user developer data

<table>
<thead>
<tr>
<th>Path To</th>
<th>From</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>Spreadsheet development knowledge</td>
<td>0.335</td>
<td>0.075</td>
<td>4.435***</td>
</tr>
<tr>
<td>Training</td>
<td>Spreadsheet development knowledge</td>
<td>0.824</td>
<td>0.200</td>
<td>4.114***</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>System quality</td>
<td>0.102</td>
<td>0.017</td>
<td>6.106***</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>Perceived system quality</td>
<td>0.025</td>
<td>0.018</td>
<td>1.418</td>
</tr>
<tr>
<td>System quality</td>
<td>Perceived system quality</td>
<td>0.361</td>
<td>0.086</td>
<td>4.201***</td>
</tr>
<tr>
<td>Involvement</td>
<td>Perceived system quality</td>
<td>0.392</td>
<td>0.072</td>
<td>5.427***</td>
</tr>
<tr>
<td>System quality</td>
<td>Individual impact</td>
<td>0.489</td>
<td>0.288</td>
<td>1.702*</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>User satisfaction</td>
<td>0.852</td>
<td>0.053</td>
<td>16.191***</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Perceived individual impact</td>
<td>1.768</td>
<td>0.078</td>
<td>22.748***</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Individual impact</td>
<td>0.849</td>
<td>0.248</td>
<td>3.420***</td>
</tr>
<tr>
<td>Individual impact</td>
<td>Perceived individual impact</td>
<td>0.015</td>
<td>0.022</td>
<td>0.687</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>Individual impact</td>
<td>0.233</td>
<td>0.057</td>
<td>3.920***</td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>57.37</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>28</td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.001</td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>2.05</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.935</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.873</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.081</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.949</td>
</tr>
</tbody>
</table>

* $p < 0.05$

*** $p < 0.001$
As can be seen from the \( t \) values in Table 9.18, ten of the twelve hypothesised paths were significant. The significant relationships did however vary greatly in strength (see the standardised path coefficients in Figure 9.4). The results for each of the hypotheses are stated below.

**H1:** Years of spreadsheet experience had a moderate significant positive influence on level of spreadsheet development knowledge. Therefore, this hypothesis was supported.
H2: Level of previous spreadsheet training had a moderate significant positive influence on level of spreadsheet development knowledge. Therefore, this hypothesis was supported.

H3: Spreadsheet development knowledge had a moderate significant positive influence on system quality. Therefore, this hypothesis was supported.

H4: System quality had a moderate significant positive influence on perceived system quality. Therefore, this hypothesis was supported.

H5: Spreadsheet development knowledge did not demonstrate a significant negative influence on perceived system quality, so this hypothesis was not supported.

H6: Involvement had a moderate significant positive influence on perceived system quality. Therefore, this hypothesis was supported.

H7: System quality had a weak and just significant positive influence on individual impact. This hypothesis was therefore cautiously supported.

H8: Perceived system quality had a strong significant positive effect on user satisfaction. Therefore, this hypothesis was supported.

H9: User satisfaction had a strong significant positive effect on perceived individual impact. Therefore, this hypothesis was supported.
H10: Individual impact did not demonstrate a significant influence on perceived
individual impact, so this hypothesis was not supported.

H11: User satisfaction had a moderate significant positive effect on individual impact.
Therefore, this hypothesis was supported.

H12: Spreadsheet development knowledge had a moderate significant positive
influence on individual impact. Therefore, this hypothesis was supported.

Table 9.19 below reports the standardised total effects (direct plus indirect) estimated
for the model tested with user developer data. In addition to a direct effect on
spreadsheet development knowledge, both training and experience had weak indirect
effects on all of the dependent variables in the model. Spreadsheet development
knowledge, a product of training and experience as well as other factors not included in
the model, also had a significant effect on all of the subsequent dependent variables in
the model, as did involvement. Spreadsheet development knowledge had a stronger
effect on individual impact than did involvement, but involvement had a stronger effect
on perceived system quality, user satisfaction, and perceived individual impact.

In summary, the proposed model appears to provide a useful model for aspects of UDA
success for end users using applications they have developed themselves. Model fit was
acceptable, and ten of the twelve proposed hypotheses were supported. However, the
model only explains a moderate portion of the variance of some constructs and two
hypotheses were not supported.
<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Experience</th>
<th>Spreadsheet dev. knowledge</th>
<th>System quality</th>
<th>Involvement</th>
<th>Perceived system quality</th>
<th>User satisfaction</th>
<th>Individual impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet dev. knowledge</td>
<td>0.302*</td>
<td>0.325**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System quality</td>
<td>0.140*</td>
<td>0.151*</td>
<td>0.463*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>0.084**</td>
<td>0.091**</td>
<td>0.280**</td>
<td>0.358*</td>
<td>0.389**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>0.072**</td>
<td>0.078**</td>
<td>0.239**</td>
<td>0.306*</td>
<td>0.332**</td>
<td>0.854*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.131*</td>
<td>0.141*</td>
<td>0.433*</td>
<td>0.224*</td>
<td>0.085**</td>
<td>0.218*</td>
<td>0.255*</td>
<td></td>
</tr>
<tr>
<td>Perceived individual impact</td>
<td>0.071*</td>
<td>0.076**</td>
<td>0.234**</td>
<td>0.291*</td>
<td>0.311**</td>
<td>0.801*</td>
<td>0.938*</td>
<td>0.026</td>
</tr>
</tbody>
</table>

* p < 0.05 (two-tailed)
** p < 0.01 (two-tailed)
9.4.2 Test of the model for end user developers using applications developed by other end users

This section describes the testing of the second structural model using the subset of data where the users of the applications were not the developers. The goodness of fit measures, model coefficients, standard errors and t values are reported in Table 9.20 below. Figure 9.5 shows the standardised coefficients for the model.

The goodness of fit measures provided conflicting information. Model $\chi^2$ was significant at $p < 0.001$ and $\chi^2$/df was also higher than recommended. RMSEA was above the recommended level at 0.118. However, the GFI (0.904), AGFI (0.818) and TLI (0.907) all indicated good fit.

The second criterion of good fit considered was the proportion of variance in dependent variables explained by the model (see Figure 9.5). Both user satisfaction ($R^2 = 0.528$) and perceived individual impact ($R^2 = 0.940$) are explained well by the model. The model explains 18.9% of the variance in spreadsheet development knowledge, 43.0% of the variance in perceived system quality and 25.8% of the variance in individual impact. Whilst the model does not explain a large proportion of the variance of some constructs, the consistency in results between the two tests of the model provides further support for it.
Table 9.20: Model coefficients, standard errors, $t$ values and goodness of fit measures for the model tested with the non user developer data

<table>
<thead>
<tr>
<th>Path</th>
<th>From</th>
<th>To</th>
<th>Estimate</th>
<th>Standard error</th>
<th>$t$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>Spreadsheet development knowledge</td>
<td>Spreadsheept development knowledge</td>
<td>0.244</td>
<td>0.085</td>
<td>2.889**</td>
</tr>
<tr>
<td>Training</td>
<td>Spreadsheet development knowledge</td>
<td>Perceived system quality</td>
<td>0.865</td>
<td>0.225</td>
<td>3.847***</td>
</tr>
<tr>
<td>Training</td>
<td>Perceived system quality</td>
<td>Spreadsheept development knowledge</td>
<td>0.026</td>
<td>0.015</td>
<td>1.771</td>
</tr>
<tr>
<td>System quality</td>
<td>Perceived system quality</td>
<td>Spreadsheept development knowledge</td>
<td>0.474</td>
<td>0.074</td>
<td>6.393***</td>
</tr>
<tr>
<td>Involvement</td>
<td>Perceived system quality</td>
<td>Spreadsheept development knowledge</td>
<td>0.416</td>
<td>0.061</td>
<td>6.780***</td>
</tr>
<tr>
<td>System quality</td>
<td>Individual impact</td>
<td>Spreadsheept development knowledge</td>
<td>0.305</td>
<td>0.242</td>
<td>1.260</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>User satisfaction</td>
<td>Spreadsheept development knowledge</td>
<td>0.724</td>
<td>0.062</td>
<td>11.616***</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Perceived individual impact</td>
<td>Spreadsheept development knowledge</td>
<td>1.807</td>
<td>0.063</td>
<td>28.653***</td>
</tr>
<tr>
<td>User satisfaction</td>
<td>Individual impact</td>
<td>Spreadsheept development knowledge</td>
<td>1.005</td>
<td>0.224</td>
<td>4.483***</td>
</tr>
<tr>
<td>Individual impact</td>
<td>Perceived individual impact</td>
<td>Spreadsheept development knowledge</td>
<td>0.008</td>
<td>0.019</td>
<td>0.426</td>
</tr>
<tr>
<td>Individual impact</td>
<td>Individual impact</td>
<td>Spreadsheept development knowledge</td>
<td>0.198</td>
<td>0.046</td>
<td>4.336***</td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ($\chi^2$)</td>
<td>92.48</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>29</td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.000</td>
</tr>
<tr>
<td>Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df)</td>
<td>3.19</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.904</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.818</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.118</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.907</td>
</tr>
</tbody>
</table>

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$
As can be seen from the $t$ values in Table 9.20, eight of the eleven hypothesised paths were significant. The significant relationships did, however, vary greatly in strength (see the standardised path coefficients in Figure 9.5). The results for each of the hypotheses are stated below (note: the same data was used to test hypothesis 1 and 2 as was used in the previous test of the model described above in Section 9.4.1).

**H1:** Years of spreadsheet experience had a moderate significant positive influence on level of spreadsheet development knowledge. Therefore, this hypothesis was again supported.

**H2:** Level of previous spreadsheet training had a moderate significant positive influence on level of spreadsheet development knowledge. Therefore, this hypothesis was again supported.
(H3: Not tested with for the non user developer data.)

H4: System quality had a moderate significant positive influence on perceived system quality. Therefore, this hypothesis was supported.

H5: Spreadsheet development knowledge did not demonstrate a significant negative influence on perceived system quality, so this hypothesis was not supported.

H6: Involvement had a moderate significant positive influence on perceived system quality. Therefore, this hypothesis was supported.

H7: System quality did not demonstrate a significant influence on individual impact. Therefore, this hypothesis was not supported.

H8: Perceived system quality had a strong significant positive effect on user satisfaction. Therefore, this hypothesis was supported.

H9: User satisfaction had a strong significant positive effect on perceived individual impact. Therefore, this hypothesis was supported.

H10: Individual impact did not demonstrate a significant influence on perceived individual impact, so this hypothesis was not supported.
H11: User satisfaction had a moderate significant positive effect on individual impact. Therefore, this hypothesis was supported.

H12: Spreadsheet development knowledge had a moderate significant positive influence on individual impact. Therefore, this hypothesis was supported.

Table 9.21 below reports the standardised total effects estimated for the model tested with the data for end users using applications developed by another end user. These results are consistent with the total effects measured for the model when tested with the end user developer data (see Section 9.4.1). The most notable difference was that training, experience and spreadsheet development knowledge had weaker effects on the subsequent dependent variables in the model when end users were using an application developed by another end user. This is consistent with the first structural model representing both developer and user training, experience and knowledge, whereas the second model only represents user characteristics. As in the model tested with the user developer data, spreadsheet development knowledge had a stronger effect on individual impact than did involvement, but involvement had a stronger effect on perceived system quality, user satisfaction, and perceived individual impact.

In summary, although developed primarily for end user developers using their own applications, the model also appears to provide the basis for a useful model of UDA success for end users using applications developed by other end user developers. Model fit was satisfactory, and eight of the 11 proposed hypotheses were supported.
Table 9.21: Standardised total effects on dependent variables estimated for the model tested with non user developer data

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Experience</th>
<th>Spreadsheet dev. knowledge</th>
<th>System quality</th>
<th>Involvement</th>
<th>Perceived system quality</th>
<th>User satisfaction</th>
<th>Individual impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet dev. knowledge</td>
<td>0.299*</td>
<td>0.224*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>0.035*</td>
<td>0.026*</td>
<td>0.118*</td>
<td>0.451**</td>
<td>0.457*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User satisfaction</td>
<td>0.026*</td>
<td>0.019*</td>
<td>0.085*</td>
<td>0.328**</td>
<td>0.332**</td>
<td>0.727**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual impact</td>
<td>0.099*</td>
<td>0.074**</td>
<td>0.331**</td>
<td>0.209*</td>
<td>0.113**</td>
<td>0.246*</td>
<td>0.339*</td>
<td>0.013</td>
</tr>
<tr>
<td>Perceived individual impact</td>
<td>0.026</td>
<td>0.019*</td>
<td>0.087</td>
<td>0.319**</td>
<td>0.322**</td>
<td>0.704**</td>
<td>0.969*</td>
<td>0.013</td>
</tr>
</tbody>
</table>

* p < 0.05 (two-tailed)
** p < 0.01 (two-tailed)
9.5 Overview

This chapter reported the results of the data collection and analyses undertaken during Study 2. Measurement models were presented for the variables considered in the structural models. These measurement models were satisfactory and hence were used to create composite variables for structural model evaluation. A comparison of the two groups: end user developers using their own applications and end users using applications developed by other end users was then provided. End user developers using their own applications were found to be both more satisfied with the application than they were with one developed by another end user, and were also found to make more correct decisions, but to take longer to do so. Possible reasons for these findings are discussed in Chapter 10.

