Cranfield University

Bassil A. Yaghi

The Moderating Effects of Performance Measurement Use on the Relationship between Organizational Performance Measurement Diversity and Product Innovation

School of Management

PhD Thesis

Cranfield University

School of Management

PhD Dissertation

2007

Bassil A. Yaghi

The Moderating Effects of Performance Measurement Use on the Relationship between Organizational Performance Measurement Diversity and Product Innovation

Supervisor: Professor Andrew Neely

Academic Year 2006-2007

ABSTRACT

This study sets out to address the question of whether the effect of organizational performance measurement diversity on product innovation will differ depending on how organizational performance measures are used.

There is strong empirical evidence that many companies who are successful today are less likely to be successful in the future because they fail to innovate. It is surprisingly then, that when everyone stresses the importance of innovation, there are many organizations adopting performance measurement systems, which may constrain their innovativeness.

Currently, there are three differing perspectives on the effect of measurement on a firm's propensity to innovate. Moreover, each of these has empirical evidence to support its argument. The first perspective views measurement as constraining innovation because it impedes creativity, experimentation, and search in firms. The second perspective views measurement as helping innovation because it triggers search, facilitates decision-making, and increases risk-taking. The third perspective views measurement as having insignificant or little impact on innovation because it is used primarily for signalling.

A possible explanation of the contradiction in the empirical findings of these studies is that they generally ignore how measurement is used. Therefore, using the behavioural theory of innovation, I argue that one possible way of resolving the contradictory findings is by incorporating measurement use as a moderating variable.

Using data from a cross-sectional, large-scale, probability sample survey of 145 UK manufacturing firms, I show that organizational performance measurement diversity interacts with performance measurement use to determine product innovation. My findings suggest that the extent to which a firm offers new products will be more positively (negatively) associated with performance

measurement diversity when diagnostic use is high (low) holding interactive use constant and will be more negatively (positively) associated with performance measurement diversity when interactive use is high (low) holding diagnostic use constant.

ACKNOWLEDGEMENTS

In the course of my PhD research, I have benefited greatly from the support of my colleagues, friends, and family. I am particularly grateful and indebted to my PhD supervisor- Professor Andy Neely whose focus on high quality, relevant research provided me with a standard for which I will continue to strive. I feel uniquely privileged to have had the opportunity to develop and refine my research skills with him and I look forward to many years of fruitful collaboration with him. I can not adequately express my appreciation to him. He has served not only as my PhD supervisor, but also as a role model, teacher and friend.

I am grateful to my colleagues at the Centre for Business Performance (CBP) and to the AIM research fellows at the Innovation and Productivity Grand Challenge (IPGC). I have benefited greatly from being part of their intellectual and social communities. I am also grateful to all the researchers whose ideas I have used in this thesis. The reference list gives some indication of how much I have gained from their work.

I am particularly grateful to my family, especially my wife, but also my children for their unconditional support and encouragement. Lastly, I owe great deal of gratitude to my parents for always being there for me.

DEDICATION

To my family,

for your unwavering encouragement, your patient understanding, and, most of all, your love.

Table of Contents

Table of Contents V
Index of Figures X
Index of Tables XIV
Index of Tables XIV
1 Introduction 1
1.1 Introduction 1
1.2 Motivation 2
1.3 Research Aims and Questions 4
1.4 Research Gaps 5
1.4.1 Theoretical Gaps 5
1.4.2 Empirical Gaps 5
1.4.3 Contextual Gaps 5
1.4.4 Substantive Gaps 6
1.5 Research Model 6
1.6 Research Contributions 7
1.7 Research Approach and Thesis Structure 7
2 Literature Review 11
2.1 Introduction 11
2.2 What is Performance Measurement? 12
2.3 What is Product Innovation? 18
2.4 What is the impact of Measurement on Innovation? 24
2.4.1 Contingency-based Innovation Models 26
2.4.2 Creativity-based Innovation Models 31
2.4.3 Practice-based Innovation Models 36
2.4.4 Summary 41
2.5 What is Performance Measurement Use? 48
2.6 Research Gaps 59
2.6.1 Theoretical Gaps 59
2.6.2 Empirical Gaps 59
2.6.3 Contextual Gaps 59
2.6.4 Substantive Gaps 60
3 Theoretical Model and Hypotheses 61
3.1 Introduction 61
3.2 Definitions of Key Constructs 62
3.2.1 Performance Measurement Systems, Performance measurement and Performance Measures 62
3.2.2 Performance Measurement Use: Diagnostic and Interactive Uses 62
3.2.3 Organizational Performance Measurement Diversity 62
3.2.4 Innovation Search 63
3.2.5 Organizational Risk Taking 63
3.2.6 Product Innovation 65
3.2.7 Organizational Attention 65
3.2.8 Aspiration Levels (Performance Targets) 65
3.2.9 Bounded Rationality 67

3.2.10 Organizational Learning 68
3.3 Theoretical Foundation: Performance Feedback Theory......................69
3.4 Research Model 75
3.5 Research Hypotheses 79
3.5.1 The relationship between performance measurement diversity and product innovation 79
3.5.2 The Moderating Role of Performance Measurement Use on the Relationship between Performance Measurement Diversity and Product Innovation 81
3.5.3 The Moderating Role of Performance Measurement Use on the Relationship between Performance Measurement Diversity and Innovation Search 83
3.5.3.1 Diagnostic Use 83
3.5.3.2 Interactive Use 85
3.5.4 The Moderating Role of Performance Measurement Use on the Relationship between Performance Measurement Diversity and Risk Taking 87
3.5.4.1 Diagnostic Use 87
3.5.4.2 Interactive Use 88
3.5.5 The mediating role of innovation search.......................................89
3.5.6 The Mediating Role of Organizational Risk Taking.........................90
3.5.7 Summary 92
4 Research Methodology 94
4.1 Introduction..94
4.2 Philosophical Perspective 96
4.2.1 Introduction..96
4.2.2 Defining Ontology and Epistemology 96
4.2.2.1 Ontology 96
4.2.2.2 Epistemology 96
4.2.3 Ontology and Epistemology and Management Research 97
4.2.4 My Ontological and Epistemological Position 100
4.3 Research Strategy..105
4.4 Research Rigour..107
4.4.1 Defining Research Rigour..107
4.4.2 Evaluating Research Rigour..107
4.4.3 Validity of Research Findings 108
4.4.3.1 Defining Validity 108
4.4.3.2 Construct Validity..109
4.4.3.3 Internal Validity..111
4.4.3.4 External Validity..112
4.4.3.5 Statistical Conclusion Validity 112
4.4.3.6 Summary 112
4.5 Research Relevance 114
4.5.1 Defining Research Relevance 114
4.5.2 Evaluating Research Relevance..114
4.6 Research Design..116
4.6.1 Introduction..116
4.6.2 Survey Research Method 120

4.6.2.1 Definition 120
4.6.2.2 Types of Survey Research 120
4.6.2.3 Requirements for Survey Research..122
4.6.3 Empirical Setting: UK Manufacturing..124
4.6.3.1 Industrial Context and Background 124
4.6.3.2 Motivation for Studying UK Manufacturing 126
4.6.3.3 Prior Studies on Innovation in UK Manufacturing..................126
4.6.4 Sample..127
4.6.4.1 Population 127
4.6.4.2 Sampling Frame 127
4.6.4.3 Sampling Method..127
4.6.4.4 Unit of Analysis 128
4.6.4.5 Key Informants 128
4.7 Operationalisation of Constructs – Measures 129
4.7.1 Dependent Variables..129
4.7.1.1 Product Innovation..129
4.7.2 Independent Variables..137
4.7.2.1 Organizational Performance Measurement Diversity.............137
4.7.3 Mediating Variables...138
4.7.3.1 Innovation Search 138
4.7.3.2 Risk 139
4.7.4 Moderator Variables 139
4.7.4.1 Diagnostic Use..141
4.7.4.2 Interactive Use 141
4.7.5 Control Variables...142
4.7.5.1 Firm Level Controls...143
4.7.5.1.1 Firm Size..143
4.7.5.1.2 Firm Age 144
4.7.5.2 Industry Level Controls 144
4.7.5.3 Country Level Controls..145
4.7.6 Summary 145
4.8 Model Estimation..146
4.8.1 Moderation Analysis...146
4.8.2 Mediation Analysis 154
4.8.3 Direct Effects Analysis 157
4.8.4 Summary 157
4.9 Summary 160
5 Data Collection 161
5.1 Introduction...161
5.2 Key Informants 162
5.3 Development of Survey Instruments 165
5.4 Administration of Survey Instruments...168
6 Data Results and Analysis 171
6.1 Introduction..171
6.2 Analysing respondents 172
6.2.1 Analysing sample demographics 172
6.2.2 Analysing sample size...175
6.2.3 Testing for Non-response Bias...176

6.2.4 Testing for Data Source Bias (Multiple rater agreement) 177
6.3 Assessing the Quality of the Data 178
6.3.1 Assessing Missing data 178
6.3.2 Assessing Outliers 179
6.3.3 Assessing Normality 180
6.3.4 Testing for common method variance 182
6.4 Validating Measures 183
6.4.1 Assessing Content (Face) Validity 185
6.4.2 Assessing Unidimensionality 188
6.4.2.1 Diagnostic Use Scale 192
6.4.2.2 Interactive Use Scale 193
6.4.2.3 Organizational Performance Measurement Diversity Scale ... 197
6.4.2.4 Search Scale 203
6.4.2.5 Risk Taking Scale 206
6.4.2.6 Product Innovation Scale 208
6.4.3 Assessing Reliability 211
6.4.4 Assessing Convergent Validity 212
6.4.5 Assessing Discrminant Validity 214
6.4.6 Overall Assessment of the Measures 218
6.5 Testing hypotheses 219
6.5.1 Descriptive Statistics 219
6.5.2 Testing moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation (Hypotheses 2a, 2b, 2c, 3a, and 3b) 220
6.5.2.1 Testing multiple regression assumptions 222
6.5.2.2 The moderating role of diagnostic use 228
6.5.2.3 The moderating role of interactive use 231
6.5.3 Testing moderating effects of performance measurement use on the relationship between performance measurement diversity and innovation search (Hypotheses 4a, 4b, 4c, 5a, and 5b) 234
6.5.3.1 Moderating role of diagnostic use 235
6.5.3.2 Moderating role of interactive use 238
6.5.4 Testing moderating effects of performance measurement use on the relationship between performance measurement diversity and risk (Hypotheses 6a, 6b, 7a, and 7b) 241
6.5.4.1 Moderating role of diagnostic use 243
6.5.4.2 Moderating role of interactive use 243
6.5.5 Testing mediating effects innovation search and risk (Hypotheses 9 and 11) 246
6.5.6 Testing direct effects of performance measurement diversity, innovation search and risk on product innovation (Hypotheses 1, 8 and 10) 249
6.5.6.1 Direct effect of performance measurement diversity on product innovation (Hypothesis 1) 250
6.5.6.2 Direct effects of innovation search on product innovation (Hypothesis 8) 251
6.5.6.3 Direct effects of organizational risk taking on product innovation (Hypothesis 10) 251

6.5.7 Summary of the results of Hypothesis Testing 251
7 DISCUSSION, CONTRIBUTION, AND LIMITATIONS 253
7.1 Introduction 253
7.2 Discussion 254
7.3 Contribution 256
7.3.1 Contribution to Theory 257
7.3.1.1 Behavioural Theory of Innovation 257
7.3.1.2 Strategic control theory 258
7.3.1.3 Theory of dynamic capabilities 259
7.3.2 Contribution to Research 260
7.3.3 Contribution to Practice 261
7.4 Research Limitations 263
7.4.1 Conceptual 264
7.4.2 Sampling 264
7.4.3 Instrumentation 264
7.4.4 Statistical Methods 266
7.4.5 Research Design 267
7.5 Summary 268
8 CONCLUSIONS 269
8.1 Introduction 269
8.2 Research Rigour & Relevance 272
8.3 Research Findings 274
8.4 Future Research 276
8.5 Summary 281
9 REFERENCES 282
10 APPENDICES 342
10.1 Appendix A: Historical Review of Performance Measurement Research 342
10.2 Appendix B: Review of Performance Measurement Frameworks 348
10.2.1 Performance Measurement Matrix 349
10.2.2 Performance Measurement Questionnaire (PMQ) 350
10.2.3 Results and Determinants Framework 352
10.2.4 SMART (Performance Pyramid) 353
10.2.5 Balanced Scorecard 354
10.2.6 Integrated Performance Measurement System 355
10.2.7 Performance Prism 357
10.3 Appendix C: Cover Letter and Mail-based Survey Instrument 359
10.4 Appendix D: Cover Letter and Web-based Survey Instrument 365
10.5 Appendix E: Existing Scales 372
10.5.1 Scales for Measuring Simons' Diagnostic and Interactive Uses 372
10.5.2 Existing Scales for Measuring Organizational Performance Measurement Diversity 374
10.5.3 Existing Scales for Measuring Product Innovation 376
10.6 Appendix F: Simons' Levers of Control – Empirical Research 378

Index of Figures

Figure 1-1: Three perspectives on the impact of measurement on innovation 3
Figure 1-2: Research model 6
Figure 2-1: Outline of chapter 2 12
Figure 2-2: Performance measurement system concept (Source: Adapted from Neely et al., 1995) 14
Figure 2-3: Innovation space (Tushman, Smith, Wood, Westerman, & O'Reilly, 2004) 20
Figure 2-4: Market-based Innovations (Source: Christensen & Raynor, 2003)...21
Figure 2-5: Components of creative performance (Source: Amabile, 1986)........33
Figure 2-6: The Componential theory of organizational creativity and innovation (Source: Amabile 1997: 53) 35
Figure 2-7: Componential model of organizational innovation (Source: Amabile, 1988: 152) 35
Figure 2-8: First generation innovation model: Technology push (Source: Rothwell, 1994) 37
Figure 2-9: Second generation innovation model: Market pull (Source: Rothwell, 1994) 37
Figure 2-10: The process of technical innovation (Source: Utterback, 1971)......38
Figure 2-11:Third generation innovation model: Coupling model (Source: Rothwell, 1992) 39
Figure 2-12: Third generation innovation model: Chain-linked model (Source: Kline & Rosenberg, 1986) 39
Figure 2-13: Fourth generation innovation model: Integrated model (Rothwell, 1992) 40
Figure 2-14: Fifth generation innovation model: System integration and network model 41
Figure 2-15: Levers of control framework (Source Simons, 1995: 157)..............57
Figure 3-1: Outline of Chapter 3 61
Figure 3-2: Determinants of organizational aspiration levels 66
Figure 3-3: Experiential learning performance feedback-based model (Source: Yaghi, 2005 adapted from Cyert & March, 1963; Greve, 2003a; March, 1994)..74
Figure 3-4: Performance feedback-based experiential learning (Adapted from Argote, 1999; Cohen & Sproull, 1996; Lomi, Larsen, & Ginsberg, 1997; Schulz, 2002) 75
Figure 3-5: Theoretical underpinnings of the research model 76
Figure 3-6: Behavioural model of innovation (Source: Fiol, 1996; Greve 2003a; Simon & March, 1958) 77
Figure 3-7: Research model 78
Figure 3-8: Diagnostic control system (source: Simons, 1995) 83
Figure 4-1: Outline of chapter 4 95
Figure 4-2: The circularity of epistemology (Source: Johnson & Duberly, 2000: 4) 97
Figure 4-3: Management research framework (Source: Yaghi, 2005) 99
Figure 4-4: The Relationship between philosophical perspective and research process (Source: Yaghi, 2005) 100

Figure 4-5: Ingredients of Realist Social Explanation (Source: adapted from Pawson & Tilley, 1997) 104
Figure 4-6: Relationship between concepts, constructs, and operational definitions (Source: Black, 1999: 36) 110
Figure 4-7: Types of validities mapped into research design (Source: Black, 1999: 58) 113
Figure 4-8: Types of validities mapped into theoretical and operational domains (Source: Adapted from Malhotra & Grover, 1998) 113
Figure 4-9: The impact frontier: rigour and relevance (Source: Adapted from Davenport & Markus, 1999) 114
Figure 4-10: The maturity cycle of research (Source: Malhotra & Grover, 1998) 122
Figure 4-11: Linkages between decisions in survey planning (Source: Adapted in Alreck & Settle, 1985 as reported in Forza, 2002) 124
Figure 4-12: Composition of the UK economy 125
Figure 4-13: The relationship between inventions, patents, and innovations (Source: Basberg, 1987) 134
Figure 4-14: Different operationalisations of diagnostic and interactive uses ... 140
Figure 4-15: The two types of moderations: moderation as form and moderation as strength (Source: Hartmann & Moers, 1999) 148
Figure 4-16: Monotonic and nonmonotonic forms of moderation relationships (Source: Adapted from Gerdin, 2005; Hartmann & Moers, 1999) 149
Figure 4-17: Direct effects model 155
Figure 4-18: Partially mediated model 156
Figure 5-1: Outline of chapter 5 161
Figure 5-2: Responses from the surveys 170
Figure 6-1: Outline of chapter 6 172
Figure 6-2: Respondents profile 173
Figure 6-3: A framework for measure validation (Sources: Bagozzi, Yi, Philips, 1991; Hair et al., 2006; Kaynak & Hartley, 2005; Ping, 2004; O'Leary-Kelly & Vokurka, 1998) 184
Figure 6-4: Scale development process (Source: Adapted from (Churhill, 1979; Froehle & Roth, 2004; Li, Rao, Ragu-Nathan, & Ragu-Nathan, 2005; Lori, Malhorta, Sharma, 2004; Menor & Roth, 2004; Moore & Benbasa, 1991; Stratman & Roth, 2002) 186
Figure 6-5: Diagnostic use measurement model 193
Figure 6-6: Interactive use scale measurement model 195
Figure 6-7: Performance measurement use (diagnostic + interactive) measurement model 197
Figure 6-8: Organizational performance measurement diversity (Stakeholder) scale measurement model 200
Figure 6-9: Organizational performance measurement diversity (Operations) scale measurement model 201
Figure 6-10: Organizational performance measurement diversity (Financial + Stakeholder + Operations) scale measurement model 203
Figure 6-11: Innovation search scale measurement model 206
Figure 6-12: Risk scale measurement model 208
Figure 6-13: Product innovation scale measurement model 210

Figure 6-14: Plot of standardized residuals versus predicted values of product innovation 222
Figure 6-15: Plot of residuals versus time sequence case ID....223
Figure 6-16: Histogram of the residuals 225
Figure 6-17: Normal P-P plot of regression standardized residual....225
Figure 6-18: Standardized residuals versus the explanatory variables 227
Figure 6-19: Partial derivative of product innovation showing the inflection point of diagnostic use 229
Figure 6-20: Effect of diagnostic use and performance measurement diversity on product innovation with interactive use held constant....230
Figure 6-21: Partial derivative of product innovation showing the inflection point of interactive use....232
Figure 6-22: Effect of interactive use and performance measurement diversity on product innovation with diagnostic use held constant....233
Figure 6-23: Partial derivative of innovation search showing the inflection point of diagnostic use 237
Figure 6-24: Effect of diagnostic use and performance measurement diversity on innovation search with interactive use held constant....238
Figure 6-25: Partial derivative of innovation search showing the inflection point of interactive use....240
Figure 6-26: Effect of interactive use and performance measurement diversity on innovation search with diagnostic use held constant....241
Figure 6-27: Partial derivative of risk showing the inflection point of interactive use 245
Figure 6-28: Effect of interactive use and performance measurement diversity on risk with diagnostic use held constant....246
Figure 7-1: Outline of chapter 7....253
Figure 7-2: Relationship between performance measurement diversity, diagnostic use, and product innovation....262
Figure 7-3: Relationship between organizational performance measurement, interactive use, and product innovation 263
Figure 8-1: Outline of chapter 8....271
Figure 8-3: Performance measurement: Spatial and temporal dimensions....280
Figure 8-4: Performance measurement: Multilevel framework....281
Figure 10-1: Performance measurement research streams: Mapping the field (Source: Adapted from Neely, 2005)....347
Figure 10-2: Classification of performance measurement frameworks (Source De Toni & Tonchia, 2001) 349
Figure 10-3: Performance measurement matrix (Source: Keegan, Eiler, Jones, 1989)....350
Figure 10-4: Performance measurement quessionaire : Part II (Source: Dixon, Nanni, Vollmann, 1990) 352
Figure 10-5: Performance measurement quessionaire : Part III (Source: Dixon, Nanni, Vollmann, 1990) 352
Figure 10-6: Results and determinants framework (Fitzgerald, Johnston, Bringall, Silvestro, & Voss, 1991)....353
Figure 10-7: SMART framework (Source: Cross & Lynch, 1991)....354
Figure 10-8: The balanced scorecrad (Source: Kaplan & Norton, 1996)....355

Figure 10-9: Integrated performance measurement systems (Source: Pun & White, 2005 based on Bittici, Carrie & McDevitt, 1997)....................................356
Figure 10-10: The performance prism (Source: Neely, Adams, & Kennerly, 2002) 358

Index of Tables

Table 1-1: Thesis structure and research activities 10
Table 2-1: Performance measurement definitions (source: adapted from Franco et al. 2004) 14
Table 2-2: Definitions of product innovation 22
Table 2-3: Formalization-innovation relationships in theories of organizational innovation (Source: Damanpour, 1991; Damanpour & Gopalakrishnan, 1998) .29
Table 2-4: Innovation models and performance measurement 42
Table 2-5: Empirical research investigating the impact of PMS on innovation ...43
Table 2-6: Summary of major classifications of PMS use 50
Table 2-7: Classification of uses of performance measurement systems 54
Table 2-8: Contrasting styles of PMS use (Source: adapted from Simons, 1994; 1995; 2000) 58
Table 3-1: Summary of hypotheses 92
Table 4-1: Contrasting positivism with social constructionism (Source: Easterby-Smith, Thorpe, & Lowe, 2002: 30) 101
Table 4-2: Domain of reality (Source: Bhaskar, 1978) 103
Table 4-3: Criteria for assessing rigour of research projects (Source: adapted from Shrivastava, 1987 and Varadarajan, 2003) 108
Table 4-4: The consequences of poor construct conceptualisation (Source: MacKenzie, 2003) 111
Table 4-5: Types of Validities (Source: Adapted from Shadish, et al: 38-39) ...112
Table 4-6: Criteria for assessing practical usefulness of research projects (Source: adapted from Shrivastava, 1987 and Thomas & Tymon, 1982) 115
Table 4-7: Evaluating different research methods against critical research requirements 116
Table 4-8: Justification of the research design 118
Table 4-9: Purposes of survey research (Source: Adapted from Forza, 2002; Pinsonneault & Kraemer, 1993) 121
Table 4-10: Requirements for different survey research purposes (Source: adapted from Forza, 2002 based on Pinsonneault & Kraemer, 1993) 123
Table 4-11: Sampling Approaches (Source: Forza, 2002) 127
Table 4-12: Measures of innovation (Source: adapted from Brouwer & Kleinknecht, 1999; Coombs, Goffin & Mitchell, 2005: 45; Narandren, & Richards, 1996) 132
Table 4-13: Approaches to measuring innovation (Source: adapted from Coombs, Narandren, & Richards, 1996; Kleinknecht, 1993; 1996) 136
Table 4-14: Summary of survey scales 146
Table 4-15: Framework for testing classical regression assumptions (Source: Hair et al., 2006; Studenmund, 2001) 152
Table 4-16: Summary of the hypotheses and the data analytics used to test them 158
Table 4-17: Assessing Survey Research Design (Source: Adapted from Malhotra & Grover, 1998) 160
Table 5-1: Key informants strategies (Source: Huber & Power, 1985) 164

Table 5-2: Dillman's (1978; 2000) TDM recommendations (Source: adapted from (Ranaweera, 2002: 150-152) 165
Table 6-1: Respondents profile 173
Table 6-2: Standard industry code (SIC) of the responding organizations 174
Table 6-3: Tests of mean equality between early versus late mail- survey respondents 176
Table 6-4: Tests of mean equality between early mail- survey and web-survey respondents 177
Table 6-5: Potential outliers 179
Table 6-6: Assessment of the normality of the items 180
Table 6-7: Results of the unrotated exploratory factor analysis (EFA) testing for common method variance 183
Table 6-8: Description of fit indices and their recommended critical values (Compiled from different sources: (Byrne, 2001; Hoyle & Panter, 1995; Hu & Bentler, 1999; Kline, 2003) 191
Table 6-9: Diagnostic scale: Descriptive statistics 192
Table 6-10: Confirmatory factor analysis of diagnostic use scale 192
Table 6-11 : Interactive scale: Descriptive statistics 194
Table 6-12: Confirmatory factor analysis of interactive use scale 195
Table 6-13: Confirmatory factor analysis of performance measurement use (diagnostic + interactive) 196
Table 6-14: Organizational Performance measurement diversity scale : Descriptive statistics 197
Table 6-15: KMO and Bartlett's test for organizational performance measurement diversity items 198
Table 6-16: Organizational performance measurement diversity pattern matrix 199
Table 6-17: Confirmatory factor analysis of organizational performance measurement diversity (stakeholder) scale 199
Table 6-18: Confirmatory factor analysis of organizational performance measurement diversity (operations) scale 201
Table 6-19: Confirmatory factor analysis of organizational performance measurement diversity (financial + stakeholder + operations) scale 202
Table 6-20: Search scale: Descriptive statistics 203
Table 6-21: KMO and Bartlett's test for innovation activities performance items 204
Table 6-22: Innovation activities pattern matrix 205
Table 6-23: Confirmatory factor analysis of innovation search scale 205
Table 6-24: Risk scale : Descriptive statistics 206
Table 6-25: Confirmatory factor analysis of risk scale 207
Table 6-26: Product innovation scale: Descriptive statistics 208
Table 6-27: Confirmatory factor analysis of product innovation scale 209
Table 6-28: Confirmatory factor analysis of product innovation scale after deleting item 3 (V3) 210
Table 6-29: Reliability measures for multi-tem scales 212
Table 6-30: Standardized factor loadings and their statistical significance 213
Table 6-31: Average Variance Extracted (AVE) from multi-tem scales 214

Table 6-32: Results of discriminant validity: Correlations exhibited left of the diagonal, squared correlations (R^2) between measures exhibited right of the diagonal and extracted variances (AVE) exhibited at the diagonal 215
Table 6-33: Results of discriminant validity pair-wise tests 216
Table 6-34: Summary results of construct validation 218
Table 6-35: Statistical descriptives and correlations of the variables used in hypotheses testing 219
Table 6-36: Casewise diagnostics 220
Table 6-37: OLS Regression results of the moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation 221
Table 6-38: Collinearity statistics 224
Table 6-39: Correlations between the standardized residuals and indepedent variables 226
Table 6-40: Location of inflection point for diagnostic use in moderating the relationship between performance measurement diversity and product innovation 229
Table 6-41: Location of inflection point for interactive use in moderating the relationship between performance measurement diversity and product innovation 232
Table 6-42: OLS regression results of the moderating effects of performance measurement use on the relationship between performance measurement diversity and innovation search...235
Table 6-43: Location of inflection point for diagnostic use in moderating the relationship between performance measurement diversity and innovation search 236
Table 6-44: Location of inflection point for interactive use in moderating the relationship between performance measurement diversity and innovation search 239
Table 6-45: OLS regression results of the moderating effects of performance measurement use on the relationship between performance measurement diversity and risk...242
Table 6-46: Location of inflection point for interactive use in moderating the relationship between performance measurement diversity and product innovation 244
Table 6-47: Test results of the mediating effects of innovation search and risk 248
Table 6-48: OLS regression results for the impact of performance measurement diversity, innovation search, and risk taking on product innovation 250
Table 6-49: summary of the results of the hypothesis testing...........................252
Table 8-1: Evaluation of research rigour (Source: adapted from Shrivastava, 1987 and Varadarajan, 2003) 272
Table 8-2: Evaluation of practical usefulness (Source: adapted from Shrivastava, 1987 and Thomas & Tymon, 1982) 273
Table 10-1: Performance measurement evolution and the development of performance measurement research streams (Source: Neely, 2005) 346
Table 10-2: Existing scales for measuring Simons' diagnostic and interactive uses..372

Table 10-3: Existing scales for measuring organizational performance measurement diversity. 375
Table 10-4: Existing scales for measuring product innovation. 376
Table 10-5: Empirical research using Simons' levers of control framework 378
Table 10-6: Research models for empirical research using Simons' levers of control framework 386

"Knowledge seems to increase questions at a faster rate than it increases answers. It provides too many qualifications, recognizes too much complexity" (March, 1994: 265).

"The pursuit of organizational [by analogy human] intelligence is an activity in which knowledge can sometimes produce power but more reliably produces humility" (March, 1999: 10), emphasis added.

1 Introduction

1.1 Introduction

The research reported in this thesis explores the impact of organizational performance measurement diversity on product innovation and tests whether the impact differs depending on how the performance measurement system is used. Organizational performance measurement diversity[1] is defined by the extent to which top management teams measure and use information related to a broad set of financial and non-financial measures and product innovation refers to the market introduction of a new good or service or a significantly improved good or service. The thesis is investigated from a Campbellian realist philosophical perspective, and is examined through the lens of behavioural theory of innovation (Greve, 2003b). This research investigates the moderating effects of diagnostic and interactive uses (Simons, 1995) of performance measurement systems on the relationship between organizational performance measurement diversity and product innovation through the mediating mechanisms of innovation search and risk taking. Diagnostic use is defined by the extent to which top management teams use performance measures to monitor organizational outcomes to correct deviations from preset standards of performance (targets) and interactive use is defined by the extent they involve themselves regularly and personally in decision activities of subordinates to focus their subordinates' search on the strategic uncertainties facing their organizations.

The research uses a cross-sectional, large-scale, probability sample survey of 145 UK manufacturing firms to test the hypotheses advanced in this study.

The aim of this chapter is not only to set the context of the research but also to indicate the shape and content of the forthcoming in-depth discussions of the relevant literature streams, development of the research model and research

[1] In this thesis, I use the terms organizational performance measurement diversity and performance measurement diversity interchangeably.

hypotheses, methodological approach and empirical findings. In order to set the scene for the study this chapter describes the motivation for the study, the research aims and questions, research gaps uncovered in the literature review, research model, research contributions, and research approach and the structure of the thesis.

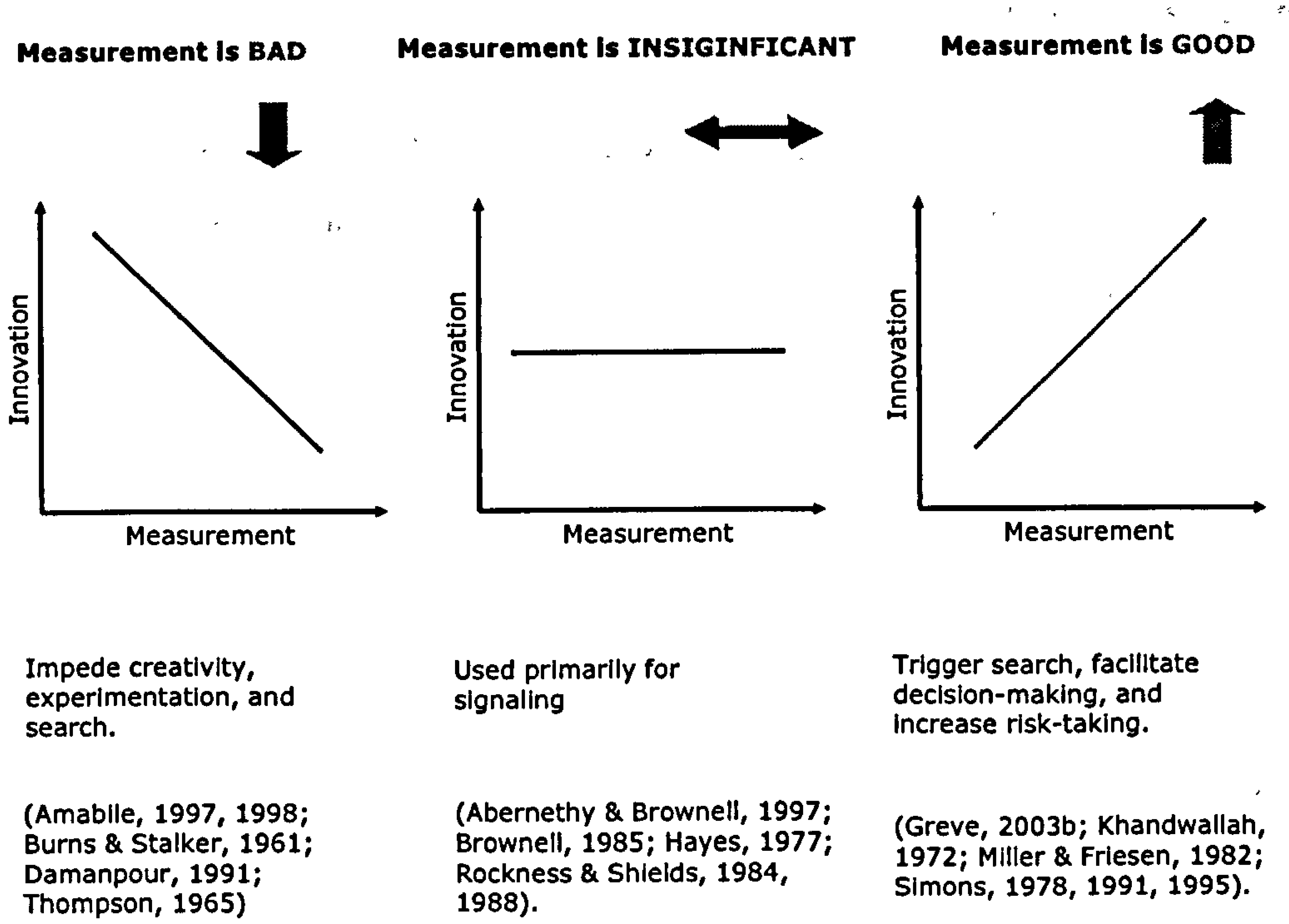
1.2 Motivation

There is a strong empirical evidence that many companies who are successful today are less likely to be successful in the future because they fail to sustain innovation. For example, Christensen's (1997) study of hard disk drive industry, Henderson & Clark's (1991) study of photolithographic alignment equipment industry, Tripsas & Gaviti's (2000) study of digital imaging, and Tushman & Anderson's (1986) study of airlines, minicomputers, and cements have all shown that firms that were once successful have either failed to survive or to sustain their performance in the long run because they could not sustain innovation. In today's dynamic environments, for organizations to survive and prosper, they must innovate (Bessant, 2003; Damanpour, 1991) because by introducing new products, organizations are able to establish new markets, develop new technologies, develop new capabilities and adapt to changes in their competitive environments (Brown & Eisenhardt, 1995; Damanpour, 1991; Danneels, 2002; Doughtery, 1992; Eisenhardt & Martin, 2000; Schoonhoven, Eisenhardt, & Lyman, 1990). As Peter Drucker sums it up, "Today, no one needs to be convinced of the importance of innovation ... How to innovate is the key question" (Goffin & Mitchell, 2005: 43). Surprisingly then, why when everyone is saying that innovation is key, are there so many organizations adopting performance measurement systems or management control systems[2] (Neely, 2005; Rigby, 2001; Silk, 1998) which may constrain their innovativeness?

[2] Management control system (MCS) is a broad term that includes performance measurement systems. Chenhall (2003) notes that the "definition of MCS has evolved over the years from one focusing on the provision of more formal, financially quantifiable information to assist managerial decision making to one that embraces a much broader scope of information. This includes external information related to markets, customers, competitors, non-financial information related to production processes, predictive information and a broad array of decision support mechanisms, and informal personal and social controls." Throughout this paper, the terms

As illustrated in Figure 1-1, there are three differing perspectives on the effect of performance measurement systems on a firm's propensity to innovate with each of these having empirical evidence to support its argument. The first perspective views performance measurement systems as constraining innovation because they impede creativity, experimentation, and search in firms (e.g. Amabile, 1983, 1988, 1996, 1997, 1998; Amabile, Conti, Lazenby, Herron, 1996; Burns & Stalker, 1961; Damanpour, 1991; Thompson, 1965). The second perspective views performance measurement systems as helping innovation because they trigger search, facilitate decision-making, and increase risk-taking (e.g. Godener & Soderquist, 2004; Greve, 2003b; Khandwallah, 1972, 1973; Miller & Friesen, 1982; Simons, 1978; 1991; 1995). The third perspective views performance measurement systems as having insignificant or little impact on innovation because they are used primarily for signalling and they do not impact innovation in high uncertainty environments (e.g. Abernethy & Brownell, 1997; Brownell, 1985; Hayes, 1977; Rockness & Shields, 1984, 1988).

Figure 1-1: Three perspectives on the impact of measurement on innovation



performance measurement system (PMS) and management control system (MCS) will be used interchangeably.

A possible explanation of the contradiction in the empirical findings of these studies is that they generally ignore how performance measurement systems are used "style of use". One useful conceptualization of styles of use is Simons' (1995) diagnostic and interactive uses (styles). Therefore, I argue that one possible way of resolving the contradictory findings of these perspectives is by incorporating performance measurement use (diagnostic and interactive) as a moderating variable. Based on the behavioural theory of innovation, I argue that the impact of performance measurement on product innovation can be either positive or negative, depending upon the way performance measures are used.

1.3 Research Aims and Questions

The aim of this PhD research is to enhance our understanding of the impact of performance measurement on product innovation. This aim is translated to the following research question:

What is the moderating effect of performance measurement use on the relationship between organizational performance measurement diversity and product innovation?

This research question in turn is translated into four sub-questions as follows:

1- What are the different styles of use of performance measurement systems?

2- To what extent does the diagnostic use of performance measurement systems moderate the relationship between organizational performance measurement diversity and product innovation?

3- To what extent does the interactive use of performance measurement systems moderate the relationship between organizational performance measurement diversity and product innovation?

4- What are the mechanisms the mediate the impact of the interaction between performance measurement uses (diagnostic and interactive) and organizational performance measurement diversity on product innovation?

1.4 Research Gaps

The literature review reported in the second chapter identified a number of research gaps in the existing performance measurement and innovation literatures. Following Atuahene-Gima's (2004) classification, the gaps are grouped under four categories: theoretical, empirical, contextual, and substantive. Theoretical gaps refer to insufficient explanation or prediction of some phenomena – the "why". Empirical gaps refer to the lack of empirical studies or inconsistent findings on some phenomena – the "where" and "when". Contextual gaps refer to the generalizability of the findings of existing research. Substantive gaps refer to the lack of managerial understanding of "how".

1.4.1 Theoretical Gaps

- The interaction between performance measurement use and organizational performance measurement diversity has not been largely theorized.
- Behavioural theory of innovation assumes that performance measures are used diagnostically.

1.4.2 Empirical Gaps

- The few empirical studies that examined the relationship between performance measurement systems and product innovation have yielded inconsistent findings.
- There are no quantitative studies that examined the impact of interactive and diagnostic use on product innovation.

1.4.3 Contextual Gaps

- There are no quantitative studies that examined the relationship between performance measurement systems and product innovation in UK setting.
- There are no quantitative studies that examined the impact of interactive and/or diagnostic use on product innovation in UK setting.

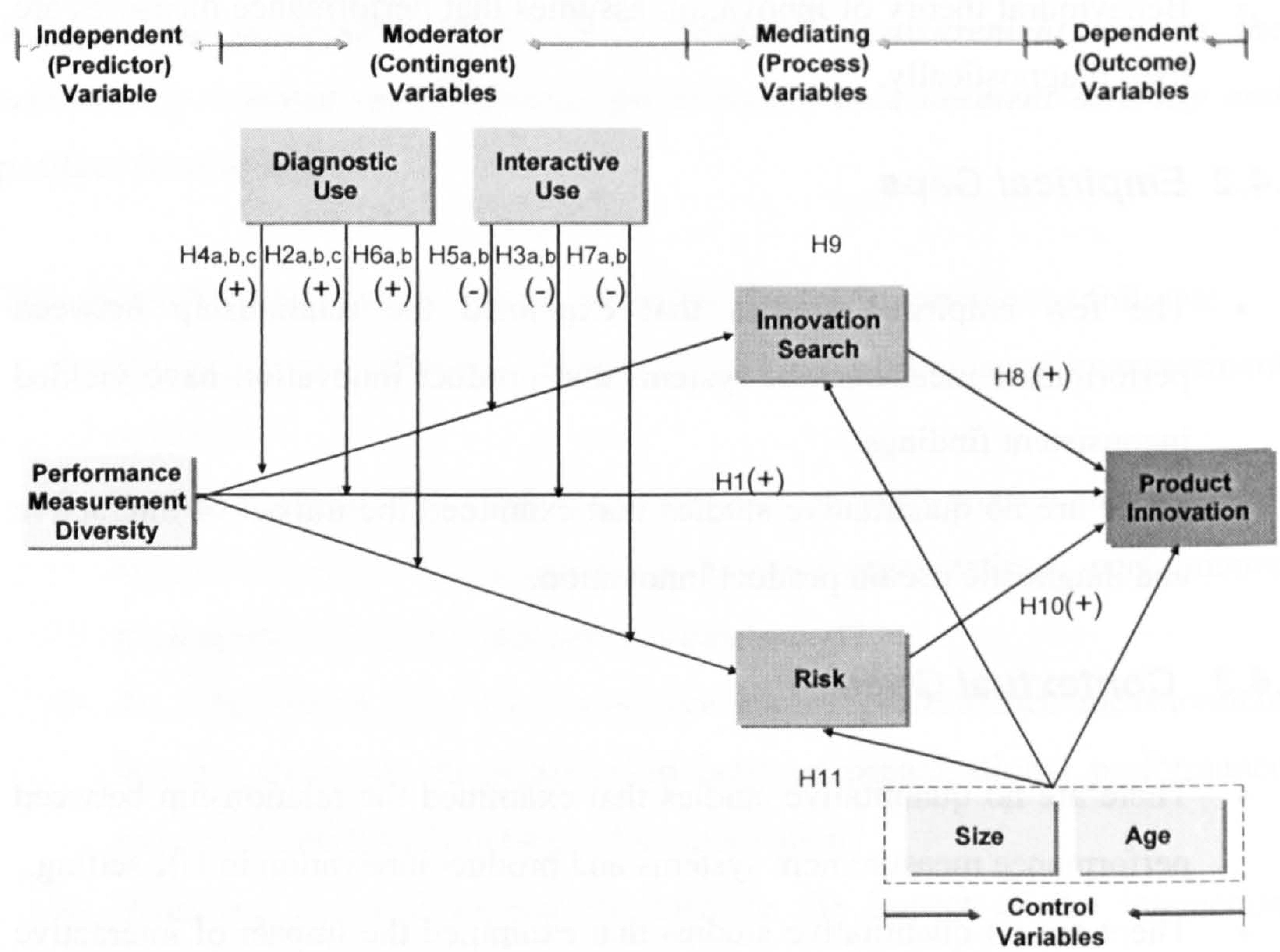
1.4.4 Substantive Gaps

- There is a managerial requirement for theoretically grounded and empirically tested innovation models that enable managers to understand how best they could use performance measurement systems to manage product innovations in their firms.

1.5 Research Model

To achieve the aim of this research, answer the research questions, and address the identified research gaps, I develop a product innovation research model and research hypotheses that are theoretically grounded on the behavioural theory of innovation (Greve, 2003b) which is a special (applied) case of the performance feedback theory (Greve, 2003a; March, 1994) as illustrated in Figure 1-2.

Figure 1-2: Research model



1.6 Research Contributions

By developing and testing a product innovation research model based on the behavioural theory of innovation that incorporates both performance measurement use and organizational performance measurement diversity in British manufacturing context, this study contributes to existing theoretical, empirical, and practice literatures in several ways.

First, the findings of this study has theoretical implications for behavioural theory of innovation (Greve, 2003b), strategic control theory (Simons, 1995), and the theory of dynamic capabilities (Teece, Pisano, & Shuen, 1997). Second, it contributes to research by testing the impact of the interaction of diagnostic and interactive uses with organizational performance measurement diversity on product innovation at the firm level in the UK manufacturing sector. Third, it contributes to practice by providing insights as to how firms might become more innovative by properly designing and using their performance measurement systems.

1.7 Research Approach and Thesis Structure

The thesis structure follows closely the research approach and it consists of ten chapters as illustrated in Table 1-1. The first chapter sets the scene and the focus of research.

I start the second chapter by reviewing the definitions of performance measurement and product innovation and justifying my choice of the definitions adopted in this study. Then, I discuss how measurement may constrain or help innovation by examining the theoretical and empirical literature from three perspectives: organizational contingency, organizational creativity, and practice (rational) perspectives. Next, I review the literature on the performance measurement use and justify my choice of the Simons' (1995) diagnostic and interactive use. In the final section of the chapter, I summarize the gaps in the literature uncovered in the literature review.

The third chapter starts with the definitions of the constructs under study followed by a review of the performance feedback theory. Next, I advance the research model that is based on the behavioural theory of innovation followed by hypotheses development. The final section summarizes the research hypotheses advanced in this study.

The fourth chapter examines the methodology of the study to ensure that its design is appropriate to provide answers to the research questions and to test the research hypotheses. In this chapter, I define the ontological and epistemological elements of philosophical perspectives and explain how they linked to management research. Then, I discuss Campbellian realism, which forms the basis of my ontological and epistemological philosophical perspective. Since philosophical perspectives influence the logic of inquiry or research strategy, I explain the research strategy adopted in this study in the next section. Then I discuss two important criteria for evaluating academic research: rigour and relevance and I develop criteria for evaluating rigour and relevance. Next, I explain the cross-sectional, probability sample, survey design employed in this study to answer the research questions. Having justified my research design, I operationalise the constructs used in this study. Then, I discuss the three statistical analyses that will be used to test the three types of the hypotheses advanced in this study: moderation analysis, mediation analysis, and direct effect analysis. The final section shows how the research design complies with all the survey research design criteria advanced by Malhotra & Grover (1998).

In the fifth chapter, I explain how the survey instruments are developed incorporating the chosen validated scales in the previous chapter and how they are administered to collect information on the variables under study. In this chapter, I define the key informants in this study and I detail the guidelines used to minimize the inaccuracy and bias of their retrospective accounts. Then, I explain the development of the mail-based and web-based versions of the survey using Dillman's (1978) total design method. In the final section of the chapter, I explain the administration of the mail- and web-based surveys using Dillman's TDM procedure and I report the results.

In the sixth chapter, the collected data is analysed. I start by analysing the respondents' profiles and their organizations' industry affiliation. Next, I demonstrate that the sample size of 145 used in this study exceed the suggested limits of the various statistical techniques undertaken in the analysis of data. I also test for the existence of non-response bias. Having analysed the respondents, I move to assessing the quality of the data by assessing missing data, outliers, normality, and common method bias. Having ensured the data meet rigorous quality standards, I move to validating the measures used in this study in the fourth section. This is accomplished by assessing the content (face) validity, unidimensionality, reliability, convergent validity, and divergent validity. Having ensured the validity and reliability of the measures, I move to testing the research hypotheses in the final section using the model estimation techniques. Finally, I summarize the results of the hypotheses testing.

I start the seventh chapter by discussing the findings reported in the previous chapter. Next, I show how these findings contribute to existing theoretical, empirical, and practice literatures in several ways. The discussion then turns to the limitations of the research.

In the eighth chapter, I tie together those that have gone before by summarising the findings of the research and identifying areas that require further work.

The ninth chapter lists the reference used in this study.

In the tenth chapter, I include the appendices. In the first two appendices, I offer a historical review of performance measurement research and performance measurement frameworks, respectively. Then, in the next two appendices, I present the mail- and web-based survey instruments and the accompanying cover letters, respectively. In the last appendix, I review the empirical evidence on Simons' (1995) strategic control theory.

Table 1-1: Thesis structure and research activities

Chapter	Title	Research Activities
1	Introduction	▪ Setting the scene and focus of the research ▪ Stating research aims ▪ Stating research questions
2	Literature review	▪ Defining performance measurement ▪ Defining product innovation ▪ Reviewing literature on the impact of measurement on innovation ▪ Reviewing empirical evidence ▪ Reviewing literature on performance measurement use ▪ Identifying research gaps
3	Theoretical Model and Hypotheses	▪ Defining key constructs. ▪ Presenting the performance feedback theory ▪ Developing the research model ▪ Developing research hypotheses
4	Research methodology	▪ Stating ontological and epistemological philosophical position ▪ Specifying research design ▪ Operationalising research constructs ▪ Translating research hypotheses into statistical models ▪ Specifying how the hypotheses are going to be tested
5	Data collection	▪ Developing mail and web survey instruments ▪ Piloting survey ▪ Administering survey
6	Data results and analysis	▪ Analysing respondents ▪ Assessing the quality of data ▪ Validating measures ▪ Testing hypotheses
7	Discussion, contribution, and, limitations	▪ Discussing the findings of the study ▪ Discussing research contributions ▪ Discussing research limitations
8	Contributions	▪ Summarising the findings of the study ▪ Suggesting future research issues
9	References	▪ Listing references
10	Appendices	▪ Reviewing performance measurement research ▪ Reviewing performance measurement frameworks ▪ Reviewing survey instruments ▪ Reviewing existing scales ▪ Reviewing empirical research on Simons' levers of control

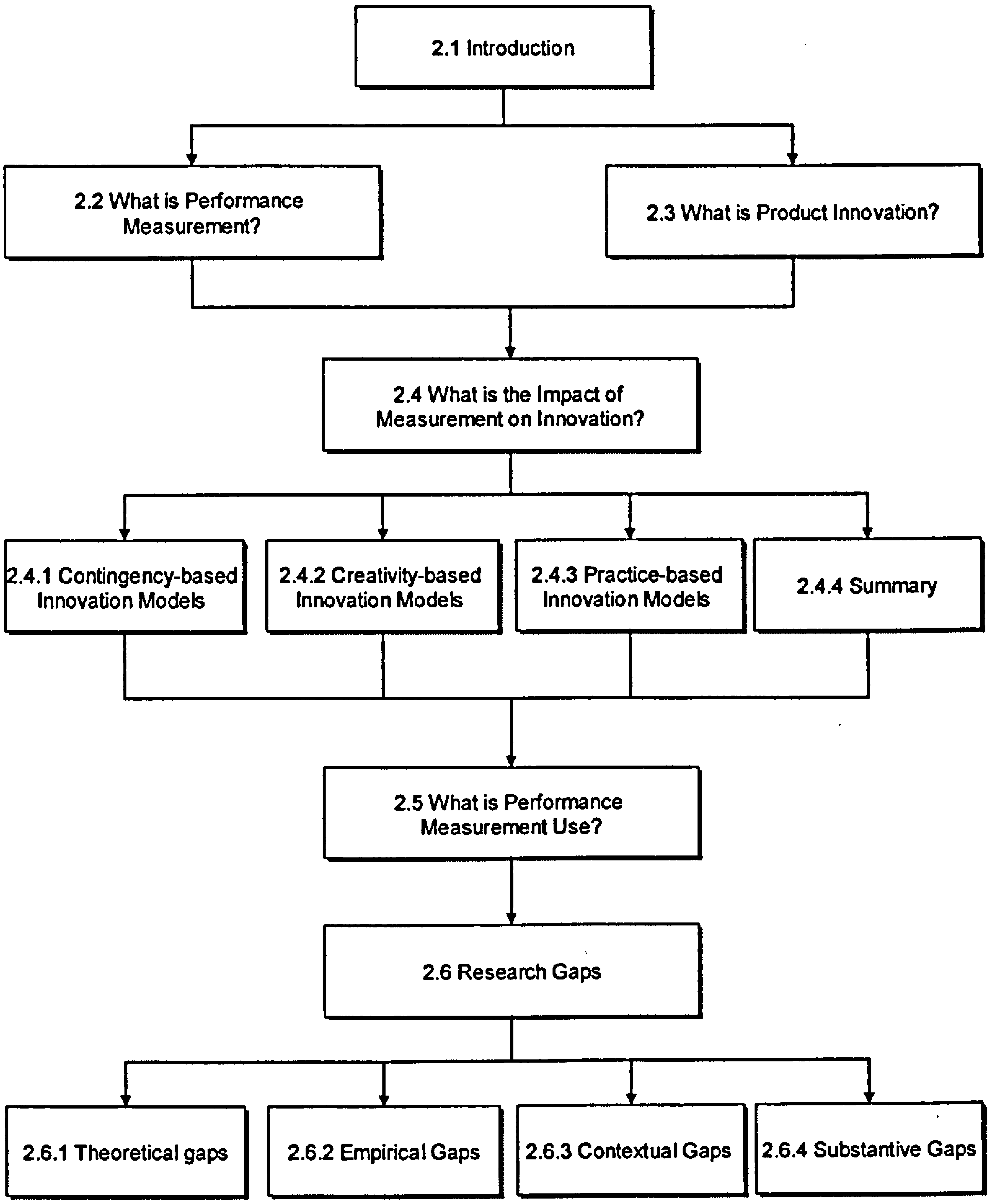
2 Literature Review

2.1 Introduction

As mentioned in the previous section the aim of this research is to enhance our understanding of the impact of performance measurement on product innovation. Furthermore, the aim is translated to the following research question: *What is the moderating effect of performance measurement use on the relationship between organizational performance measurement diversity and product innovation?* Therefore, this chapter reviews the literature streams that have investigated the impact of performance measurement on innovation.

There are six sections in this chapter, as illustrated in Figure 2-1. The first section introduces the chapter. Since both performance measurement and product innovation literature streams are very diverse with a broad range of authors, papers, and disciplines contributing to it, I start by reviewing the various definitions of performance measurement and product innovation in the second and third sections, respectively. Out of the so many definitions, I choose and justify my choice of the performance measurement and product innovation definitions used in this study. Having defined the dependent (outcome) and independent (explanatory) variables of the study, I move to review the literature on the relationship between performance measurement and innovation. I accomplish this by discussing how measurement may constrain or help innovation by examining the theoretical and empirical literature from three perspectives: organizational contingency, organizational creativity, and practice (rational) perspectives. In the fifth section, I review the literature on performance measurement use (moderating) variable and justify my choice of the Simons' (1995) diagnostic and interactive use. In the final section, I summarize the research gaps in the literature uncovered in the literature review. Theses research gaps are classified into four types: theoretical, empirical, contextual, and substantive.

Figure 2-1: Outline of chapter 2



2.2 What is Performance Measurement?

Performance measurement research[3] is very diverse with a broad range of authors, papers, and disciplines contributing to it (Marr & Schiuma, 2003; Neely, 2002). This diversity is reflected in the many definitions offered for performance measurement systems. In Table 2-1, I reproduce and update the extensive review

[3] See appendix A for a historical review of performance measurement research.

of performance measurement definitions carried out by the researchers at the Centre for Business Performance, Cranfield School of Management.

In this paper, I adopt the definitions of performance measurement advanced by Neely, Gregory & Platts (1995). They are as follows:

- **Performance measurement systems** can be defined as the set of metrics used to quantify both the efficiency and effectiveness of actions.
- ***Performance measurement*** can be defined as the process of quantifying the efficiency and effectiveness of action.
- ***Performance measure*** can be defined as a metric used to quantify the efficiency and/or effectiveness of an action.

There are at least four reasons for choosing the definition of performance measurement system advanced by Neely et al. (1995):

1) The definition is broad enough to encompass the two general approaches to designing performance measurement: organizational performance measurement diversity and alignment with strategy and value drivers (Ittner, Larcker, Randall, 2003). The first general approach emphasizes "supplementing traditional financial measures with a diverse mix of non-financial measures that are expected to capture key strategic performance dimensions that are not accurately reflected in short-term accounting measures... [The] second general approach emphasizes the implementation of performance measurement systems that are more closely linked to the firm's *specific* strategy and value drivers." (Ittner et al., 2003: 717-718).
2) The definition is not framework specific. This allows for a wider coverage of performance measurement frameworks (See appendix B for a review of performance measurement frameworks) used in practice such as Kaplan & Norton's Balanced Scorecard (1992; 1996; 2001; 2004) and Neely, Adams, & Kennerly's Performance Prism (2002).
3) The definition allows for different levels of analysis: measure, set of measures (e.g., performance category), and whole system. In addition, implicit in the definition the possibility of creating links between

measures and set of measures (See Figure 2-2 for an illustration of the concept).

4) The definition is widely adopted in performance measurement research. Citation analysis undertaken by Neely (2005) and Marr & Schiuma (2003) revealed that the Neely et al's (1995) article is one of the most cited articles in performance measurement research.

Figure 2-2: Performance measurement system concept (Source: Adapted from Neely et al., 1995)

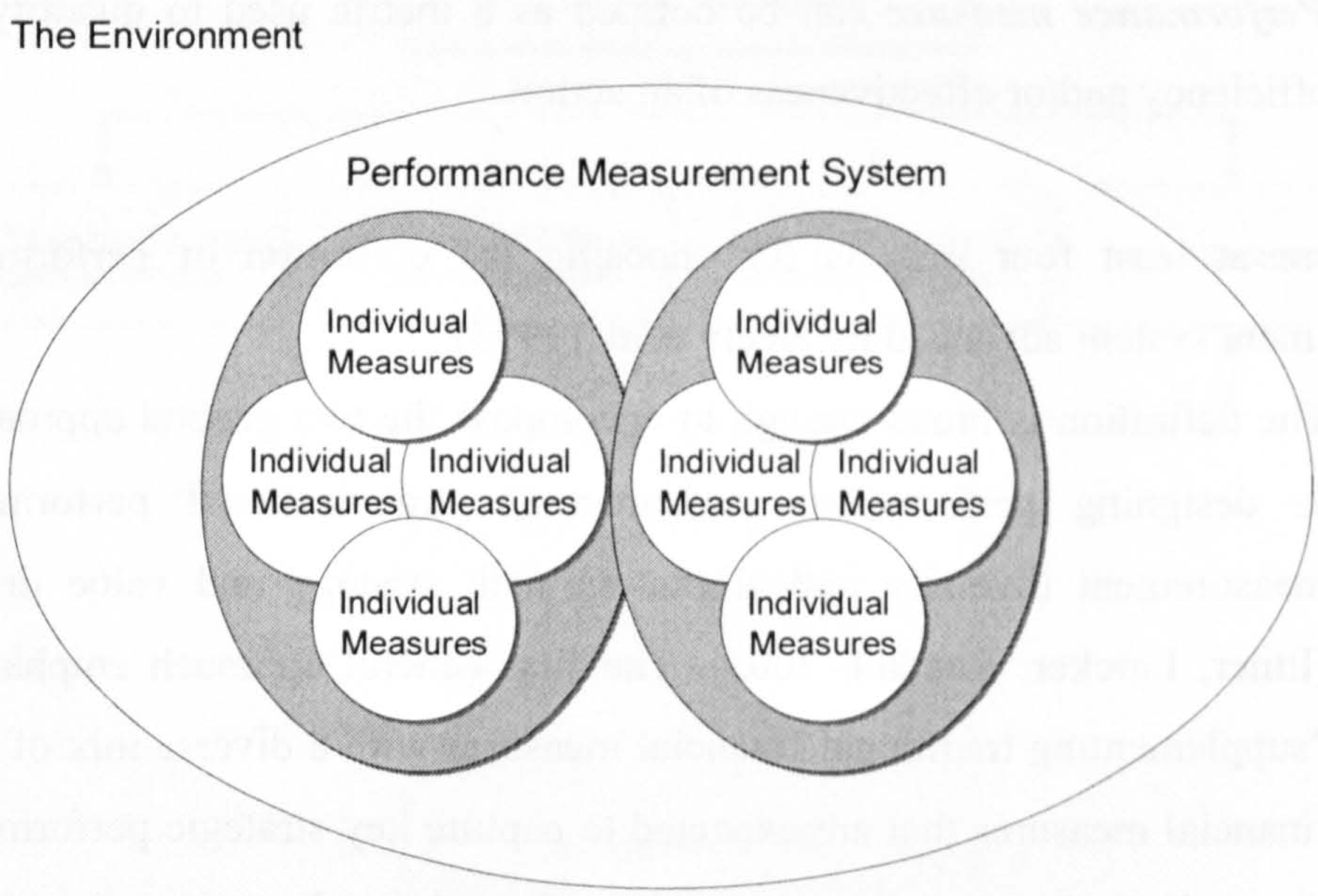


Table 2-1: Performance measurement definitions (source: adapted from Franco et al. 2004)

Study	Definition
(Rogers, 1990)	A business performance measurement system can be characterized as "... an integrated set of planning and review procedures which cascade down through the organization to provide a link between each individual and the overall strategy of the organization."
(Lynch & Cross, 1991)	A strategic performance measurement system is based on concepts of total quality management, industrial engineering, and activity accounting. A 2-way communications system is required to institute the strategic vision in the organization.

Study	Definition
(McGee, 1992)	The components of a strategic performance measurement system are: (1) performance metrics - defining evaluation criteria and corresponding measures that will operate as leading indicators of performance against strategic goals and initiatives. (2) Management process alignment - designing and reengineering core management processes to incorporate new performance metrics as they evolve, and balancing the various management processes of the organization so that they reinforce one another. The processes include: planning and capital allocation, performance assessment, management compensation and rewards, and stakeholder relationships. (3) Measurement and reporting infrastructure: establishing processes and supporting technology infrastructures to collect the raw data needed for all of an organization's performance metrics and to disseminate the results throughout the organization as needed.
(Lebas, 1995)	A performance measurement system should include a component that will continuously check the validity of the cause-and-effect relationships among the measures.
(Neely, Gregory & Platts, 1995)	A performance measurement system (PMS) is the set of metrics used to quantify both the efficiency and effectiveness of actions [...] A PMS can be examined at three different levels. (1) At the level of individual performance measures, the PMS can be analysed by asking questions such as: What performance measures are used? What are they used for? How much do they cost? What benefit do they provide? (2) At the next higher level, the performance measurement system as an entity, can be analysed by exploring issues such as: Have all the appropriate elements (internal, external, financial, nonfinancial) have been covered? Have measures which relate to the rate of improvement been introduced? Have measures which relate to both the long and the short term objectives of the business been introduced? Have the measures been integrated, both vertically and horizontally? Do any of the measures conflict with one another? (3) And at the level of the relationship between the performance measurement system and the environment within which it operates. At this level the system can be analysed by assessing: Whether the measures reinforce the firm's strategies; whether the measures match the organization's culture; whether the measures are consistent with the existing recognition and reward structure; whether some measures focus on customer satisfaction; whether some measures focus on what the competition is doing.
(Kaplan & Norton, 1996; 2001; 2004)	A balanced scorecard is a comprehensive set of performance measures defined from four different measurement perspectives (financial, customer, internal, and learning and growth) that provides a framework for translating business strategy into operational terms.
(Atkinson, Waterhouse & Wells, 1997)	The performance measurement system is the tool the company uses to monitor contractual relationships.
(Bititci, Carrie & Mcdevitt, 1997)	A performance measurement system is the information system which is at the heart of the performance management process and it is of critical importance to the effective and efficient functioning of the performance management system.

Study	Definition
(Atkinson, 1998)	The give and take between the organisation and its critical stakeholders will define the organisation's secondary objectives, which are the focus of the strategic performance measurement system. This is the critical attribute of a strategic performance measurement system; namely that it focuses attention on what planners believe is critical to promote the organisation's success.
(Neely, 1998)	A performance measurement system enables informed decisions to be made and actions to be taken because it quantifies the efficiency and effectiveness of past actions through the acquisition, collation, sorting, analysis, interpretation, and dissemination of appropriate data. [...]
(De Haas & Kleingeld, 1999)	A set of performance indicators with procedures for periodic data gathering and the group of organizational actors they relate to, form the elements of a PM system.
(Gates, 1999)	A strategic performance measurement system translates business strategies into deliverable results. Combine financial, strategic and operating measures to gauge how well a company meets its targets
(Otley, 1999)	A performance measurement system is a system that provides the information that is intended to be useful to managers in performing their jobs and to assist organizations in developing and maintaining viable patterns of behaviour. Any assessment of the role of such information requires consideration of how managers make use of the information being provided to them. Main components of a PMS: (1) objectives, (2) strategy, (3) targets, (4) rewards, (5) information flows (feedback and feed-forward).
(Forza & Salvador, 2000)	On one hand a business performance measurement system feeds forward the various process owners with goals have to meet, on the other it gives them feedback on the outcome of their activities, and therefore on their progress towards the goals set.
(Simons, 2000)	This book focuses on performance measurement and control systems, which are the formal, information-based routines and procedures managers use to maintain or alter patterns in organizational activities. Performance measurement systems assist managers in tracking the implementation of business strategy by comparing actual results against strategic goals and objectives. A performance measurement system typically comprises systematic methods of setting business goals together with periodic feedback reports that indicate progress against those goals.
(Maisel, 2001)	A business performance measurement system enables an enterprise to plan, measure, and control its performance and helps ensure that sales and marketing initiatives, operating practices, information technology resources, business decision, and people's activities are aligned with business strategies to achieve desired business results and create shareholder value
(Bourne, Neely, Mills & Platts, 2003)	A business performance measurement system refers to the use of a multi-dimensional set of performance measures for the planning and management of a business.

