COMPARISON OF THE MAIN METHODS FOR EVALUATING THE USABILITY OF COMPUTER SOFTWARE

A thesis presented in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy in Psychology
at Massey University.

Ronald Derek Henderson

1992

Abstract

The aim of this thesis is to examine the dominant computer software usability evaluation methods. Four evaluation methods (logged data, questionnaire, interview, and verbal protocol analysis) were used to evaluate three different business software types (spreadsheet, word processor, and database) using a between groups design, involving 148 individuals of both genders. When each evaluation method was examined individually, the results tended to support findings from previous research. Comparisons were made to examine the efficiency of each evaluation method, in terms of its ability to highlight usability problems (both between and within the evaluation strategy). Here support for the efficiency of the verbal protocol analysis method was found. The efficiency of using two evaluation methods was also examined, where it was found that no significant improvement was obtained over the verbal protocol analysis used by itself. A comparison addressing the practicality of using these methods was also conducted. It seems that each method has differing strengths and weaknesses depending on the stage of the evaluation. From these results a theory for the effectiveness of evaluation strategies is presented. Suggestions for improving the methods commonly used, are also made. The thesis concludes by discussing the software evaluation domain and its relationship to the wider evaluation context.

Acknowledgements

There have been many individuals who have significantly influenced my thoughts and behaviour throughout the development and refinement of this thesis. Each deserves my heartfelt gratitude.

Special mention must go to my family. In particular, my wife Karen, who has encouraged, supported and helped me throughout the research process. My parents, Derek and June, and my parents-in-law, Ray and Patricia, have also provided immeasurable support and encouragement throughout this period.

Special mention must also go to my supervisors, Doctors Mike Smith, John Podd and Hugo Varela. All have influenced and supported me significantly throughout the development of this thesis. My chief supervisor, Mike, has continually supported and encouraged me throughout both a significant period before undertaking this dissertation, and through the dissertation process. Mike has also been party to innumerable discussions on the nature of applied research, stimulating thought provoking enquiry on my part. John, in turn, has provided thoughtful discussion on the research process, and his insights have been both stimulating and beneficial. Deep gratitude must be expressed towards Hugo, who patiently coaxed me through the pains of learning to program in PASCAL, the virtues of log linear analysis, and trout fishing. I am deeply indebted to all; thank you.

Not least, gratitude must be expressed to each and all of the research subjects who undertook the study and provided the wealth of information that form the basis of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES