

Contents

Part I Background

1	Introduction	3
1.1	Software Engineering Context	3
1.2	Science and Software Engineering	5
1.3	Exercises	8
2	Empirical Strategies	9
2.1	Overview of Empirical Strategies	10
2.2	Surveys	12
2.2.1	Survey Characteristics	12
2.2.2	Survey Purposes	13
2.2.3	Data Collection	13
2.3	Case Studies	14
2.3.1	Case Study Arrangements	15
2.3.2	Confounding Factors and Other Aspects	15
2.4	Experiments	16
2.4.1	Characteristics	17
2.4.2	Experiment Process	18
2.5	Empirical Strategies Comparison	18
2.6	Replications	19
2.7	Theory in Software Engineering	21
2.8	Aggregating Evidence from Empirical Studies	22
2.9	Empiricism in a Software Engineering Context	24
2.9.1	Empirical Evaluation of Process Changes	24
2.9.2	Quality Improvement Paradigm	26
2.9.3	Experience Factory	27
2.9.4	Goal/Question/Metric Method	29
2.10	Empirically-Based Technology Transfer	30
2.11	Ethics in Experimentation	33
2.12	Exercises	36

3 Measurement	37
3.1 Basic Concepts	38
3.1.1 Scale Types	39
3.1.2 Objective and Subjective Measures	40
3.1.3 Direct or Indirect Measures	41
3.2 Measurements in Software Engineering	41
3.3 Measurements in Practice	42
3.4 Exercises	43
4 Systematic Literature Reviews	45
4.1 Planning the Review	45
4.2 Conducting the Review	46
4.3 Reporting the Review	51
4.4 Mapping Studies	52
4.5 Example Reviews	52
4.6 Exercises	54
5 Case Studies	55
5.1 Case Studies in Its Context	56
5.1.1 Why Case Studies in Software Engineering?	57
5.1.2 Case Study Research Process	58
5.2 Design and Planning	58
5.2.1 Case Study Planning	58
5.2.2 Case Study Protocol	60
5.3 Preparation and Collection of Data	61
5.3.1 Interviews	62
5.3.2 Observations	64
5.3.3 Archival Data	65
5.3.4 Metrics	65
5.4 Data Analysis	65
5.4.1 Quantitative Data Analysis	65
5.4.2 Qualitative Data Analysis	66
5.4.3 Validity	68
5.5 Reporting	69
5.5.1 Characteristics	69
5.5.2 Structure	71
5.6 Exercises	72
6 Experiment Process	73
6.1 Variables, Treatments, Objects and Subjects	74
6.2 Process	76
6.3 Overview	81
6.4 Exercises	81

Part II Steps in the Experiment Process

7 Scoping	85
7.1 Scope Experiment	85
7.2 Example Experiment	87
7.3 Exercises	88
8 Planning	89
8.1 Context Selection	89
8.2 Hypothesis Formulation	91
8.3 Variables Selection	92
8.4 Selection of Subjects	92
8.5 Experiment Design	93
8.5.1 Choice of Experiment Design	93
8.5.2 General Design Principles	94
8.5.3 Standard Design Types	95
8.6 Instrumentation	101
8.7 Validity Evaluation	102
8.8 Detailed Description of Validity Threats	104
8.8.1 Conclusion Validity	104
8.8.2 Internal Validity	106
8.8.3 Construct Validity	108
8.8.4 External Validity	110
8.9 Priority Among Types of Validity Threats	111
8.10 Example Experiment	112
8.11 Exercises	116
9 Operation	117
9.1 Preparation	117
9.1.1 Commit Participants	118
9.1.2 Instrumentation Concerns	119
9.2 Execution	120
9.2.1 Data Collection	120
9.2.2 Experimental Environment	120
9.3 Data Validation	121
9.4 Example Operation	121
9.5 Exercises	122
10 Analysis and Interpretation	123
10.1 Descriptive Statistics	123
10.1.1 Measures of Central Tendency	124
10.1.2 Measures of Dispersion	126
10.1.3 Measures of Dependency	127
10.1.4 Graphical Visualization	128
10.2 Data Set Reduction	131

10.3	Hypothesis Testing	132
10.3.1	Basic Concept	132
10.3.2	Parametric and Non-parametric Tests	135
10.3.3	Overview of Tests	136
10.3.4	t-Test	138
10.3.5	Mann-Whitney	139
10.3.6	F-Test	140
10.3.7	Paired t-Test	140
10.3.8	Wilcoxon	141
10.3.9	Sign Test	142
10.3.10	ANOVA (ANalysis Of VAriance)	143
10.3.11	Kruskal-Wallis	144
10.3.12	Chi-2	145
10.3.13	Model Adequacy Checking	148
10.3.14	Drawing Conclusions	149
10.4	Example Analysis	150
10.5	Exercises	151
11	Presentation and Package	153
11.1	Experiment Report Structure	153
11.2	Exercises	157

Part III Example Experiments

12	Experiment Process Illustration	161
12.1	Scoping	161
12.1.1	Goal Definition	161
12.1.2	Summary of Scoping	163
12.2	Planning	163
12.2.1	Context Selection	163
12.2.2	Hypothesis Formulation	163
12.2.3	Variables Selection	165
12.2.4	Selection of Subjects	165
12.2.5	Experiment Design	165
12.2.6	Instrumentation	166
12.2.7	Validity Evaluation	166
12.3	Operation	167
12.3.1	Preparation	167
12.3.2	Execution	168
12.3.3	Data Validation	168
12.4	Analysis and Interpretation	169
12.4.1	Descriptive Statistics	169
12.4.2	Data Reduction	172
12.4.3	Hypothesis Testing	172
12.5	Summary	173
12.6	Conclusion	174

13	Are the Perspectives Really Different? Further Experimentation on Scenario-Based Reading of Requirements	175
13.1	Introduction	176
13.2	Related Work	177
13.3	Research Questions	181
13.4	Experiment Planning	182
13.4.1	Variables	182
13.4.2	Hypotheses	182
13.4.3	Design	184
13.4.4	Threats to Validity	184
13.5	Experiment Operation	186
13.6	Data Analysis	187
13.6.1	Individual Performance for Different Perspectives	187
13.6.2	Defects Found by Different Perspectives	189
13.6.3	Is the Sample Size Large Enough?	192
13.6.4	Experience of Subjects	194
13.7	Interpretations of Results	194
13.8	Summary and Conclusions	195
13.9	Data on Individual Performance	197
13.10	Data on Defects Found by Perspectives	198
13.10.1	PG document	198
13.10.2	ATM document	199

Appendices

A	Exercises	203
A.1	Training	203
A.1.1	Normally Distributed Data	204
A.1.2	Experience	204
A.1.3	Programming	205
A.1.4	Design	209
A.1.5	Inspections	212
A.2	Reviewing	212
A.3	Assignments	213
A.3.1	Unit Test and Code Reviews	214
A.3.2	Inspection Methods	214
A.3.3	Requirements Notation	215
B	Statistical Tables	217
	References	223
	Index	233