Study	Definition
(Ittner, Larcker & Randall, 2003)	A strategic performance measurement system: (1) provides information that allows the firm to identify the strategies offering the highest potential for achieving the firm's objectives, and (2) aligns management processes, such as target setting, decision-making, and performance evaluation, with the achievement of the chosen strategic objectives.
(Kerssens-van Drongelen & Fisscher; 2003)	Business performance measurement and reporting takes place at 2 levels: (1) company as a whole, reporting to external stakeholders, (2) within the company, between managers and their subordinates. At both levels there are 3 types of actors: (a) evaluators (e.g. managers, external stakeholders), (b) evaluate (e.g. middle managers, company), (c) assessor, which is the person or institution assessing the effectiveness and efficiency of performance measurement and reporting process and its outputs (e.g. controllers, external accountant audits).
(Cavalluzzo & Ittner, 2004)	"We refer to performance measure development or performance measurement systems more generally as a collection of performance metrics that are reported on a regular basis through the organization's information systems."
(Franco et al., 2004)	"The set of processes an organization uses to manage its strategy implementation, communicate its position and progress, and influence its employees' behaviours and actions. It requires the identification of strategic objectives, multidimentional performance measures, targets and the development of a supporting infrastructure."
(Chenhall, 2005)	"A distinctive feature of these strategic performance measurement systems (SPMS) is that they are designed to present managers with financial and non-financial measures covering different perspectives which, in combination, provide a way of translating strategy into a coherent set of performance measures...This system of associated measures has the potential to identify the cause-effect linkages that describe the way operations are related to the organization's strategy. The aim is to provide a rational framework to formulate and implement strategies."
(GAO, 2005)	Performance measurement is the ongoing monitoring and reporting of program accomplishments, particularly progress towards preestablished goals. It is typically conducted by program or agency management. Performance measures may address the type or level of program activities conducted (process), the direct products and services delivered by a program (outputs), and/or the results of those products and services (outcomes). A "program" may be any activity, project, function, or policy that has an identifiable purpose or set of objectives.
(Tuomela, 2005)	Performance measurement systems are collections of financial and/or non-financial performance indicators that managers use to evaluate their own or their unit's performance or the performance of their subordinates... In strategic performance measurement systems, performance indicators describe either the critical success factors with regard to strategy implementation or to the outcomes that the strategy is expected to yield.

2.3 What is Product Innovation?

Innovation has been defined in many different ways (see Table 2-2). However, common to all of these definitions is the concept of newness, which distinguishes innovations from other types of change (Knight, 1967). Beyond agreeing on that innovations entail some newness, the definitions differ in many aspects. Based on a review of innovation research in economics, organizational sociology, and technology management, Gopalakrishnan & Damanpour (1997) argue that researchers should define innovations in terms of three dimensions: stages of the innovation process, levels of analysis, and types of innovation. Garcia & Calantone (2002) adds innovativeness as a fourth dimension based on their review of innovation research in marketing, management, and engineering.

Next, before advancing the definition, I adopted for product innovation and how does it address the proposed dimensions, I briefly define the dimensions. First, innovation may refer to the generation or the adoption process. The generation of innovation refers to problem solving and decision-making involved in the development of new products and services whilst the adoption of innovation refers to the process of accepting innovation by the adopting organization (Gopalakrishnan & Damanpour, 1997).

Second, innovations could be studied at different levels of analysis: national, industry, organizational, sub-unit, and individual. This dimension attempts to answer the question of which unit of generation or adoption is newness defined?

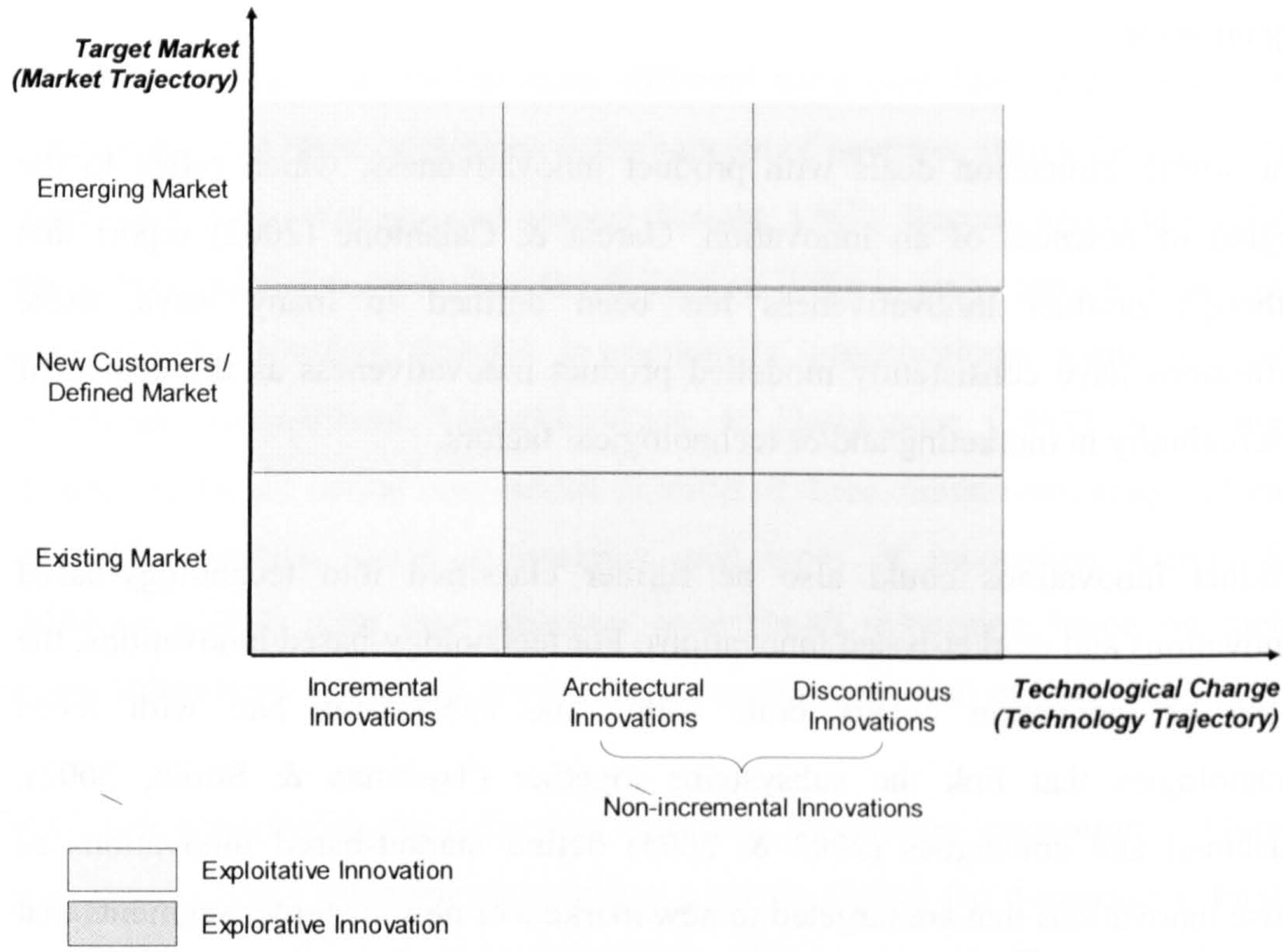
The third dimension pertains to the nature of the innovation outcome. Innovations could be technical or administrative. Damanpour (1991) defines technical innovations as involving products, services, and production process technologies. Technical innovations can be further classified into product or process innovations. Administrative innovations involve organizational structures and administrative processes. He adds that technical innovations are related directly to the basic work activities of the organization whereas administrative innovations are related to the management of the basic activities of the organization.

The fourth dimension deals with product innovativeness, which refers to the degree of newness of an innovation. Garcia & Calantone (2002) report that although product innovativeness has been defined in many ways, these definitions have consistently modelled product innovativeness as the degree of discontinuity in marketing and/or technological factors.

Product innovations could also be further classified into technology-based innovations and market-based innovations. For technology-based innovations, the locus of innovation occurs both within the subsystems and with those technologies that link the subsystems together (Tushman & Smith, 2002). Tushman and colleagues (2002 & 2004) define market-based innovations as those innovations that are targeted to new markets or new customer segments that are often technically simple products are often missed by incumbents.

There are a number of innovation frameworks that capture technological and market innovations such as the product innovation grid by Meyer & Roberts (1986). Tushman and colleagues (O'Rielly & Tushman, 2004; Tushman & Smith, 2002; Tushman, Smith, Wood, Westerman, & O'Reilly, 2004) build on the innovation grid and advance a framework they call innovation space. It is based on the two dimensions of technology and market where the origin is the focal firm's existing product/market choices as illustrated in Figure 2-3. These two dimension or trajectories could also be conceptualised as technology and market risk, respectively.

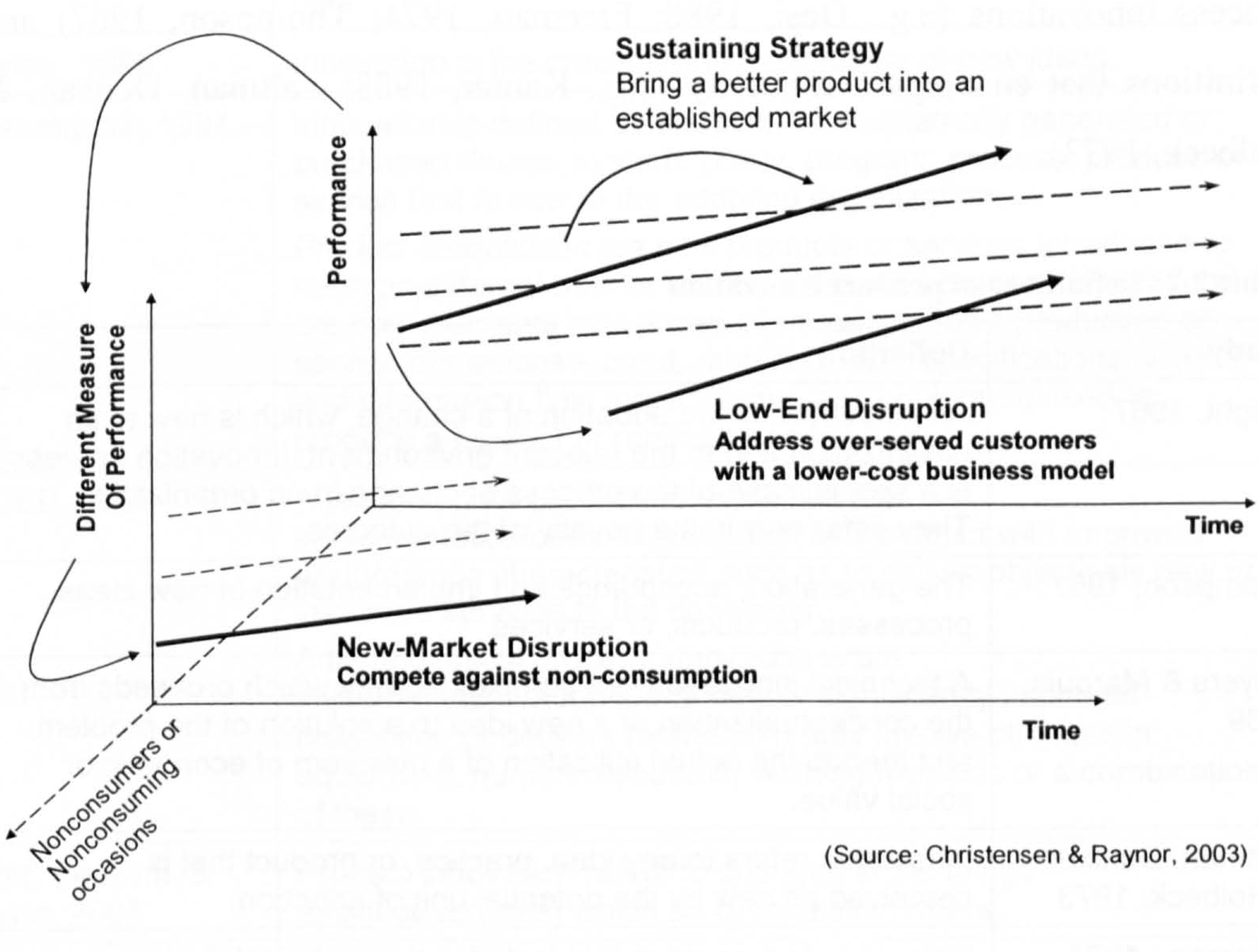
Figure 2-3: Innovation space (Tushman, Smith, Wood, Westerman, & O'Reilly, 2004)



Based on the locus of innovation, product innovations could be classified into three types of innovations: incremental, architectural, and discontinuous innovations. Incremental innovations improve price/performance along current technological trajectory for existing subsystems and linking mechanisms (Gatignon, Tushman, Smith, & Anderson, 2002; Hollander, 1965; Myers & Marquis, 1969; Tushman & Smith, 2002). Architectural innovations modify the linkages between the subsystems, add, or subtract subsystems (Baldwin & Clark, 2000; Henderson & Clark, 1990). Discontinuous innovations involve major technical change in the core subsystems and disrupt the existing technological trajectory because they trigger cascading changes in other less core subsystems and linking mechanisms (Ahuja & Lambert, 2001; Dosi 1982; Gatignon et al. 2002; Tushman & Smith, 2002; Tushman, et al., 2004). Architectural and discontinuous innovations have the severest impact on current industry players and they are referred to as non-incremental innovations (Tushman et al., 2004).

Tushman and colleagues (2002 & 2004) define market-based innovations as those innovations that are targeted to new markets or new customer segments that are often technically simple products that are often missed by incumbents (See Figure 2-4). Christensen (1997) and Christensen & Raynor, (2003) further differentiate market-based innovation based on their disruption of the incumbent's organizational architecture. Christensen & Raynor (2003) define two types of market-based innovations: sustaining innovation and disruptive innovation[4]. The sustaining innovation brings better products into an established market. The disruptive innovations are classified into two types: low-end disruption and new-market disruption. The former addresses over-served customers with a lower cost business model and the latter competes against non-consumption as illustrated in Figure 2-4.

Figure 2-4: Market-based Innovations (Source: Christensen & Raynor, 2003)



[4] Christensen (1997) initially labelled one type of market-based innovations as disruptive technologies then he relabelled them as disruptive innovations (Christensen & Raynor, 2003). These innovations are not technologically disruptive but they are disruptive to an incumbent's existing organizational architecture.

In this study, I adopt the definition advanced by DTI (2005a) which addresses the four dimensions suggested above. Product innovation refers to the market introduction of a *new* good or service or a *significantly* improved good or service. The innovation must be new to the organization, but it does not need to be new to the market and it does not matter if the innovation was originally developed by the organization or by other organizations. In this definition, innovation refers to product innovations but it does not distinguish between the different levels of innovativeness. It also corresponds to the generation process and organizational level of analysis. Innovation generation is defined here from the perspective of the innovating organization.

As shown in Table 2-2, innovation has been defined in many ways. The outcome-based definition adopted in this study is broad enough to encompass the different types of product innovations but it is also limited enough to differentiate it from process innovations (e.g., Dosi, 1988; Freeman, 1974; Thompson, 1967) and definitions that encompass creativity (e.g., Kanter, 1988; Zaltman, Duncan, & Holbeck, 1973).

Table 2-2: Definitions of product innovation

Study	Definition
Knight, 1967	An innovation is the adoption of a change, which is new to an organization and to the relevant environment. Innovation process is a special case of the process of change in an organization. They differ only in the novelty of the outcome.
Thompson, 1967	The generation, acceptance and implementation of new ideas, processes, products, or services.
Meyers & Marquis, 1969	A technical innovation is a complex activity which proceeds from the conceptualization of a new idea to a solution of the problem and then to the actual utilization of a new item of economic or social value.
Zaltman, Duncan, & Holbeck, 1973	Innovation refers to any idea, practice, or product that is perceived as new by the potential unit of adoption.
Freeman , 1974	Innovation is a process that includes the technical design, manufacturing management and commercial activities involved in the marketing of a new (or improved) product or the first use of a new (or improved) manufacturing process or equipment.
Utterback & Abernethy, 1975	Product innovations are new outputs or services that are introduced for the benefit of customers or clients.
Downs & Mohr,	Innovation refers to the adoption of means or ends that are new

Study	Definition
1976	to the adopting unit.
Nelson & Winter, 1982	Innovation is produced by a firm for sale to customers who will use it.
Van de Ven, 1986	Innovation is a new idea, which may be a recombination of old ides, a scheme that challenges the present order, a formula, or a unique approach which is perceived as new by the individuals involved. The process of innovation is defined as the development and implementation of new ideas by people who over time engage in transactions with others within as institutional context.
Amabile, 1988	Creativity is the production of novel and useful ideas by an individual or small group of individuals working together. Organizational innovation is the successful implementation of creative ideas within an organization.
Burgelman & Sayles, 1988	Innovation refers to a company's efforts in instituting new methods of production and/or bringing new products or services to market.
Dosi, 1988	Innovation concerns the search for, and the discovery of, experimentation, development, imitation, and adoption of new products, new production processes and new organisational set-ups.
Kanter, 1988	Innovation is the creation and exploitation of new ideas.
Damanpour, 1991	Innovation is defined as adoption of an internally generated or purchased device, system, policy, program, process, product, or service that is new to the adopting organization. *Product innovations* are new products or services introduced to meet an external user or market need, and *process innovations* are new elements introduced into organization's production or service operations – input, materials, task specifications, work and information flow mechanisms, and equipment used to produce a product or render a service.
OCED, 1997	A technological product innovation is the implementation/commercialisation of a product with improved performance characteristics such as to deliver objectively new or improved services to the consumer. A technological process innovation is the implementation/adoption of new or significantly improved production or delivery methods. It may involve changes in equipment, human resources, working methods or a combination of these.
Tidd, Bessant, & Pavitt, 2005	Product innovation refers to changes in the things (product/services) which an organization offers.
Trott, 2002	Innovation is the management of all the activities involved in the process of idea generation, technology development, manufacturing and marketing of a new (or improved) product or manufacturing process or equipment. Innovation = theoretical conception + technical invention + commercial exploitation
Burgelman,	Innovations are the outcome of the innovation process, which

Study	Definition
Christensen, & Wheelwright, 2004	can be defined as the combined activities leading to new, marketable products and services and/or new production and delivery systems.
Afuah, 2003	Innovation is the use of new knowledge to offer a new product or service that customers want. It is invention + commercialization.

2.4 What is the impact of Measurement on Innovation?

Innovation is a broad topic that has been studied by different disciplines including marketing, operations management, strategic management, technology management, organizational behaviour, and economics (Hauser, Tellis, & Griffin, 2005). There have been many recent literature reviews on innovation management as follows:

- Historical (e.g., Marinova & Phillimore, 2003)
- Disciplinary - technology management (e.g., Dodgson, 2000; Tidd, Bessant, Pavitt, 2005); marketing (e.g., Hauser, Tellis, & Griffin 2005); and sociology (e.g., Hage, 1999)
- Topical - innovation management measurement (e.g., Adams, Bessant, & Phelps, 2006); product development (e.g., Brown & Eisenhardt, 1995); product development decisions (e.g., Krishnan & Ulrich, 2001); and innovation search strategies (e.g., Katila, 2000; Mehdi, 2002)
- Theoretical - addressing only one theoretical perspective - network (e.g., Pittaway, Robertson, Munir, Denyer, & Neely, 2004) and organizational learning (e.g., Yaghi, 2005)
- Theoretical - addressing more than one theoretical perspective (e.g., Galende, 2006; Lam, 2005)

- Innovation research streams in general (e.g., Downs & Mohr, 1976; Drazin & Schoonhoven, 1996; Fiol, 1996; Slappendel, 1996; Wolfe, 1994).

To inform the research question of how performance measurement affects innovation, I review how measurement may constrain or help innovation by examining the measurement role in innovation models based on different theoretical perspectives. Although this approach is challenging, it ensures that the research is grounded in theory. The challenge in conducting this review stems from two issues.

First, there have been few reviews based on the theoretical foundations of innovation models (For exceptions, refer to Galende, 2006 and Lam, 2005) and none of these reviews examined the role of measurement in innovation.

Second, there have been at least 14 theoretical perspectives through which innovation was examined: evolutionary economics (e.g., Savotti, 1996); organizational learning (e.g., Greve, 2003b); transaction cost economics (e.g., Robertson & Gatignon, 1998); resource-based view (e.g., Verona, 1999); knowledge-based view (e.g., Gopalakrishnan & Bierly, 2001); dynamic capabilities (e.g., Tushman et al., 2004); contingency (e.g., Zaltman, Duncan, & Holebeck, 1973); network (e.g., Tsai & Ghoshal, 1998); industrial organizations (e.g., Roberts, 1999); agency theory (e.g., Hoskisson & Hitt, 1988); institutional (e.g., Doughtery & Heller, 1994); organizational ecology (e.g., Burgelman, 1983, 2002); rational perspective (e.g., Rothwell, 1992); and complexity perspective (e.g., Frenken, 2006).

In this study, I limit my review to innovation models[5] from three theoretical perspectives that assign a major role to measurement in the innovation process: organizational learning from performance feedback, contingency theory and

[5] In this study, I follow Lave & March (1975) in not differentiating between theories and models. Therefore, I will use these terms interchangeably in this thesis.

organizational creativity. I also review the practice-based innovation models because they are widely used in the technology and innovation management literature although they tend to ignore or minimize the role of measurement in innovation. Since I adopt the organizational learning perspective in this study, I review the innovation model in the next chapter instead of this section.

2.4.1 Contingency-based Innovation Models

Contingency theory began in early 1960s as an offshoot of systems theory in response to administrative theorists seeking one best way to organize; for example, the bureaucratic organization by Weber (1947) and M-form organization by Chandler (1962). Galbraith (1973) stated two underlying assumptions of contingency theory: (1) There is no one best way to organize and (2) Any way of organizing is not equally effective. Scott (1981) added a third assumption: "The best way to organize depends on the nature of the environment to which the organization relates." As the contingency perspective gained currency, the contingency variables used by researcher expanded from environmental contingencies (Burns & Stalk, 1961; Lawrence & Lorsch, 1967) to include among others: strategy (Chandler, 1977) and technology (Woodward, 1958). One of the early dependent variables investigated by contingency work was innovation (Burns & Stalker, 1961).

Studies investigating the impact of measurement on innovation using the contingency perspective could be grouped into two streams. One stream consists mainly of organizational and innovation management researchers and the second stream consists mainly of management accounting and performance measurement researchers.

One of the major differences between the two streams is how they conceptualise and operationalise performance measurement. Researchers in the first stream tend to use the formalization concept, which is broader than the performance measurement concept in their studies whereas researchers in the second stream tend to be more specific in their conceptualisation and operationalisation of performance measurement. Formalization usually refers to the controls organizations employ that consist of codified body of rules, procedures or behaviour prescriptions used in conducting organizational activities (Pierce & Delbecq, 1976). Both research streams develop structural contingency theories of innovations that aim to specify organizational design characteristics that lead to innovation. Damanpour & Gopalakrishan (1998) note that these theories of innovation can be grouped into two sets: uni-dimensional and middle range theories.

The first research stream tends to develop Unitarian and middle range theories of organizational innovation. Uni-dimensional theories of innovation postulate relationships between structural variables and innovation. For example, Burns and Stalker (1961) in their study of the British & Scottish electronics industry after World War II found that firms could be grouped based on the conditions of relative stability and change of their environments into two polar types: mechanistic or organic. Mechanistic organizations emphasized production. They were rigid and hierarchical and were suitable to stable environments. On the other hand, organic organizations emphasized adaptation. They were flexible and flatter and were more suitable to conditions of rapid change and innovation. Burns & Stalker (1961) saw formalization process as being characteristic of mechanistic organizations and therefore incompatible with innovation in environments characterized by rapid change. Pierce & Delbecq (1976) note that high formalization was generally postulated as incompatible with innovation because of its emphasis on predictability of performance whereas low formalization was more compatible with innovation because it permitted openness which is a pre-condition for generating ideas that may eventually translate into innovations.

However, uni-dimensional theories of innovation yielded inconsistent results (Downs & Mohr, 1976) which prompted innovation scholars to develop middle-range theories to account for these consistencies. Damanpour & Gopalakrishan (1998) note that middle-range theories of innovations were based on three distinctions: (1) the distinction between types of innovation (Daft, 1978); (2) the distinction between radicalness of innovation (Dewar & Dutton, 1986); and (3) the distinction between stages of innovation process (Duncan, 1976). The first two theories focus on the output of the innovation and divide it into two types and the third theory focuses on the process of innovation and divides it into two stages.

Daft (1978) advanced the first middle-range theory reviewed here and he called it dual-core model of organizational innovation. The theory distinguishes between technological and administrative innovations. Daft (1978) postulated that these two innovations follow different innovation processes which are in turn managed in two different organizational cores: technical and administrative. Technical core is mainly concerned with technical innovations, which typically follow a bottom up process whereas the administrative core is mainly concerned with administrative innovations, which typically follow a top down process. The dual-core theory suggest that formalization will be positively related to administrative innovations because it facilitates the top down process of innovation and negatively related to technical innovations because it does not facilitate the bottom up process of innovation.

The second middle-range theory distinguishes between two types of innovation based on the extent of their novelty. Damanpour & Gopalakrishan (1998) report that researchers offered inconsistent hypotheses regarding the impact of formalization on innovation in the theory of innovation radicalness. Therefore, they proposed similar prediction to the dual-core theory of innovation where it is hypothesized that formalization will be associated with high incremental innovation and with low radical innovation.

The third middle-range theory is the ambidextrous theory of innovation (Duncan, 1976; Zaltman, Duncan, & Holebeck, 1973), which distinguishes between two stages of innovation: initiation and implementation. The initiation stage involves the activities pertaining to the knowledge awareness of innovation, formation of attitudes toward innovation and decision to adopt the innovation. The implementation sage involves the activities pertaining to initial implementation and continued-sustained implementation. The ambidextrous theory of innovation postulates that the initiation stage requires an organization structure that is characterized by high complexity, lower formalization, and lower centralization whereas the implementation stage requires an organization structure that is characterized by low complexity, higher formalization, and higher centralization. The predictions of the impact of formalization on innovation as advanced by innovation scholars belonging to the first stream are summarized in Table 2-3. However, Damanpour (1991) reports in his meta-analysis studies that none of the predicted relationship was statistically significant as shown in Table 2-3.

Table 2-3: Formalization-innovation relationships in theories of organizational innovation (Source: Damanpour, 1991; Damanpour & Gopalakrishnan, 1998)

	Uni-dimensional Theories	Middle-range Theories					
		Dual-core Theory		Radicalness Theory		Ambidextrous Theory	
Structural Variable	Organizational Innovation	Administ. Innov.	Technical Innov.	Incremental Innov.	Radical Innov.	Initiation	Implementation
Formalization	-	High	Low	High	Low	Low	High
Meta-analysis Results		-0.07	0.04	-0.04	-0.10	-0.11	-.0.01

The second research stream tends to develop Unitarian theories of organizational innovation. They developed models to examine the impact of management control systems (MCS) on innovation using different units of analysis: R&D department, new product development project/team and organization. However, except for few studies that used organizations as their unit of analysis, most of the innovation models used R&D and project development as their units of analysis (Bisbe & Otley, 2004).

Studies examining the role of measurement in R&D departments operationalised management control systems in two ways: narrow (financial) and broad (multiple controls/financial & non-financial).

Studies that examined the roles of financial measures in R&D departments (Brownell, 1985; Hayes, 1977; Rockness & Shields, 1988) revealed that financial measures do not have an important role in R&D departments other than signalling the commitment of the organization to its R&D efforts (Davila, 2000). For example, Rockness & Shields (1988: 571) state that the perceived importance of budgets in R&D departments "decreases monotonically from planning to monitoring, monitoring to evaluating, and evaluating to rewarding". Brownell (1985) reports that "accounting information is used in R&D more as a measure of the organization's financial commitments to the activity than as a surrogate for planned or targeted output."

The evidence from the studies that have adopted a broader view of management control systems (Abernethy & Brownell, 1997; Kamm, 1980; Rockness & Shields, 1984) suggests that management control systems have, at most, a minor role in innovation. For example, Abernethy & Brownell (1997: 245) found that "reliance on accounting controls has significant positive effects on performance only where task uncertainty is lowest" while "behaviour controls appear to contribute to performance in no situation". However, Nixon (1998) and Godener & Soderquist (2004), in contrast to the other studies report that performance measurement has a positive effect. Nixon (1998) notes that "notwithstanding the practical difficulties and limitations of R& D evaluation, the use of financial appraisal methods can assist communication, build consensus and provide a context which decision-makers can use to evaluate the more subjective dimensions of projects." Godener & Soderquist (2004) found that performance measurement systems are used at two levels: the project level and the functional level. At these levels, they identified five uses of the results of performance measurement and two impacts on the organization and behaviour of the company. They found that managers use the measures at the project level to: (1) reorient projects before failure or cancelling product platforms and families; (2) decide corrective actions at the right moment in running projects; and (3) select promising and relevant project for launching. At the functional level, managers used measures to: (1) ensure better coherence and relevance of product portfolios and (2) improve processes and adherence to processes. These uses resulted in enhancing the research and new product development (R&NPD) staff motivation at the project level and facilitating well-balanced arbitration and decision-making at the functional level.

At the new product development project/team level of analysis, the studies yielded inconsistent findings as well. Ditillo (2004) notes that the literature on product development suggests that when management control systems provide information directed to coordination and learning, they affect performance in a positive way (Koga and Davila, 1998 and Nixon, 1998). However, alternative arguments and evidence (Eisenhardt & Tabrizi, 1995) propose that such a relationship does not exist or is negative. He explains that Davila (2000) tried to explain these inconsistencies by suggesting that these contradictory results might be the result of a different interpretation of the role of management control systems that should be considered as information tools to face uncertainty rather than control mechanisms to reduce goal divergence.

At the organizational level of analysis, the studies yielded inconsistent findings as well. Khandwalla (1972) finds that reliance on formal control systems increases with the intensity of competition. Similarly, Simons (1987) reports that high performing prospectors rely on the information provided by frequently updated formal control systems to drive organizational learning. Contrary to other studies, Bisbe & Otley (2004) do not find support to the proposition that the more interactive the use of MCS by top managers, the higher the product innovation. The findings of their study suggest that interactive use may favour innovation in low-innovating firms, while the effect appears to be in the opposite direction in high-innovating firms.

2.4.2 Creativity-based Innovation Models

Organizational creativity literature makes a distinction between the process of creativity and the process of innovation. Creativity is defined as the production of novel and useful ideas in any domain and innovation is defined as the successful implementation of creative ideas within an organization (Amabile, 1988). Amabile offers two criteria for judging creativity, "[a] product or response will be judged as creative to the extent that (a) it is both a novel and appropriate, useful, correct or valuable response to the task at hand, and (b) the task is heuristic rather than algorithmic[6]" (1996: 35). Kanter (1988) notes that creativity is associated mainly with individuals whilst innovations are associated mainly with organizations.

Sundgren, Dimenäs, Gustafsson, & Selart (2005) note that many empirical studies established a link between creative environments and innovations (e.g. Abbey & Dickson, 1983; Bommer & Jalajas, 2002; Ekvall, 1996, 1997). For example, Bharadwaj & Menon (2000) found that higher levels of individual and organizational creativity lead to higher levels of innovations. Although creativity and innovation are separate constructs, they are highly linked. Hence, the impact of measurement on creativity translates to an impact on innovation[7].

Amabile (1983a; 1988; 1996; 1998) advanced the componential theory of organizational creativity and innovation that links individual (small-group) creativity and organizational innovation processes in one model. The model consists of a creativity element and an innovation element. As illustrated in Figure 2-5, the creative element consists of three components: domain-relevant skills, creativity-relevant skills, and task motivation.

[6] Algorithmic tasks consist of those tasks for which a known formula or path to the solution exists whereas heuristic problems are those tasks not having a clear or easily identifiable path to a solution.

[7] Woodward, Sawyer, & Griffin (1993) view organizational creativity as a subset of the domain of innovation and view innovation as a subset of the domain of organizational change.

Figure 2-5: Components of creative performance (Source: Amabile, 1986)

Domain-Relevant Skills	Creativity-Relevant Skills	Task Motivation
Includes: ▪ Knowledge about the domain ▪ Technical skills required ▪ Special domain-relevant "talent"	**Includes:** ▪ Appropriate cognitive style ▪ Implicit or explicit knowledge of heuristics for generating novel ideas ▪ Conducive work style	**Includes:** ▪ Attitudes toward the task ▪ Perceptions of own motivation for undertaking the task
Depends On: ▪ Innate cognitive abilities ▪ Innate perceptual and motor skills ▪ Formal and informal education	**Depends On:** ▪ Training ▪ Experience in idea generation ▪ Personality characteristics	**Depends On:** ▪ Initial level of intrinsic motivation toward the task ▪ Presence or absence of salient extrinsic constraints ▪ Individual ability to cognitively minimize extrinsic constraints

The central proposition of the theory is summarized in what Amabile calls "The Intrinsic Motivation Principle of Creativity: People will be most creative when they feel motivated primarily by the interest, enjoyment, satisfaction, and challenge of the work itself – and not by external pressures. That is, people who are intrinsically motivated will be more likely to generate truly creative ideas than people who are extrinsically motivated" (188: 142-143).

To explain the mechanism through which intrinsic motivation may impact idea generation, Amabile notes that "Simon (1967) postulates that the most important function of motivation is the control of attention" and "he proposes that motivation determines which goal hierarchy will be activated at any given time, and suggests that the more intense the motivation to achieve an original goal, the less attention will be paid to aspects of the environment that are irrelevant (or seemingly irrelevant) to achieving that goal" (1988: 143). Therefore, she argues that motivational states influence idea generation by influencing the likelihood that alternative solutions will be explored. External motivation focuses the individuals on external goals rather than on the intrinsic aspects of the task. Thus, individuals who are externally motivated will be more likely single-mindedly pursuing the external goals and will be less creative. Performance measurement is viewed as an external motivational technique that focuses the individual on achieving the goal. Therefore, when performance measurement is employed by organizations, it may create the conditions where the creativity heuristics of exploration, set breaking, and risk taking are least likely to be used (Amabile, 1988, 1986).

In summary, the theory predicts that elements of the work environment will impact individuals' creativity primarily through the intrinsic motivation component of the individual creativity process and it also predicts that the creativity produced by individuals and teams of individuals serves as a primary source for innovation within organizations (See Figure 2-6 and Figure 2-7). Amabile argues that creativity and innovation are fostered by allowing a considerable degree of freedom or autonomy in the conduct of one's work and empirical evidence supports this claim[8] (experimental study- Amabile, 1983a; 1983b; survey-based study - Amabile et al., 1996; qualitative study - Amabile & Gryskiewics, 1987). Therefore, organizational creativity generally attributes a negative role to measurement on creativity (Amabile, 1999).

The investment theory of creativity (Sternberg & Lubart, 1991, 1995) and the interactionist model of organizational creativity (Woodman & Schoenfeldt, 1989; 1990; Woodman et al., 1993), which are elaborations of the componential theory of organizational creativity also attribute a negative role to measurement on creativity.

---

[8] In a recent study, Webster (2006) did not find a relationship between interactive use of performance measurement and creativity at the individual level.

Figure 2-6: The Componential theory of organizational creativity and innovation (Source: Amabile 1997: 53)

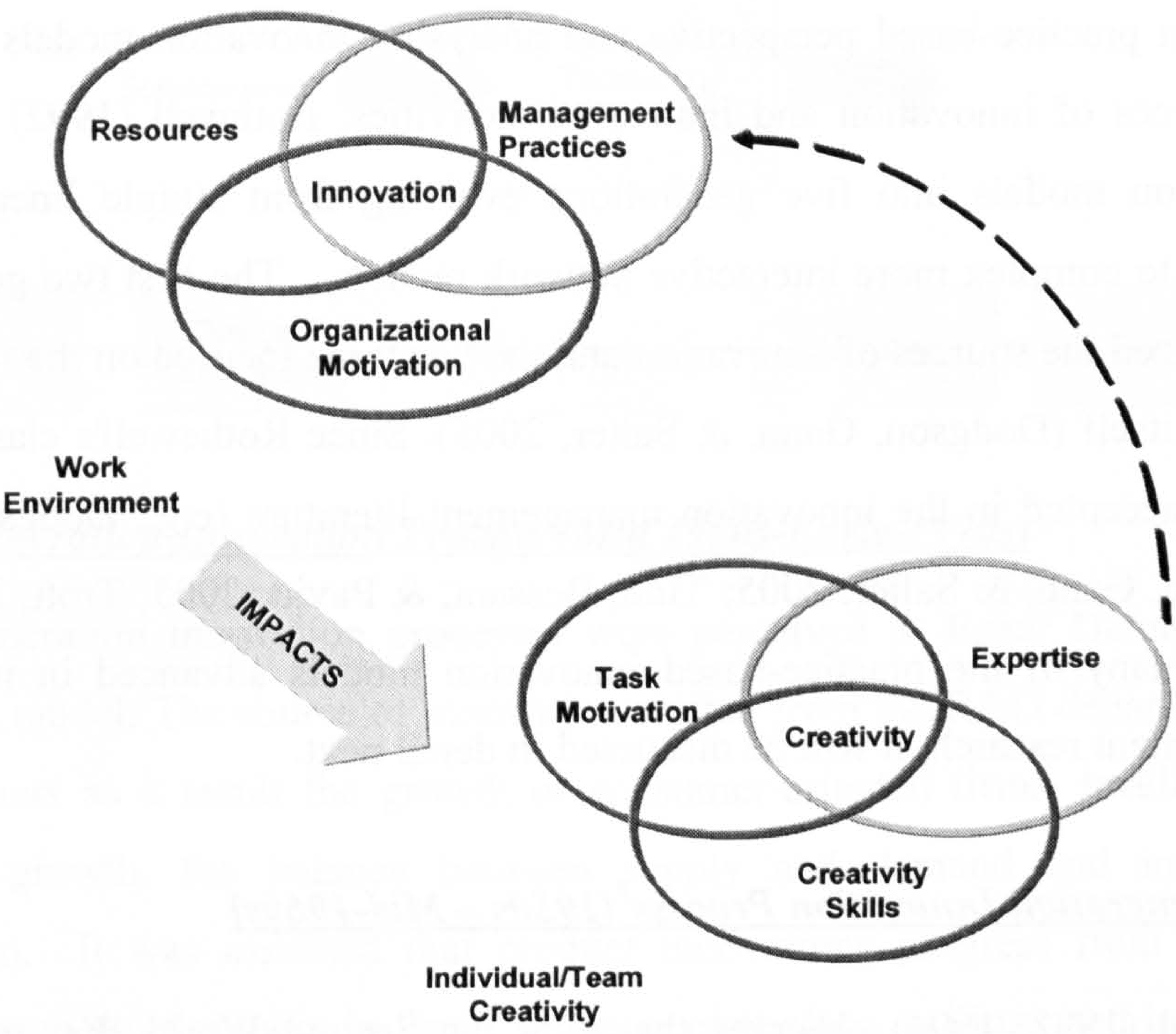
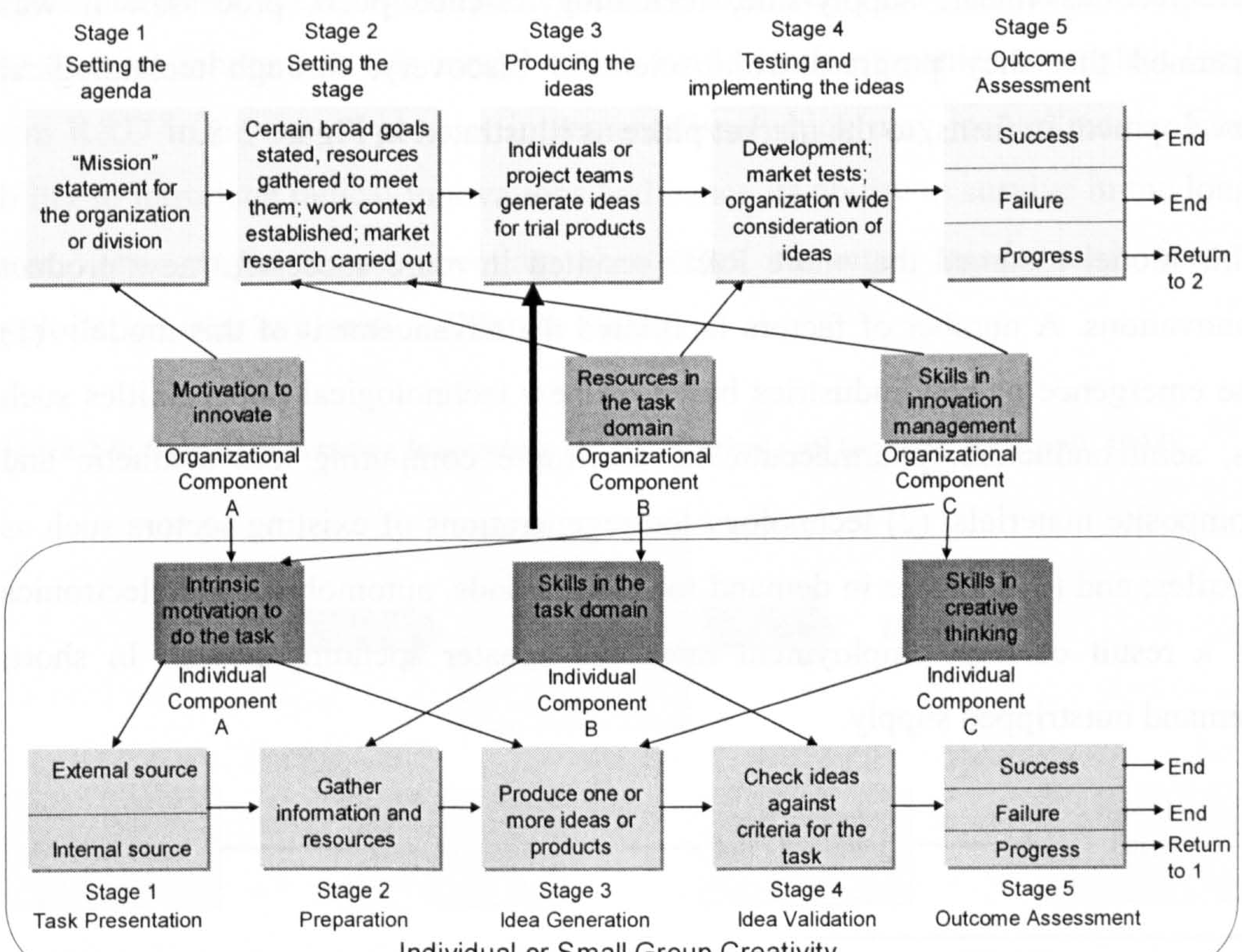


Figure 2-7: Componential model of organizational innovation (Source: Amabile, 1988: 152)



2.4.3 Practice-based Innovation Models

Taking a practice-based perspective and analysing innovation models based on the sources of innovation and innovation activities, Rothwell (1992) classified innovation models into five generations evolving from simple linear models through to complex more interactive network models. The first two generations emphasized the sources of innovation and the last three focused on the innovation process itself (Dodgson, Gann, & Salter, 2005). Since Rothewell's classification is well accepted in the innovation management literature (e.g., Dodgson, 2000; Dodgson, Gann, & Salter, 2005; Tidd, Bessant, & Pavitt, 2005; Trott, 2002) and covers many of the practice-based innovation models advanced in innovation management research, it will be discussed in detail next.

First Generation Innovation Process[9] (1950s – Mid-1960s)

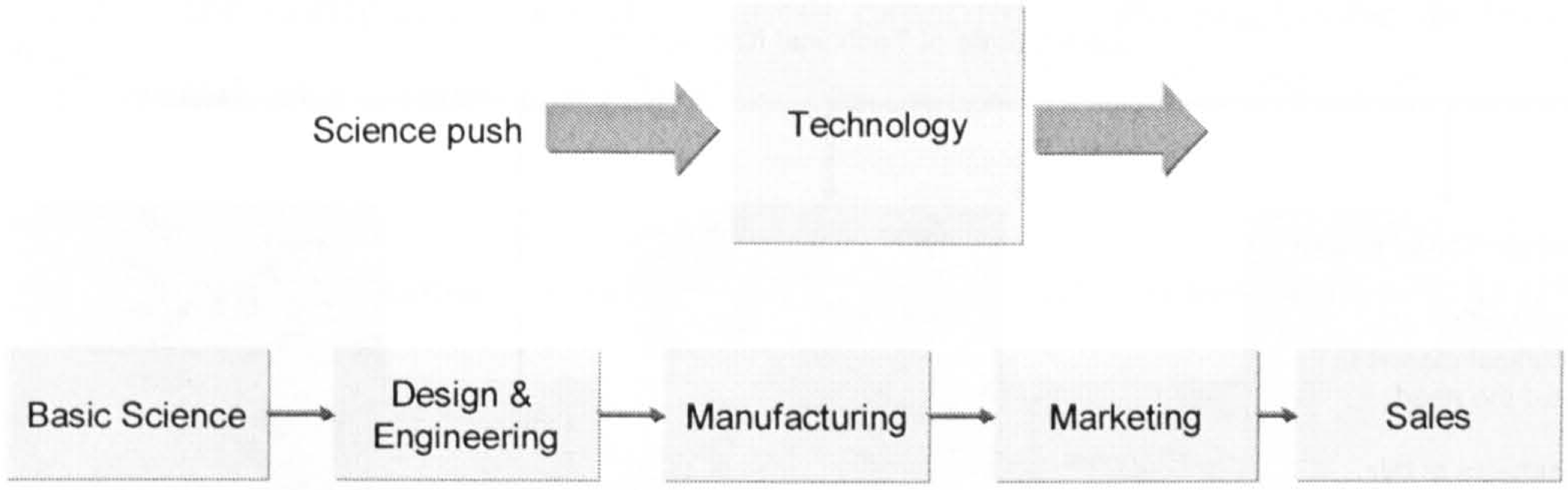
Rothwell (1992; 1994) observes that after the Second World War to the mid 1960s, innovation processes (commercialization of technological change) were perceived as linear, supply-side, technology/science push, processes. It was assumed that they progress from scientific discovery, through technological development in firms, to the market place as illustrated in Figure 2-8.

This model assumed that more R&D resulted in more successful new product innovations. A number of factors facilitated the advancement of this model: (1) the emergence of new industries based on new technological opportunities such as, semiconductors, pharmaceuticals, electronic computing and synthetic and composite materials; (2) technology led regenerations of existing sectors such as textiles; and (3) increase in demand for white goods, automobiles and electronics as a result of high employment rates and greater spending power. In short, demand outstripped supply.

---

[9] Godin (2005a) notes that Bush was the first to advocate this model in 1945 in "Science: The Endless Frontier".

Figure 2-8: First generation innovation model: Technology push (Source: Rothwell, 1994)



Second Generation Innovation Process (Mid 1960s-Early-1970s)

Second generation innovation processes were perceived as linear Demand-pull innovation model. The source of innovation shifted from the R&D department to the customers as a result the growth of consumer-oriented firms, levelling off economic growth, the balance between supply and demand and increased competition. It was assumed that product innovations progress from market needs, through development in firms, to the market place as illustrated in Figure 2-9. The model developed by Utterback (1971) shown in Figure 2-10 is representative of second-generation models. Rothwell (1994) notes that one of the dangers inherent in this model was it could lead companies to neglect long term R&D in favour of product incrementalism which may jeopardize the firm ability to introduce radical innovations and hence its ability to survive in the long term. The emphasis on meeting short-term customer needs may favour exploitation over exploration (March, 1991).

Figure 2-9: Second generation innovation model: Market pull (Source: Rothwell, 1994)

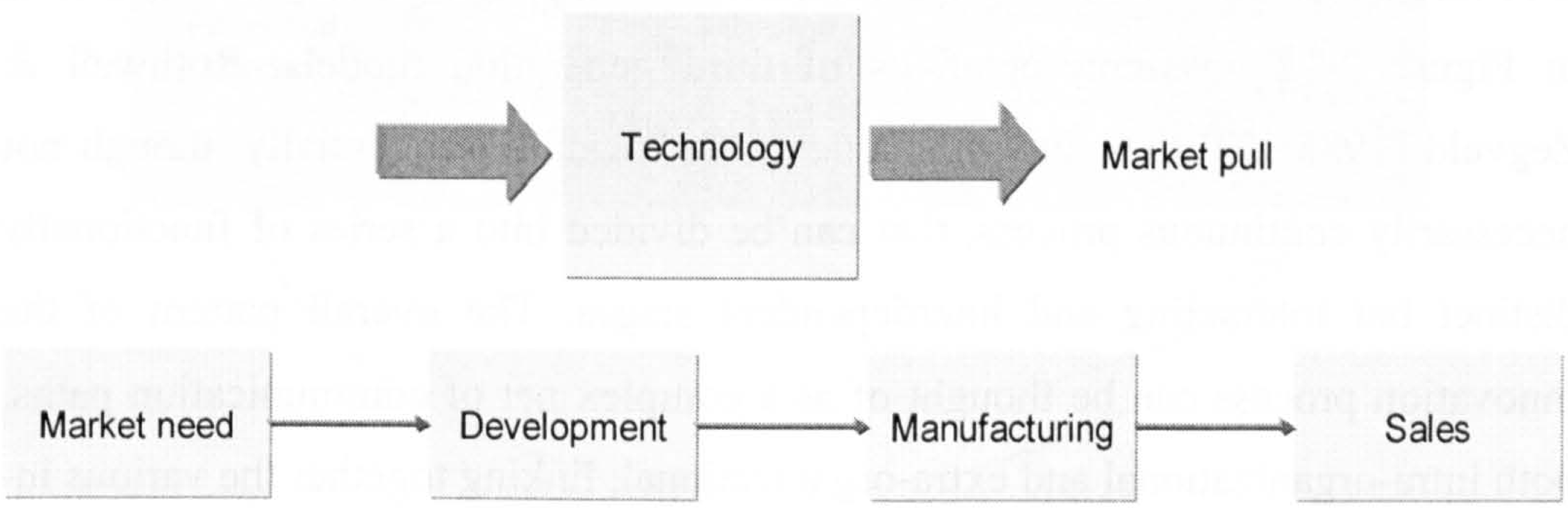
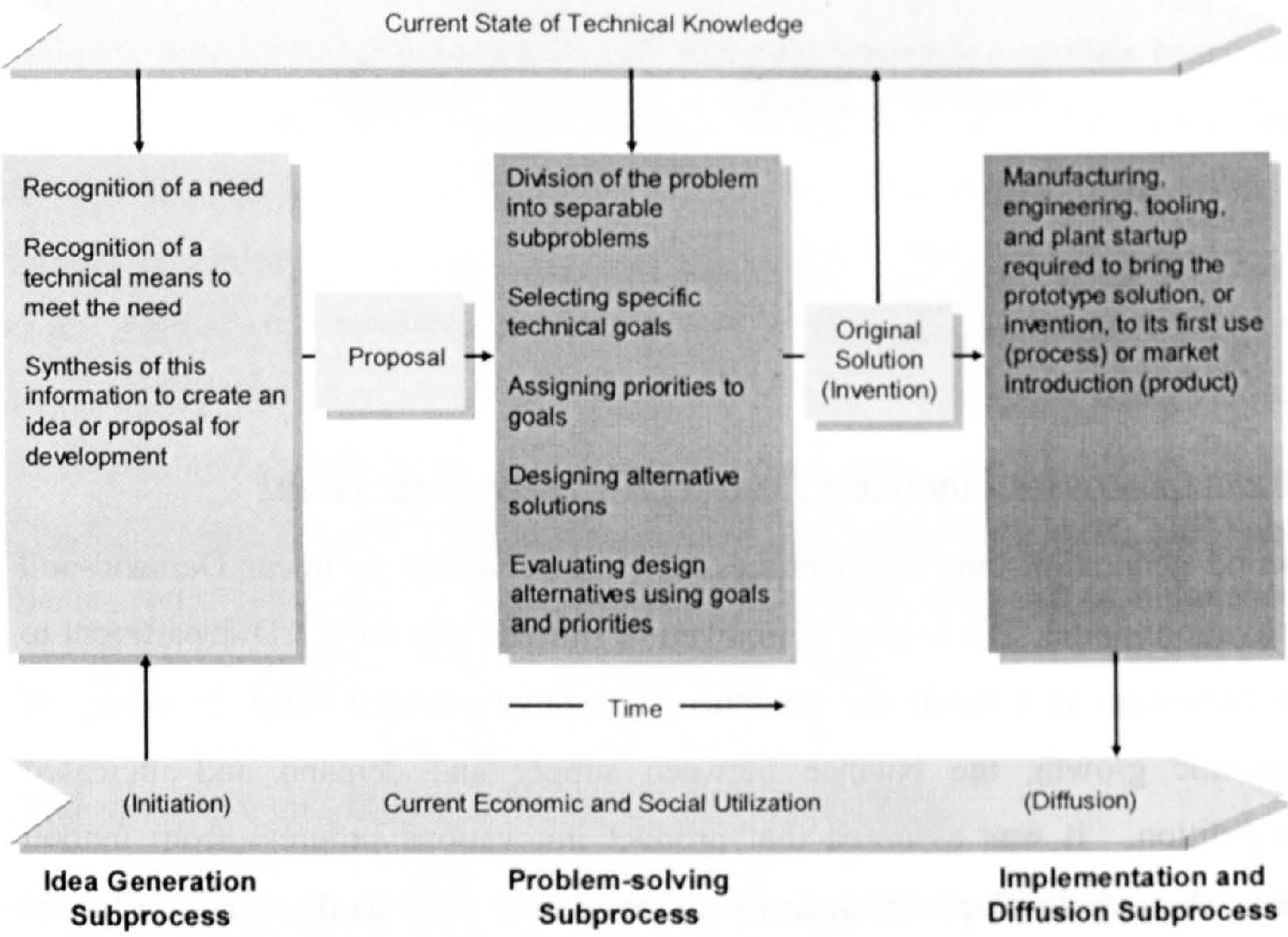


Figure 2-10: The process of technical innovation (Source: Utterback, 1971)



Third Generation Innovation Process (Early 1970s – Mid 1980s)

Third generation innovation models moved from the simple linear models that emphasized the source of innovation into the nonlinear interactive models with feedback loops that emphasized the innovation activities. They combined supply-push and demand-pull models with complex interactions.

The coupling model developed by Rothwell & Zegveld (1985) shown in Figure 2-11 and the chain-linked model developed by Kline & Rosenberg (1986) shown in Figure 2-12 are representatives of third generation models. Rothwell & Zegveld (1985: 50) explain their model as "a logically sequentially, though not necessarily continuous process, that can be divided into a series of functionally distinct but interacting and interdependent stages. The overall pattern of the innovation process can be thought of as a complex net of communication paths, both intra-organizational and extra-organizational, linking together the various in-house functions and linking the firm to the broader scientific and technological community and to the market place. In other words the process of innovation represents the confluence of technological capabilities and market-needs within the framework of the innovating firm"

Figure 2-11:Third generation innovation model: Coupling model (Source: Rothwell, 1992)

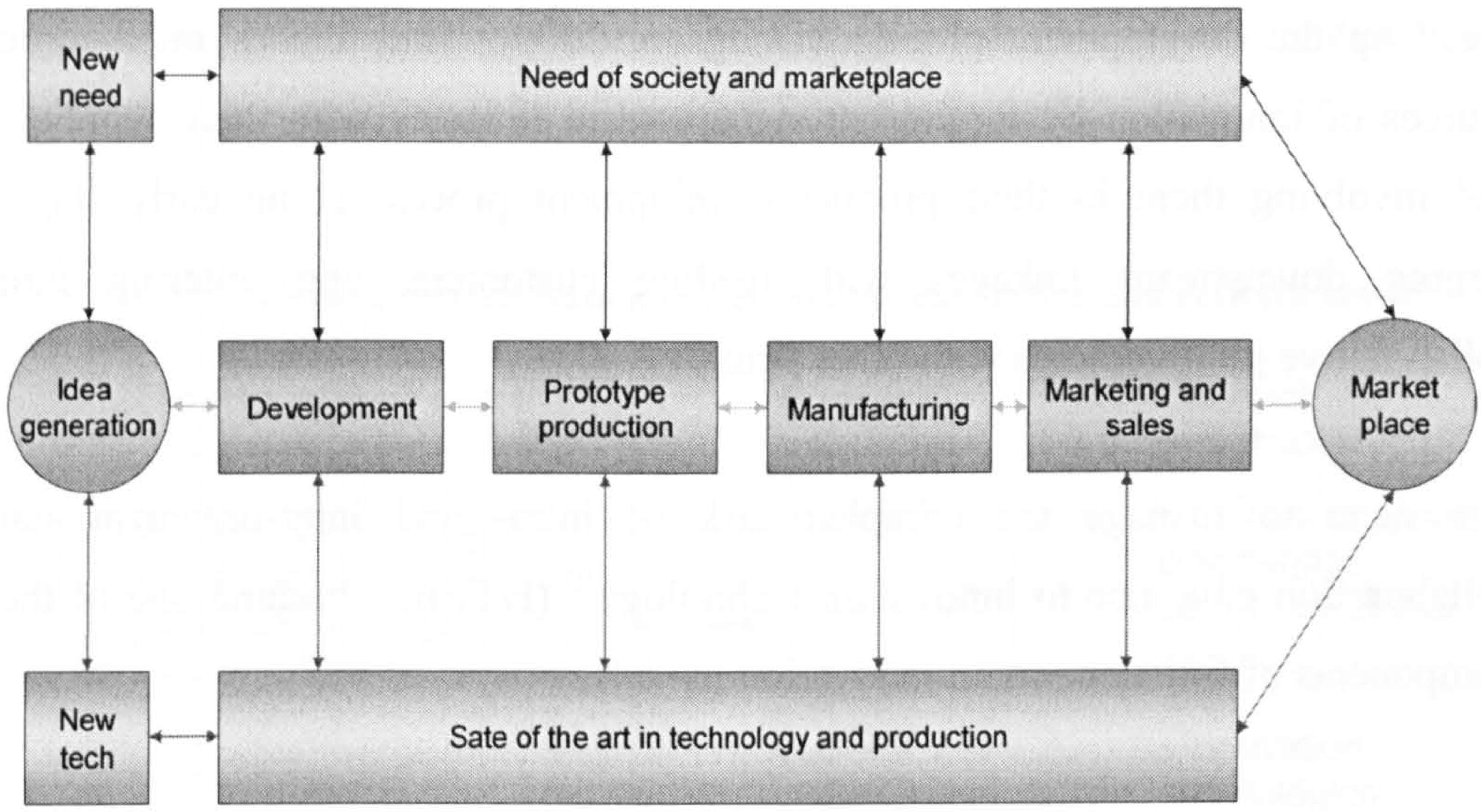
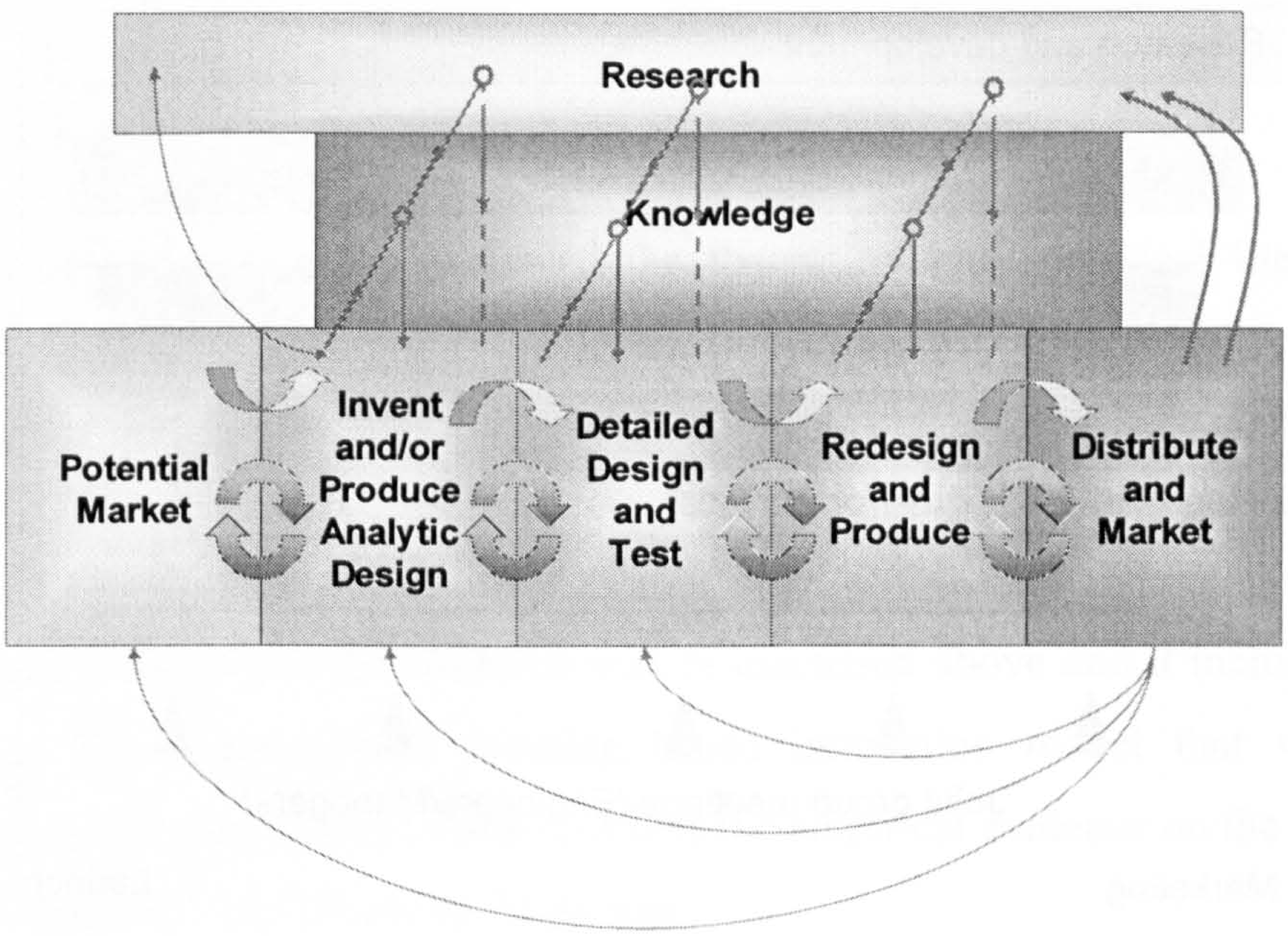


Figure 2-12: Third generation innovation model: Chain-linked model (Source: Kline & Rosenberg, 1986)

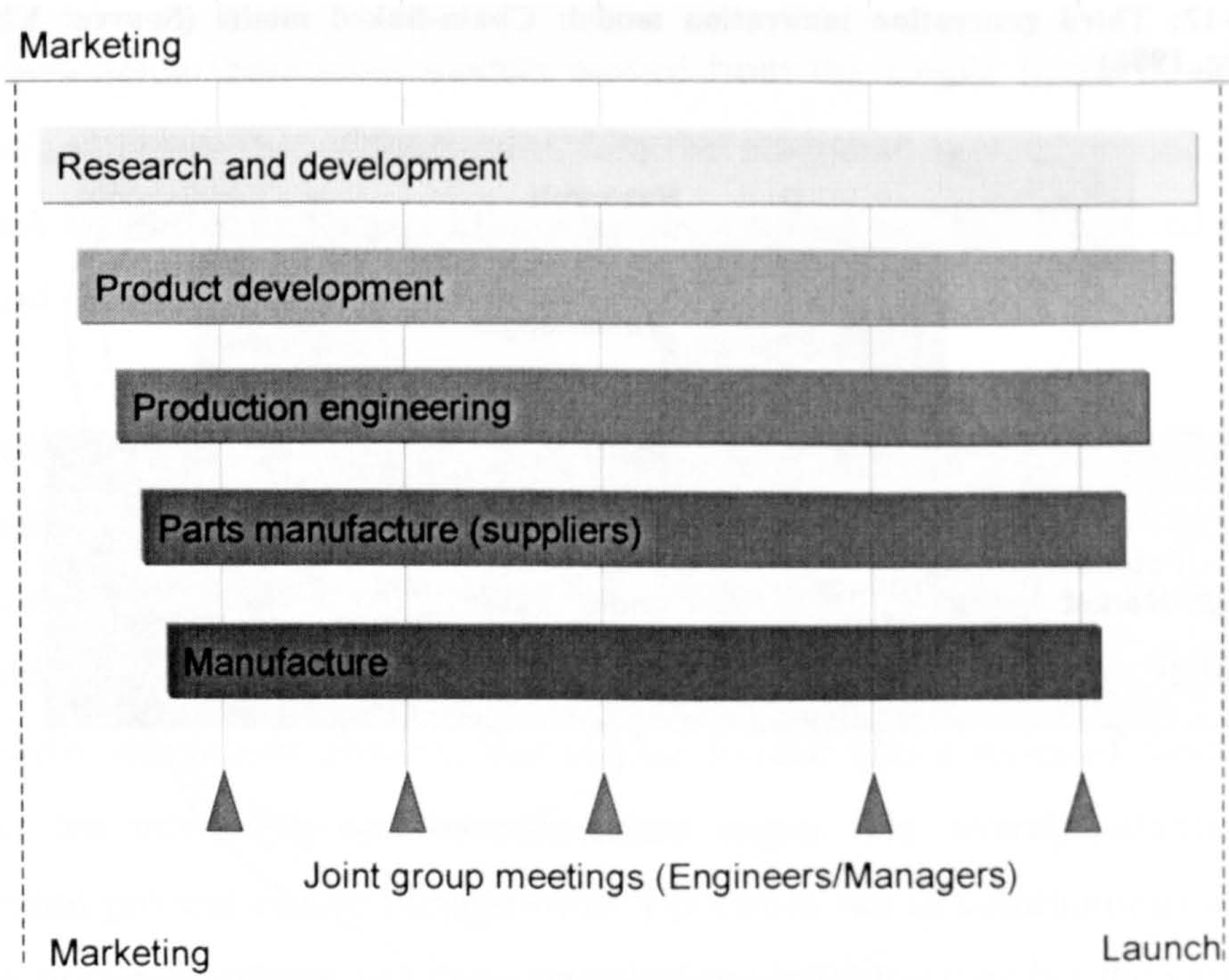


Fourth Generation Innovation Process (Early 1980s – Early 1990s)

Fourth generation innovation models were integrated models that incorporated two Japanese features: integrated development teams and parallel development. Firms using these models organized around business process instead of functions and formed cross-departmental teams that worked in simultaneously in parallel to speed up the development process (Dodgson, 2000). They incorporated more sources of innovation by forging strong upstream linkages with their suppliers and involving them in their product development process at an early stage, forging downstream linkages with leading customers, and entering into collaborative joint ventures with other firms.