The results of the structural model evaluation were also presented in detail. Whilst the model did not meet all of the criteria established for evaluation, it proved to be satisfactory when tested with data for each of the groups. Ten of the twelve hypotheses were supported when the model was tested with the data from end user developers using their own applications for decision making, and eight out of eleven hypotheses were supported when the model was tested for end users developers using applications developed by other end users. The results of the structural model evaluations reported in this chapter are discussed in detail in Chapter 10.
Chapter 10
Study 2 Discussion

10.1 Introduction

This thesis is concerned with understanding the UDA success process, specifically addressing three research questions relating to the role of system quality in UDA success and the ability of end user developers to judge whether applications they develop will have a positive impact on their performance of tasks. Study 2 was designed to pursue some of the issues raised in Study 1 and hence to provide further insight into the research questions for the thesis. In particular, Study 2 was carried out to:

- Clarify the roles of both perceived individual impact and independently measured individual impact in UDA success.
- Investigate possible reasons for the lack of relationship between perceived system quality and system quality in Study 1, and to explore the impact of this lack of relationship on the ability of end user developers to judge whether the applications they develop will have a positive impact on their performance of tasks.
- Investigate where the hypothesised relationships between experience and both perceived and independently measured system quality, and training and both perceived and independently measured system quality might be breaking down.
- Explore whether involvement plays a role in influencing the UDA success process.

The results of Study 2 were presented in Chapter 9 and this chapter provides a discussion and interpretation of the results. The chapter first considers the results of the comparison between end user developers using their own applications and end user
developers using applications developed by other end users. The chapter then considers the results of each of the hypotheses tested for the structural models and seeks explanations for the results obtained. In each case, the results of Study 2 are compared with the results of Study 1 and implications are discussed. This is followed by a discussion of the overall implications of the findings and the contribution of the model. The chapter then assesses what progress has been made towards answering the research questions posed for this thesis and, finally, the limitations of the research are discussed.

10.2 Comparison between user developers and users who were not the developer

The results of the comparison between end user developers using their own applications and end users using applications developed by other end users suggests that the process of developing an application not only predisposes an end user developer to be more satisfied with the application than they would be if it were developed by another end user, but also leads them to perform better with the application than they would if it were developed by another end user.

The end user developers in Study 2 had significantly higher levels of involvement, user satisfaction and perceived individual impact when using their own applications than they did when using applications developed by another end user with approximately the same level of spreadsheet development knowledge. They also perceived their applications to be of higher system quality. These results are consistent with the results in the literature on user involvement in the development of organisational systems. For example, Doll and Torkzadeh (1988) found user participation in design to be positively correlated with end user computing satisfaction, and Lawrence and Low (1993) found
that the more a user felt involved with the development process, the more satisfied they were with the system. The results are also consistent with McGill et al.’s (1998) study in the UDA domain, where end user developers were found to be more satisfied with their own applications than with those developed by other end users.

The results also confirm Cheney, Mann and Amoroso’s (1986) claim that end user development can be considered as the ultimate user involvement. The higher levels of perceived system quality for end users’ own applications highlight the subjectivity of system quality for end users. In contrast to ‘software engineering’ definitions of system quality (e.g. Boehm et al., 1978; Cavano & McCall, 1978), Amoroso and Cheney (1992) implicitly acknowledge how difficult it is for end users to assess the quality of applications by defining UDA quality as a combination of end user information satisfaction and application utilisation. This definition, however, ignores the underlying necessity for the more technical dimensions of system quality to be taken account of in order to have reliable and maintainable applications.

End user developers made significantly more correct decisions when using their own applications than when using an application developed by another end user. None of the participants was particularly familiar with the problem to be solved and they had the same background knowledge when using each application, so domain knowledge was not a factor. The improved performance could be due to a greater familiarity with the application itself, achieved through the development process. Successful use of user developed spreadsheet applications appears to require substantial end user knowledge because of the lack of separation of data and processing that is commonly found (Hall, 1996; Ronen, Palley, & Lucas, 1989). Developing an application appears to allow the
user to develop a robust understanding of it that makes it easier to use and makes it possible for the user developer to successfully adjust aspects of it when necessary.

The improved performance could also be due to a greater determination to achieve the correct answers, because of the higher levels of involvement. This explanation receives support from the additional time user developers spent making the decisions. On average the user developers spent an extra two and a half minutes trying to answer the 10 questions. This was unexpected, but may be due to the end user developers’ greater commitment to succeeding with their own applications. Comments from participants during the sessions support this possible explanation.

McGill et al. (1998) questioned the usefulness of user satisfaction as a measure of UDA success, finding that end users were significantly more satisfied with applications they had developed themselves than they were with applications developed by other end users. However, no measures of performance were included in that study. The results of Study 2 suggest that raised levels of user satisfaction and other perceptual variables are appropriate for end user developers, as they are consistent with better levels of individual performance.

The results of this comparison between end user developers using their own applications and end users using applications developed by other end users has implications for staff movement in organisations. If an end user develops an application for his or her own use, and its use has a positive impact on performance, this does not guarantee that the same will be true if another end user starts to use it. Organisations should recognise that the use of UDAs by end users other than the developer may carry
with it greater risks. In addition, if users are developing applications for others to use, particular attention must be paid to ensure that these applications are of sufficient quality for successful use not to rely on additional insight gained during the development process.

### 10.3 Discussion of hypotheses

This section discusses the results of the hypotheses relating to the model proposed in Chapter 7 (see Figure 7.1 for a diagram of the proposed relationships). Table 10.1 below summarises these relationships and indicates which were supported and which were not. It also indicates the strength of each of the supported relationships.

<table>
<thead>
<tr>
<th>Relationships supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience $\rightarrow$ spreadsheet development knowledge (moderate)</td>
</tr>
<tr>
<td>Training $\rightarrow$ spreadsheet development knowledge (moderate)</td>
</tr>
<tr>
<td>Spreadsheet development knowledge $\rightarrow$ system quality (moderate)</td>
</tr>
<tr>
<td>System quality $\rightarrow$ perceived system quality (moderate)</td>
</tr>
<tr>
<td>Involvement $\rightarrow$ perceived system quality (moderate)</td>
</tr>
<tr>
<td>Perceived system quality $\rightarrow$ user satisfaction (strong)</td>
</tr>
<tr>
<td>User satisfaction $\rightarrow$ perceived individual impact (strong)</td>
</tr>
<tr>
<td>User satisfaction $\rightarrow$ individual impact (moderate)</td>
</tr>
<tr>
<td>Spreadsheet development knowledge $\rightarrow$ individual impact (moderate)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relationships partially supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>System quality $\rightarrow$ individual impact (weak and significant only for the user developer data)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relationships not supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet development knowledge $\rightarrow$ perceived system quality</td>
</tr>
<tr>
<td>Individual impact $\rightarrow$ perceived individual impact</td>
</tr>
</tbody>
</table>
10.3.1 Experience influences spreadsheet development knowledge

It was hypothesised that experience would have a significant positive relationship with level of spreadsheet development knowledge (hypothesis H1). This hypothesis was supported. The longer the end users had been using spreadsheets, the greater their spreadsheet development knowledge. In terms of the results of Study 1, this finding suggests that the lack of relationship between experience and system quality was not due to lack of spreadsheet development knowledge acquired through experience.

However, an exploratory study on the nature of spreadsheet knowledge has suggested that the relationship between experience and spreadsheet knowledge is a complex one (McGill & Dixon, 2001), with experience being more strongly related to knowledge of spreadsheet functionality than to knowledge of spreadsheet development processes or knowledge of spreadsheet quality assurance. As mentioned in Chapter 7, this might be because although regular spreadsheet use should provide opportunities to encounter advanced software features, it will not necessarily ensure that end users learn about development processes or quality assurance. This issue should be investigated further in future research.

The fact that the relationship between experience and spreadsheet development knowledge was not strong might be partially due to the operationalisation of experience in this study. As discussed in Chapter 2, experience has been measured in different ways in different studies in the literature. Operationalising experience as the number of years an end user has used spreadsheets, whilst being simple and probably reliable, neglects the intensity of use over that period and hence may have weakened the relationship.
The lack of strength in the relationship also serves as a reminder that other factors not included in this model may play a role in influencing an end user’s level of spreadsheet development knowledge (e.g. individual differences in cognitive ability (Simon, Grover, Teng, & Whitcomb, 1996)).

10.3.2 Training influences spreadsheet development knowledge

It was hypothesised that level of previous spreadsheet training would positively influence level of spreadsheet development knowledge (hypothesis H2). This hypothesis was supported, although the relationship was not strong. Those end users with higher levels of previous spreadsheet training exhibited higher levels of spreadsheet development knowledge. This finding is inconsistent with the results of Chan and Storey (1996) who found no relationship between having had training and spreadsheet proficiency, but consistent with the findings of Nelson and Cheney (1987) who concluded that there is generally a positive relationship between the computer-related training that a user receives and his or her ability to use the computer resource. However, both these previous studies used self-report measures of knowledge/ability leaving open the possibility that end users’ perceptions of their own skill are not reliable. Therefore, the positive relationship identified in Study 2 should provide encouragement to organisations that invest in end user development training.

In terms of the results of Study 1, the Study 2 finding about the relationship between training and spreadsheet development knowledge suggests that the lack of relationship between training and system quality in Study 1 was not due to a lack of spreadsheet development knowledge developed via training. However, as discussed in Section 6.2.2,
this relationship may not be a simple one. Whilst training appears to increase overall
levels of spreadsheet knowledge, these increases may be concentrated on knowledge of
spreadsheet features rather than on spreadsheet development knowledge that will
increase aspects of system quality (McGill & Dixon, 2001).

As discussed above with respect to the relationship between experience and spreadsheet
development knowledge, the lack of strength in the relationship also serves as a
reminder that other factors not included in this model may influence an end user’s level
of spreadsheet development knowledge.

10.3.3 Spreadsheet development knowledge influences system quality

As hypothesised, spreadsheet development knowledge had a significant positive
influence on system quality (hypothesis H3). End user developers with greater levels of
spreadsheet development knowledge developed spreadsheets of higher system quality.
This finding is consistent with the results of Agboola (1998) who found a strong
relationship between data modelling knowledge and the quality of database
implementations. However, the findings differ from those of Kreie (1998) who found no
relationship between knowledge of spreadsheets and spreadsheet quality. Kreie
speculated that the lack of relationship in her study was because all of her participants
had relatively high levels of spreadsheet knowledge (an average of 86% on her
knowledge test, with little variability). She concluded that beyond a certain level,
knowledge of spreadsheet software is probably not a determining factor in the quality of
spreadsheets. The instrument used to measure spreadsheet knowledge in Kreie’s study
differs from the instrument used in Study 2 in that it focuses on only one aspect of
spreadsheet development knowledge: knowledge of spreadsheet features. A consideration of other facets of spreadsheet development knowledge could perhaps have provided her with a clearer insight into the relationship.

Hypotheses H1 to H3 were designed to provide further insight into the relationships between experience and system quality, and training and system quality. In Study 1, neither amount of previous spreadsheet experience nor amount of previous spreadsheet training influenced system quality. Hence, one of the aims of Study 2 was to determine why the proposed relationship was not found. Possible answers were that end users do not acquire the necessary knowledge and skill to develop good quality applications and/or that they do not apply the knowledge they have. Taken together with the results of the previous two hypotheses the result for hypothesis H3 provides a different picture from that obtained in Study 1. End user developers with higher levels of experience and training appeared both to acquire the necessary knowledge and skill to develop good quality applications, and to apply the knowledge they have to develop good quality applications.

One reason for these conflicting findings might be the nature of the applications developed. The spreadsheets developed by the participants in Study 1 were much larger and more complex than the applications developed by the participants in Study 2. The average file size in Study 1 was 182K with the largest file being 814K. Many of these applications consisted of linked worksheets and linked workbooks, increasing the complexity. In Study 2, the average file size was 16K and there was no linking between workbooks. The applications considered in Study 1 are more typical of those found in the workplace (Hall, 1996). Application size has been found to be correlated with
degree of application risk (Schultheis & Sumner, 1994). The relationship between spreadsheet development knowledge and system quality may be direct when the required applications are of a size and complexity achievable by the end users. However, it is possible that when the task to be supported is very complex or very simple the relationship does not hold, given the typical levels of end user developer knowledge. This would explain both Kreie’s (1998) results and the results of Study 1, and is consistent with the concerns expressed by Jarvenpaa, Dickson and DeSanctis (1985) about the impact of overly complex tasks on results obtained in decision making studies.

The task used in Kreie’s (1998) study was trivial, requiring very little spreadsheet knowledge. It was thus not surprising that no relationship was found between spreadsheet knowledge and system quality in that study. However, the decision support systems developed in Study 1 needed to be quite complex and sophisticated to meet the requirements of the BPG. A large proportion of the participants in Study 1 may have been operating beyond their zone of capability, and hence a relationship between spreadsheet knowledge and system quality was not detected. In Study 2, although the applications developed were small, the task was of medium complexity. This may have provided a setting in which the relationship between spreadsheet development knowledge and system quality could be observed.

The applications developed in Study 2 were smaller than many UDAs created in organisational settings, yet those developed in Study 1 appeared to be of about average size. This raises concerns about whether end users are currently capable of developing quality applications of the types they appear to require. If typical levels of spreadsheet
development knowledge are not sufficient for ‘normal’ applications then organisations should intervene either to improve levels of knowledge, or to provide guidelines about what kinds of applications are suitable for end user development. Several authors have explored the role of domain knowledge in the success of end user development, but have concluded that the limiting factor is usually knowledge of the development tool rather than domain knowledge (Agboola, 1998; Galletta et al., 1993; Mackay & Elam, 1992).

Another factor that may have contributed to the difference in the relationships observed between experience and system quality, and training and system quality between Study 1 and Study 2 is the levels of motivation and interest of the participants. Cronan and Douglas (1990) identified large improvements in productivity due to user development of applications following a training program. They noted that the large effect they had found might have been partially due to the fact that the participants were particularly enthusiastic. The same could be speculated about the participants in Study 2. Despite having slightly less experience and training than those in Study 1, all were volunteers, and appeared eager to participate and demonstrate their knowledge. In Study 1 the participants had less commitment to the study.

10.3.4 System quality influences perceived system quality

As hypothesised, in Study 2 system quality had a significant positive influence on perceived system quality for both end user developers and end users using applications developed by another end user (hypothesis H4). This finding supports the implicit assumption in much of the organisational IS literature that user perceptions of system quality reflect ‘actual’ system quality (DeLone & McLean, 1992). This finding is,
however, in contrast to the lack of relationship found in Study 1, and to the concerns expressed in the literature about the ability of end users to make realistic judgements of system quality (Edberg & Bowman, 1996; Kreie et al., 2000; Shayo, et al., 1999). The difference in results between Study 2 and Study 1 may again be due to the size and complexity of applications being considered. Increased feedback may also have played a role. It should be noted that the measurement model for perceived system quality was developed Study 2 using the same approach as was used in Study 1. Therefore, the concern raised in Section 6.7.2, that measurement problems may have contributed to the lack of relationship found in Study 1, seems unlikely to be true.

As discussed in Chapter 6, Huitfeldt and Middleton (2001) argued that the standard system quality criteria are oriented towards information technology maintenance staff rather than end users, and that it is difficult for an end user to evaluate the quality of an application. Although the instruments used to measure perceived system quality in both Study 1 and Study 2 were derived from an instrument designed specifically for end users (Rivard et al., 1997), informal feedback from participants suggested that they found system quality assessment a difficult task. Yet, despite these perceptions of difficulty, end user assessments of system quality were correlated with the independent assessments in this study. The applications developed in Study 2 may have been of a size and complexity that made it possible for end users to make relatively accurate assessments of system quality. The difficulty of assessing system quality is likely to be magnified as the size and complexity of applications increases.

In a study of user evaluation of task-technology fit, Goodhue, Klein and March (2000) commented on the role of performance feedback. They noted that the link between user
evaluations and objective measures might be stronger when feedback is present. In Study 2, participants received informal feedback about their applications when answering the 10 car rental questions. Whilst participants were not told whether their answers were correct, they quickly realised if their application was unable to support a particular task or if it was cumbersome for it to do so. This may have increased their awareness of issues such as functionality, ease of use and maintenance, and hence may have raised the accuracy of system quality perceptions. In Study 1, because the applications were larger and more interlinked, and because outcomes were dependent upon the combined efforts of a team, individual participants may not have received the same degree of feedback about their own applications.

10.3.5 Spreadsheet development knowledge does not influence perceived system quality

It was hypothesised that end users with more spreadsheet development knowledge would perceive applications to be of lower system quality (hypothesis H5). This hypothesis was developed based on the results of Study 1, and justified by the assumption that end users with low levels of spreadsheet development knowledge may not recognise system quality problems and hence may have inflated perceptions of quality, whereas end users with high levels of knowledge may be more critical. However, this hypothesis was not supported either for the end users using applications they had developed themselves or for end users with applications developed by someone else. In both cases, there was a slight positive relationship but it was only significant in the case of end users assessing the quality of applications developed by someone else.
Two possible reasons why this hypothesis was not supported involve the size of the applications and the role of performance feedback on perceptions of system quality. As previously discussed, the applications in Study 2 were smaller and less complex than those in Study 1. It is likely that system quality problems were more visible in these applications so that even those with little spreadsheet development knowledge became aware of them. As discussed above, this awareness process was also facilitated by the informal feedback participants received when answering the car rental questions.

Marcolin et al. (2000) found that subjects who completed a test of software knowledge before rating their self-efficacy rated themselves lower than those who completed the test after rating their self-efficacy. They found that the knowledge test was an anchoring stimulus that resulted in assessments that were more accurate. In Study 2 both the experience of answering the spreadsheet development knowledge questions (before answering the perceived system quality items) and of trying to answer the car rental questions using either their or another’s UDA may have moderated the proposed negative influence of lack of knowledge.

10.3.6 Involvement influences perceived system quality
As described in Section 9.2.4, the importance subscale of involvement was found to be less well fitting than the personal relevance subscale. This may have been because of the artificial nature of the task participants were asked to do. However, the results of the structural model evaluation suggest that the scale was suitable to include. As hypothesised, involvement had a significant positive influence on perceived system quality for both the end users using applications they had developed themselves, and for end users with applications developed by another end user (hypothesis H6).
These findings are consistent with the literature linking involvement to user satisfaction (Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Blili, Raymond, & Rivard, 1998; Doll & Torkzadeh, 1988; Lawrence & Low, 1993), but suggest that this relationship might be mediated via perceived system quality. Whilst not formally explored in this study, post hoc analysis was used to compare whether a direct relationship between involvement and user satisfaction provided a better fit than the indirect relationship tested in the model used in Study 2. An additional direct path between involvement and user satisfaction proved to have a non-significant $t$ value (estimate = 0.012, S.E. = 0.08, $t = 0.222$, $p > 0.05$) demonstrating that the relationship is more likely to be mediated via perceived system quality. This is in contrast to Seddon and Kiew’s (1994; 1996) suggestion that perceived system quality and perceived information quality are unlikely to be influenced by involvement.

The relationship between involvement and perceived system quality was of approximately the same strength as that between system quality and perceived system quality, indicating just how important end user involvement is to the success of an application.

10.3.7 System quality influences individual impact if the user is also the developer

It was hypothesised that system quality would have a significant positive influence on individual impact for both groups (hypothesis H7). The hypothesis was supported for the end user developers using their own applications, however the relationship was weak. Although there was a weak positive relationship between system quality and
individual impact for the end users using an application that was developed by someone else, this relationship was not significant. The first result is consistent with Hubona and Cheney’s (1994) study in which they found a significant direct relationship between system characteristics and user performance, but believed it to be less important than the effect mediated by user satisfaction.

The relationship between system quality and user satisfaction may not have held for the group of end users using an application that was developed by someone else, because of the role of the users’ spreadsheet knowledge. Unlike compiled applications, spreadsheet applications may be easily corrected and enhanced, unless heavily protected. This is particularly true for relatively small and straightforward applications such as the ones in Study 2. It may be that knowledgeable users of both their own and other applications were able to overcome the defects in lower quality applications to achieve good performance on the car rental questions by making minor adjustments to the UDAs (see Section 10.3.12 below). Whilst participants were instructed not to attempt to redesign spreadsheets (see Section 8.4) they were told they could make corrections and adjustments to the spreadsheets, and did so.

10.3.8 Perceived system quality influences user satisfaction

As hypothesised, perceived system quality had a strong positive effect on user satisfaction for both the end user developers using their own applications and those using applications developed by someone else (hypothesis H8). This is consistent both with previously published results (Rivard et al., 1997; Seddon & Kiew, 1996) and with the results of Study 1.
10.3.9 User satisfaction influences perceived individual impact

As hypothesised, user satisfaction had a strong positive effect on perceived individual impact for both the end user developers using their own applications and those using applications developed by someone else (hypothesis H9). This is consistent both with previously published results (Gatian, 1994; Gelderman, 1998) and with the results of Study 1.

10.3.10 Individual impact does not influence perceived individual impact

It was hypothesised that individual impact would have a significant positive influence on perceived individual impact for both groups. This hypothesis was not supported for either group. The lack of direct influence of individual impact on perceived individual impact is consistent with concerns expressed in the literature about the use of measures such as perceived individual impact as surrogates for independent measures of impact on individual performance (Goodhue, 1995; Goodhue et al., 2000). Goodhue, Klein and March (2000) obtained conflicting results in their attempt to address the relationship between user evaluations and objective performance. They found a significant relationship with one measure of performance (time to complete a task), but not with the other one that they used (accuracy of decision making).

The lack of direct relationship found between individual impact and perceived individual impact may again highlight the role of feedback. Whilst end users in Study 2 were aware if their application was unable to provide an answer to a question, they did not receive formal feedback on the accuracy of their decisions. In this artificial situation, participants may not have had sufficient awareness of the quality of their
decision making for individual impact and perceived individual impact to be directly related. The relationship proposed by Seddon (1997b) in which net benefits to individuals directly impact upon perceived usefulness, is therefore likely to be contingent upon strong feedback to users about benefits received. Presumably, in the workplace end users have higher levels of feedback than that provided in Study 2. However, organisations should be mindful of the need for this feedback, and try to provide for it in task design.

The lack of direct relationship between individual impact and perceived individual impact needs to be considered with respect to the interpretation of the results of Study 1. No relationship was found between perceived individual impact (no measure of individual impact was available) and organisational impact in Study 1. It was speculated that perceived individual impact may not be a good indicator of individual impact in terms of decision making performance, and that therefore its relationship with organisational impact may have been distorted. The results of Study 2 can not, however, either confirm or rule out this possible explanation. This is because although individual impact was not found to directly influence perceived individual impact in Study 2, individual impact and perceived individual impact were significantly correlated (user developer data: \( r = 0.354, p < 0.001 \); non user developer data: \( r = 0.411, p < 0.001 \)). This correlation was because of the influence of user satisfaction on each construct. Thus, perceived individual impact may nevertheless provide some indication of independently measured individual impact.
10.3.11 User satisfaction influences individual impact

As hypothesised, user satisfaction had a significant positive effect on individual impact for both the end user developers using their own applications and those using applications developed by someone else (hypothesis H11). These results are consistent with those of Hubona and Cheney (1994) who found both a direct relationship between system characteristics and user performance, and a relationship mediated by user satisfaction. The results are also consistent with those of Gatian (1994) and Gelderman (1998) as these studies both included direct measures of individual impact.

Davis and Srinivasan (1988) commented that although user satisfaction is commonly used to judge IS success because it is a more convenient measure than performance related measures, there is also an implicit assumption that user satisfaction with an IS results in some positive change in user behaviour resulting in increased effectiveness. The results for Study 2 provide support for this assumption, however the relationship was only of moderate strength, indicating that other factors also play a major role in determining individual impact.

Several authors have questioned the idea that individual impact is dependent upon user satisfaction (Etezadi-Amoli & Farhoomand, 1996; Hufnagel, 1990), and several have gone further, suggesting that it is more likely that the relationship is in the other direction (Ballantine et al., 1998; Seddon, 1997b). Hubona and Cheney (1994) claimed that their results rule this option out. Post hoc analysis on the data from Study 2 is not conclusive. Changing the direction of the relationship between user satisfaction and individual impact had no impact on model fit (see Appendix J for model fit information). However, if it is assumed that perceived individual impact influences user
satisfaction instead of vice versa, the model is much less well fitting, with none of the
guidelines for good fit being met (see Appendix J for model fit information). This
suggests that the relationship is unlikely to be in this direction.

10.3.12 Spreadsheet development knowledge influences individual impact

As hypothesised, spreadsheet development knowledge had a significant positive effect
on individual impact for both the end user developers using their own applications and
for those using applications developed by someone else (hypothesis H12). Thus,
successful use of user developed spreadsheet applications in this study appeared to
require substantial end user knowledge. This is likely to be because of the lack of
separation of data and processing that is commonly found in spreadsheet applications
(Hall, 1996; Ronen et al., 1989). Development of good quality UDAs should however
reduce this dependence on the development tool knowledge of the user. Users in this
study required spreadsheet development knowledge for successful use because the
applications were generally of low quality. Application development tool knowledge
may also play an important role in the use of other kinds of UDAs and should be
investigated in future research.

10.4 Support for the Study 2 model of UDA success

Study 2 was designed to provide further insight into the research questions for the thesis
by pursuing some of the issues raised in Study 1. In order to do this a model was
developed and tested. The proposed model appears to provide a useful model of the
relationships under consideration. Model fit was satisfactory, and the model explained a
large proportion of the variance in user satisfaction and perceived impact, but only a
moderate amount of the variance in the other dependent variables. This suggests that there are additional constructs that should perhaps be included in the model (these are discussed in Section 10.4.1 below). The majority of hypotheses tested were supported (see Figure 10.1 below), but several were not. Nevertheless, this model can be considered to have provided a useful contribution to research on UDA success.

![Figure 10.1: Relationships supported in Study 2](image)

The results of the structural model evaluation suggest that end user developers with more experience and training attain higher levels of spreadsheet development knowledge and consequently develop better quality spreadsheets. Training and experience appear to contribute about equally to spreadsheet development knowledge. Those who develop better quality spreadsheets also appear to perceive them to be of better quality. However, as there was a low correlation between spreadsheet development knowledge and perceived system quality, level of spreadsheet development knowledge does not appear to directly influence the ability of end users to assess the quality of their applications.
End user developers who are more involved with their applications, do however, perceive the applications to be of higher quality. System quality and involvement appear to have influences of approximately equal strength on perceived system quality. This higher perceived quality influences end users’ satisfaction with the applications, which is in turn translated into both greater perceived individual impact and greater directly measured individual impact. Individual impact is also directly influenced by both system quality and spreadsheet development knowledge. The direct effect of spreadsheet development knowledge appears to be slightly larger than that of user satisfaction, but the effect of system quality is weak. However, when indirect effects are also considered the role of system quality strengthens to be approximately the same as that of user satisfaction.

As discussed above, the hypothesised model paths that were supported in Study 2 suggest that for small to moderate size applications, increases in spreadsheet development knowledge lead to increases in system quality. They also suggest that for these kinds of applications, end users have perceptions of system quality that are consistent with expert perceptions. Therefore, organisations should be able to rely on end users’ judgements for these kinds of applications. However, if the applications developed in Study 1 are more typical of those being developed in the workplace than those in Study 2, then the results of these two studies highlight the need to either increase end users’ levels of development knowledge via training so that they can cope with applications of greater complexity, or to provide other forms of support for development such as intelligent tool support (Shah & Lawrence, 1996).
As mentioned earlier, Jarvenpaa, Dickson and DeSanctis (1985) obtained inconclusive results in an experimental study of the effectiveness of graphical output in decision making, and partially attributed this outcome to using an experimental task that was too difficult. They provided cautions about the choice and development of experimental tasks in IS research and urged researchers to ensure that task difficulty is appropriate. This caution about the need to ensure appropriate task difficulty also applies to organisations. Unfortunately, it may be the case that in end user development end users are trying to do too much. They may be attempting to develop applications that are too difficult for them given their current levels of training and hence levels of development knowledge. Salchenberger (1993) provided a set of conditions that should be met if end user development is to be appropriate. Conditions that were violated by the kinds of applications participants developed in Study 1 include:

- System scope should be limited (not organisation wide)
- Level of decision-making supported should be operational or tactical
- System should not be not overly complex
- Data significance to other departments should be limited.

These and other conditions may routinely be violated in organisations where end user developers are overly ambitious in their UDA development.

Nevertheless, the fact that end users’ perceptions of the quality of small to moderate UDAs are consistent with those of independent assessors for smaller applications should provide some reassurance to organisations that rely heavily upon end users assessments of quality and fitness for use. The results should, however, be treated as a reminder that organisations should provide guidelines as to the types of applications that are suitable for end user development.
The results of Study 2 provided new insight into the role of user involvement in end user development. In addition to confirming that the beneficial effects of involvement discussed in the organisational IS literature (e.g. Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Doll & Torkzadeh, 1988; Lawrence & Low, 1993) apply in the UDA domain, this study has clarified the process by which the benefits are obtained. The results suggest that the influence of involvement on user satisfaction is mediated via perceived system quality.

The other major contribution of the model tested in Study 2 is the insight it has provided into the importance of spreadsheet development knowledge for successful use of a spreadsheet application. Users of organisational systems do not require knowledge of the tools with which systems are developed, but for user developed spreadsheet applications, successful use appears to require sufficient knowledge to understand and, if necessary, alter the application. This relationship between tool knowledge and individual impact should be explored for other kinds of end user development tools in future research.

The relationships found between perceived system quality, user satisfaction and perceived individual impact in Study 2 are consistent with those found in Study 1. The finding that user satisfaction also positively influences individual impact is encouraging. It suggests that for small to medium sized applications, the two separate UDA success subsystems alluded to in Section 6.3 (one mediated by user satisfaction and one that involves a direct influence of system characteristics on performance) are closely linked. This means that performance is not being unduly influenced by system quality problems
that end users are unaware of. User satisfaction appears to be a reasonable indicator of success for organisations to use for this kind of application, although as a large proportion of the variance in individual impact remained unexplained, it can not be totally relied upon.

In summary, Study 2 has made a valuable contribution to research on UDA success. It has provided additional insight into the way in which UDAs influence individual performance and has identified several factors that can influence this success.

### 10.4.1 Additional constructs that might be relevant

As mentioned above, whilst the model proposed in Study 2 explained a large proportion of the variance of user satisfaction and perceived individual impact, the other dependent variables were only partially explained by the model. Several issues were raised in the discussion of the individual hypotheses that may require the inclusion of additional constructs in the model.

Only approximately a quarter of the variance in spreadsheet development knowledge was explained by the influences of experience and training. Whilst this may be partially due to the operationalisations of these constructs, other factors such as cognitive ability (Simon et al., 1996) may play a role and should be investigated in future research.

Neither system quality nor individual impact was well explained by the model (21.5% and 28.5% of variance explained for the user developer data), so other factors that might influence these variables should be considered. As both the development of a UDA and successful use of it are tasks, Campbell’s (1990) theory of individual task performance
may be of use. This theory suggests that individuals make choices relating to the degree of effort to invest in a task and this may influence their performance.

A measure of the nature of the task being accomplished by the UDA also might be a valuable addition to models of UDA success. Brancheau and Brown (1993) included task as one of four individual level factors of importance in their model of EUC management, and Raymond and Bergeron (1992) included task variety in their model of decision support system success. In previous research, lack of task structure has been shown to have negative effects on user satisfaction with IS (Guimaraes & Igbaria, 1997), and task complexity has been shown to influence user satisfaction (Blili et al., 1998). However, most of the research relating to tasks has looked at variables such as the number of software packages used and the frequency of use and has not dealt specifically with applications developed by end users (e.g. Kim, Suh, & Lee, 1998). Hence, further investigation of the role of task complexity and structure seems warranted.

The role of domain knowledge was discussed briefly in Section 10.3.3. Although research suggests that development tool knowledge (rather than domain knowledge) is usually the limiting factor in determining UDA success (Agboola, 1998; Galletta et al., 1993; Mackay & Elam, 1992), domain knowledge is an important prerequisite for successful UDA development, and should be considered when investigating task complexity and structure.

The size and complexity of applications developed in response to task demands also appears to influence the proposed relationships in the models considered in this thesis.
Therefore, this should be considered in future models. Little has been written about the role of application size or complexity, and although Raymond and Bergeron included application type in their model of decision support system success, they did not consider the size or complexity of the applications. Reithel, Nichols, and Robinson (1996) noted that the probability of errors is higher in larger spreadsheets, yet end user developers were more confident about large spreadsheets than about small spreadsheets. The size and complexity of applications developed by end users is growing (McLean, Kappelman, & Thompson, 1993; Taylor, Moynihan, & Wood-Harper, 1998), so a better understanding of the role of these factors in UDA success is required.

Another issue raised by the results of this study is the role of performance feedback. Feedback reduces uncertainty about user performance (Larson, 1984). Martocchio (1992) showed that feedback could increase both self-efficacy and the performance of users. In a study of user evaluation of task-technology fit, Goodhue, Klein and March (2000) commented on the role of performance feedback. They noted that participants in their study lacked feedback on their performance and suggested that the link between user evaluations and performance measures might be stronger when feedback is present. This suggests that raising the awareness of the quality and performance of applications may influence the relationship between perceptual measures of success and measures that are more direct. The role of feedback should be investigated in future research.

Information quality was not considered in the model tested in Study 2, despite its importance in the models tested in Study 1. This was partially because the focus of the thesis is on the role of system quality in UDA success, and partially because of difficulties in obtaining independent measurements of information quality. However,
the model from Study 2 should be extended to incorporate the role of information quality.

10.5 Contribution of Study 2 to answering the research questions

Chapter 1 stated the research questions for this thesis, and Chapter 6 discussed the contribution that Study 1 made towards answering these questions. This section discusses the additional contribution of Study 2 towards answering the research questions.

The first research question for the thesis was:

*How does UDA quality contribute to user performance on tasks?*

Neither of the models tested in Study 1 provided a complete explanation of how UDA quality contributes to user performance on tasks. The lack of relationship between system quality and perceived system quality was of particular concern. If perceived system quality does not reflect system quality, it raises the issue of how UDA success is mediated. In Study 2, perceived system quality reflected system quality, and the mechanism by which system quality contributes to user performance was demonstrated to be both mediated through user satisfaction and direct. Study 2 supported the finding from Study 1 that perceived system quality influences user satisfaction, which in turn influences perceived individual impact. In addition, Study 2 demonstrated that user satisfaction with UDAs can directly influence individual impact (when measured independently). This is a valuable contribution to the study of UDA success.
However, the difference in results between the two studies suggests that factors not considered in the models may influence the relationship between system quality and user performance. Application size and complexity, and task structure and complexity may be amongst these variables, but further research is required to explore these possibilities.

The second research question for this thesis was:

*Do end user developers have any misconceptions about the quality of their applications? If so, how do these misconceptions impact upon their ability to judge whether the applications they develop will have a positive impact on their performance of tasks?*

In Study 1, no relationship was found between perceived system quality and system quality, and concerns were raised that user satisfaction with UDAs may be based upon erroneous perceptions. Study 2 provided an opportunity to compare more closely end user perceptions of system quality with independent measures. In Study 2 system quality and perceived system quality were significantly correlated. This consistency suggests that misconceptions need not always occur and the possibility was raised that misconceptions about system quality are more likely when end users are developing applications that are complex and/or large. Study 1 did not include an independent measure of individual impact, but in Study 2 both perceived individual impact and individual impact were directly influenced by user satisfaction and were significantly correlated. This highlights the importance of user perceptions for UDA success and suggests that for applications of the size and complexity considered in Study 2 user perceptions of impact can be depended upon. It should be noted however, that the relationships between system quality and perceived system quality and between

275
perceived individual impact and individual impact were not a particularly strong, indicating that while user developers have some understanding of the quality and impact of their applications, it is far from perfect. In summary, the study provided some reassurance that misconceptions of the type experienced in Study 1 need not occur if end users are developing applications for which they have sufficient levels of spreadsheet development knowledge.

The third research question asked in this thesis was:

*What characteristics of end user developers influence their ability to judge whether the applications they develop will have a positive impact on their performance of the tasks the UDA is designed to support?*

Study 1 provided an initial investigation of the roles of experience and training in UDA success, however, the results were inconclusive. The negative relationship between experience and perceived system quality suggested that lack of experience might impede the ability of user developers to assess the quality of their applications. Study 2 was designed to extend the investigation undertaken in Study 1 to try and determine more clearly the roles played by experience and training. Study 2 also introduced two additional constructs: spreadsheet development knowledge and involvement. These constructs were included to help explain where the hypothesised relationships between experience and both perceived and independently measured system quality, and training and both perceived and independently measured system quality might be breaking down. Involvement was also considered to be another factor that might be influencing perceived system quality.
The anticipated negative relationship between experience and perceived system quality was not detected in Study 2, again suggesting that when end users are developing applications consistent with their levels of knowledge, they are able to judge the success of their applications. The problems may occur when they operate outside their zone of capability. However, this possibility will need to be explored further, as it was not tested explicitly in the two studies undertaken for this thesis.

The results of Study 2 suggest that the user characteristics of involvement and spreadsheet development knowledge play important roles in the UDA success process. Level of user involvement influences perceived system quality, which in turn influences user satisfaction. The comparison between end user developers using their own applications and end user developers using UDAs developed by other end users provides support for the importance of involvement in the UDA domain. End user developers had both higher levels of involvement with their own applications and performed better with them. This suggests that the process of developing an application leads to significant advantages for the end user developer. Conversely, this finding should also serve as a caution for organisations. If a user developer has developed an application for his or her own use and then leaves the position or organisation, it can not be assumed that another end user will necessarily be able to use it successfully.

Spreadsheet development knowledge was shown to be important in two ways. It influences the quality of the system being developed, but also acts directly upon the individual impact of the application. Thus, problems with the system quality of spreadsheet applications can be partially balanced by the spreadsheet development
knowledge of the user. The generalisability of this finding to UDAs developed using other tools needs to be explored in future research.

10.6 Limitations of the research described in this thesis

Several limitations of the research are apparent and must be considered in future investigations of the models presented in this thesis. The relatively small sample size in Study 1 can be considered a limitation of that study, though it was partially addressed with the larger sample size used in Study 2. The possibility that some of the differences between the results of Study 1 and Study 2 were due to lack of power because of the smaller sample size in Study 1 needs to be further explored.

The reliance on MBA students as participants in Study 1 and the use of a simulated business environment can also be considered limitations of that study. Whilst the limited range of end users considered and the artificial nature of the organisational impact measures are limitations, the strong internal validity of the approach should provide a strong foundation for future studies with a wider range of end user developers.

The goodness of fit of several of the measurement models was marginal in both studies (in particular perceived system quality). This may have influenced the results of the structural model evaluations. Whilst a conscious decision was made to fix the measurement of perceived system quality to that of system quality in order to enable direct comparisons to be made, it would be valuable to explore differences between end user and expert perceptions of system quality further.
Another limitation of the research is the fact that the only application development tool considered was spreadsheets. Whilst spreadsheets have been the most commonly used end user application development tool (Taylor et al., 1998), the generalisability of the results to users of other development tools, such as database management systems and Web development tools, needs to be investigated in future research. In particular, the role of tool and application development knowledge in the success of other types of UDA is an area for future research. Whilst spreadsheet development knowledge was shown to play an important role in successful use of end user developed spreadsheet applications, the importance of application development knowledge with other kinds of applications needs to be determined.

In addition to extending the research to the use of other development tools, the roles of task structure and complexity, and application size and complexity also need to be considered further. The conclusions drawn in this thesis are based on only two research situations and the research would benefit from consideration of a wider range of applications. Whilst the research situations chosen provided the benefit of control of external variability and hence internal validity, they were not ideal in terms of providing external validity. It would be valuable to undertake a field study in a range of organisations to extend the external validity of the research. This would address concerns relating to the nature of applications developed, and to the representativeness of the participants.
10.7 Overview

This chapter provided a discussion of the results of Study 2. Study 2 was designed to provide further insight into the research questions for the thesis by addressing issues raised in Study 1. A new model was developed for Study 2, which enabled the roles of perceived individual impact and independently measured individual impact to be further explored. The model also facilitated further investigation of the roles of training and experience and the relationship between system quality and perceived system quality. It also introduced two additional constructs: spreadsheet development knowledge and involvement.

The chapter first considered the results of the comparison between user developers using their own applications and user developers using applications developed by other end users. End user developers were found to be both more satisfied with their own applications and to make more accurate decisions with them. This may be attributable to the higher levels of involvement end user developers have with their applications. It was also speculated that a greater familiarity with their own applications, derived from the development process, enhances their performance when using them. Developing an application may help the user to develop a robust understanding of it that makes it easier to use successfully.

The proposed model appeared to provide a useful model of the relationships under consideration, for both end user developers and those using applications developed by other end users. The results of the structural model evaluation suggested that end user developers with more experience and training develop higher levels of spreadsheet development knowledge and consequently develop better quality spreadsheets. Those
who develop better quality spreadsheets also appear to perceive them to be of better quality. However, level of spreadsheet development knowledge does not appear to affect the ability of end users to assess the quality of their applications. End user developers who are more involved with their applications, do however, appear to perceive the applications to be of higher quality. This higher perceived quality influences their satisfaction with the applications, which in turn is translated into both greater perceived individual impact and greater independently measured individual impact. Individual impact was also found to directly influenced by both system quality and spreadsheet development knowledge.

