GUIDELINES FOR DISSERTATIONS AND THESES

IN EMPIRICAL SOFTWARE ENGINEERING

By

Edward B. Allen

A Thesis Guideline
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Computer Science
in the Department of Computer Science and Engineering

Mississippi State, Mississippi

December 2003

GUIDELINES FOR DISSERTATIONS AND THESES

IN EMPIRICAL SOFTWARE ENGINEERING

By

Edward B. Allen

Approved:

Julia E. Hodges Professor of Computer Science and Engineering, and Department Head (Major Professor)	Rayford B. Vaughn Associate Professor of Computer Science and Engineering (Committee Member)
Susan M. Bridges Professor of Computer Science and Engineering, and Graduate Coordinator Department of Computer Science and Engineering	Thomas Philip Professor of Computer Science and Engineering (Committee Member)

A. Wayne Bennett
Dean of the James Worth Bagley College of Engineering

Name: Edward B. Allen

Date of Degree: December 15, 2003

Institution: Mississippi State University

Major Field: Computer Science

Major Professor: Dr. Julia E. Hodges

Title of Study: GUIDELINES FOR DISSERTATIONS AND THESES IN EMPIRICAL SOFTWARE ENGINEERING

Pages in Study: 14

Candidate for Degree of Master of Science

This thesis guideline explains an example outline for a dissertation or thesis in empirical software engineering.

This document is in the form of a hypothetical master's thesis with a hypothetical Graduate Committee.

Limit the length of the abstract: (1) less than 150 words for a master's thesis, (2) less than 350 words for a dissertation, (3) 100–150 words for IEEE journals, and (4) less than 250 words for conference papers.

Write a "report in miniature", which means it has a self-contained message that does not expect the reader to look at the paper. Answer the questions, "What?", "How?", and "Why?"

Use no citations, footnotes, or equations. Use no superlatives; you are not selling anything. Minimize use of acronyms and mathematical symbols.

DEDICATION

To Angie.

ACKNOWLEDGMENTS

At a minimum, you should acknowledge financial support, permission to use copyrighted materials, trademarks and service marks, and personal assistance. The following is a hypothetical example; edit it with your information, and add your personal feelings.

This work was supported in part by grant CCR-999999 from the National Science Foundation. The findings and opinions in this thesis guideline belong solely to the author, and are not necessarily those of the sponsor.

Permission to reproduce the MSU logo was given by Mississippi State University.

Unix® is a registered trademark of The Open Group.

I thank Angie Allen for helpful discussions regarding book design. I thank my committee for their comments on this thesis guideline, and I thank Dr. Julia E. Hodges for directing this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS, AND NOMENCLATURE

This section is optional, and may be formatted in an appropriate manner. The title `\listofsymbolsname` may be redefined as needed.

g A variable representing gnats

G A variable representing gnus

NASA National Aeronautics and Space Administration

gnat A small animal, found in the North Woods, that causes no end of trouble.

gnu A large animal, found in crossword puzzles, that causes no end of trouble.

CHAPTER I

INTRODUCTION

The goal of this thesis guideline is to suggest an outline for dissertations, and theses in empirical software engineering. Zelkowitz and Wallace present a taxonomy of experimental and empirical studies [6].

This document is itself an example of a master's thesis. The Office of Graduate Studies publishes the *Guidelines for Preparing Dissertations and Theses* [2], which we refer to as the *Guidelines*. In the following, we uses the word "document" to mean a dissertation or thesis in the Department of Computer Science and Engineering. The University requires that all dissertations and theses comply with the *Guidelines*, latest edition.

The outline suggested here is based on a synthesis of suggestions by Singer [4], Kitchenham et al. [1], and Sim et al. [3].

1.1 Hypothesis

This section presents the research hypothesis of this work.

> Grammatically, the hypothesis is stated as a fact, even though we are uncertain whether it is true or not. A statement of the hypothesis is formatted in block style.

Other paragraphs may explain the context of the hypothesis and special terms.

1.2 Research Questions

This section lists research questions. Answers to the questions will provide significant evidence regarding whether the hypothesis is true or not. The format is a numbered list.

The following are research questions.

1. What is some evidence for the hypothesis?
2. Another question?

1.3 Relevance

This section explains the potential impacts of this research on its field and on society. This section provides motivation for performing the research.

1.4 Overview

This section presents an overview of the remainder of this document.

CHAPTER II

RELATED WORK

This chapter summarizes work in the technical literature that is relevant to the research. Unpublished research results by others, such as local team members, can also be discussed here.

Sections may be useful to categorize the discussion. For example, a document in the software metrics field might have the following sections.

- Software Engineering
- Software Metrics
- Statistics

CHAPTER III

TOOLS

This chapter explains tools used in data collection, experiments, and quantitative analysis.

Off-the-shelf tools must have citations. Development of custom tools must be explained. Authors of tools who are not the author of the document should be thanked in the Acknowledgments section.

CHAPTER IV

METHODOLOGY

This chapter presents the methods of the research. This outline assumes only one case study. If multiple case studies are included in the document which use different methods, then one may prefer a different outline to associate each method with its case study.

4.1 Procedure

This section presents a procedure for conducting the experiment(s), stated in a way that the procedure could be reproduced by another case study.

1. Each step of the procedure is numbered.
2. etc.

4.2 Experimental Design

This section presents the experimental design in terms of statistical techniques. Most studies will include the following information where applicable.