The need to manage the complex task of intra- and inter-organizational collaboration gave rise to innovation technology[10] (IvT) that became one of the components of fifth generation innovation models.

Figure 2-13: Fourth generation innovation model: Integrated model (Rothwell, 1992)

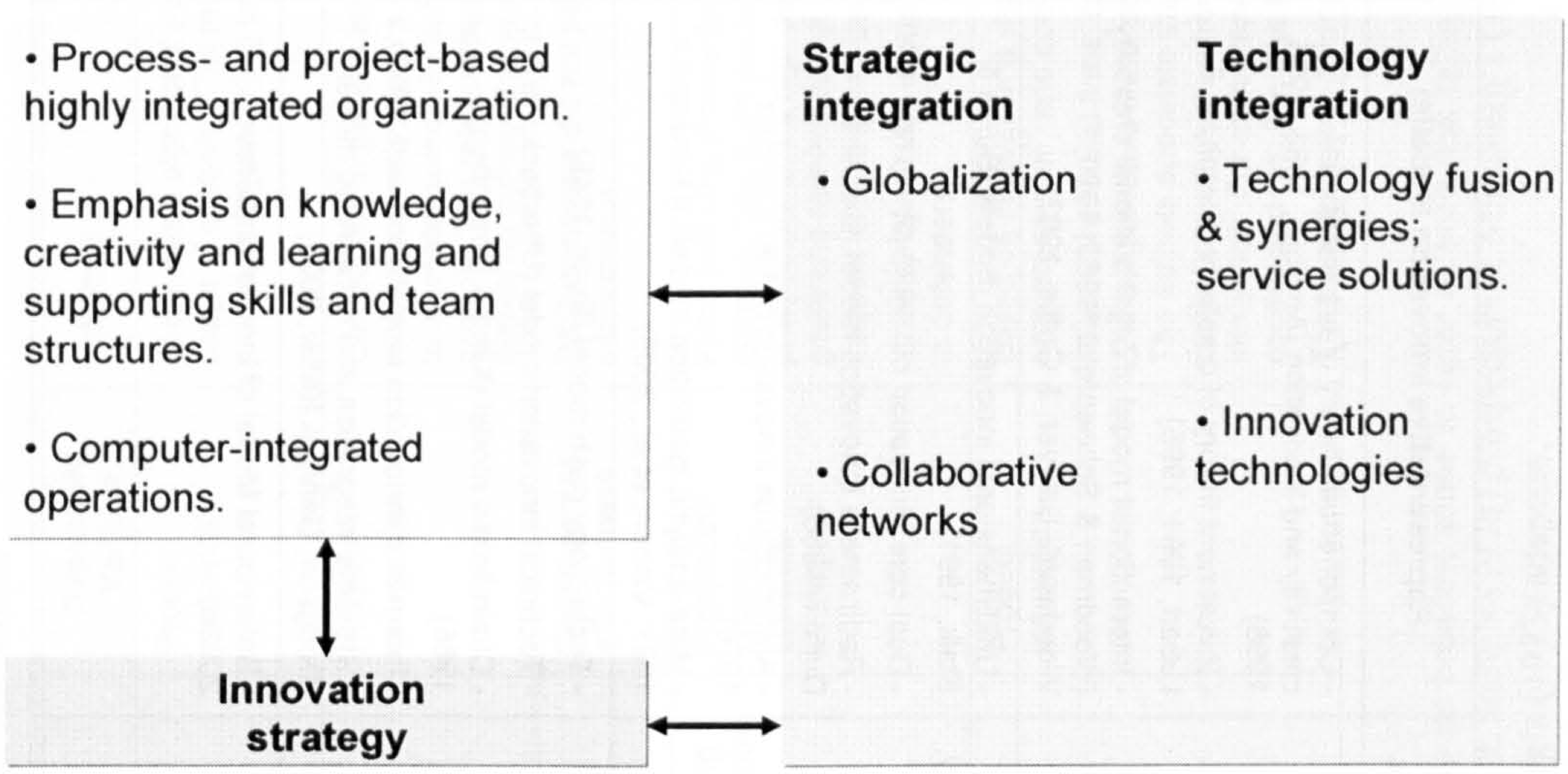


[10] Innovation technologies are technologies used in the innovation process that include simulation, modelling tools, virtual reality, data mining, and rapid prototyping Dodgson et al (2005).

Fifth Generation Innovation Process (Mid 1990s – Present)

Fifth generation innovation process, the systems integration and networking model (SIN), is emergent and still developing (Dodgson et al., 2005). As illustrated in Figure 2-14. SIN includes integrated parallel development; strong linkages with leading users; strategic integration with key suppliers including co-development of new products; horizontal linkages that include joint ventures and collaborative research; and use of innovation technology.

Figure 2-14: Fifth generation innovation model: System integration and network model



2.4.4 Summary

In the previous sections, I reviewed the theoretical and empirical literatures examining the impact of measurement on innovation. More specifically, I used innovation models from three perspectives to examine this relationship. Table 2-4 synthesizes the empirical evidence on the relationship between performance measurement and innovation by linking the innovation models according to their theoretical perspective with measurement as discussed above and it includes the findings from organizational learning based innovation model that will be discussed in the next chapter. Table 2-5 lists the empirical evidence on the impact of performance measurement on innovation.

Table 2-4: Innovation models and performance measurement

Perspective	Rationality Assumptions of Managers	Agency	Level of Analysis	Unit of Analysis	Representative Innovation Model(s)	Empirical Evidence on the Impact of Measurement on Innovation	Empirical Studies
Creativity based innovation models	Bounded rationality	Yes	Organization, Workgroup, Individual	Project	- Componential theory of organizational creativity and innovation (Amabile, 1988, 1996, 1998) - Investment theory of creativity (Sternberg & Lubart, 1991; 1995) - Interactionist model of organizational creativity (Woodman & Schoenfeldt, 1989; 1990; Woodward, Sawyer, & Griffin, 1993)	Measurement may be harmful to creativity and innovation because it limits freedom.	Abbey & Dickson, 1983; Amabile, 1988; Amabile et al., 1996; Webster, 2006
Contingency-based Innovation Models	Rational to bounded rationality	No for some & limited for others.	Organization, organizational sub-units, workgroups	Structure, technology, strategy	- Unidimensional innovation models (Burns & Stalk, 1961) - Dual-core innovation model (Daft, 1978). - Radicalness innovation model (Dewar & Dutton, 1986) - Ambidextrous innovation model (Duncan, 1976). - Meta-analytic innovation models (Damanpour, 1991; Vincent et al., 2005)	Measurement has no or little significance. Measurement has a positive impact on innovation. Measurement has a negative impact on innovation.	Abernethy & Brownell, 1997; Bisbe & Otley, 2004; Brownell, 1985.Burns & Stalk, 1961; Damanpour, 1991; Davilla 2000; Nixon, 1998; Rockness & Shields, 1984, 1988; Vincent et al., 2005
Practice-based Models	Rational	Yes	Organization, organizational sub-units, workgroups	Practices, NPD projects	- Technology push model (Bush, 1945). - Technical innovation model (Utterback, 1971). - Chain-linked model (Kline & Rosenberg, 1986). - Integrated innovation model (Rothwell, 1992). - Systems integration and networking model (Dodgson, Gann, & Salter, 2005).	Generally, these models either ignore the role of performance measurement or attribute little impact to it.	Dodgson et al., 2005; Kline & Rosenberg, 1986; Rothwell & Zegveld, 1985; Utterback, 1971
Organizational learning-based innovation models	Bounded rationality	Yes	Organization	Routines, Decisions	- Behavioural Model of Innovation (Greve, 2003b)	This model postulates that increased levels of measurement may lead to increase in performance gaps that may result in increased innovation rates.	Antonelli, 1989; Bolton, 1993; Greve 2003, 2004

Table 2-5: Empirical research investigating the impact of PMS on innovation

Study	Method	Sample	Unit	MCS	Dependent Variable	Findings
Khandawalla, 1972	Survey	92 US Manufacturing firms	Firm	Multiple controls	Product competition	There was a significant positive relationship between overall competition and usage of controls. This relationship was strongest for product competition, moderate for marketing competition and weakest for price competition.
Khandawalla, 1973	Survey	96 US Manufacturing firms	Firm	Multiple controls	Product competition	Competition in general tends to lead : (a) to the decentralization of decision making authority; (b) to a greater use of sophisticated management controls; & (c) to a greater selectivity in the delegation of authority as well as the employment of controls. This relationship was strongest for product competition, moderate for marketing competition and weakest for price competition.
Hayes, 1977	Cross-sectional survey	70 R&D departments	R&D department	Financial measures	Departmental effectiveness	R&D managers did not perceive financial measures to measure their department's performance.
Simons, 1978	Cross-sectional survey & Interviews	108 Canadian manufacturing firms	Firm	Accounting based control systems	Prospector strategy; Defender strategy	Firms pursuing prospector strategies use management control systems more intensely than firms pursuing defender strategies. More specifically, successful prospectors use a high degree of forecast data in control reports, set tight budget goals and monitor output carefully. Moreover, large prospectors emphasize frequent reporting and use uniform control systems that are modified frequently. Industry dynamism is positively related to ROI for prospector and negatively related to defenders.
Miller & Friesen, 1982	Cross-sectional survey	52 Canadian firms in multiple industries	Firm	Multiple controls	Product innovation	Based on innovation & risk taking, they split the sample into two groups: entrepreneurial[11] and conservative[12]. In entrepreneurial firms, controls were negatively correlated with product innovation & in conservative firms; controls were positively correlated with product innovation. The authors speculate that conservative firms use formal control systems to signal market opportunities when results are declining results and result in more innovation. For entrepreneurial firms, however, control systems flag innovative excess and result in less innovation. They conclude that determinants of product innovations in firms are to a very great

[11] Miller & Friesen (1982) note that entrepreneurial is used here in its broadest sense to refer to bold risk taking and high levels of innovation and not to owner-managed or small centralized companies. They correspond to Miles & Snow's (1978) prospectors & Mintzberg's (1973) entrepreneurial organizations.

[12] They correspond to Miles & Snow's (1978) defenders & Mintzberg's (1973) adapters.

Study	Method	Sample	Unit	MCS	Dependent Variable	Findings
						extent a function of strategy.
Rockness & Shields, 1984	Cross-sectional survey	76 R&D workgroup leaders from 10 organizations in the USA	R&D workgroup	Ouchi's (1977; 1979) organisational control framework: input (social + budget), behaviour (rules & SOP + technical scheduling methods & Output (market)		The results provide little support for the notion that the importance of controls varies with task characteristics, measurability of the output, dependence or complexity. The study finds that output controls are used although it was not possible to measure the output of a task in a timely fashion. One possible explanation for this result is that some controls are used to legitimize the task rather than evaluate it.
Brownell, 1985	Cross-sectional survey and un-structured	21personnel in R&D and 40 personnel from marketing in multinational electronics & computer company	Department	Financial measures	Managerial performance	Accounting information is used in R&D more as a measure of the organization's financial commitments to the activity than as a surrogate for planned or targeted output. Budget participation was found to have significantly greater positive effects on managerial performance in R&D than in marketing.
Rockness & Shields, 1988	Cross-sectional survey	76 R&D workgroup leaders from 10 organizations in the USA	R&D workgroup	Ouchi's (1977; 1979) organisational control framework: input (social + budget), behaviour (rules & SOP + technical scheduling methods & Output (market)		The perceived importance of the expenditure budget for management control in R&D monotonically decreases across planning, monitoring, evaluating and rewarding.

Study	Method	Sample	Unit	MCS	Dependent Variable	Findings
Damanpour, 1991	Meta analysis	23 empirical studies: 21 articles & 2 books	Firm	Formalization [13]	Rate of adoption of innovations; Organizational innovativeness	No relationship was found between formalization & innovation is not significant. For the initiation stage of innovation the relationship is barely significant & for the implementation stage the relationship is non-significant. However, when the studies were divided into two subgroups: low & high innovation scope[14], formalization showed stronger relationship with innovation in the negative direction.
(Amabile, Conti, Coon, Lazenby, & Herron, 1996)	Survey	306 Projects from US electronics company	Project	Controls	Creativity	Five work environment dimensions do consistently differ between high- and low-creativity projects, and thus these dimensions may play an important role in influencing creative behaviour in organizations: challenge, organizational encouragement, work group supports, supervisory encouragement, and organizational impediments. Particularly surprising, challenge & work group supports showed the highest effect sizes and resources, workload pressures, & freedom played less important role in creativity.
Abernethy & Brownell, 1997	Cross-sectional survey	1 Australian industrial firm & 1 US scientific firm	R&D Workgroup	Accounting, behaviour, & personnel controls	Managerial performance	Where task uncertainty is highest, reliance on personnel controls is significantly and positively related to performance, while reliance on either accounting or behaviour controls has significantly poorer implications for performance. Reliance on accounting controls has significant positive effects on performance only where task uncertainty is lowest. Behaviour controls appear to contribute positively to performance in no situation, and contribute negatively where task uncertainty is highest.
Nixon, 1998	Case study	NPD project in one manufacturer of casting machines (CCM Ltd.) in the UK	NPD Project	Performance measures	Project performance	The measures used by CCM Ltd. are consistent with the literature in so far as they are driven by customer needs, are quantitative and objective, and relate to variables that the project team can influence; they also have a strategic orientation, reflect CSFs, are balanced between financial and non-financial metrics, are simple and transparent, and support the collaboration needed for a multi-disciplinary team approach to NPD." Notwithstanding the practical difficulties and limitations of R& D evaluation, the

[13] "This reflects the emphasis on following rules and procedures in conducting organizational activities. Formalization is typically measured by the presence of rule manuals and job descriptions, or more generally, by the degree of freedom available to organizational members as they pursue their functions and responsibilities versus the extent of rules that precisely define their activities" (Damanpour, 1991).

[14] Scope of innovations refers to the number of innovations adopted in a given time period.

Study	Method	Sample	Unit	MCS	Dependent Variable	Findings
						use of financial appraisal methods can assist communication, build consensus and provide a context which decision-makers can use to evaluate the more subjective dimensions of projects.
Davila, 2000	Case study & Cross-sectional survey	Case studies: 12 business units in 7 companies Survey: 56 project managers from 11 companies. The companies in both studies are from medical devices industry in Europe & US	Project manager	- Project management system - Interactive - Diagnostic	MCS; project performance	Management control systems' design is related to performance, in contrast to other variables including product strategy that have no such relationship. Cost and design information has a positive effect upon performance. In contrast, time information hinders performance.
Hoque & James, 2000	Cross-sectional survey	67 Australian manufacturing firms	Firm	Usage of BALANCED SCORECARD	Organizational performance	Product lifecycle is positively associated with BSC usage. Furthermore, firms that have higher proportion of new products have a greater tendency to make use of measures related to new products.
Bisbe & Otley, 2004	Cross-sectional survey	58 medium sized, mature Spanish manufacturing firms.	Firm	- Balanced scorecard - Budgets - Project management systems	Product innovation	The evidence does not support the proposition that the more interactive the use of MCS by top managers, the higher the product innovation (and acting through innovation, the better the performance). The evidence suggests that interactive use may favour innovation in low-innovating firms, while the effect appears to be in the opposite direction in high-innovating firms. The evidence supports the proposition that the impact of product innovation on performance is moderated by the style of use of MCS.
Ditillo, 2004	Case study	Multinational software firm in the UK.	Project team (Software develop-ment)	- Action, results (output/performance measures), & values.	Knowledge coordination & integration	This study introduces extends to the concept of uncertainty to knowledge complexity and differentiate between three types of knowledge complexity: computational, technical, & cognitional. It finds that MCSs could be effective in knowledge intensive organizations because they may play a role in coordinating & integrating knowledge. More specifically, contexts involving knowledge characterized by computational complexity tend to be regulated through action oriented controls whereas contexts involving knowledge characterized by computational complexity tend to be regulated through results oriented controls

Study	Method	Sample	Unit	MCS	Dependent Variable	Findings
						and contexts involving knowledge characterized by cognitional complexity tend to be regulated through personnel/cultural forms of controls
Godener & Soderquist, 2004	Case study	3 large companies in the electronics industry in France.	Division	R&D and new product development (NPD) performance measures	Uses of performance measures and their impacts	They found performance measurement systems are used at two levels: the project level and the functional level. At these levels, they identified five uses of the results of performance measurement and two impacts on the organization and behaviour of the company. They found that managers use the measures at the project level to: (1) reorient projects before failure or cancelling product platforms and families; (2) decide corrective actions at the right moment in running projects; & (3) select promising and relevant project for launching. At the functional level, managers used measures to: (1) ensure better coherence and relevance of product portfolios & (2) Improve processes and adherence to processes. These uses resulted in enhancing the R&NPD staff motivation at the project level and facilitating well-balanced arbitration and decision-making at the functional level.
Vincent, Bharadwaj, & Challagalla, 2005	Meta analysis	Meta-analytic database of 155 independent samples from 108 studies from the period of 1970 through 2004.	Firm	Formalization	Innovation and organizational performance	It was hypothesized that formalization limits organizational flexibility and stifles the creativity of employees because of its focus on rules and procedures within the organization. Formalization results in standardized behaviour from employees (Robbins 1990), thereby inhibiting innovation. However, the results do not support this hypothesis and find no significant relationship between formalization and innovation. There is however, a significant positive relationship between formalization and financial performance.
Webster, 2006	Cross-sectional survey	114 managers in large Australian manufacturing firms.	individual	- Interactive use.	Creativity	The study hypothesized a positive relationship between the interactive use of PMS and individual creativity but the results failed to find a relationship between interactive use and individual creativity.

2.5 What is Performance Measurement Use?

The literature review undertaken in the past section revealed that there are three differing perspectives on the effect of performance measurement systems on a firm's propensity to innovate with each of them having empirical evidence to support its argument. The first perspective views performance measurement systems as constraining innovation because they impede creativity, experimentation, and search in firms (e.g. Amabile, 1997, 1998; Amabile et al., 1996; Burns & Stalker, 1961; Damanpour, 1991; Thompson, 1965). The second perspective views performance measurement systems as helping innovation because they trigger search, facilitate decision-making, and increase risk-taking (e.g. Godener & Soderquist, 2004; Greve, 2003b; Khandwallah, 1972, 1973; Miller & Friesen, 1982; Simons, 1978; 1991; 1995). The third perspective views performance measurement systems as having insignificant or little impact on innovation because they are used primarily for signalling and they do not impact innovation in high uncertainty environments (Abernethy & Brownell, 1997; Brownell, 1985; Hayes, 1977; Rockness & Shields, 1984, 1988).

A possible explanation of the contradiction in the empirical findings of the studies discussed in the previous section is that they generally ignore how performance measurement systems are used "style of use". Ferreira & Otley (2005) remark that "use can be more significant than the existence and actual design of the control system."

In Table 2-6, I summarize twelve studies that classified the various uses of performance measurement starting with a study by Simon, Guetzkow, Kozmetsky, and Tyndall in 1954.

Based on these studies, the uses of performance measurement systems could be reduced to ten uses as illustrated in Table 2-7. These uses are monitoring, learning, decision-making, legitimization, compliance, external communication, alignment, rationalization, and motivational. The uses of performance measurement systems for legitimization, compliance, and external communication are targeted mainly to external audiences and the rest of the uses are targeted mainly to internal audiences.

A closer examination of Table 2-7 reveals that all of the twelve studies agree on two uses of performance measurement: monitoring (control) and learning. Therefore, these two uses will be used as moderating variable in this study.

Table 2-6: Summary of major classifications of PMS use

No	Study	Classification	Description
1	(Simon, Guetzkow, Kozmetsky, & Tyndall, 1954)	Score card	In this role, control systems answer the question - How am I doing? Score carding involves preparing standardized report on performance measures.
		Problem solving	In this role, control systems answer the question - What problems should I look into? Problem solving involves using data to solve problems.
		Attention directing	In this role, control systems answer the question - of the several ways of doing the job, which is best? They call attention to problems. This role is closely related to the principle of exceptions. In this role, managers accept the validity of the standards and manage by exception.
2	(Hrebiniak, 1978)	Legitimization	To legitimize the organization. Legitimacy is granted formally (e.g., by a corporate charter) and informally (e.g. as a result of market mechanisms).
		Constraints	To serve as constraints on individual and organizational behavior.
		Creation of discontent	To create discontent; to provide opportunities to achieve or targets at which to aim.
		Development of commitment	To help develop commitment to the organization.
		Reduction of uncertainty	To reduce uncertainty and provide facts for organizational members to use in decision making.
		Organizational learning	To aid in the process of learning and adaptation.
3	(Burchell, Clubb, Hopwood, Hughes, & Nahapiet, 1980)	Answer machine	Given low uncertainty of objectives and low/high uncertainty over causation, accounting systems may play the role of answering machines. In this role, they provide standard techniques, methods and systems (e.g. stock control systems, investment appraisal methods)
		Learning machine	Given low uncertainty over objectives and high uncertainty over causation, accounting systems may play the role of learning machines. In this role, they provide assistance through decision support systems. They perform ad hoc analyses, what-if models and sensitivity analysis.
		Ammunition machine	Given high uncertainty of objectives and low uncertainty over causation, standards for appraisal and criteria for guiding the organizational task are inherently problematic. In these situations, accounting systems may be used as ammunition machines by managers to promote their own particular positions.
		Rationalization machine	Given high uncertainty of objectives and high uncertainty over causation, accounting systems may be used as rationalization machines. In this role, they seek to legitimize and justify actions that already have been decided upon.

No	Study	Classification	Description
4	(Earl & Hopwood, 1980)	Answer machine	Given low uncertainty of objectives and low/high uncertainty over causation, accounting systems may play the role of answering machines. In this role, they provide standard techniques, methods and systems (e.g. stock control systems, investment appraisal methods)
		Learning machine	Given low uncertainty over objectives and high uncertainty over causation, accounting systems may play the role of learning machines. In this role, they provide assistance through decision support systems. They perform ad hoc analyses, what-if models and sensitivity analysis.
		Dialogue machine	Given low uncertainty over objectives and high uncertainty over causation, accounting systems may play the role of dialog machines stimulating debate and communications in organizations.
		Idea machine	Given low uncertainty over objectives and high uncertainty over causation, accounting systems may play the role of idea machines by providing multiple streams of ideas that may trigger creativity in organizations.
5	(Boland & Pondy, 1983)	Rational	Rational use emphasizes model-based analysis that encompasses relevant causal factor and selects desired outcomes based on a comprehensive understanding.
		Natural	Natural use it emphasizes understanding as a basis of problem solving. Understanding is reached through interactions and adjustment within organizational contexts.
6	(Simons, 1995)	Belief systems	Belief systems refer to the explicit set of shared beliefs that define basic values, purpose, and direction.
		Boundary systems	Boundary systems refer to the formally stated limits and rules that must be respected
		Diagnostic	The diagnostic use refers to the formal feedback systems used to monitor predictable goal achievement.
		Interactive	The interactive use focuses attention and force dialogue throughout the organization by reflecting signals sent by top managers.
7	(Atkinson, Waterhouse, & Wells, 1997)	Coordination (Alignment)	In the coordinating role, "[p]erformance measurement focuses and coordinates the decision-making activities ... by designing, monitoring, and evaluating coherent, comprehensive performance measures. This requires *horizontal* coordination, which a company achieves by focusing on its primary and secondary objectives[15], to ensure that decision makers consider all facets of the decision. It also requires vertical coordination, which ensures that each member understands how he or she contributes to achieving the secondary objectives and, ultimately, the primary objectives."

[15] Primary objectives are defined by the organization's owners. Secondary objectives are what the organization expects from and gives to each stakeholder group, so that it could achieve its primary objectives.

No	Study	Classification	Description
		Monitoring	In the monitoring role, performance measurement systems measure and report performance in meeting stakeholder requirements. 'Monitoring information identifies: (1) realized performance levels on primary and secondary objectives; (2) opportunities to improve process performance and to suggest what improvements are required; (3) whether existing processes are meeting expectations in terms of how they are contributing to the primary objectives... ; and (4) whether existing processes are achieving their performance potential."
		Diagnosing	In the diagnosing role, performance measurement systems support the understanding of how process performance affects organizational learning and performance. Performance measurement systems monitor and compare results on primary and secondary objectives in order to evaluate the nature of their assumed causal relationship.
8	(Neely, 1998)	Comply	"In all organisations there are certain performance parameters which can be described as non-negotiable... So how can organizations use their performance parameters are not infringed? In this context the organisation's measures are effectively a radar system. They are designed to provide an early-warning signal that something is about to go wrong and that corrective action should be taken" (178-179).
		Check	Performance measurement system should allow the following questions to be answered: "How do we look to our shareholders? How do we look to our customers? What we must excel at internally? How can we continue to innovate and create value" (179-180)
		Challenge	Measurement plays a role in testing the validities underlying the assumptions or hypotheses underlying the performance measures. The question that is being asked is our strategy correct? Challenging assumptions leads to double loop learning[16] (180)
9	(Vandenbosch, 1999)	Score keeping	Performance measurement systems can be used to produce standardized reports that are relatively consistent overtime.
		Problem solving	Performance measurement systems can be used to support problem solving
		Focusing attention	Performance measurement systems can be used to focus attention by defining issues to be dealt with in organizations.

[16] "Single-loop learning occurs when matches are created, or when mismatches are corrected by changing actions. Double-loop learning are corrected by first examining and altering governing variables and then action" (Argyris, 1992: 8-9).

No	Study	Classification	Description
		Legitimizing decisions	Performance measurement systems can be used to legitimize decisions in organizations.
10	(Simons, 2000)	Decision making	Managers rely on information to improve decision processes. Managers use information for decision making in two broad categories: (1) information for planning and (2) information for coordination.
		Control	Managers use information for control when they use feedback to ensure that inputs, processes, and outputs are aligned to achieve organizational goals. Managers use output goals to motivate and evaluate employees. So the control function serves as a motivation and evaluation tool.
		Signalling	Mangers use information for signalling when they send cues throughout the organization about their preferences.
		Education and learning	Managers use inf[illegible] for education and learning to train managers and employees and to enable the entire [illegible]tion to understand changes in the internal and external environment.
		External communication	Managers comm[illegible] [illegible]formation to external stakeholders, such as suppliers, customers, and capital providers.
11	(Ittner, Larcker, Randall & 2003)	Problem identification	Performance measurement systems are used to identify problems in organizations.
		Capital investment	Performance measurement systems are used to evaluate capital investment decisions in organizations.
		Performance evaluation	Performance measurement systems are used to evaluate managerial performance.
		External disclosure	Performance measurement systems are used to disclose information to external stakeholders.
12	(Henri, 2006a)	Monitoring	Performance measures are used to provide feedback regarding expectations and to communicate with various stakeholders (monitoring).
		Attention focusing	In addition, top managers use performance measures to send signals throughout the firm (attention focusing).
		Strategic decision making	During the decision-making process, they are employed as a facilitator (strategic decision making)
		Legitimization	Justify decisions or actions (legitimization).

Table 2-7: Classification of uses of performance measurement systems

Use/Study	1	2	3	4	5	6	7	8	9	10	11	12
Monitoring	Scorecard	Constraint	Answer machine	Answer machine	Rational	Diagnostic	Monitoring	Check	Score keeping	Control	Performance evaluation	Monitoring
Learning	Attention	Organizational learning Reduction of uncertainty	Learning machine Idea machine	Dialogue machine	Natural	Interactive	Diagnosing	Challenge	Focus attention	Education and learning signalling	Problem identification	Attention focusing
Decision making	Problem solving								Problem solving	Decision making	Capital investment	Strategic decision making
Legitimization		Legitimization							Legitimizing decisions			Legitimization
Compliance								Comply				
External communication										External communication	External disclosure	
Alignment (coordination)							Coordination					
Rationalization			Rational-ization									
Motivation		Creation of discontent Development of commitment	Rational-ization									

Studies:

1- (Simon, Guetzkow, Kozmetsky, & Tyndall, 1954); 2- (Hrebiniak, 1978); 3- (Burchell, Clubb, Hopwood, Hughes, & Nahapiet, 1980); 4- (Earl & Hopwood, 1980);

5- (Boland & Pondy, 1983); 6- (Simons, 1995); 7- (Atkinson, Waterhouse, & Wells, 1997); 8- (Neely, 1998);

9- (Vandenbosch, 1999); 10- (Simons, 2000); 11- (Ittner, Larcker, Randall & 2003); 12- (Henri, 2006a)

In this study, I chose Simons' (1995) diagnostic (control) and interactive (learning) uses to conceptualise the style of use of performance measures. The rationale for choosing Simons' conceptualisation is two fold. First, it provides one of the most comprehensive treatments of the concept of performance measurement use. Ferreira & Otley (2005) remark that "[a]part from Hopwood's categories (now often discussed in terms of 'rigid' and 'flexible' use) the only substantial contribution is that made by Simons (1995) in terms of his four 'levers of control' categories, and his concept of 'interactive' use". Second, there is a substantial empirical literature stream that has built on Simons' conceptualization (e.g. Abernethy & Brownell, 1999; Abernethy, Bouwens, & Van Lent; Bisbe & Otley, 2004; Bruining, Bonnet, Wright, 2004; Collier, 2005; Davila, 2000; Ferreira & Otley, 2005; Gil et al., 2003; Henri, 2006b; Marginson, 2002; Tuomela, 2005; Widener, 2005). I review this literature in Table 10-5 and Table 10-6 in appendix F.

Simons (1995: 10) advanced the levers of control framework as an "integrated theory for the control of strategy" that involved controlling the processes of strategy formulation and implementation. As shown in Figure 2-15, his framework is based on four levers or management control systems: belief systems, boundary systems, diagnostic control systems, and interactive control systems. He defined management control systems as "the formal, information-based routines and procedures managers use to maintain or alter patterns in organizational activities" (Simons, 1978; 1994; 1995). Simons (2000) posited that for an organization to control its strategy, it should balance five organizational tensions: (1) balance profit, growth, and control; (2) balance short-term results against long-term capabilities and growth opportunities; (3) balance performance expectations of different constituencies; (4) balance opportunities and attention; and (5) balance the motives of human behaviour.