In discussing each of these relationships, the results of Study 2 were compared with the results of Study 1 and implications discussed. The major differences in results between the two studies appear to relate to the relationships between individual characteristics, and both system quality and perceived system quality. The possibility is raised that differences in the size and complexity of tasks and applications between the two studies contributed to the differences in results, and the need to further explore the role of these characteristics in the success of end user development is highlighted. The results of Study 2 also made a valuable contribution to the study of UDA success by providing insight into the role of user involvement.
Chapter 11
Conclusions

11.1 Summary of the research and its contributions
This chapter concludes the investigation of UDA success described in this thesis. The research centred around the role of system quality in UDA success and enabled comparison of end user developer perceptions of system quality and individual impact with independent measures of these variables. This provided some insight into the ability of end user developers to judge whether the systems they develop have a positive impact on their performance of tasks. The research investigated factors that might impact upon this ability. The research objective and the associated research questions were addressed through two empirical studies.

Two models of UDA success provided the starting point for this research. The first model was DeLone and McLean’s (1992) model of IS success, and the second model was a version of DeLone and McLean’s model that was modified to address concerns about the DeLone and Mclean model (Seddon, 1997b) and to reflect current research about UDA success (e.g. Al-Shawaf, 1993; Edberg & Bowman, 1996; Hobbs, McGill, & Rowe, 1998; Igbaria, 1990; Janvrin & Morrison, 2000). This model was also extended to include the roles of user training and experience in UDA success.

Study 1 was a test of the two initial research models, using a field study. This field study involved business people who were also completing an MBA and were participating in a business policy simulation game where they developed their own
spreadsheet applications to assist in decision making. This approach allowed testing of the models in full, including measurement of a number of organisational outcome measures. Structural equation modelling was used to test the models.

Neither of the models was well supported by the data. Of the nine hypothesised relationships in the DeLone and McLean (1992) model, four were found to be significant and the remainder not significant. The analysis provided strong support for relationships between perceived system quality and user satisfaction, information quality and user satisfaction, user satisfaction and intended use, and user satisfaction and individual impact. Of the 13 relationships hypothesised in the alternate model only six were found to be significant. It is notable that the model paths that were supported in Study 1 were primarily those that reflect user perceptions. This study highlighted that user perceptions of IS success play a significant role in the UDA domain. The results of the test of the alternate model did, however, suggest that there may also be a direct relationship between system quality and impact on task performance.

The second study was designed to specifically address several issues arising from the first study. In particular, it was intended to:

- Clarify the roles of both individual impact as perceived by end user developers and independently measured individual impact in UDA success
- Investigate possible reasons for the lack of relationship demonstrated between perceived system quality and system quality in Study 1
- Investigate where the hypothesised relationships between experience and both perceived system quality and system quality, and training and both perceived system quality and system quality might be breaking down.
Study 2 was a laboratory experiment, and the participants were end users from a range of public and private organisations. A revised research model was developed based on both the findings from Study 1, and the UDA and IS success literature, and it includes two concepts that were not included in the Study 1 models: spreadsheet development knowledge and involvement. Structural equation modelling was again used to test the model.

The proposed model appeared to provide a useful model of the relationships under consideration. Ten of the twelve hypotheses were supported when the model was tested with the data from end user developers using their own applications for decision making. The results of the structural model evaluation suggest that end user developers with more experience and training build up higher levels of spreadsheet development knowledge and consequently develop better quality spreadsheets. Those who develop better quality spreadsheets also appear to perceive them to be of better quality. End user developers who are more involved with their applications also perceive them to be of higher quality. This higher perceived quality influences their satisfaction with the applications, which in turn is translated into both greater perceived individual impact and greater independently measured individual impact. Individual impact was also found to be directly influenced by both system quality and spreadsheet development knowledge.

There were, however, several notable differences in the results of the two studies. Firstly, experience was shown to have a moderate negative influence on perceived system quality in Study 1, but no similar effect was observed in Study 2. Also,
perceived system quality did not reflect independently measured system quality in Study 1, but did so in Study 2.

Substantial progress was made towards answering the research questions for this thesis. The first research question for this thesis was:

*How does UDA quality contribute to user performance on tasks?*

The mechanism by which UDA system quality contributes to user performance was demonstrated to be both direct and mediated through user satisfaction. In ideal situations, perceived system quality reflects system quality and influences user satisfaction, which in turn influences perceived individual impact. User satisfaction directly influences impact on individual task performance (when measured independently). In addition, UDA system quality has a weak direct influence on individual task performance. This clarification of the way in which UDA quality contributes to user performance is a valuable contribution to the study of UDA success. However, perceived system quality does not appear to reflect system quality in all situations. This suggests that factors not considered in this thesis may influence the relationship between system quality and user performance. Application size and complexity, and task structure and complexity may be amongst these variables.

The second research question for this thesis was:

*Do end user developers have any misconceptions about the quality of their applications? If so, how do these misconceptions impact upon their ability to judge whether the applications they develop will have a positive impact on their performance of tasks?*
As mentioned above, while end user developers were shown to have misconceptions about system quality in some situations, they were not as evident in others. This suggests that user satisfaction with UDAs may sometimes be based upon erroneous perceptions and hence will influence end users’ ability to judge whether the applications they develop will have a positive impact. However, the research provided some reassurance that misconceptions about system quality need not occur if end users are developing applications for which they have appropriate levels of spreadsheet development knowledge and task knowledge.

The third research question asked in this thesis was:

What characteristics of end user developers influence their ability to judge whether the applications they develop will have a positive impact on their performance of the tasks the UDA is designed to support?

The user characteristics of experience, training, spreadsheet development knowledge and involvement were all shown to play important roles in the UDA success process. Experience and training influence spreadsheet development knowledge, which in turn influences system quality. Spreadsheet development knowledge was shown to be important in two ways. It influences the quality of the system being developed, but also acts directly upon the individual impact of the application. Thus, problems with the system quality of spreadsheet applications can be partially balanced by the spreadsheet development knowledge of the user.

The research undertaken for this thesis suggests that when end users are developing applications of appropriate sophistication for their level of knowledge, they are able to judge the success of their applications. Problems may occur when they operate outside
their zone of capability, however, this possibility will need to be explored further, as it was not tested explicitly in this thesis.

Level of user involvement was shown to influence perceived system quality, which in turn influences user satisfaction, which influences individual impact. Thus, those end user developers with higher levels of involvement are more satisfied with their applications and perform better with them.

11.2 Implications of these studies for UDA success research

The research described in this thesis has added to the existing research on UDA success. It has provided insight into the role of user involvement in end user development. In addition to confirming that the beneficial effects of involvement discussed in the organisational IS literature (e.g. Amoako-Gyampah & White, 1993; Barki & Hartwick, 1994; Doll & Torkzadeh, 1988; Lawrence & Low, 1993) apply in the UDA domain, the process by which the benefits are obtained was clarified. The influence of involvement on user satisfaction was shown to be mediated via perceived system quality.

Whilst the importance of application development knowledge has been assumed in the UDA literature, this research has provided evidence of its role. Spreadsheet development knowledge was shown to play two roles in the successful use of spreadsheets. It not only influences the quality of the application being developed, but also directly influences the task performance of the end user developer. The relationship between application development tool knowledge and individual impact needs to be explored for other kinds of end user development tools.
This research has also confirmed the importance of user satisfaction as the key mediating variable the UDA success process. User satisfaction was shown to be directly influenced by perceived system quality and to directly influence both perceived individual impact and independently measured impact on task performance. In Study 2, where the applications developed were relatively small, user satisfaction was shown to be indirectly influenced by experience, training, spreadsheet development knowledge, system quality and involvement. The fact that approximately half of the effect of system quality on individual impact is mediated through user satisfaction is of particular interest. Overall, these findings suggest that for applications of small to moderate size and complexity, user satisfaction is an appropriate indicator of UDA success, consistent with its use as the most important success indicator for organisational systems (Gatian, 1994; Gelderman, 1998; Seddon, 1997b). However, the role of user satisfaction needs to be further explored with a range of tasks and applications of different sizes and complexities.

Whilst perceived individual impact and independently measured individual impact were shown to be correlated in this research, no direct relationship between them was established. This correlation was because of the direct influence of user satisfaction on each construct. This thesis has highlighted that the nature of the relationship between them needs further research.

The DeLone and McLean (1992) model of IS success was the starting point for the research discussed in this thesis. DeLone and McLean clearly identified separable measures of IS success and proposed relationships between them. This thesis has shown
that while their model provided a useful initial framework for an investigation of UDA success it does not satisfactorily model UDA success. The research described in this thesis has demonstrated that unless applications are small and not complex, there is a need to distinguish between user perceptions of system quality and independent expert assessments of system quality. It also suggests that use does not play the central role in UDA success indicated by DeLone and McLean’s model. This thesis has also highlighted a number of factors that should be considered in future models of UDA success. These factors include those explicitly tested in Study 1 and Study 2 (experience, training, spreadsheet development knowledge and involvement) and others proposed in the discussion of the results (e.g. application size and complexity, task structure and complexity, domain knowledge, cognitive ability, degree of effort invested, and feedback on performance).

11.3 Practical implications of the research

The results described in this thesis have practical implications for the management of user development of applications in organisations. Despite early concerns about its risks (e.g. Alavi & Weiss, 1985-1986; Davis, 1988), end user development has become an integral part of organisational information provision (McLean, Kappelman, & Thompson, 1993; Shayo et al., 1999). With the increasing availability of World Wide Web technology to end user developers has come an expanded set of opportunities and risks for organisations. UDAs can now have more profound effects on business processes, partners, and customers than ever before, as Web technology permits end users to design applications that are immediately accessible by unlimited numbers of people from anywhere in the world (Nelson & Todd, 1999).
Training has been proposed as a valuable means to improve the quality of UDAs and to increase the impact of UDAs. For example, Nelson (1991) suggested that training is perhaps the most effective tool for minimising the risks associated with EUC. Cragg and King (1993) also called for an increase in end user training. In addition to the role of training in facilitating a direct improvement in system quality, they recognised the importance of training to help end user developers bridge the gap between being competent assessors of the quality and impact of small applications to being competent assessors of large and complex applications.

This thesis has shown that end users with more training have higher levels of spreadsheet development knowledge and develop better quality applications. This should provide encouragement to organisations that already invest in end user development training. However, if the applications developed in Study 1 are more typical of those being developed in the workplace than those developed in Study 2, then the levels of training currently received by end users are not sufficient. The results of the two studies highlight the need to enable end users to successfully develop larger and more complex applications. This might be achieved by increasing end user levels of development knowledge via training so that they can cope with applications of greater complexity, or by providing other improved forms of support for development such as the intelligent end user development tools that have been suggested by several authors (Isakowitz, Schocken, & Lucas, 1995; Shah & Lawrence, 1996).

The results of Study 1 suggest that end users may be attempting to develop applications that are too complex for them. Whilst increased training may prepare end user developers to cope with more complex applications, the role of organisational standards
and guidelines should also be considered. In concluding their field study of spreadsheet development practices, Cragg and King (1993) stated that there is a need for setting and enforcing organisational spreadsheet standards. This is despite evidence that users are less satisfied when subject to greater application development control (Bergeron & Berube, 1988; Bowman, 1988). The results described in this thesis suggest that there is a particular need for guidelines on the kinds of applications that are suitable for end user development. Whilst the model tested in Study 2 suggests that end users are able to make reliable assessments of the quality and impact of small systems, and hence may have less need for organisational intervention, when larger and more complex applications are developed users this may not be the case. Several authors have proposed guidelines recommending what kinds of applications are appropriate for end user development (Salchenberger, 1993), and what kinds are not (Bowman, 1990). These types of guidelines need to be researched further so that more detailed guidance can be provided to prospective end user developers. In particular, the ability to tailor recommendations on what types of applications are appropriate to individual end users’ backgrounds would be very valuable.

Given that previous research has found low compliance with end user development standards and procedures (Bowman, 1988; Taylor et al, 1998), and that organisational support policies and procedures appear to be more influential in determining user development success than control policies and procedures (Crawford, 1986), guidelines such as those discussed above might best be introduced via training and other forms of organisational support. However, in a recent study of organisational support for end user development of Web applications, Nelson and Todd (1999) found that most firms
seemed to be relying on generation of standards as a means of control. They also found that there was limited support for end user development.

The possible role of intelligent end user development tools was mentioned earlier in this section. In recent years there has been a proliferation of new features in each new release of spreadsheet software. However, anecdotal evidence suggests that these new features have not led to increases in the ability of end user developers to develop high quality spreadsheets. Rather, the fundamentals of ‘what’ the end user is attempting to do are being obscured by the multiplicity of ways ‘how’ to achieve it. The results of the research described in this thesis suggest that the attentions of spreadsheet designers should be focused on providing support to end user developers in the process of application development. Rather than requiring more advanced functions they appear to require support in planning, developing and testing their applications.

The findings about the role of involvement in the UDA success process also have practical implications for organisations. Involvement was shown to influence perceived system quality, which in turn has an influence on user satisfaction and performance. Involvement with UDAs therefore appears to be something that should be supported. This could perhaps be accomplished by formalising the role of user application development in job descriptions. By recognising its importance, organisations may enhance the involvement of the end users.

End user developers were also shown to be more involved with their applications than were end users who use applications developed by other end users, and to perform better with them. The implication of this finding also needs to be considered by
organisations. If an end user develops an application for his or her own use, and its use has a positive impact on performance, this does not guarantee that the same will be true if another end user starts to use it. Organisations should recognise that the use of UDAs by end users other than the developer may carry with it greater risks. In addition, if users are developing applications for others to use, particular attention must be paid to ensure that these applications are of sufficient quality for successful use not to rely solely on the additional insight gained during the development process.

This research has provided insight into the role of user satisfaction in UDA success. This is important in terms of the management of user development of applications, because organisations rely on user perceptions of quality and fitness for use. The research described in this thesis suggests that for small to moderate applications, when managers can be confident that end user developers have the appropriate application development knowledge, user satisfaction is an appropriate indicator of the impact of UDAs on task performance. However, this conclusion is preliminary and needs to be verified with a wide range of end users, tasks and applications.