- Sampling method
- Randomization techiques
- Counterbalancing techniques
- Within subject *vs.* between subject data collection

- List of all variables and their respective levels (i.e. possible values)
- Statistical experimental design (i.e. statistical techniques to be used with special emphasis on modeling techniques and inferential statistics)

CHAPTER V

CASE STUDY

This chapter describes the case study and its results.

5.1 System under Study

This section describes the source of statistical observations. Usually, this is a computer system.

If human subjects are used, this section should be titled “Participants” and should describe their characteristics. This may include a statistical summary of the subjects and the population they represent.

5.2 Apparatus, Materials, and Artifacts

The title of this section will depend on the nature of the experiments.

- *Apparatus* refers to equipment that is part of experimental conditions. For example, software execution experiments should describe the host computer system that runs the software.
- *Materials* refers to other items that affect the experimental conditions. For example, instructions and questionaires for human subjects. For software execution experiments, test data should be characterized including a citation for details. A statistical summary of characteristics of test data may be appropriate.
- *Artifacts* refers to documents and files from a real-world setting that are the subject of study. For example, in a software metrics study, source code files are often studied.

5.3 Data Collection Details

This section explains the results of actual data collection activity, noting any deviations from planned methods.

5.4 Descriptive Statistics

This section presents statistics summarizing collected data.

As an example table, Table 5.1 lists meat prices for gnats, gnus, and other animals. See `examplethesis.ps` for further details on how to create tables.

Table 5.1 Meat prices (portrait)

Item		Price ($)
gnats	gram	13.65
gnats	each	.01
gnu	stuffed	92.50
emur	stuffed	33.33
armadillo	frozen	8.99

As an example figure, a lion, as shown in Figure 5.1, eats gnus. Gnats eat lions. See `examplethesis.ps` for further details on how to create figures.

5.5 Modeling

This section presents statistical models derived from collected data, including results and statistics regarding the quality of the models.

Figure 5.1 Lion

5.6 Inferential Statistics

This section presents statistics resulting from the modeling. These are the basis for inferences regarding the research questions.

CHAPTER VI

ANALYSIS

This chapter discusses each research question.

Subsections may correspond to each research question.

6.1 What is some evidence for the hypothesis?

This section points out relevant evidence in the case study and derives the answer to the question. A discussion of interesting aspects may be appropriate.

6.2 Another question?

Similarly, each question is discussed.

6.3 Threats to Validity

This section discusses factors that mitigate the significance of the experiments. Votta and Porter discuss some common threats to validity for empirical software engineering research [5]. If various questions face different threats, then a different outline may be desirable.

6.3.1 Internal Threats to Validity

This section discusses aspects of the methodology and modeling that limit their accuracy, precision, or significance.

6.3.2 External Threats to Validity

This section discusses aspects of the case study that limit the generalization of conclusions to other situations.

CHAPTER VII

CONCLUSIONS

The concluding chapter synthesizes the chapters in the body of the thesis, explaining how the research questions contribute evidence regarding the research hypothesis.

7.1 Evaluation of Hypothesis

This section points out how the analysis of the research questions supports the hypothesis, or not. Restate the hypothesis here for convenient reference.

> Grammatically, the hypothesis is stated as a fact, even though we are uncertain whether it is true or not. A statement of the hypothesis is formatted in block style.

7.2 Contributions

This section summarizes the contributions of this research to science, emphasizing new knowledge gained.

This example thesis made no contribution to knowledge of science and technology.

7.3 For Further Research

This section explains opportunities for further research based on this work.

Gnats and gnus is a facinating subject with many possible topics for further research.

REFERENCES

[1] B. A. Kitchenham, S. L. Pfleeger, L. M. Pickard, P. W. Jones, D. C. Hoaglin, K. El Emam, and J. Rosenberg, "Preliminary Guidelines for Empirical Research in Software Engineering," *IEEE Transactions on Software Engineering*, vol. 28, no. 8, Aug. 2002, pp. 721–734.

[2] Office of Graduate Studies, *Guidelines for Preparing Dissertations and Theses*, 4th edition, Mississippi State University, Mississippi State, Mississippi, Nov. 2001.

[3] S. E. Sim, J. Singer, and M.-A. Storey, *Beg, Borrow, or Steal: Using Multidisciplinary Approaches in Empirical Software Engineering Research*, Workshop report, International Conference on Software Engineering, Limerick, Ireland, June 2000, Available from National Research Council of Canada.

[4] J. Singer, "Using the American Psychological Association (APA) Style Guidelines to Report Experimental Results," *Proceedings of Workshop on Empirical Studies in Software Maintenance*, Oxford, England, Sept. 1999, pp. 71–75, Position paper. Available from Canadian National Reserach Council as Technical Report NRC-44130.

[5] L. G. Votta and A. A. Porter, "Experimental Software Engineering: A Report on the State of the Art," *Proceedings of the Seventeenth International Conference on Software Engineering*, Seattle, WA, Apr. 1995, IEEE Computer Society, pp. 277–279.

[6] M. V. Zelkowitz and D. R. Wallace, "Experimental Models for Validating Technology," *Computer*, vol. 31, no. 5, May 1998, pp. 23–31.

APPENDIX

AN EXAMPLE

Appendices are written like the body of the paper using `\chapter`. LaTeX numbers them with capital letters. If there is only one appendix, then use `\oneappendix` instead of `\chapter` so that the table of contents is properly formatted.