Two of Simons' (1995) levers of control are diagnostic and interactive uses of performance measurement systems. Diagnostic control systems are defined as the "formal information systems that managers use to monitor organizational outcomes and correct deviations from preset standards of performance. (Simons, 1995: 59), whereas, interactive control systems are defined as the "formal information systems managers use to involve themselves regularly and personally in decision activities of subordinates" (1995: 95). See Table 2-8 for a contrast between these two uses. Furthermore, Simons (1995) explains, "[t]he difference between diagnostic and control systems is not in their technical design features. A diagnostic control system may look identical to an interactive control system. The distinction between the two is solely in the way that mangers *use* these systems" (Simons, 2000: 208).

In making the distinction the between these two styles of use, Simons (1995) draws from the psychology literature to liken diagnostic use to the mindless construct and interactive use to the mindfulness construct. "Mindlessness refers to individual behaviors that are automatic and unthinking. They are learned through habit, rules, and accepted classification categories... Mindfulness, in contrast, is concerned with the creation of new categories, openness to new information, and awareness of multiple perspectives (Langer, 1989, 62). Langer argues that a preoccupation with *outcomes* can make us mindless, whereas mindfulness is an orientation to the *processes* that lead to outcomes (p. 75)... Recall that diagnostic control monitors outcomes, while interactive control focuses on process" (Simons, 1995: 103-104).

Figure 2-15: Levers of control framework (Source Simons, 1995: 157)

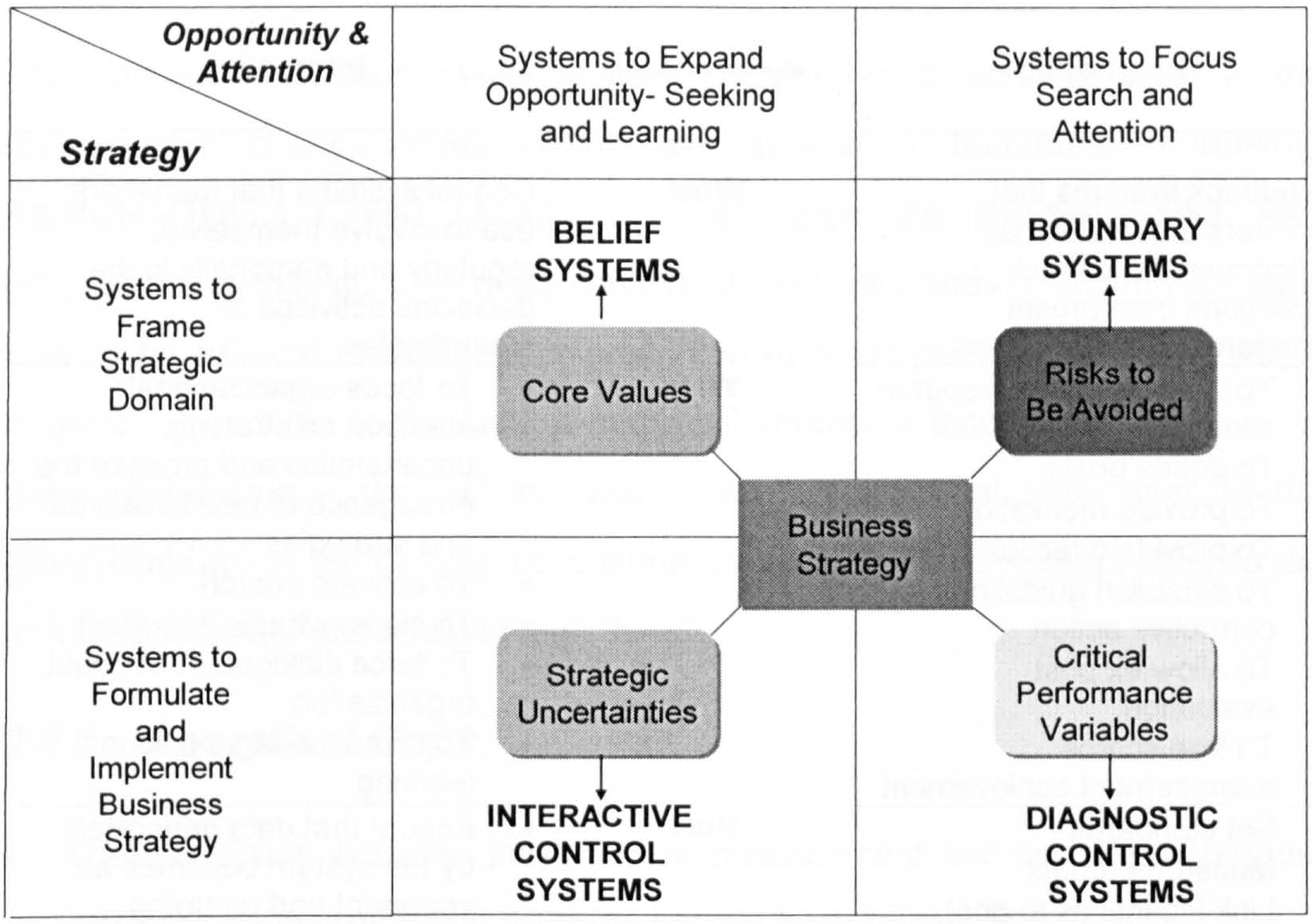


Table 2-8: Contrasting styles of PMS use (Source: adapted from Simons, 1994; 1995; 2000)

Diagnostic	⟷	Interactive
Feedback systems that monitors organizational outcomes and correct deviations from preset standards of performance	**What**	Control systems that managers use to involve themselves regularly and personally in the decisions activities of subordinates
• To allow effective resource allocation • To define goals • To provide motivation • To provide direction • To establish guidelines for corrective action • To allow ex post evaluation • To free scarce management achievement	**Why**	• To focus organizational attention on strategic uncertainties and provoke the emergence of new initiatives and strategies • To activate search • To signal what is important • To force dialogue throughout organization • To stimulate organizational learning
• Set standards • Measure outputs • Link incentives to goal achievement	**How**	• Ensure that data generated by the system becomes an important and recurring agenda in discussions with subordinates • ensure that the system is the focus of regular attention by managers throughout the organization • Participate in face-to-face meetings with subordinates • Continually challenge and debate data, assumptions, and action plans
Performance standards can be preset Outputs can be measured Feedback information can be used to influence or correct deviations from standard Process or output is critical performance variable	**When**	Strategic uncertainties require search for disruptive change and opportunities
Single loop	**Learning**	Double loop
Critical performance variables	**Focus**	Strategic uncertainties
Deductive (flying by instrument)	**Analytical Reasoning**	Inductive , sensory (flying by feel)
No surprises	**Goal**	Creative search
Negative	**Feedback**	Positive
Input or process	**Adjustment to**	Double loop learning
Past and present	**Time Frame**	Present and future
Eliminate need for talk	**Communication**	Provide common language

2.6 Research Gaps

The extensive literature review identified a number of research gaps in the existing performance measurement and innovation literatures. Following Atuahene-Gima's (2004) classification, the gaps are grouped under four categories: theoretical, empirical, contextual, and substantive. Theoretical gaps refer to insufficient explanation or prediction of some phenomena – the "why". Empirical gaps refer to the lack of empirical studies or inconsistent findings on some phenomena – the "where" and "when". Contextual gaps refer to the generalizability of the findings of existing research. Substantive gaps refer to the lack of managerial understanding of "how".

2.6.1 Theoretical Gaps

- The interaction between performance measurement use and organizational performance measurement diversity has not been largely theorized.
- Behavioural theory of innovation assumes that performance measures are used diagnostically.

2.6.2 Empirical Gaps

- The few empirical studies examining the relationship between performance measurement systems and product innovation have yielded inconsistent findings.
- There are no quantitative studies that examined the impact of interactive and diagnostic use on product innovation.

2.6.3 Contextual Gaps

- There are no quantitative studies that examined the relationship between performance measurement systems and product innovation in UK setting.
- There are no quantitative studies that examined the impact of interactive and/or diagnostic use on product innovation in UK setting.

2.6.4 Substantive Gaps

- There is a managerial requirement for theoretically grounded and empirically tested innovation models that enable managers to understand how best they could use performance measurement systems to manage product innovations in their firms.

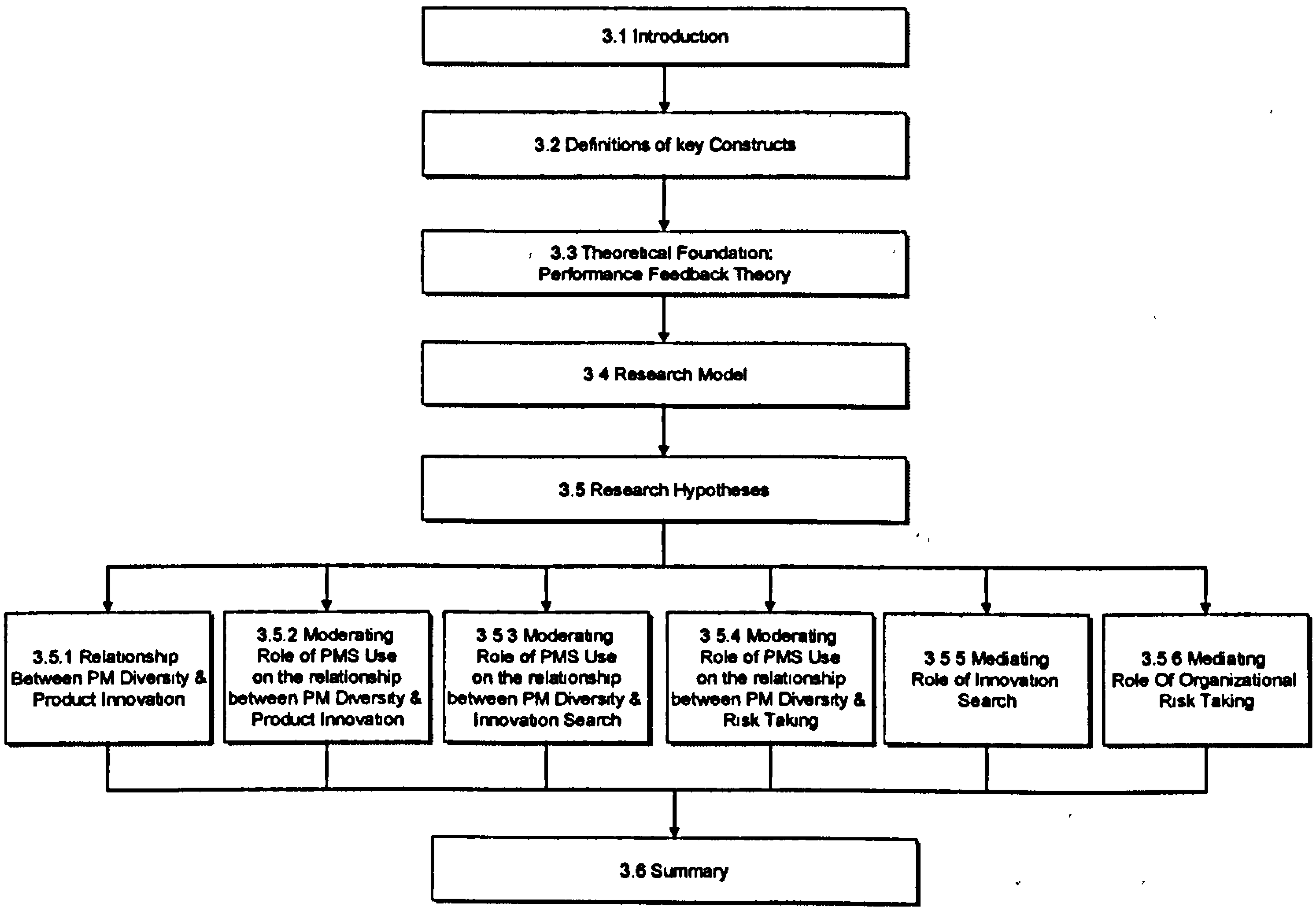
3 Theoretical Model and Hypotheses

3.1 Introduction

In the previous chapter, my review of the literature on the impact of performance measurement on product innovation from three perspectives pointed to a number of research gaps that I try to address in this chapter.

There are six sections in this chapter, as illustrated in Figure 3-1. The first section introduces the chapter followed by a section that defines the key constructs in the study. The third section introduces the performance feedback theory as the theoretical foundation of the research undertaken in this study. The fourth section presents the behavioural theory of innovation and the research model of the study. The fifth section consists of five subsections and presents the arguments for advancing the research hypotheses. The final section lists the hypotheses advanced in this study.

Figure 3-1: Outline of Chapter 3



3.2 Definitions of Key Constructs

3.2.1 Performance Measurement Systems, Performance measurement and Performance Measures

In this study, I adopt the definitions of performance measurement advanced by Neely, Gregory & Platts (1995). They are as follows:

- **Performance measurement systems** can be defined as the set of metrics used to quantify both the efficiency and effectiveness of actions.
- ***Performance measurement*** can be defined as the process of quantifying the efficiency and effectiveness of action.
- ***Performance measure*** can be defined as a metric used to quantify the efficiency and/or effectiveness of an action.

3.2.2 Performance Measurement Use: Diagnostic and Interactive Uses

Simons' (1995) conceptualisation of diagnostic and interactive uses of performance measurement system is adopted in this study. Diagnostic use is defined by the extent to which top management teams use performance measures to monitor organizational outcomes to correct deviations from preset standards of performance (targets) and interactive use is defined by the extent they involve themselves regularly and personally in decision activities of subordinates to focus their subordinates' search on the strategic uncertainties facing their organizations.

3.2.3 Organizational Performance Measurement Diversity

In this study, I adopt the definition advanced by Henri (2006b) that is based on the work of Hoque & James (2000), Ittner et al (2003), and Scott & Tisen (1999). Organizational performance measurement diversity refers specifically to the extent to which top management teams measure and use information related to a broad set of financial and non-financial measures. Organizational performance measurement diversity is a major design element and it is the building block of all performance measurement systems and frameworks (See review of performance measurement frameworks in appendix B).

3.2.4 Innovation Search

Building on the concepts introduced by March & Simon (1958), Nelson & Winter (1982) and Winter (1984), I follow the approach taken by Katila (2000; 2002) and Katila & Ahuja (2002) in defining innovation search as the firm's problem solving activities that involves the creation of new products. This conceptualisation is consistent with the behavioural theory of the firm (Bromiley, 2005), organizational learning perspective (Huber, 1991) and builds on previous research that have taken the same approach (e.g., Ahuja & Katila 2004; Doughtery & Hardy, 1996; Katila, 2000; 2002; Katila & Ahuja, 2002; Mahdi, 2002; 2003).

3.2.5 Organizational Risk Taking

In this study, I adopt the approach developed by March & Shapira (1987; 1992) that extended behavioural theory of the firm (Cyert & March, 1962) rather than the approach used by classical decision theory because it is consistent with how managers see risk. Managers associate risk[17] with negative outcomes and view a risky choice as one that contains a threat of a very poor performance (March & Shapira, 1987).

Shapira (1994: 4) notes that "[t]he classic treatment of risk in decision theory (Luce and Raiffa, 1957) distinguishes among three types of decision-making situations:

(a) *Certainty*, where each action is known to lead invariably to a particular outcome.

(b) *Risk*, where each action leads to a few known outcomes, each of which occurs with a specific probability.

[17] For extensive reviews of the definitions and measures of risk in strategic management research, please refer to the following studies: Baird & Thomas, 1990; Bromiley, Miller, & Rau, 2001; Ruefli, Collins, Lacugna, 1999; and Collins & Ruefli, 1996: 23-45.

(c) *Uncertainty,* where each action may lead to a set of consequences, yet the probabilities of these outcomes are unknown" (Emphases are in the original text).

However, empirical studies by MacCrimmon & Wehrung (1986) and Shapira (1994) found that managers do not equate risk of an alternative with the variance of the probability distribution of the possible outcomes that might follow the choice of the alternative, as the classical decision theory would suggest (March & Shapira, 1987). According to March & Shapira (1987), there are three differences in the way managers perceive risk compared to the treatment of risk in the classical decision theory: (1) Most managers do not treat uncertainty about positive outcomes as an important part of risk; (2) Managers do not see risk primarily as a probability concept; (3) Managers do not reduce risk to a single quantifiable construct.

The organizational risk taking approach that is taken in this study builds on the prospect theory (Kahneman & Tversky, 1979) but it also differs from it in at least three aspects. First, Bromiley, Miller, & Rau (2001) observe that prospect theory attempts to explain risk taking at the individual level using experiments that eliminate extraneous factors making the results less applicable at the organizational level. Second, they note that prospect theory asserts risk aversion in the domain of gains and risk seeking in the domain of losses but the theory does not predict minimum risk aversion and risk seeking near the reference point as does the organizational risk taking theory adopted in this study. Third, March & Shapira (1987) note that managers perceive opportunities that they have less control over their outcomes as riskier than opportunities that they have more control over their outcomes. This moderating effect of outcome control on risk taking in the loss and gain domains has not been theorized by the prospect theory (Forlani, 2002).

3.2.6 Product Innovation

In this study, I adopt the definition advanced by DTI (2005a). Product innovation refers to the market introduction of a *new* good or service or a *significantly* improved good or service. The innovation must be new to the organization, but it does not need to be new to the market and it does not matter if the innovation was originally developed by the organization or by other organizations.

3.2.7 Organizational Attention

Consistent with Simon's (1947) conceptualisation of attention as an organizational level construct that encompasses encoding, I adopt the definition advanced by Ocasio (1998). *Attention* is here defined to encompass the noticing, encoding, interpreting, and focusing of time and effort by organizational decision-makers on both (a) *issues*; the available repertoire of categories for making sense of the environment: problems, opportunities, and threats and (b) *answers*: the available repertoire of action alternatives: proposals, routines, projects, programs, and procedures.

Building on the work of March & Olson (1976), Ocasio notes that although individuals perform the act of attention, it is justified as an organizational-level construct because attention is situated in the context of the firm's activities and procedures that are linked to the issues, solutions, and organizational decisions makers that are distributed throughout the organization.

3.2.8 Aspiration Levels (Performance Targets)

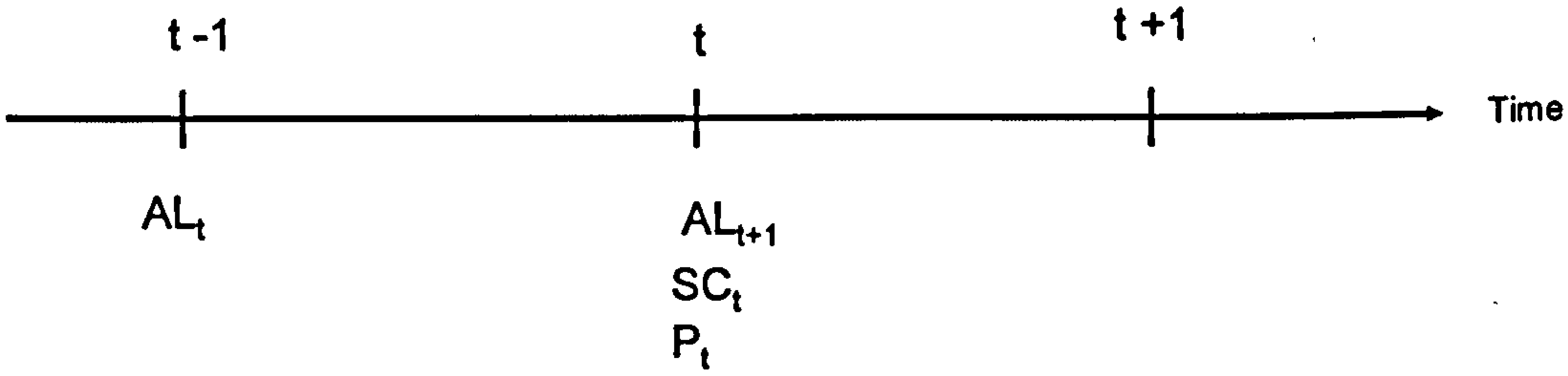
Aspirations levels are the targets set for performance measures. They are reference points that identify the boundaries between success and failure (March & Simon, 1958). Aspiration levels serve as targets or goals for organizational performance (Mezias, Chen, Murphy, 2002). Some of the other definitions compiled by Greve (2003a) are "the level of future performance in a familiar task which an individual ... explicitly undertakes to reach" (Frank, 1935), "reference point that is psychologically neutral" (Kameda & Davis, 1990), and "the smallest outcome that would be deemed satisfactory by the decision maker" (Schneider, 1992). In all of the definitions, aspiration levels are conceptualised as references against which performance is compared.

Consistent with the bounded rationality assumption (Simon 1955), it has been argued that decisions makers use aspiration levels because of the ambiguous relationships between organizational actions and organizational performances, limited cognitive abilities of managers, limited information available to them, and the need to economize on the amount of information needed to make decisions (Lant 1991; Mezias et al., 2002).

Aspiration levels are adaptive and respond to performance feedback (Cyert & March; 1963, Lant, 1992; Mezias et al., 2002).

Cyert & March (1963) identified three variables that determine organizational aspiration levels: previous aspiration levels, organization's previous performance, and the previous performance of comparable competitors (See Figure 3-2). These determinants of organizational aspiration levels have been supported by empirical evidence (Glynn, Lant, & Mezias, 1991; Lant 1992; Mezias et al., 2002).

Figure 3-2: Determinants of organizational aspiration levels



AL_{t+1} = Function (P_t, AL_t, SC_t)

AL_{t+1} = Aspiration level for (*t +1*) is set at (*t*)
P_t = Performance of focal organization at (*t*)
AL_t = Aspiration level at (*t -1*)
SC_t = Social comparison (competitors performance) at (*t*)

3.2.9 *Bounded Rationality*

One of central assumptions of the research model advanced in this study is that managers are intendedly rational but only limitedly so (Simon, 1947).

Bounded rationality is a variant of rational choice that takes into account the cognitive limitations of the decision-makers: limitations of both knowledge and computational capacity (Simon, 1955; 1956; 1997). Decision-makers have limited information on the available alternatives and their consequences and even if this information is available, they cannot computationally process it.

Two consequences follow from the bounded rationality assumptions. First, decision-makers staisfice – look for a course of action that is simply satisfactory rather than maximize – look for the best course of action as posited by rational choice theory. Second, decision-makers simplify the complex world instead of approaching it with its full complexity as posited by the rational choice theory because the capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behaviour in the real word (Simon, 1947).

Bounded rationality could be contrasted with the rationality assumption in the rational choice theories. Simon (1947) notes that "1) Rationality requires complete knowledge and anticipation of the consequences that will follow on each choice. In fact, knowledge of consequences is always fragmentary. 2) Since these consequences lie in the future, imagination must supply the lack of experienced feeling in attaching value to them. But, values can only be imperfectly anticipated. 3) Rationality requires a choice among all possible alternative behaviours. In actual behaviour, only a very few of all these possible alternatives ever come to mind".

3.2.10 *Organizational Learning*

In this paper, following Greve (2003a) and Schulz (2002), I adopt Levitt & March's (1988) definition of organizational learning as being (1) routine-based, (2) history-dependent and adapted to experience, and (3) oriented to goals.

In this definition, organizational routines are conceptualised as repositories of organizational knowledge and form the basis of organizational memory and learning is conceptualised as the creation and modification of routines through the encoding of lessons from experiences that are goal oriented into routines that guide individual and group behaviour (Schulz, 2002). Levitt & March (1988) defines routines as including "the forms, rules, procedures, conventions, strategies, and technologies around which organizations are constructed and through which they operate. It also includes the structure of beliefs, frameworks, paradigms, codes, cultures, and knowledge the buttress, elaborate, and contradicts the formal routines".

There are also a number of other characteristics that differentiate this definition from others: (1) it is a process definition showing how organizations learn, (2) the conceptualisation of routines as organizational knowledge that are independent of the individual actors that execute them and are capable of surviving considerable turnover in individual actors establishes a supra-individual basis of organizational learning, retaining notions of limited-rational and rational adaptation and locates routine-based learning on an organizational level, above the level of individual learning (Levitt & March, 1988; Shulz, 2002), (3) the emphasis on the ecologies of learning where organizations are viewed as collections of subunits learning in an environment of other collections of subunits (Cangelosi & Dill, 1965; Levitt & March, 1988).

3.3 Theoretical Foundation: Performance Feedback Theory

Organizational theories assume that organizations and individuals learn and adjust their behaviour in response to past experiences and thus they exhibit adaptive behaviour (Bromiley 1991; Cyert & March, 1963; Herriot, Levinthal, & March 1983; Lant 1991; Lant & Montgomery 1987; Levitt & March, 1988; March, 2006; March & Simon, 1958; Mezias, 1988; Mezias, Chen, & Murphy, 2002; Morecroft, 1985; Sterman, 1989). More specifically, the behavioural theory of the firm (Cyert & March, 1963) state that organizations adapt their behaviour in response to performance feedback against aspiration levels. Therefore, Behavioural theory of the firm (BTOF) serves as a useful perspective to describe organizational change (Levitt & March, 1988). Consequently, BTOF is well suited to predict firms' propensity to innovate because innovations are special case of change (Bolton, 1993; Knight, 1967).

Performance feedback theory combines BTOF and organization risk theory (March, 1994; March & Shapira, 1987; 1992; Shapira, 1994) as illustrated in Figure 3-3. It emphasizes the goal orientation of organizations. Performance feedback models are based on the idea that organizations learn when they experience problems (March, 2006).

Shulz (2002) explains that performance feedback models treat organizational learning as a two-stage process. In the first stage, organizations adapt their behaviour when performance fall short of aspirations (*behavioural adjustment)*, and in the second stage, they would adapt their aspirations to achieved performance (*aspiration level adjustment)* as illustrated in Figure 3-4.

Lant & Mezias (1990) articulate three basic components of a performance feedback-based learning model. "First, organizations have a target level of performance or aspiration level to which they compare their actual performance....Second, performance above or below aspiration level affects the likelihood of observable organizational change....Third, a learning model suggests that the acquisition and processing of information about alternatives takes place in a relatively costly process of search".

Expanding the five steps by Greve (2003c), the performance feedback model illustrated in Figure 3-3 consists of seven steps that lead to organizational change.

1. *Setting goals.*

   Gross (1969) notes that "[t]he central concept in the study of organizations is that of organizational *goal.* He adds, "[o]ne might even claim that the notion of goal is coincidental with that of an organization." Scott (2003) remarks that '[t]he concept of goals is among the most slippery and treacherous of all those employed by organizational analysts."

   The perspective on organizational goals undertaken in this study is that of the BTOF (Cyert & March, 1963; Simon, 1965). According to this perspective, organizational goals are set by the dominant coalition of the firm through a bargaining process (Cyert & March, 1963; March, 1962). Scott (2003) notes that this conception of organizational goals avoids many problems that have plagued earlier explanations and he gives the following reasons for adopting this conception:

   - The problem of reification is avoided: individuals and groups have interests, and the process by which these preferences come to be imposed on the organization is specified.
   - It is recognized that although individuals and groups specify the goals of the organization, there is no presumption that they do so on an equal footing, nor is it assumed that they hold common objectives.
   - It is recognized that although individuals and groups impose goals on the organization, in most cases no single individual or group is powerful enough to determine completely the organization's goals;

hence, the organization's goals are distinct from those of any of its participants.

- Allowance is made for differences in interests among participants. Some, but not all, of these differences may be resolved by negotiation, so at any time, conflicting goals may be present.
- It is recognized that the size and the composition of the dominant coalition differ from one organization to another and vary within the same organization over time.

He also suggests two additions to the list:

- While interests are seated in specific individuals and groups, it is an oversimplification to assume that organizational goals represent simply a negotiation among and an aggregation of exiting interests, because new interests continually emerge in the course of the interaction (Wallace 1975: 127).
- It is essential to emphasize that the dominant coalition may include and represent interests of constituencies or "stakeholders" outside of the formal boundaries of the organization.

2. ***Designing performance measures.***

   Once organizational goals have been determined, organizations quantify these goals through performance measures (Neely, 1998).

3. ***Setting targets (aspirations levels).***

   The aspiration level is determined by the past performance of the focal organization or the performance of its competitors, and helps a boundedly rational decision maker interpret performance by dividing it into a success range above the aspiration level and a failure range below the aspiration level (March & Simon, 1958).

4. ***Evaluating performance.***

The organizational performance is compared with an aspiration level that specifies what level of performance the decision maker would view as satisfactory. Performance below the aspiration level creates an organizational problem, which causes search for solutions. This problemistic search is driven by heuristic rules such as searching for solutions that are proximate to the symptom and the current organizational strategy (Cyert & March, 1963). Problemistic search means "search that is stimulated by a problem ... and is directed toward finding a solution to that problem" (Cyert & March, 1963: 121). Organizations also search when they have slack resources, such as extra time and funds that can be used for generating innovations or investing in future competitiveness (Penrose, 1959).

5. ***Taking risks.***

   Solutions generated by problemistic and slack search are evaluated for their risk and fit to the organizational strategy. Risk evaluation is done by comparing the potential loss in a given project with the manager's risk tolerance (March & Shapira, 1987; 1992; Shapira, 1994), which is influenced by the organizational performance. Performance below the aspiration level increases the risk tolerance, and performance above the aspiration level decreases the risk tolerance (Kahneman & Tversky, 1979).

6. ***Making decisions.***

   Decisions are made based on the available problems and solutions and the decision-maker risk tolerance.

7. ***Making strategic changes.***

   The adoption of new solutions by the decision makers in response to performance feedback results in organizational change. Thus, the effects of performance feedback on organizational search and managerial risk combine to yield the effect on the rate of making organizational change (Greve, 2003a).

The performance feedback theory have been successful in predicting many strategic outcomes: innovations (Singh, 1986; Greve, 2003b), facility investment decisions (Greve, 2003c), R&D spending (Antonelli, 1989; Hundley, Jacobson, & Park, 1996; Kamien & Schwartz, 1982), new ways of doing R&D (Bolton, 1993), market niche change (Greve, 1998b), firm risk taking (Bromiley, 1991), factory expansion (Audia & Greve, 2006) and strategic reorientation (Audia, Locke, & Smith, 2000; Greve, 1998a; Lant, 1992; Lant, Milliken, & Batra, 1992).

It has also been applied in many industries: airlines and trucking (Audia et al., 2000), furniture and software (Lant et al, 1992), radio stations (Greve, 1998b), semiconductors (Boeker, 1997), computer workstations (Audia & Sorenson, 2001), shipping building (Greve, 2003b). Furthermore, it has been applied in US context (Greve, 1998b) and Japanese context (Greve, 2003b).