In conclusion, the research described in this thesis represents significant progress towards understanding the UDA success process. The importance of end users’ perceptions of their applications has been highlighted, and the roles of spreadsheet development knowledge and involvement clarified. The implications for the management of user development of applications have been discussed and several directions for future research have also been identified.
Appendices
Appendix A
Study 1 questionnaire
Thank you for agreeing to participate in this study! As you know, Associate Professor Lanny Entrekin has approved my interaction with you to help in my research. All information provided will be strictly confidential. Reporting will be in statistical terms only.

Please complete this questionnaire and bring it to the C606 Strategic Management and Policy boardroom meeting on Wednesday 7 October. In addition, please make a copy (on the disk provided) of the part of your BPG decision support system that you have had the most involvement with. For example, it may be the marketing spreadsheet or the production spreadsheet or a series of interlinked worksheets. This copy should be as at the end of Year 4 (when you make your 8th decision).

If you would like to have a summary of the results of this study about the impact of decision support system quality on organisational performance please indicate in the box below.

Please send me a summary of the results  

If you have any questions about this project please don’t hesitate to contact me. Tanya McGill (ph. 335 5085).
**Background information**

1. How long have you been using computers? _______ Years

2. Apart from work relating to the Business Policy Game, on average how frequently do you use a computer? Please tick the box which best describes your average computer use
   - [ ] Less than once a month
   - [ ] Once a month
   - [ ] A few times a month
   - [ ] A few times a week
   - [ ] About once a day
   - [ ] Several times a day

3. How long have you been using spreadsheets? _______ Years

4. Please rate your skill with spreadsheets compared to others in this class by circling the number that best indicates your relative skill level
   - [ ] Little or no skill
   - [ ] 1
   - [ ] 2
   - [ ] 3
   - [ ] 4
   - [ ] 5
   - [ ] Extremely skilful

5. Which of the following categories best describes the level of training you have had in the use of spreadsheets. Please circle the number which best describes your level of training in each of the categories
   - (a) University or TAFE courses
   - (b) Training provided by vendors
   - (c) In-company courses
   - (d) Self study
   - [ ] None
   - [ ] 1
   - [ ] 2
   - [ ] 3
   - [ ] 4
   - [ ] 5
   - [ ] Extremely intensive

6. How old are you? _______ Years

7. What sex are you?
   - [ ] Female
   - [ ] Male

8. Which component of your company’s decision support system have you had the most involvement with? PLEASE TICK ONLY ONE
   - [ ] Marketing
   - [ ] Production
   - [ ] Finance
   - [ ] Other

8a. What is the file name (and worksheet name(s) if applicable) of the part of your company’s decision support system that you have had the most involvement with? For example MARKETING.XLS Sheet1 or MARKETING.XLS SalesPeople
   - File name: ________________
   - Worksheet: ________________

9. On average how many hours per week do you use the system you have the most involvement with? _______ Hours
   - rarely
   - often

10. Overall, how would you rate your intended use of the system over the next year of the BPG? 1 2 3 4 5

The remainder of the questions relate to the part of your company decision support system (e.g., marketing spreadsheet or production spreadsheet) that you have had the most involvement with in Year 4 of the Business Policy Game (BPG).

This will be “the system” that you discuss throughout the questionnaire. Please make your answers to ALL the following questions specific to this spreadsheet.

8. Which component of your company’s decision support system have you had the most involvement with? PLEASE TICK ONLY ONE
   - [ ] Marketing
   - [ ] Production
   - [ ] Finance
   - [ ] Other

8a. What is the file name (and worksheet name(s) if applicable) of the part of your company’s decision support system that you have had the most involvement with? For example MARKETING.XLS Sheet1 or MARKETING.XLS SalesPeople
   - File name: ________________
   - Worksheet: ________________

9. On average how many hours per week do you use the system you have the most involvement with? _______ Hours
   - rarely
   - often

10. Overall, how would you rate your intended use of the system over the next year of the BPG? 1 2 3 4 5
Could you please describe your involvement in the development and use of this part of your company decision support system (e.g. did you set it up alone or with someone else? have you made adjustments to it? did you use the output from it to make recommendations on decisions? etc.)

Please indicate on the scales below your perceptions of the information quality of the system by circling the most appropriate number for each item.

12 Do you think the output from your system is presented in a useful format? never 1 2 3 4 5 6 7 always
13 Are you satisfied with the accuracy of your system? 1 2 3 4 5 6 7
14 Is the information provided by your system clear? 1 2 3 4 5 6 7
15 Is the system accurate? 1 2 3 4 5 6 7
16 Does the system provide sufficient information? 1 2 3 4 5 6 7
17 Does the system provide up to date information? 1 2 3 4 5 6 7
18 Do you get the information you need in time? 1 2 3 4 5 6 7
19 Does the system provide output that seems to be just about exactly what you need? 1 2 3 4 5 6 7
20 Does the system provide the precise information you need? 1 2 3 4 5 6 7
21 Does the system’s information content meet your needs? 1 2 3 4 5 6 7
22 Is the information provided by your system understandable? 1 2 3 4 5 6 7
23 Is the information produced by your system valid? 1 2 3 4 5 6 7
24 Is the information provided by your system verifiable? 1 2 3 4 5 6 7
25 Is the information provided by your system complete? 1 2 3 4 5 6 7

Please indicate on the scales below your satisfaction with the system and your assessment of its impact on your performance by circling the most appropriate number for each item.

26 How adequately do you feel the system meets the information processing needs of your area of responsibility in the BPG? adequately 1 2 3 4 5 6 7 inadequately
27 How efficient is the system used for your area of responsibility? inefficient 1 2 3 4 5 6 7 efficient
28 How effective is the system? effective 1 2 3 4 5 6 7 ineffective
29 Overall, are you satisfied with the system? dissatisfied 1 2 3 4 5 6 7 satisfied
30 The system has a large, positive impact on my effectiveness and productivity in my role in the BPG agree 1 2 3 4 5 6 7 disagree
31 The system is an important and valuable aid to me in the performance of my role in the BPG agree 1 2 3 4 5 6 7 disagree
Please indicate on the scales below your perceptions of the system quality of the spreadsheet you have the most involvement with by circling the most appropriate number for each item. Note that these questions relate specifically to the system developed to support your decision making rather than to the spreadsheet package you used (e.g. Microsoft Excel).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>strongly agree</th>
<th>strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>The system is easy to use</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>The system is user friendly</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Compared to other systems I have used, the system is easy to learn to use</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>I find it easy to get the system to do what I want it to</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>It is easy for me to become skilful at using the system</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>I believe that the system is cumbersome to use</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Using the system requires a lot of mental effort</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Using the system is often frustrating</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
</tbody>
</table>

The following questions are more specific but also relate to your perceptions of the system quality of the spreadsheet you have the most involvement with. Please answer by circling the most appropriate number for each item.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>strongly agree</th>
<th>strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>Using the system is easy, even after a long period of non-utilisation</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Unauthorised users could easily access all the data or a part of it</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>The use of this system has resulted in the reduction in the number of errors I would make in performing my role in the BPG</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Each user owns a unique password</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Unauthorised access is controlled in several parts of the system</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Errors in the system are easy to identify</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>The system can be run on computers other than the one presently used</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>The system increased my data processing capacity</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>The system is easy to learn by new users</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>The system could be used in other similar organisational environments, without any major modification</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Each password limits the access to specific parts of the system</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>This system provides the capability to import data from other applications</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>This system (rather than the spreadsheet package) automatically corrects certain types of errors, at data-entry time</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>It is possible to copy parts of the system (outputs or data) into other systems or to link with other systems</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>Should an error arise, the system provides the capability to perform some checking in order to locate the source of error</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Decisions made on the basis of the information provided by the system are central to my role in the BPG</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>Processing of the various operations is fast (backup, calculations,…)</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>The system brings about benefits that are more important than its costs (time and money)</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>strongly agree</td>
<td>strongly disagree</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>58</td>
<td>The system provides all the information it should</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>For me, using this system has been a source of work improvement</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>This system (rather than the spreadsheet package) always issues an error message when it detects an error</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>All outputs provided by this system are required</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>The data entry sections provide the capability to easily make corrections to data</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>The outputs are easy to understand</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>The same terminology is used throughout the system</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>The terms used in data-entry sections are familiar to users</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>The information contained in the outputs always matches the actual facts</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>Data entry sections are organised in such a way that the data elements are logically grouped together</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>The data entry areas clearly show the spaces reserved to record the data</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>Outputs provided by this system are comprehensive</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>The system contains all the information required to produce comprehensive outputs</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>The system does not destroy any information without asking for a confirmation and getting a positive response</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>The system provides default values in the data-entry section</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>The format of a given piece of information is always the same, wherever it is used in the system</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>The system performs an automatic backup of the data</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>Data is labelled so that it can be easily matched with other parts of the system</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>The data contained in the system are always up-to-date</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>Headings provide information related to the nature of data in the system (e.g.: emp-no = employee number)</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>The system never modifies a cell without asking for a confirmation and getting a positive response</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>The system is broken up into separate and independent sections</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>Recovery and retrieval procedures are available in case of a system malfunction</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>Each section has a unique function</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Each section includes enough information to help you understand its functioning</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>Queries are easy to make</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>The wording of any user documentation matches user terminology</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>The documentation we have created for the system is comprehensive</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>The documentation describes, step by step and with examples, how to use the application</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
</tbody>
</table>
87 The documentation structure is appropriate (index, table of contents, chapter organisation,…)  

88 The documentation is easy to use

89 The documentation provides all the information required to use the system

90 Message presentation is always the same (position, terminology, style…)

91 The documentation explains the functioning of the system

92 Corrections to errors in the system are easy to make

Please bring the completed questionnaire and floppy disk to your C606 Strategic Management and policy boardroom meeting on Wednesday 7 October. They will be collected during the session. If you can’t attend this then please bring them to your C606 lecture on Monday 12 October.

Thank you for your help!!!
Appendix B
Items retained after Study 1 measurement model
development
Information quality

Q18 Do you get the information you need in time?
Q19 Does the system provide output that seems to be just about exactly what you need?
Q20 Does the system provide the precise information you need?
Q21 Does the system’s information content meet your needs?
Q22 Is the information provided by your system understandable?
Q25 Is the information provided by your system complete?

System quality and perceived system quality

Economy
Q47 The system increased my data processing capacity

Portability
Q46 The system can be run on computers other than the one presently used
Q49 The system could be used in other similar organisational environments, without any major modification

Reliability
Q45 Unauthorised access is controlled in several parts of the system
Q62 The data entry sections provide the capability to easily make corrections to data
Q92 Corrections to errors in the system are easy to make

Understandability
Q64 The same terminology is used throughout the system
Q67 Data entry sections are organised in such a way that the data elements are logically grouped together
Q68 The data entry areas clearly show the spaces reserved to record the data
Q75 Data is labelled so that it can be easily matched with other parts of the system
Q79 The system is broken up into separate and independent sections
Q81 Each section has a unique function
Q82 Each section includes enough information to help you understand its functioning
Q89 The documentation provides all the information required to use the system
Q91 The documentation explains the functioning of the system

Userfriendliness
Q40 Using the system is easy, even after a long period of non-utilisation
Q48 The system is easy to learn by new users
Q65 The terms used in data-entry sections are familiar to users
Q83 Queries are easy to make
**User satisfaction**

Q27  How efficient is the system used for your area of responsibility? (inefficient …efficient)
Q28  How effective is the system? (effective……ineffective)
Q29  Overall, are you satisfied with the system? (dissatisfied……..satisfied)

**Use**

Q10  Overall, how would you rate your intended use of the system over the next year of the BPG? (rarely….often)

**Individual impact**

Q30  The system has a large, positive impact on my effectiveness and productivity in my role in the BPG
Q31  The system is an important and valuable aid to me in the performance of my role in the BPG
Appendix C
Advertisement used to recruit participants for Study 2
Have You Used Microsoft Excel?

Do You Want to Learn More?

I need your help in a research project on spreadsheet use. Basic or high level expertise is desired, with the research project taking the form of a practical activity. In return, you will get a free training session, which will focus on developing spreadsheets. Flexible training sessions will be available and will be tailored to suit your needs.

Further Information

If you are interested, please call Tanya McGill, Senior Lecturer, School of Information Technology, on (08) 9360 2798 or (08) 9360 6120 or via email mcgill@murdoch.edu.au

Web: http://www.murdoch.edu.au

great minds at work
Appendix D
Training session provided for participants in Study 2
Developing Spreadsheet Applications

Training Session

Tanya McGill
School of Information Technology
Murdoch University
Phone: 9360 2798
Email: mcgill@murdoch.edu.au
Developing Spreadsheet Applications

Spreadsheet software is a very popular tool in businesses today, however studies have shown that even experienced spreadsheet users make errors when developing spreadsheet applications. The spreadsheet below shows the kinds of errors that are common.

You can make sure that your spreadsheet is accurate by following certain steps when you develop a spreadsheet. These steps are:
Steps in creating a spreadsheet
1. Define the inputs, outputs and calculations that are needed
2. Design a layout for your spreadsheet
3. Construct your spreadsheet
4. Test your spreadsheet
5. Document what your spreadsheet does and how to use it

1. Define the inputs, outputs and calculations that are needed

To develop a useful error-free spreadsheet you need to do some planning before you start using the computer.

1. Think about the purpose of the spreadsheet – what exactly do you want it to do?
2. Define the outputs – what results should the spreadsheet produce?
3. Define the inputs – what data will you need to enter to calculate the results
4. Determine how the outputs will be generated – what calculations will be needed?

Salary Increase Example:
Suppose that all the staff in your department at work have been awarded a pay rise. The pay rise consists of a flat increase of $1,000 for everyone plus 2% of current salary. You want to develop a spreadsheet to analyse the salary increases. You might define the outputs, inputs and calculations as shown below.
To do – Greeting Card Pricing Spreadsheet:

Assume that you have a small business that sells novelty greeting cards to local newsagents and card shops. The price of each card varies according to the cost of materials and how long it takes you to make it. You charge cost of materials plus either $1.00 or $1.50 depending on the complexity of the design. You decide to set up a spreadsheet to keep track of the number of each type of cards sold and the income from those cards.