Figure 3-3: Experiential learning performance feedback-based model (Source: Yaghi, 2005 adapted from Cyert & March, 1963; Greve, 2003a; March, 1994)

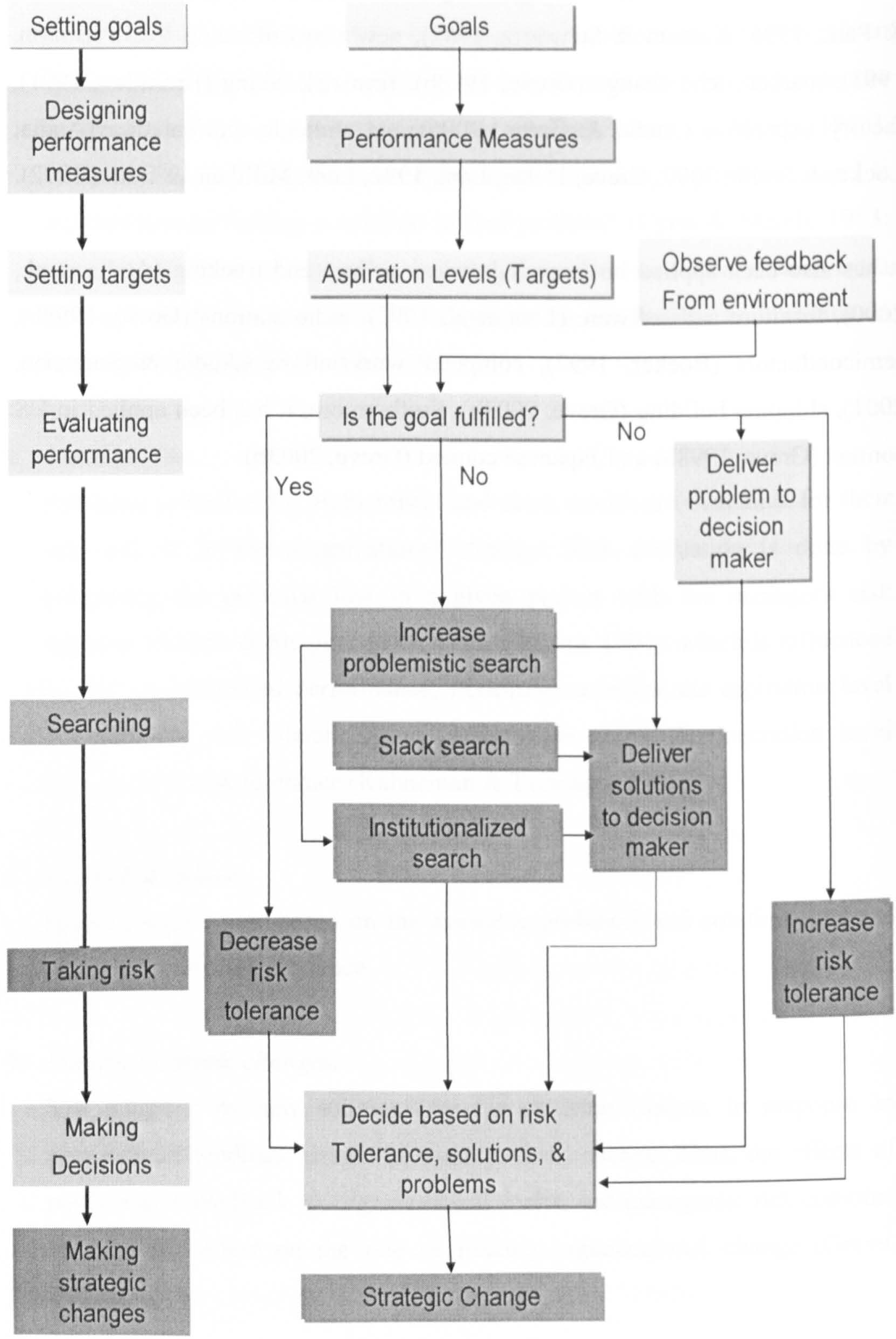
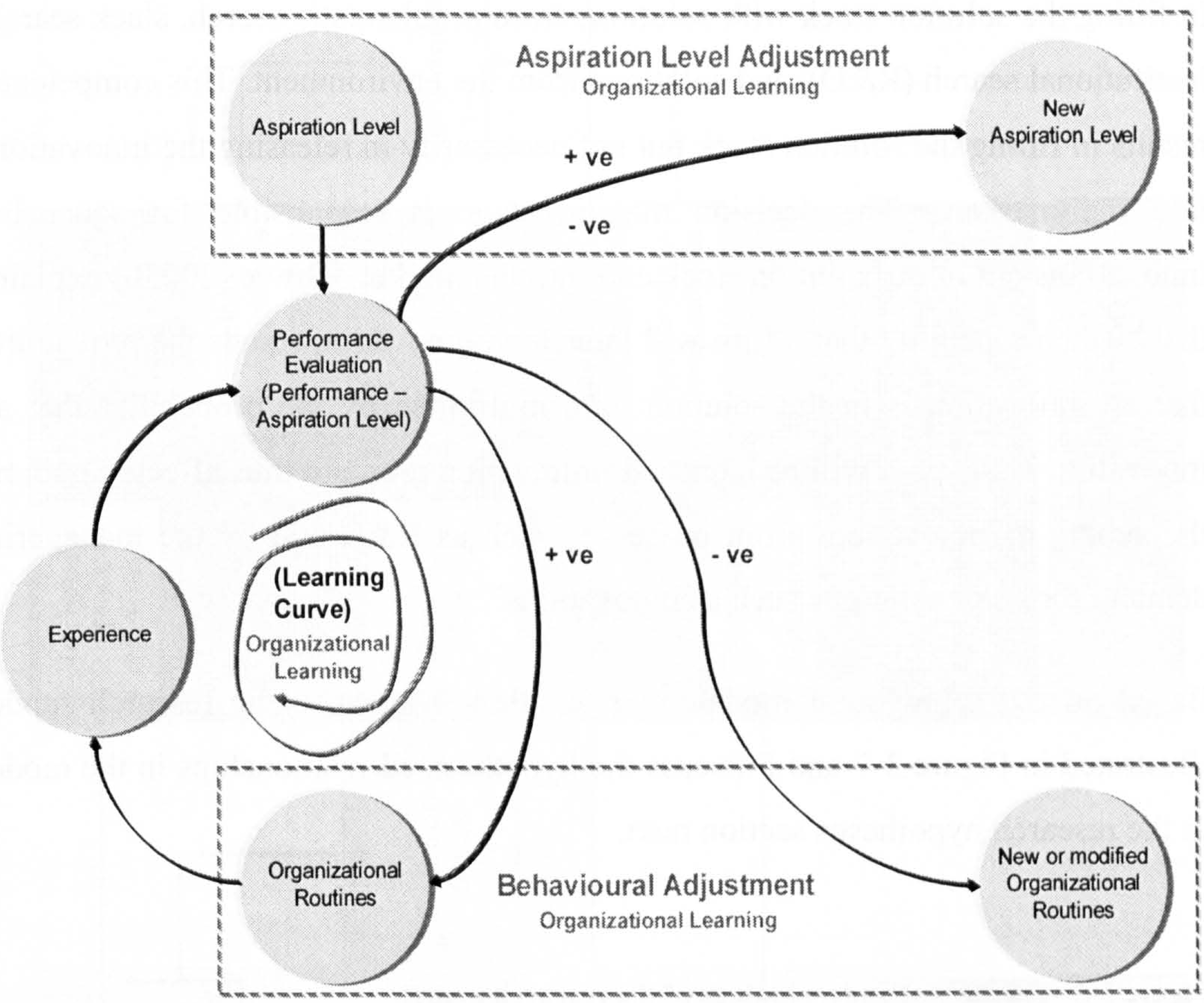


Figure 3-4: Performance feedback-based experiential learning (Adapted from Argote, 1999; Cohen & Sproull, 1996; Lomi, Larsen, & Ginsberg, 1997; Schulz, 2002)



3.4 Research Model

The research model advanced in this thesis is based on the behavioural theory of innovation (Greve, 2003b) which is a special (applied) case of the performance feedback theory (Greve, 2003a; March, 1994). The theoretical underpinnings of the research model are illustrated in Figure 3-5.

As illustrated in Figure 3-6, the behavioural model of innovation[18] is based on the work of Greve (2003b) and Fiol (1996). In this model, organizational capacity to innovate is conceptualised as two sponges. The first sponge is the product

---

[18] Yaghi (2005) undertook a systematic literature review (SLR) to document the behavioural model of product innovation. SLR adopts a scientific technique for searching, evaluating, extracting, and synthesizing literature with an audit trail that remedies many of the shortfalls of narrative literature review (Tranfield, Denyer, & Smart, 2003)

development sponge. Organizational competency in product development is a necessary condition but not sufficient for launching product innovations. It results in filling the solution stock with solutions from problemistic search, slack search, institutional search (R&D), and solutions from the environment. This competency results in filling the solution stock but not necessarily in releasing the innovations into the market. The decision making stage is responsible for squeezing innovations out of the solution stock and into the market. Greve (2003b) explains that "[t]he probability that a firm will launch an innovation equals the probability that an innovation is in the solution pool multiplied by the probability that an innovation in the pool will be launched. Innovation rates are thus affected both by the supply of innovations from processes such as R&D and by the managerial demand for risky solutions such as innovations".

Based on the behavioural model of innovation, I advance the research model illustrated in Figure 3-7 and I discuss the hypothesized relationships in the model in the research hypotheses section next.

Figure 3-5: Theoretical underpinnings of the research model

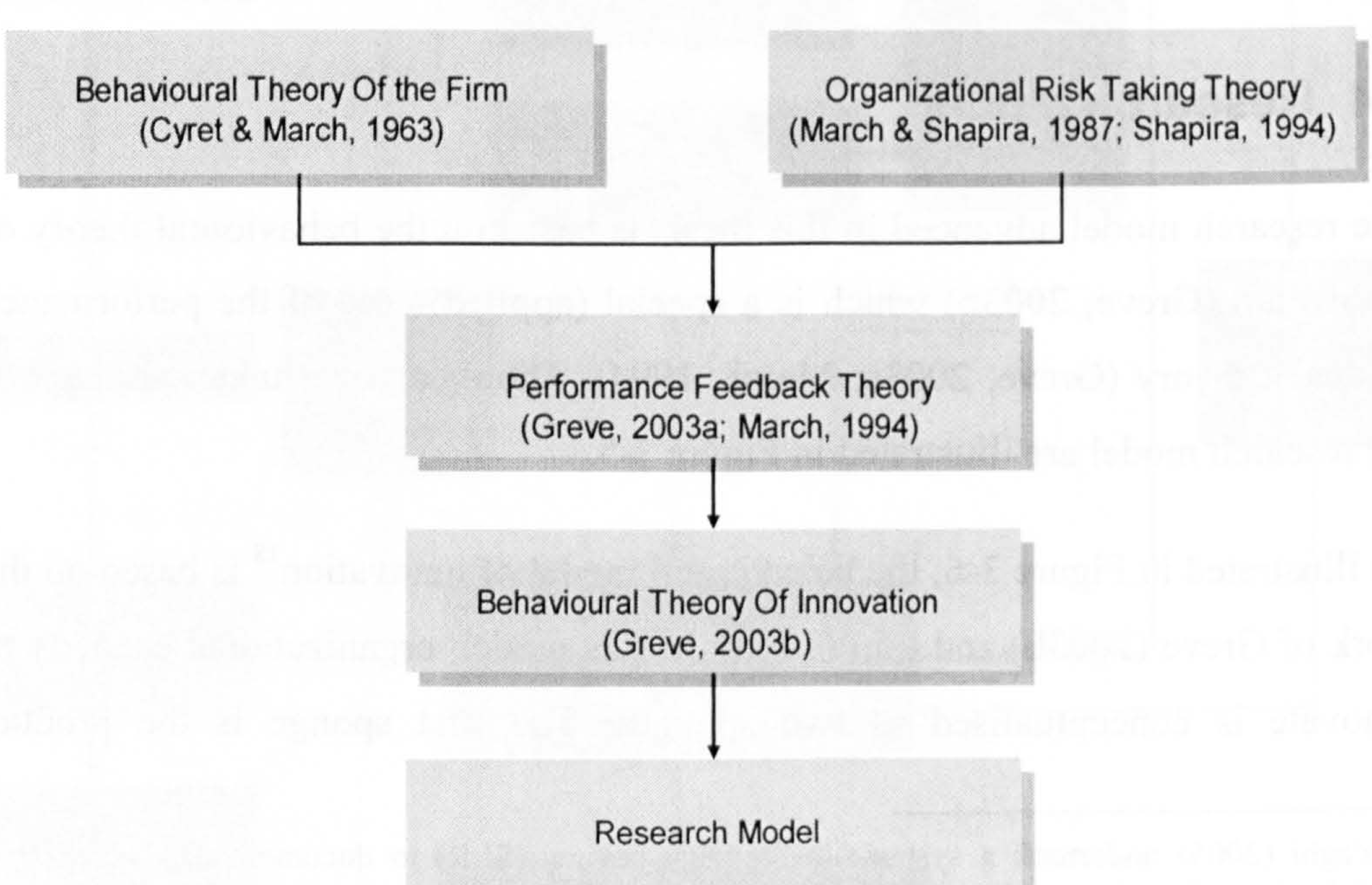


Figure 3-6: Behavioural model of innovation (Source: Fiol, 1996; Greve 2003a; Simon & March, 1958)

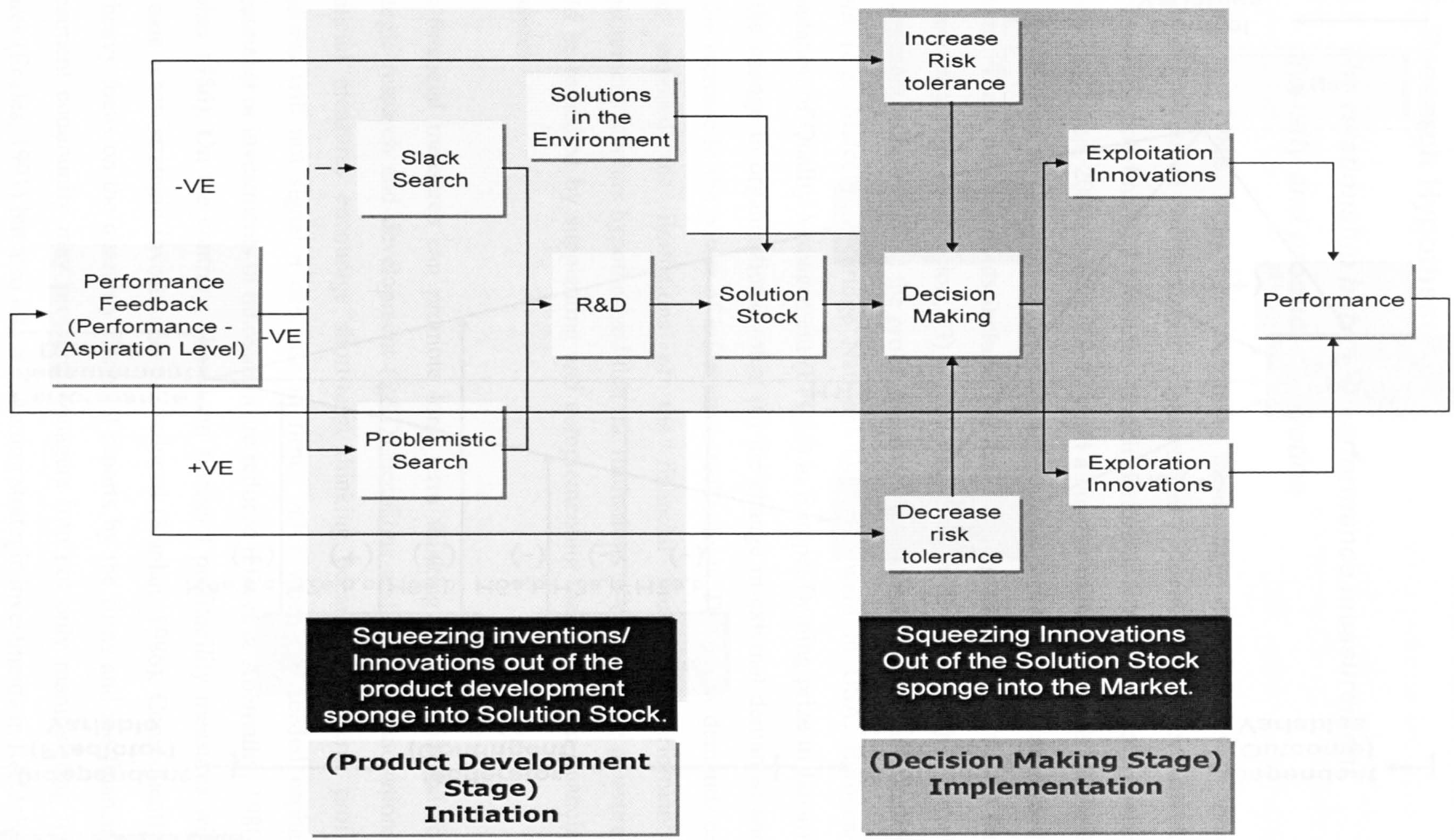
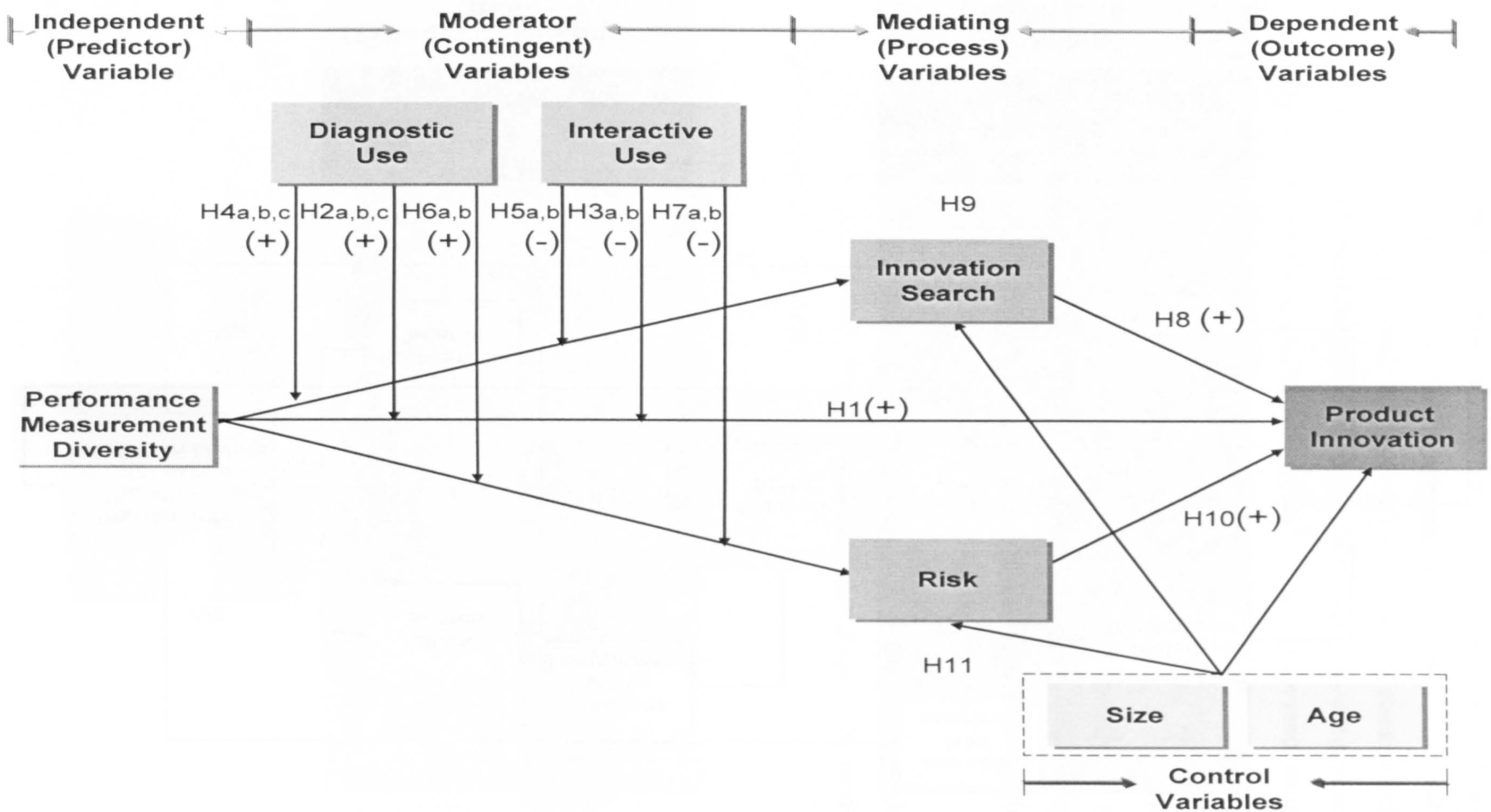


Figure 3-7: Research model



3.5 Research Hypotheses

3.5.1 The relationship between performance measurement diversity and product innovation

By the late 1980s, firms were becoming increasingly dissatisfied with their financially oriented performance measurement systems and they started incorporating non-financial measures in their performance measurement systems (Neely & Bourne, 2000). Neely (1998) cited a number of new demands on firms that contributed to the demise of financially based performance measurement systems. These new demands include: (1) the changing nature of work; (2) the increase in global competition; (3) the implementation of quality improvement management techniques; (4) the proliferation of national and international quality award (e.g. Malcolm Baldridge National Quality Award in USA, European Foundation of Quality Management (EFQM) in Europe, Deming prize in Japan); (5) the change in organizational roles; (6) the change in external demands; and (7) the increase in the power of information technology. These new demands on firms exposed the limitations of the financial measures. Performance measurement scholars hypothesized that the limitations of performance measures could be remedied by supplanting and complementing them with non-financial measures.

Non-financial measures can promote long term thinking that is essential for strategic research and development (R&D) decisions and product innovations. Financial measures encourage short-term thinking because short-term profit measures will not signal a decrease in firms' value when the needed capital investments or investments in innovation are reduced (Hayes & Abernathy, 1980; Kaplan 1986). On the contrary, short-term financial profitability measures will increase when strategic investments are reduced (Kaplan, 1986). Consequently, the heavy focus on the quarterly financial reports by the firms and the financial investment community may pressure managers into not only manipulating the figures (Eccles, 1991) but also into abandoning strategic investments in R&D that will create the new product pipeline for the future. Furthermore, under investment in innovative activities in the short run may result in lack of absorptive capacity (Cohen & Levinthal, 1990) which may then deprive firms in the long run from the capability to recognize new knowledge, assimilate it and use it to produce product innovations.

Kaplan & Norton (1992) note that income-based measures are better at measuring the consequences of past decisions than predicting future performance and non-financial measures may serve as leading indicators to future performance. Therefore, the use of non-financial measures may promote innovations because they could be used to measure the drivers of product innovations.

They could also provide strategic focus because traditional financial measures lack strategic focus as they fail to provide information on quality, responsiveness, and flexibility (Neely, 1998). Non-financial measures are also more practical than financial measures for the innovating firms because they allow managers the flexibility to design measures that are relevant to innovation[19] (e.g., new product launch times; number of ideas from outside R&D) without being constrained to quantifying them in financial terms (Davila, Epstein, & Shelton, 2006; Ghalayini & Noble, 1996; Kaplan, 1986).

However, the use of non-financial measures is not without cost (Dina, 2006). Neely (2006b) draws our attention to a number of challenges in using non-financial measures: (1) increased performance measurement diversity may lead to data overload; (2) non-financial measures are more difficult to design, measure, and collect compared to financial measures; and (3) if non-financial measures are not clearly defined, they may lead organizations to end up measuring the same concept in multiple different ways.

---

[19] Several authors offered frameworks/strategy maps/success maps to measure innovation. See for example: Adams, Bessant, & Phelps, 2006; Davila, Epstein, & Shelton, 2006: 143-178; Kaplan & Norton, 2004: 135-191; Neely, 2004; Neely, Filippini, Forza, Vinelli, & Hii, 2001.

Based on the above analysis, the following positive linear relationship is suggested:

Hypothesis 1. *There is a positive relationship between performance measurement diversity and product innovation.*

3.5.2 The Moderating Role of Performance Measurement Use on the Relationship between Performance Measurement Diversity and Product Innovation

As explained in the previous chapter, there are three differing perspectives on the effect of performance measurement systems on a firm's propensity to innovate with each one of them having empirical evidence to support its argument. The first perspective views performance measurement systems as constraining innovation because they impede creativity, experimentation, and search in firms (e.g. Amabile, 1997, 1998; Amabile et al, 1996; Burns & Stalker, 1961; Damanpour, 1991; Thompson, 1965). The second perspective views performance measurement systems as helping innovation because they trigger search, facilitate decision-making, and increase risk-taking (e.g. Greve, 2003b; Khandwalla, 1972; 1973; Miller & Friesen, 1982; Simons, 1978; 1991; 1995). The third perspective views performance measurement systems as having insignificant or little impact on innovation because they are used primarily for signalling and they do not affect innovation in high uncertainty environments (Abernethy & Brownell, 1997; Brownell, 1985; Hayes, 1977; Rockness & Shields, 1984; 1988).

One possible explanation of the contradiction in the empirical findings of these studies is that they generally ignore how performance measurement systems are used "style of use". As explained in the previous chapter, one useful conceptualisation of styles of use is Simons' (1995) diagnostic and interactive uses (styles). Therefore, I argue that one possible way of resolving the contradictory findings of these perspectives is by incorporating performance measurement use (diagnostic and interactive) as a moderating variable. Based on the behavioural theory of innovation, I argue that the impact of performance measurement on product innovation can be either positive or negative, depending upon the way performance measures are used.

However, instead of arguing in this section how hypotheses (H2a, H2b, H2c, H3a, and H3b) were advanced, I use the approach recommended by Bromiley & Johnson (2005) which entails moving beyond advancing aggregate predictions of the behavioural model of innovation to advancing hypotheses about the underlying mechanisms where a mechanism here refers to a plausible account of the process that causes a systematic relationship between variables.

Testing the underlying explanations allows me to understand if the mechanisms the research model postulates operate in the empirical world. In this study, two mechanisms are postulated that explain the research model: innovation search and organizational risk taking. Therefore, in the next sections, I will breakdown the aggregate hypotheses given below by showing that innovation search and organizational risk taking are the mediating mechanisms for the moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovations.

Hypothesis 2a. *Diagnostic use moderates the form of the relationship between performance measurement diversity and product innovation.*

Hypothesis 2b. *When levels of diagnostic use are high, performance measurement diversity will be positively related to product innovation.*

Hypothesis 2c. *When levels of diagnostic use are low, performance measurement diversity will be negatively related to product innovation.*

Hypothesis 3a. *Interactive use moderates the form of the relationship between performance measurement diversity and product innovation.*

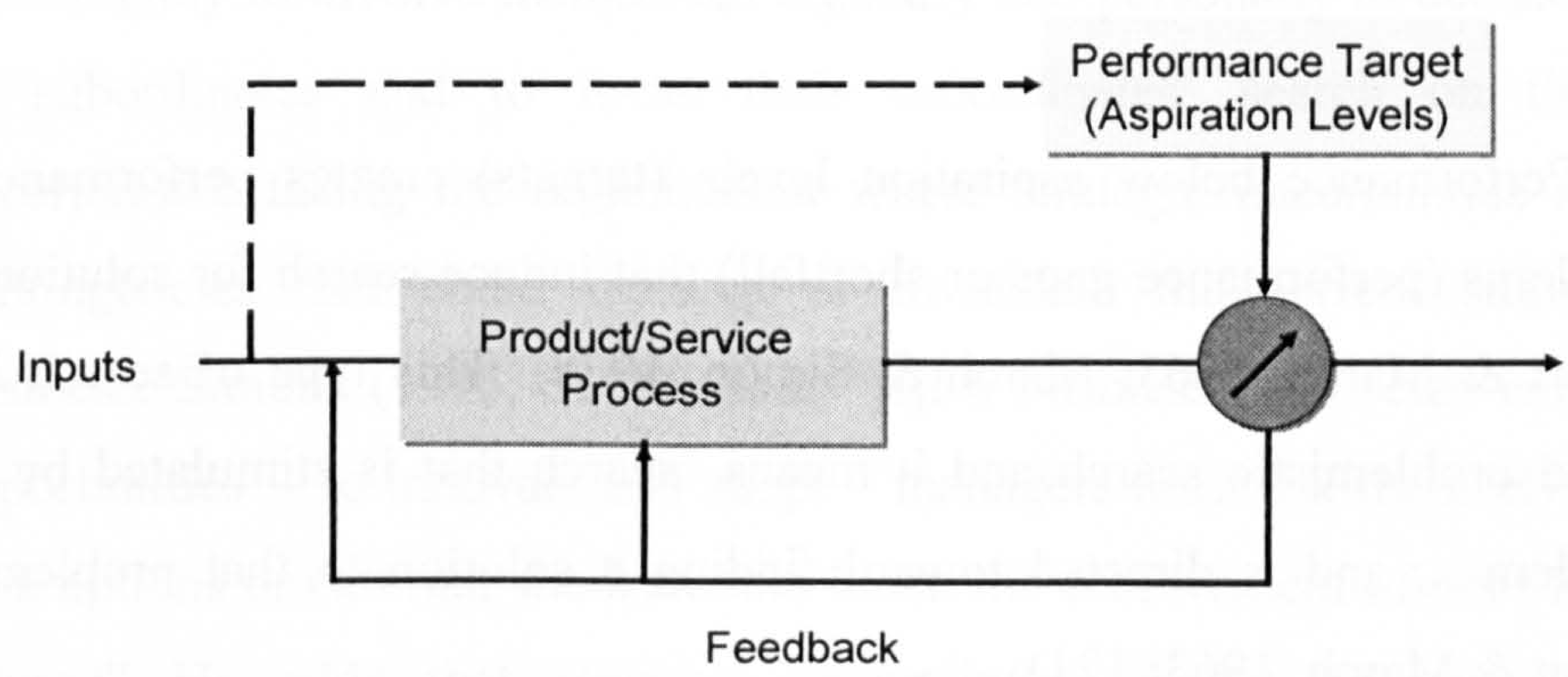
Hypothesis 3b. *When levels of interactive use are high, performance measurement diversity will be more negatively related to product innovation.*

3.5.3 The Moderating Role of Performance Measurement Use on the Relationship between Performance Measurement Diversity and Innovation Search

3.5.3.1 Diagnostic Use

Simons (1995) notes that managers use performance measurement systems diagnostically to monitor organizational outcomes and correct deviations from preset standards of performance. Figure 3-8 is a pictorial illustration of diagnostic control system.

Figure 3-8: Diagnostic control system (source: Simons, 1995)



Building on the behavioural theory of attention allocation formulated by Simon (1947), March & Simon (1958), Cyert & March (1963), and March (1988), Simons (1995) hypothesizes that managers use performance measures diagnostically to conserve attention. This assertion has found empirical support (Widener, 2005). March (1988) notes that time and attention are scarce resources because managers who are intendedly rational, but only limitedly so (Simon, 1955; 1956; 1997) do not have information on all the alternatives and their consequences and they have limited computational capacities. He adds these limitations make the organization of attention a central process out of which decisions arise because actions are determined less by choices among alternatives than by decisions with respect to search.

Diagnostic use encompasses the process of evaluating performance against their targets (aspiration levels). This evaluation process is a fundamental process in all theories of performance feedback-based experiential learning and behavioural model of innovation as illustrated in Figure 3-4 and Figure 3-6 and is the building block of behavioural theories of organizational search (March, 1988; 2006).

March (1988) notes that behavioural theories of organizational search are built on two ideas that proven remarkably durable:

(1) Organizations distinguish between meeting a performance (success) or not meeting it (failure) more than they distinguish between the various levels of success or failure (Cyert & March, 1963; March & Simon, 1958).

(2) Performance below aspiration levels (targets) creates performance problems (performance gaps or shortfall) that induce search for solutions (Cyert & March, 1963; March & Simon, 1958). This type of search is called problemistic search and it means "search that is stimulated by a problem ... and is directed toward finding a solution to that problem" (Cyert & March, 1963: 121).

Greve (2003b) notes that empirical "evidence suggests that low performance leads to increased R&D (Antonelli, 1989; Hundley, Jacobson, Park, 1996; Kaimen & Schwartz, 1982) and new ways of doing R&D (Bolton, 1993)."

Therefore, the higher levels of diagnostic use and the increased levels of performance measurement diversity are most likely to lead to increased performance gaps. This performance measurement intensity could result in increased innovation search (Greve, 2003b). Conversely, lower levels of diagnostic use will decrease the likelihood of performance gaps and will raise serious concerns about the value of performance measures as they increase in number. Therefore, the following hypotheses are proposed:

Hypothesis 4a. *Diagnostic use moderates the form of the relationship between performance measurement diversity and innovation search.*

Hypothesis 4b. *When levels of diagnostic use are high, performance measurement diversity will be positively related to innovation search.*

Hypothesis 4c. *When levels of diagnostic use are low, performance measurement diversity will be negatively related to innovation search.*

3.5.3.2 Interactive Use

Simons (1995) notes that managers use performance measurement systems interactively to involve themselves regularly and personally in decision activities of subordinates and to focus their subordinates' search on the strategic uncertainties facing the organization where strategic uncertainties refer to the contingencies that could threaten or invalidate the current strategy of the business. Simons (1995: 93-94) adds "[i]f a business is to seize emerging new opportunities - to innovate and adapt – managers must ask themselves ... What assumptions or external shocks could block the achievement of our vision for the future?" He adds that strategic uncertainties cannot be programmed and monitored on a diagnostic basis it must be managed interactively because they are in a constant state of flux.