Before doing anything else you need to define the required outputs more precisely. Jot down the outputs you think will be needed in the box below.

OUTPUTS:

Next you need to work out what information you will need to produce the outputs you have listed above. Jot down the inputs you will need in the box below.

INPUTS:

You should now be able to decide upon the calculations that will be needed to turn these inputs into the required outputs. Write down the calculations you think will be needed in the box below. Don't worry how these will be specified in your spreadsheet yet.

CALCULATIONS:

2. Design a layout for your spreadsheet

It is easiest to design the layout of your spreadsheet on paper before creating it in Microsoft Excel or other spreadsheet software. That way you will know that it will
work and won’t have to move things around and risk making undesired changes to formulas.

In general it is best to divide your spreadsheet into sections, keeping the different areas separate so that they can be clearly distinguished. For example, the area where you type in numbers should be separate from the area where calculations are done. The different sections can be on the same worksheet (if the spreadsheet is small) or on separate linked worksheets within the one workbook.

A spreadsheet should include sections for:
- \textit{Documentation} – This should always list the spreadsheet title, the purpose of the spreadsheet, the name of the developer and the date written or revised. Complex spreadsheets may be many pages long and include dozens of formulas. They should include an explanation of what the spreadsheet does, how it works and any assumptions made.
- \textit{Constants (or parameters)} – values that are used repeatedly in calculations. These should be located separately so that they can be clearly identified. This means that if you need to modify the value of a constant it only needs to be changed in one place rather than in each formula that uses it. For example, in the salary increase spreadsheet both the $1,000 flat increase figure and the 2\% figure should be treated as constants.
- \textit{Input data} – this section is where you type in data when you use the spreadsheet.
- \textit{Calculations} – this section might include both intermediate calculations and the final output information you need.

\textbf{Salary Increase Example:}
The spreadsheet for the salary increase example is relatively small and straightforward. There isn’t really any need to split it over more than one worksheet. The layout you design would probably be something like the one shown below. Note that the $1,000 and 2\% figures are located in a separate highly visible section.
To do – Greeting Card Pricing Spreadsheet:

Look back on your plans for the greeting card pricing spreadsheet. The next thing to do is to design the layout for the spreadsheet.

First answer the following questions:

1. Are there any values that are used repeatedly in calculations? ___________

2. If so, what are they? ______________________________________________

These values should be located in a separate section.

3. Do you think that this spreadsheet should have more than one worksheet? ______

4. If so, describe how you wish to divide the sections __________________________

____________________________

Sketch the layout for your spreadsheet in the box below (use the next page if you need more space).
3. Construct your spreadsheet

Once you are satisfied with your planning you can create your spreadsheet by typing in the text, formulas and data. Remember to format appropriately – make good use of fonts, sizes, number formats, column widths etc.

Salary Increase Example:
The spreadsheet you create should look something like this:

<table>
<thead>
<tr>
<th>Name</th>
<th>Current salary</th>
<th>Increase in salary</th>
<th>New salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis</td>
<td>$36,000</td>
<td>$1,750</td>
<td>$37,750</td>
</tr>
<tr>
<td>Carson</td>
<td>$25,500</td>
<td>$1,510</td>
<td>$27,010</td>
</tr>
<tr>
<td>Graves</td>
<td>$22,000</td>
<td>$1,440</td>
<td>$23,440</td>
</tr>
<tr>
<td>Jensen</td>
<td>$32,500</td>
<td>$1,850</td>
<td>$34,350</td>
</tr>
<tr>
<td>Killam</td>
<td>$26,600</td>
<td>$1,530</td>
<td>$28,130</td>
</tr>
<tr>
<td>Marsh</td>
<td>$28,500</td>
<td>$1,570</td>
<td>$30,070</td>
</tr>
<tr>
<td>Total</td>
<td>$173,000</td>
<td>$9,480</td>
<td>$182,480</td>
</tr>
</tbody>
</table>
To do – Greeting Card Pricing Spreadsheet:

It is finally time to start using Microsoft Excel! Using the layout you planned before, create your spreadsheet. You should use the sample data shown below, but to save typing you can cut and paste it from the Excel file called *GreetingCardsData.xls* on your disk.

<table>
<thead>
<tr>
<th>Item</th>
<th>Materials cost</th>
<th>Markup</th>
<th>Number sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small birthday</td>
<td>$1.00</td>
<td>$1.00</td>
<td>20</td>
</tr>
<tr>
<td>Large birthday</td>
<td>$1.50</td>
<td>$1.00</td>
<td>16</td>
</tr>
<tr>
<td>Gold birthday</td>
<td>$2.00</td>
<td>$1.50</td>
<td>7</td>
</tr>
<tr>
<td>Small get well</td>
<td>$1.00</td>
<td>$1.00</td>
<td>5</td>
</tr>
<tr>
<td>Large get well</td>
<td>$1.25</td>
<td>$1.00</td>
<td>5</td>
</tr>
<tr>
<td>Small wedding</td>
<td>$1.00</td>
<td>$1.00</td>
<td>8</td>
</tr>
<tr>
<td>Large wedding</td>
<td>$1.50</td>
<td>$1.00</td>
<td>9</td>
</tr>
<tr>
<td>Gold wedding</td>
<td>$2.00</td>
<td>$1.50</td>
<td>4</td>
</tr>
</tbody>
</table>

4. Test your spreadsheet

Testing is extremely important. Every spreadsheet you create should be thoroughly tested. It has been estimated that approximately 60% of spreadsheets used in business contain errors. Though spreadsheet output may look professional, it can hide errors that make it useless in practice.

When you test a spreadsheet it is a good idea to:

1. Make a printed copy of your spreadsheet and formulas. In Microsoft Excel you can display the formulas (as in my example above) by selecting Tools; Options; View; Formulas.
2. Check that all the necessary cells are included in calculations.
3. Check the calculations independently from the spreadsheet.

Salary Increase Example:

Verify the calculations for Marsh:

✅ Increase in salary: $1,000 + $28,500 X 0.02 = $1,570

✅ New salary: $28,500 + $1,570 = $30,070

Check the calculations of the totals:

✅ All figures in column included in calculation
To do – Greeting Card Pricing Spreadsheet:

1. Printed a copy of your spreadsheet and one showing the formulas. If you haven’t had time to complete your spreadsheet you can print out the spreadsheet called TestGreetingCards.xls.

2. Check the following calculations:

☐ Price for **Small Birthday Card**

☐ Price for **Gold Wedding Card**

☐ Total income from **Large Get Well Cards**

☐ Total monthly income formula - all figures in column included in calculation?

5. **Document what your spreadsheet does and how to use it**

If you are going to be the only user of a spreadsheet you may only write a brief documentation section. The information included in the Salary Increase spreadsheet above would be sufficient if you were to be the only user.

If others will use your spreadsheet, documentation is essential. You may not always be available to answer questions, so write enough documentation to ensure that other users will know exactly what the spreadsheet does and how to use it.

**To do - Greeting Card Example:**

What information might be useful to someone else who was going to use the greeting Card spreadsheet?
Appendix E
Study 2 participant consent form
The development of computer applications by non-computing professionals is of increasing importance to organisations. Typical applications developed by end users include budgeting systems, product pricing models and workload planning systems developed using software tools such as spreadsheets and database management systems; these applications are known as user developed applications (UDAs).

Little formal evaluation of the outcomes of user development of applications has been undertaken either by managers or by researchers and in view of the growing strategic importance of end user computing to many organisations, this lack of evaluation indicates the need for more research on UDAs.

The purpose of this project is to examine the nature of success of end user application development. It will attempt to clarify the relationships among previous experience and training, the quality of spreadsheets, and the satisfaction of users with the spreadsheets they have developed for themselves. It will also explore the impact of spreadsheet quality and end user satisfaction on individual performance.

As an end user developing spreadsheets your participation in the study will be of great value. All information provided will be strictly confidential. You will be asked to create a small spreadsheet, answer some questions using it and another spreadsheet then complete a questionnaire about your perceptions of the quality and usefulness of spreadsheets you have just used. You will also be given a free training session on developing spreadsheets.

If you have any further questions about this project please don’t hesitate to contact me at any time.
Tanya McGill
phone. (wk) 93602798
phone (hm) 93355085
e-mail: mcgill@murdoch.edu.au

I (the participant) have read the information above and any questions I have asked have been answered to my satisfaction. I agree to participate in this activity, realising that I may withdraw at any time without prejudice.

I agree that the research data gathered for this study may be published provided my name or other identifying information is not used.

Participant ___________________________ Date ___________________________

Investigator ___________________________ Date ___________________________
Appendix F
Study 2 questionnaire 1

Questionnaire to collect demographic information and assess spreadsheet knowledge
Spreadsheet Knowledge Questionnaire

Name: ____________________________  Research number: ___________(leave out)

Date: ____________________________  Group : ____________ (leave out)

Thank you for participating in the study. All the information you provide will be treated as confidential. Reporting will be in statistical terms only.

Background information

1 How long have you been using computers? _____ Years

2 On average how frequently do you use a computer? Please tick the box which best describes your average computer use
   □ Less than once a month
   □ About once a month
   □ A few times a month
   □ A few times a week
   □ About once a day
   □ Several times a day

3 How long have you been using spreadsheets? _____ Years

4 How skilful do you think you are with spreadsheets? Please rate your skill with spreadsheets by circling the number that best indicates your relative skill level
   Little 1 2 3 4 5 Very or no skillful skill

5 How much spreadsheet training have you had? Please circle the number which best describes your level of training in each of the following categories
   (a) University or TAFE courses None 1 2 3 4 5 Extremely intensive
   (b) Training provided by vendors 1 2 3 4 5
   (c) In-company courses 1 2 3 4 5
   (d) Self study 1 2 3 4 5

6 What have you used spreadsheets for?
   □ Work
   □ Recording information or making decisions at home
   □ Other

7 How old are you? _____ Years

8 What sex are you?
   □ Female
   □ Male
Spreadsheet Knowledge

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

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

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

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

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

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

6. If you want the numbers in your spreadsheet to appear as currency (that is with $ signs, etc), you would use the:
   a. Edit feature.
   b. Data feature.
   c. Format feature.
   d. Label feature.
   e. I am not familiar with this spreadsheet feature.
7. A spreadsheet is more likely to be useful over a long period of time if:
   a. Errors are easy to identify.
   b. It is easy to understand the calculations it uses.
   c. It has detailed documentation.
   d. All of the above are true.
   e. I don’t know.

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

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

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

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

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

13. In order to determine what input data is required for a spreadsheet you need to:
    a. Know what problem the spreadsheet will be used to solve.
    b. Know what questions the spreadsheet will be used to answer.
    c. Know what outputs are required from the spreadsheet.
    d. All of the above.
    e. I don’t know.
14. Which of the following ISN’T a section that spreadsheets should normally include:
   a. Documentation section.
   b. Input section.
   c. Development section.
   d. Output section.
   e. I don’t know.

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

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

17. How many errors does the spreadsheet below have in its formulas?
   a. 0.
   b. 1.
   c. 2.
   d. 5.
   e. I don’t know.
18. Which of the following is NOT a reason for documenting a spreadsheet:
   a. It helps other people to understand how to use the spreadsheet.
   b. It helps other people to understand what the spreadsheet does.
   c. It saves other people from having to use your spreadsheet.
   d. It helps you to remember what the spreadsheet does.
   e. I don’t know.

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

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

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

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

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

24. Which of the following is NOT a characteristic of a well-designed spreadsheet?
    a. Each section of the spreadsheet has a unique function.
    b. It can be printed out on one page.
    c. Corrections are easy to make.
    d. All headings and labels provide clear information about the data they relate to.
    e. I don’t know.
25. What would **MOST** improve the quality of the spreadsheet below?
   a. Naming the worksheet.
   b. Adding information about spreadsheet purpose.
   c. Increasing the column widths.
   d. Removing the blank line.
   e. I don’t know.
Appendix G
Study 2 decision questions
# Car Rental Decisions

Name: ___________________________  Research number: _________ (leave out)

Date: ___________________________  Group: ____________ (leave out)

File name of the spreadsheet you are using: ____________________________

Please use the spreadsheet you have just been given to answer the following set of questions. If the spreadsheet does not help you obtain the answers you may make minor adjustments to it, or try and answer the questions without using the spreadsheet.

**QUESTION SET 1:**

Start time  __________

1. What would you be charged by each of the rental companies if you wish to hire a car for 10 days and drive approximately 500 km with it?

<table>
<thead>
<tr>
<th>Company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>$</td>
</tr>
<tr>
<td>OnRoad</td>
<td>$</td>
</tr>
<tr>
<td>Prestige</td>
<td>$</td>
</tr>
</tbody>
</table>

2. What would you be charged by each of the rental companies if you wish to hire a car for 4 days and drive approximately 300 km with it?

<table>
<thead>
<tr>
<th>Company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>$</td>
</tr>
<tr>
<td>OnRoad</td>
<td>$</td>
</tr>
<tr>
<td>Prestige</td>
<td>$</td>
</tr>
</tbody>
</table>

3. Which rental company is the cheapest if you wish to hire a car for 8 days and drive approximately 900 km with it?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

4. Which rental company is the cheapest if you wish to hire a car for 6 days and drive approximately 1500 km with it?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>
5. Which rental company is the cheapest if you wish to hire a car for 3 days and drive approximately 600 km with it?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>$</th>
</tr>
</thead>
</table>

6. How many kilometers would you need to drive over a 7 day period before Prestige Rent-A-Car became the cheapest option?

<table>
<thead>
<tr>
<th>Kilometers</th>
<th></th>
</tr>
</thead>
</table>

7. What would OnRoad Rentals have to change its cutoff for free kilometers to, in order to be the cheapest company for a 1200 km trip over 4 days?

<table>
<thead>
<tr>
<th>OnRoad cutoff for free kilometres</th>
<th></th>
</tr>
</thead>
</table>

8. If the daily rate for Advantage Car Rentals is changed from $35 to $38 per day, which rental company is the cheapest if you wish to hire a car for 8 days and drive approximately 900 km with it? Assume that the OnRoad Rentals cutoff for free kilometers is 200 as originally stated.

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>$</th>
</tr>
</thead>
</table>

9. You intend to hire a car for 4 days and travel 600 km. Assume the daily rates for all three companies are as originally stated, but that OnRoad Rentals changes its cost per excess kilometer to 0.20. Which company now provides the best deal?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>$</th>
</tr>
</thead>
</table>

10. What would Prestige-Rent-A-Car now have to change its daily rate to, if it were to be the cheapest in the above scenario (i.e. 600 km over 4 days, with OnRoad Rentals excess cost per kilometer remaining at 0.20)?

<table>
<thead>
<tr>
<th>New Prestige daily rate</th>
<th></th>
</tr>
</thead>
</table>

Finish time \[\ \]
Car Rental Decisions

Name: ____________________ Research number: ________ (leave out)
Date: ____________________ Group: __________ (leave out)

File name of the spreadsheet you are using: __________________________

Please you the spreadsheet you have just been given to answer the following set of questions. If the spreadsheet does not help you obtain the answers you may make minor adjustments to it, or try and answer the questions without using the spreadsheet.

QUESTION SET 2:

Start time __________

1. What would you be charged by each of the rental companies if you wish to hire a car for 6 days and drive approximately 300 km with it?

<table>
<thead>
<tr>
<th>Company</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>$</td>
</tr>
<tr>
<td>OnRoad</td>
<td>$</td>
</tr>
<tr>
<td>Prestige</td>
<td>$</td>
</tr>
</tbody>
</table>

2. What would you be charged by each of the rental companies if you wish to hire a car for 8 days and drive approximately 700 km with it?

<table>
<thead>
<tr>
<th>Company</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>$</td>
</tr>
<tr>
<td>OnRoad</td>
<td>$</td>
</tr>
<tr>
<td>Prestige</td>
<td>$</td>
</tr>
</tbody>
</table>

3. Which rental company is the cheapest if you wish to hire a car for 4 days and drive approximately 600 km with it?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged $</th>
</tr>
</thead>
</table>

4. Which rental company is the cheapest if you wish to hire a car for 8 days and drive approximately 2000 km with it?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged $</th>
</tr>
</thead>
</table>
5. Which rental company is the cheapest if you wish to hire a car for 4 days and drive approximately 400 km with it?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

6. How many kilometers would you need to drive over a 6 day period before Prestige Rent-A-Car became the cheapest option?

<table>
<thead>
<tr>
<th>Kilometers</th>
</tr>
</thead>
</table>

7. What would Advantage Car Rentals have to change its cutoff for free kilometers to, in order to be the cheapest company for a 1300 km trip over 6 days?

<table>
<thead>
<tr>
<th>Advantage cutoff for free kilometres</th>
</tr>
</thead>
</table>

8. If the daily rate for OnRoad Rentals is changed from $41 to $38 per day, which rental company is the cheapest if you wish to hire a car for 8 days and drive approximately 900 km with it? Assume that the Advantage Car Rentals cutoff for free kilometers is 100 as originally stated.

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

9. You intend to hire a car for 4 days and travel 900 km. Assume the daily rates for all three companies are as originally stated, but that Advantage Car Rentals changes its cost per excess kilometer to 0.20. Which company now provides the best deal?

<table>
<thead>
<tr>
<th>Cheapest company</th>
<th>Amount charged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

10. What would Prestige-Rent-A-Car now have to change its daily rate to, if it were to be the cheapest in the above scenario (i.e. 900 km over 4 days, with Advantage Car Rentals excess cost per kilometer remaining at 0.20)?

| New Prestige daily rate |

Finish time ____________
Appendix H
Study 2 questionnaire 2

Questionnaire to collect end user perceptions of system quality, involvement, user satisfaction and individual impact
Car Rental Spreadsheet Questionnaire

The following questions relate to the spreadsheet that you have just used to answer a set of questions.

**Spreadsheet file name:**

The following questions relate to your perceptions of the quality of the spreadsheet you have just used to answer a set of questions about hire car firms. Please answer by circling the most appropriate number for each item.

<table>
<thead>
<tr>
<th></th>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Using the spreadsheet would be easy, even after a long period of not using it</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Errors in the spreadsheet are easy to identify</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The spreadsheet increased my data processing capacity</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The spreadsheet is easy to learn by new users</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Should an error occur, the spreadsheet makes it straightforward to perform some checking in order to locate the source of error</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>The data entry sections provide the capability to easily make corrections to data</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>The same terminology is used throughout the spreadsheet</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>This spreadsheet does not contain any errors</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>The terms used in the spreadsheet are familiar to users</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Data entry sections of the spreadsheet are organised so that the different bits of data are grouped together in a logical way</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>The data entry areas clearly show the spaces reserved to record the data</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>The format of a given piece of information is always the same, wherever it is used in the spreadsheet</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Data is labelled so that it can be easily matched with other parts of the spreadsheet</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>The spreadsheet is broken up into separate and independent sections</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Use of this spreadsheet would reduce the number of errors you make when choosing a rental car</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Each section has a unique function or purpose</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Each section includes enough information to help you understand what it is doing</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Queries are easy to make</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>The spreadsheet provides all the information required to use the spreadsheet (this is called documentation)</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Corrections to errors in the spreadsheet are easy to make</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
</tbody>
</table>

345
The following questions relate to your perceptions of the importance and personal relevance of the spreadsheet you have just used. Please record your impressions by marking an X on each item scale below.

This car rental spreadsheet is:

21. unimportant 1 2 3 4 5 6 7 important
22. not needed 1 2 3 4 5 6 7 needed
23. nonessential 1 2 3 4 5 6 7 essential
24. trivial 1 2 3 4 5 6 7 fundamental
25. insignificant 1 2 3 4 5 6 7 significant
26. means nothing to me 1 2 3 4 5 6 7 means a lot to me
27. unexciting 1 2 3 4 5 6 7 exciting
28. of no concern to me 1 2 3 4 5 6 7 of concern to me
29. not of interest to me 1 2 3 4 5 6 7 of interest to me
30. irrelevant to me 1 2 3 4 5 6 7 relevant to me
31. doesn’t matter to me 1 2 3 4 5 6 7 matters to me

The following questions relate to your satisfaction with the spreadsheet you have just used and to its impact on your decision making. Please answer by circling the most appropriate number for each item.

32. How adequately do you feel the spreadsheet meets your information processing needs when answering car rental queries? inadequately 1 2 3 4 5 6 7 adequately
33. How efficient is the spreadsheet? inefficient 1 2 3 4 5 6 7 efficient
34. How effective is the spreadsheet? ineffective 1 2 3 4 5 6 7 effective
35. Overall, are you satisfied with the spreadsheet? dissatisfied 1 2 3 4 5 6 7 satisfied
36. The spreadsheet has a large, positive impact on my effectiveness and productivity in answering car rental queries disagree 1 2 3 4 5 6 7 agree
37. The spreadsheet is an important and valuable aid to me in answering car rental queries disagree 1 2 3 4 5 6 7 agree
Appendix I
Confirmation of dimensionality of involvement
Table I.1 below shows the results of the factor analysis of the involvement data from Study 2. The data was examined using principal components analysis as the extraction technique and varimax as the method of rotation. Two factors with eigenvalues of greater than one emerged and these equate to the Personal Relevance and Importance factors identified by Barki and Hartwick (1994). Total variance explained by the rotated model was 82.09%.

Table I.1: Rotated component matrix of involvement items

<table>
<thead>
<tr>
<th>Item</th>
<th>Personal Relevance</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV21</td>
<td>0.381</td>
<td>0.835</td>
</tr>
<tr>
<td>INV22</td>
<td>0.293</td>
<td>0.894</td>
</tr>
<tr>
<td>INV23</td>
<td>0.314</td>
<td>0.873</td>
</tr>
<tr>
<td>INV24</td>
<td>0.395</td>
<td>0.829</td>
</tr>
<tr>
<td>INV25</td>
<td>0.440</td>
<td>0.819</td>
</tr>
<tr>
<td>INV26</td>
<td><strong>0.761</strong></td>
<td>0.445</td>
</tr>
<tr>
<td>INV27</td>
<td><strong>0.681</strong></td>
<td>0.398</td>
</tr>
<tr>
<td>INV28</td>
<td><strong>0.827</strong></td>
<td>0.350</td>
</tr>
<tr>
<td>INV29</td>
<td><strong>0.874</strong></td>
<td>0.332</td>
</tr>
<tr>
<td>INV30</td>
<td><strong>0.850</strong></td>
<td>0.310</td>
</tr>
<tr>
<td>INV31</td>
<td><strong>0.884</strong></td>
<td>0.299</td>
</tr>
</tbody>
</table>
Appendix J
Possible directions of the relationship between user satisfaction and (perceived) individual impact – model fit information
Table J.1: Model fit if the direction of the relationship between user satisfaction and individual impact is reversed (II \(\rightarrow\) US)

<table>
<thead>
<tr>
<th>Path</th>
<th>To</th>
<th>User developer data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experience</td>
<td>To</td>
</tr>
<tr>
<td></td>
<td>Spreadsheet development knowledge</td>
<td>Estimate</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>0.335</td>
</tr>
<tr>
<td></td>
<td>Training</td>
<td>0.824</td>
</tr>
<tr>
<td></td>
<td>Spreadsheet development knowledge</td>
<td>System quality</td>
</tr>
<tr>
<td></td>
<td>Spreadsheet development knowledge</td>
<td>Perceived system quality</td>
</tr>
<tr>
<td></td>
<td>System quality</td>
<td>Perceived system quality</td>
</tr>
<tr>
<td></td>
<td>Involvement</td>
<td>Perceived system quality</td>
</tr>
<tr>
<td></td>
<td>System quality</td>
<td>Individual impact</td>
</tr>
<tr>
<td></td>
<td>Perceived system quality</td>
<td>User satisfaction</td>
</tr>
<tr>
<td></td>
<td>User satisfaction</td>
<td>Perceived individual impact</td>
</tr>
<tr>
<td>Individual impact</td>
<td>User satisfaction</td>
<td>\textbf{0.048}</td>
</tr>
<tr>
<td>Individual impact</td>
<td>Perceived individual impact</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Spreadsheet development knowledge</td>
<td>Individual impact</td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square ((\chi^2))</td>
<td>59.07</td>
</tr>
<tr>
<td>Degrees of freedom (df)</td>
<td>28</td>
</tr>
<tr>
<td>Probability (p)</td>
<td>0.001</td>
</tr>
<tr>
<td>Ratio of (\chi^2) to degrees of freedom ((\chi^2/df))</td>
<td>2.11</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.935</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.872</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.084</td>
</tr>
<tr>
<td>Tucker-Lewis index (TLI)</td>
<td>0.946</td>
</tr>
</tbody>
</table>

* \(p < 0.05\)
** \(p < 0.01\)
*** \(p < 0.001\)
Table J.2: Model fit if the direction of the relationship between user satisfaction and individual impact is reversed such that individual impact influences perceived individual impact which in turn influences user satisfaction

<table>
<thead>
<tr>
<th>Path From</th>
<th>To</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>Spreadsheet development knowledge</td>
<td>0.335</td>
<td>0.075</td>
<td>4.435***</td>
</tr>
<tr>
<td>Training</td>
<td>Spreadsheet development knowledge</td>
<td>0.824</td>
<td>0.200</td>
<td>4.114***</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>System quality</td>
<td>0.102</td>
<td>0.017</td>
<td>6.107***</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>Perceived system quality</td>
<td>0.024</td>
<td>0.018</td>
<td>1.356</td>
</tr>
<tr>
<td>System quality</td>
<td>Perceived system quality</td>
<td>0.384</td>
<td>0.087</td>
<td>4.423***</td>
</tr>
<tr>
<td>Involvement</td>
<td>Perceived system quality</td>
<td>0.372</td>
<td>0.073</td>
<td>5.120***</td>
</tr>
<tr>
<td>System quality</td>
<td>Individual impact</td>
<td>0.728</td>
<td>0.285</td>
<td>2.558**</td>
</tr>
<tr>
<td>Perceived system quality</td>
<td>User satisfaction</td>
<td>0.289</td>
<td>0.034</td>
<td>8.484***</td>
</tr>
<tr>
<td><strong>Perceived individual impact</strong></td>
<td><strong>User satisfaction</strong></td>
<td><strong>0.380</strong></td>
<td><strong>0.019</strong></td>
<td><strong>20.399</strong>*</td>
</tr>
<tr>
<td>Individual impact</td>
<td>Perceived individual impact</td>
<td>0.214</td>
<td>0.043</td>
<td>4.943***</td>
</tr>
<tr>
<td>Spreadsheet development knowledge</td>
<td>Individual impact</td>
<td>0.240</td>
<td>0.059</td>
<td>4.070***</td>
</tr>
</tbody>
</table>

**Goodness of fit measures**

- Chi-square ($\chi^2$) | 178.51
- Degrees of freedom (df) | 29
- Probability (p) | 0.000
- Ratio of $\chi^2$ to degrees of freedom ($\chi^2$/df) | 6.16
- Goodness of fit index (GFI) | 0.857
- Adjusted goodness of fit index (AGFI) | 0.729
- Root mean square error of approximation (RMSEA) | 0.181
- Tucker-Lewis index (TLI) | 0.748

* p < 0.05
** p < 0.01
*** p < 0.001
References


Publications arising from this research

The following refereed publications have resulted either directly from the research reported in this thesis or from research which explores aspects of it in more detail than possible in this thesis.