Mangers use performance measures interactively to activate search and encourage experimentation and learning in their organizations (Simons, 1995). Therefore, interactive use clearly promote explorative search and is mainly driven by opportunities and diagnostic use promote problemistic search and is mainly driven by problems. However, an increase in organizational performance measurement diversity coupled with an increase in interactive use is likely to reduce innovation search because interactive use is costly. Using performance measures interactively is costlier than using them diagnostically for three reasons: economic, cognitive, and strategic. "In economic terms, interactive control systems are costly... In cognitive terms, the ability of individuals to process large amounts of disparate information is limited... In strategic terms, the primary reason for using a control system interactively is to activate learning and experimentation. Attempting to focus intensively on too many management control systems at the same time risks information overload, superficial analysis, a lack of perspective, and potential paralysis" (Simons, 1995: 115-116). A recent empirical study by Widener (2005), confirmed that interactive use of performance measurement systems consumes management's attention, which is consistent with the BTOF and levers of control framework.

March (1991) notes that organizations that "engage in exploration to the exclusion of exploitation are likely to find that they suffer the cost of experimentation without gaining many of its benefits. They exhibit too many underdeveloped new ideas and too little distinctive competence". Two recent empirical studies support the theorizing that increased level of interactive use (exploration) may be harmful as advocated by Levinthal & March (1993) and March (1991). The first study by Atuahene-Gima, Slater, Olson (2005) found that too much proactivity in seeking latent and emerging customers' needs negatively impacted product development performance. They attribute this finding to two possible causes: (1) many exploratory projects reduce the chances of building experience with a specific new knowledge base and (2) there may be cognitive barriers in processing new market information and the amount of time spent in its dissemination may generate substantial costs relative to the expected gains. The findings of the second study by Siggelkow & Rivkin (2006) was contrary to intuition that suggests that broadening low-level exploration and coupling it with higher- level coordination will broaden the exploration conducted by a firm as a whole and consistent with my argument that increased interactive use coupled with increased performance measurement diversity will negatively impact innovation search. They[20] state "[T]he more extensively low-level managers consider alternatives, the more effectively they can screen out options that do not serve their parochial interests. Ironically, then, more extensive exploration at a low level can reduce exploration for a firm as a whole and become a source of inertia". Thus, the following hypotheses are posited:

Hypothesis 5a. *Interactive use moderates the form of the relationship between performance measurement diversity and innovation search.*

Hypothesis 5b. *When levels of interactive use are high, performance measurement diversity will be more negatively related to innovation search.*

3.5.4 The Moderating Role of Performance Measurement Use on the Relationship between Performance Measurement Diversity and Risk Taking

3.5.4.1 Diagnostic Use

Risk theory (Kahneman & Tversky, 1979; Lopes, 1987; March & Shapira, 1987; 1992) predicts that risk preferences change in response to performance feedback. Performance below the aspiration level increases risk tolerance and performance above the aspiration level decrease risk tolerances. Using the same measurement instrument that I am using in this study, Singh (1986) found that performance gaps increase risk taking. As argued above, increased levels of performance measurement diversity and diagnostic use are more likely to produce performance gaps, which may lead to increased levels of risk taking.

[20] They also note, "[a]n enduring belief among management scholars and managers is that unleashing the low-level members of an organization to explore widely will broaden the exploration conducted by the organization as a whole. This sentiment has been expressed most passionately in the popular management literature. There, observers have argued that, when needing innovation, companies should "liberate" low-level managers (Peters, 1992), adopt "federal" structures in which "power belongs to the lowest possible point" (Handy, 1992: 62), and promote "activists" and "rebels" at low levels (Hamel, 2000). The freedom to explore that is granted to individual units, it is quietly implied, will aggregate to generate initiative and innovation for the entire organization."

Hypothesis 6a. *Diagnostic use moderates the form of the relationship between performance measurement diversity and organizational risk taking.*

Hypothesis 6b. *When levels of diagnostic use are high, performance measurement diversity will be positively related to organizational risk taking.*

3.5.4.2 Interactive Use

Organizational performance diversity is likely to increase risk taking by managers because it creates performance pressure on them to try to achieve their performance targets. Interactive use of performance measures is also likely to increase risk taking by managers because the personal involvement of managers and the continual dialogue across the organizations fosters a supportive organizational context (Gibson & Birkinshaw, 2004)

Although organizational performance measurement diversity and interactive use of performance measures are likely to increase risk taking when considered separately, the interplay between them is expected to decrease risk taking by managers. March & Shapira (1987) note that managers perceive opportunities that they have less control over their outcomes as riskier than opportunities that they have more control over their outcomes. This moderating effect of outcome control on risk taking in the loss and gain domains has not been theorized by the prospect theory (Kahneman & Tversky, 1979). This assertion found empirical support (Forlani, 2002). Since Simons (1995) asserts that managers use performance measures interactively to focus their subordinates on strategic uncertainties which are in state of influx, this leads me to hypothesize that an increase in the diversity of performance measures coupled with an increase in interactive use with will lead to decreased risk taking because managers will perceive less control over the outcomes of some of these strategic uncertainties and they will have less time to focus on them. Therefore, the following hypotheses are proposed:

Hypothesis 7a. *Interactive use moderates the form of the relationship between performance measurement diversity and organizational risk taking.*

Hypothesis 7b. *When levels of interactive use are high, performance measurement diversity will be more negatively related to organizational risk.*

3.5.5 The mediating role of innovation search

Innovation search is one form of organizational search. Katila & Ahuja (2002) note that "[o]rganizations engage in a wide variety of searches: they search for superior organizational designs (Bruderer & Singh, 1996), for optimal manufacturing methods (Jaikumar & Bohn, 1992), and for best ways to implement new innovations (von Hippel & Tyre, 1995)".

In this study, I do not differentiate between the different strategies of innovation search[21] because the behavioural theory of the firm (BTOF) predicts the extent of the search but not the direction of search (Bromiley, Miller, & Rau, 2001). Innovation search has many dimensions: scope, depth, age, science, and geography. Search scope (local versus distant) refers to the degree to which it entails the exploration of new knowledge (Katila & Ahuja, 2002). Local search refers to the firm's search for solutions in the neighborhood of its current expertise or knowledge (Helfat, 1994; Martin & Mitchell, 1998; Rosenkopf & Nerkar, 2001; Stuart and Podolny, 1996). Search depth refers to the degree to which it entails revisiting a firm's prior knowledge (Katila & Ahuja, 2002). The age dimension of the search addresses the question of how firms search over time (Katila, 2000). Science search refers to firm's search of the science base to overcome the limitations of their current technology base and geographic search refers to the firm's search across geographical boundaries to expand their technology base and to solve local technological problems (Ahuja & Katila, 2004). Empirical evidence suggests that innovation search results in product innovations (Ahuja & Katila, 2004; Cyert & March, 1963; Katila, 2000; Katila &

[21] Katila, 2000 provides an extensive literature review of innovation search strategies from an organizational theory perspective while Mahdi (2002; 2003) provides an extensive literature review of innovation search strategies from an innovation management perspective.

Ahuja, 2002; Rosenkopf & Nerkar, 2001; Stuart and Podolny, 1996). Therefore, the following hypotheses are proposed:

Hypothesis 8. *There is a positive relationship between innovation search and product innovations.*

Hypothesis 9. *Innovation search mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.*

3.5.6 The Mediating Role of Organizational Risk Taking

Organizational risk taking is an important building block for the behavioural theory of innovation because the adoption of solutions does not only depend on the availability of solutions and problems but also on the risk preferences of managers (Bromiley 1991; Greve, 2003a). Organizational risk taking is an important predictor of launching innovations because product innovations represent risky solutions and risky change to managers (Tushman & O'Rielly, 1997; Wan, Ong, Lee, 2000). Product innovations are risky for at least two reasons: (1) outcomes of product innovations are unknown and (2) product innovations are special cases of change, which make them susceptible to the forces of organizational inertia.

March (1997) notes that "[T]he process of innovation generates failures much more frequently than it does successes". Branscomb & Auerswald (2001) remind us that there are many more ways to fail than there is to succeed when it comes to innovation because unlike inventions where their returns may be reputational as well as monetary, returns from innovations are fundamentally linked to their market acceptance. Based on empirical evidence, Berggren & Nacher (2001) note that new-product failures are estimated to be as high as 95 percent in some industries, and the rate has not improved despite considerable academic research and management resources devoted to the issue.

Product innovations are also special case of change (Poole, Van de Ven, Dooley, & Colmes, 2000; Van de Ven, Angle, & Poole, 2000). Knight 1967 notes that "[i]nnovation process is a special case of the process of change in an organization. They differ only in the novelty of the outcome". To launch new product innovations, managers need to overcome organizational inertia that partially cancels out the greater probability of change. Hill & Rothaermel (2003) cites some of the following as forces causing organizational inertia:

1- Structural inertia. Hannan & Freeman (1977; 1984; 1989) define structural inertia as an organizational tendency to maintain its internal structure regardless of other factors or concerns. Structural inertia denotes the inability of an organization to adapt to its environment; basically, the stronger the pressures of structural inertia, the lower the organization's adaptive flexibility. They cited four factors limiting managerial discretion: organization's form constrains the options available for managers; scarcity of resources; the pattern of competition within and between populations limit choice; and bounded rationality.

2- Cognitive inertia. Top management teams tend to develop stable mental frames, logic and paradigms of how business should be run when their environment or their organizations are successful leading them to dismiss change when it is a necessity (Tripsas & Gaviti, 2000).

3- Lack of absorptive capacity. Organizations that lack the capability to recognize new knowledge, assimilate it and use it are less likely to change (Cohen & Levinthal, 1990).

4- Political pressure. The formation of coalition teams (Cyert & March, 1963) and the exercise of power in organizations to control scare resource lead to resistance to change (Pfeffer, 1992).

5- Institutional isomorphism. DiMaggio & Powell (1983) proposed three mechanisms generating isomorphic conformity, convergence around a single form, thereby reducing variation within industries & organizational fields. Coercive isomorphism results from both formal and informal pressures exerted on organizations by other organizations upon which they are dependent and from cultural expectations in the society within which the organization functions. Normative isomorphism results from the professionalisation of organizations. Professionalisation is the collective struggle of members of an occupation to define conditions and methods of their work, to control "the production of producers" and to establish a cognitive base and legitimization for their occupational autonomy. Mimetic isomorphism results from organizations copying or mimicking each other, often because of uncertainty.

To sum up, March & Shapira 1987 note "risk taking is valued, treated as essential to innovation and success." Risk taking is required to overcome organizational inertia. Hence, I advance the following hypotheses.

Hypothesis 10. *There is a positive relationship between organizational risk taking and product innovations.*

Hypothesis 11. *Organizational risk taking mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.*

3.5.7 Summary

Table 3-1 lists the research hypotheses postulated in the previous section.

Table 3-1: Summary of hypotheses

No.	Hypothesis
H1	There is a positive relationship between performance measurement diversity and product innovation.
H2a	Diagnostic use moderates the form of the relationship between performance measurement diversity and product innovation.
H2b	When levels of diagnostic use are high, performance measurement diversity will be positively related to product innovation.
H2c	When levels of diagnostic use are low, performance measurement diversity will

No.	Hypothesis
	be negatively related to product innovation.
H3a	Interactive use moderates the form of the relationship between performance measurement diversity and product innovation.
H3b	When levels of interactive use are high, performance measurement diversity will be more negatively related to product innovation.
H4a	Diagnostic use moderates the form of the relationship between performance measurement diversity and innovation search.
H4b	When levels of diagnostic use are high, performance measurement diversity will be positively related to innovation search.
H4c	When levels of diagnostic use are low, performance measurement diversity will be negatively related to innovation search.
H5a	Interactive use moderates the form of the relationship between performance measurement diversity and innovation search.
H5b	When levels of interactive use are high, performance measurement diversity will be more negatively related to innovation search.
H6a	Diagnostic use moderates the form of the relationship between performance measurement diversity and organizational risk taking.
H6b	When levels of diagnostic use are high, performance measurement diversity will be positively related to organizational risk taking.
H7a	Interactive use moderates the form of the relationship between performance measurement diversity and organizational risk taking.
H7b	When levels of interactive use are high, performance measurement diversity will be more negatively related to organizational risk taking.
H8	There is a positive relationship between innovation search and product innovations.
H9	Innovation search mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.
H10	There is a positive relationship between organizational risk taking and product innovations.
H11	Organizational risk taking mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.

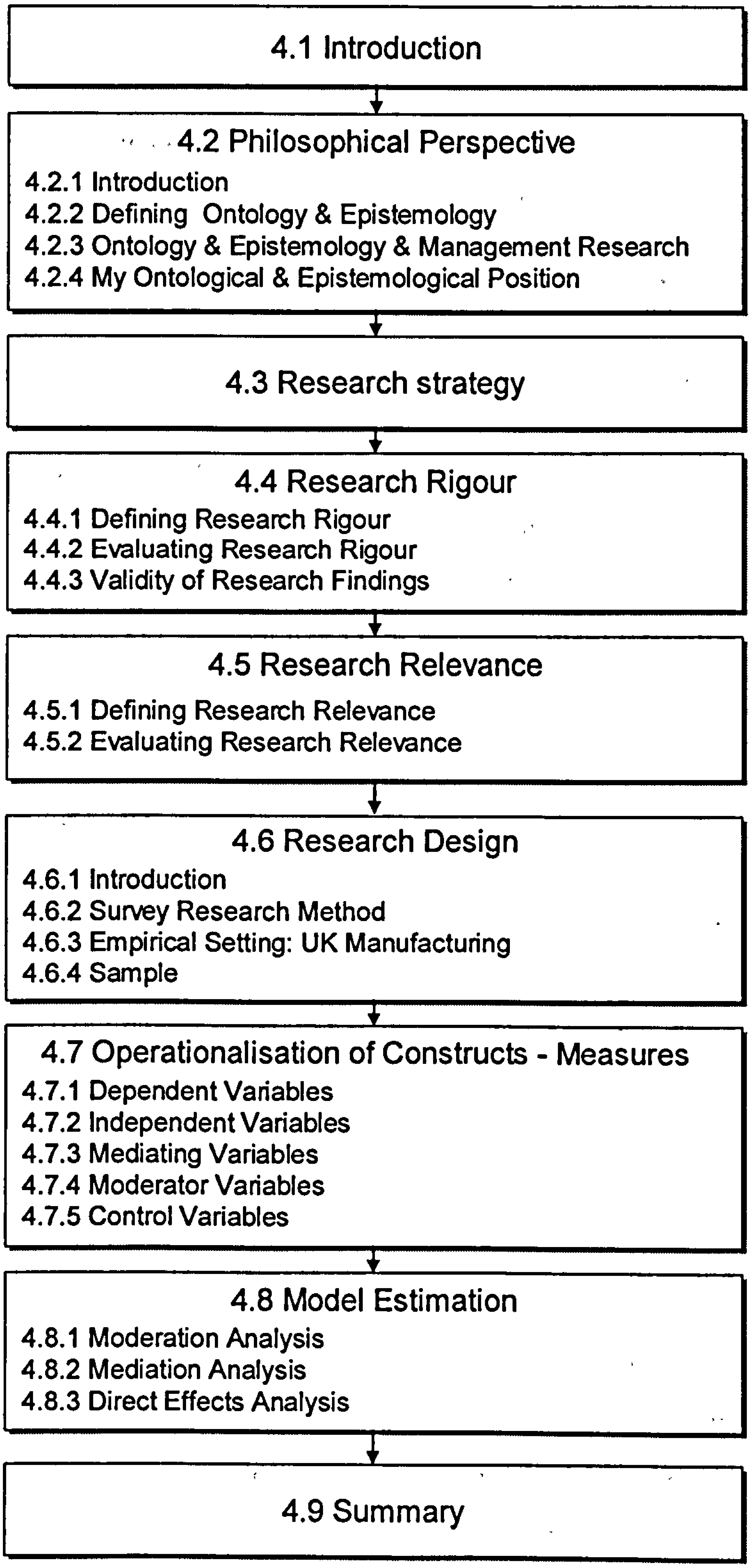
4 Research Methodology

4.1 Introduction

The previous chapter proposed eleven research hypotheses to answer the research questions based on the behavioural theory of innovation. This chapter examines the methodology of the study to ensure that its design is appropriate to provide answers to the research questions and to test the research hypotheses.

There are nine sections in this chapter, as illustrated in Figure 4-1. The first section introduces the chapter and subsequent sections. In the second section, I define the ontological and epistemological elements of philosophical perspectives and explain how they are linked to management research. Then, I discuss Campbellian realism, which forms the basis of my ontological and epistemological philosophical perspective. Since philosophical perspectives influence the logic of inquiry or research strategy, I explain the research strategy adopted in this study in the next section. Given that, research strategies provide logic or a set of procedures for answering research questions, in this section I show that the adopted deductive research strategy as prescribed by Popper (1972) is consistent with my Campbellian realist position and answers the research questions. The fourth and fifth sections of this chapter discuss two important criteria for evaluating academic research: rigour and relevance. I develop criteria for evaluating rigour and relevance, and then I apply them to this study to ensure that it conforms to the highest standards of rigour and relevance. The sixth section explains the cross-sectional probability sample, survey design employed in this study to answer the research questions. After reviewing several other research designs, I justify why the chosen design is more appropriate in this study. The seventh section operationalises the constructs used in this study. It divides the variables into five groups: dependent, independent, mediating, moderating, and control variables. The eighth section discusses the three statistical analyses that will be used to test the three types of hypotheses: moderation, mediation, and direct effect. The final section shows how the research design complies with all the survey research design criteria advanced by Malhotra & Grover (1998).

Figure 4-1: Outline of chapter 4



4.2 Philosophical Perspective

4.2.1 Introduction

I start by defining the key terms used to explicate my philosophical perspective followed by a discussion on the importance of understanding one's ontological and epistemological assumptions and how they relate to management research. In the third section, I state my ontological and epistemological position and show that they are consistent with my research strategy, design, and methods.

4.2.2 Defining Ontology and Epistemology

4.2.2.1 Ontology

"The root definition of *ontology* is the 'science or study of being'... *ontology* refers to the claims or assumptions that a particular approach to social inquiry makes about the nature of social reality – claims about what exists, what it looks like, what units make it up and how these units interact with each other" (Blaikie, 1993: 6).

4.2.2.2 Epistemology

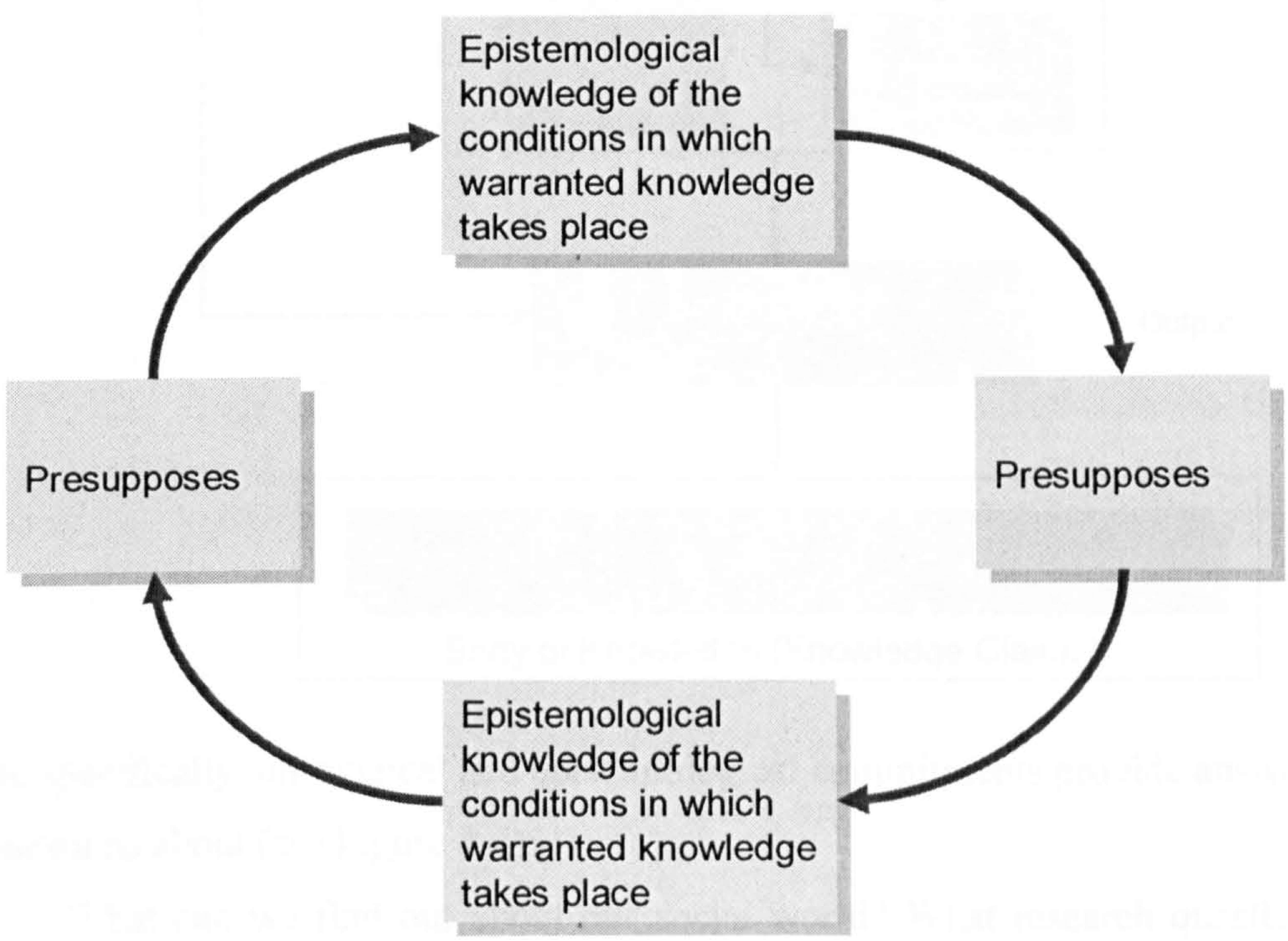
"The root definition of *epistemology* is the 'theory of science of the methods or ground of knowledge'... *epistemology* refers to the claims or assumptions made about the ways in which it is possible to gain knowledge of this reality, whatever it is understood to be; claims about what exists may be known" (Blaikie, 1993: 6-7)

Johnson and Duberley (2000: 2) break the word into its constituent Greek words: "*'episteme'* which means knowledge or 'science'; and *'logos'* which means 'knowledge', 'information, 'theory', or 'account'." Epistemology is the theory of knowledge about knowledge (Blaikie, 1993: 7; Johnson and Duberley, 2000: 2-3). Stated differently by Johnson and Duberley (2000: 2-3), "epistemology is the study of criteria by which we can know what does and does not constitute warranted, or scientific, knowledge."

4.2.3 Ontology and Epistemology and Management Research

All management researchers must start from a philosophical position that is either implicit or explicit that is known or not known. There are two core elements that define the philosophical position: ontology and epistemology (Blaikie, 1993: 6). However, in defining our epistemological assumptions, we are faced with the paradox of circularity. As noted by Johnson & Duberley (2000: 3-4), "[T]here are no secure or incontestable foundations from which we can begin any consideration of our knowledge of knowledge - rather what we have are competing philosophical assumptions about knowledge that lead us to engage with management and organizations in particular ways... [E]pistemology confronts a fundamental problem of circularity, from which it can not escape, in that any theory of knowledge (i.e. any epistemology) presupposes knowledge of the conditions in which knowledge takes place"(See Figure 4-2).

Figure 4-2: The circularity of epistemology (Source: Johnson & Duberly, 2000: 4)

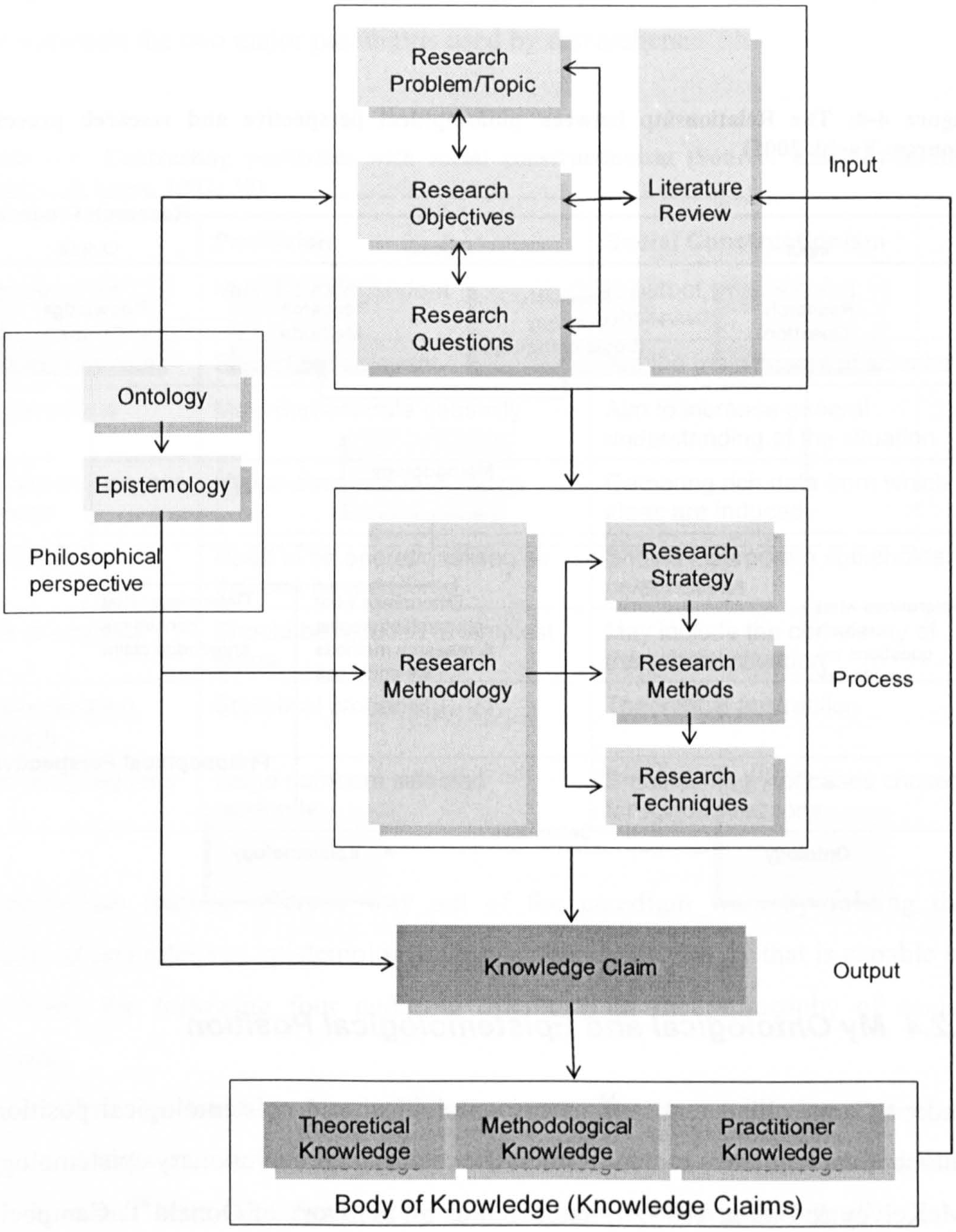


As illustrated in Figure 4-3, our ontological and epistemological positions impact the input to the research process, the process of research itself and the outcome of research. It is worth noting that methodology and methods have different meanings. Blaikie (1993) defines methodology as dealing with how research is done, or should be done, and to the critical evaluation alternative research strategies and methods and defines the methods as the actual techniques or procedures used to gather and analyse data related to some research question or hypothesis. Put differently, methodology refers to the theory of methods and methods deal with the execution of the research project.

In addition, it is worth noting that management research stocks its knowledge in four areas (Tranfield, 2004; Tranfield & Starkey, 1998):

1- Disciplinary: (psychology, sociology, economics, anthropology, …etc.)
2- Functional: (Human resources, operations, marketing, …etc)
3- Thematic: (Change management, innovation management, technology management, …etc)
4- Sectoral : (Financial services, information communications technologies, … etc.)

Figure 4-3: Management research framework (Source: Yaghi, 2005)



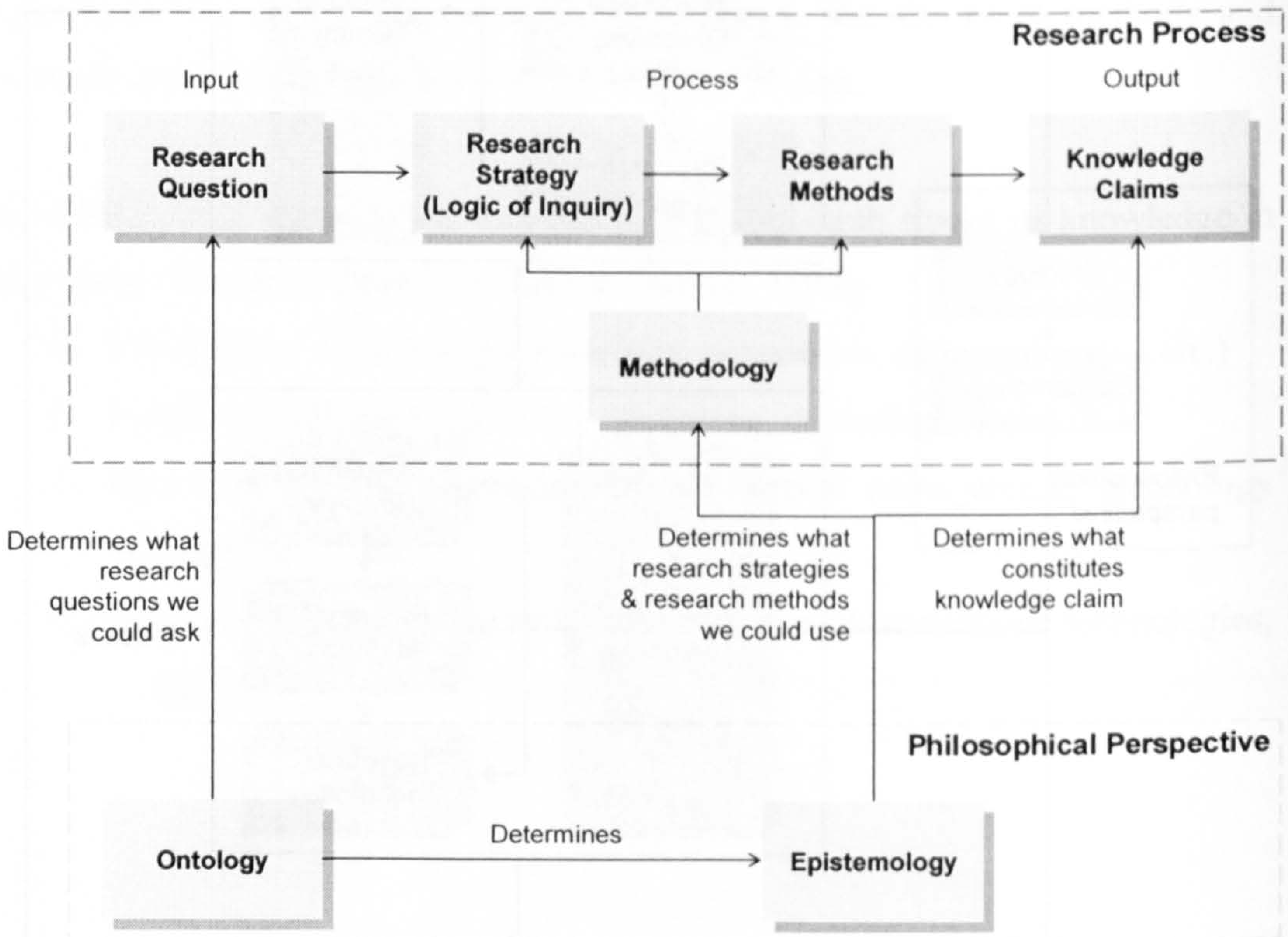
More specifically, ontological and epistemological commitments provide answers to questions about (see Figure 4-4):

1- What can we find out about our social world? What research questions could we ask?
2- What are the research strategies and methods required to answer the research questions?

3- What constitutes a knowledge claim?

4- How does scientific knowledge progress and accumulates?

Figure 4-4: The Relationship between philosophical perspective and research process (Source: Yaghi, 2005)



4.2.4 My Ontological and Epistemological Position

I adopt Campbellian realism[22] as my ontological and epistemological position. Campbellian realism is rooted in scientific realism and evolutionary epistemology (McKelvey & Baum, 1999: 5) and it builds on the work of Donald T. Campbell. McKelvey (1999: 383-384) argues that Campbellian realism offer a way out of paradigm wars[23] by providing a dynamic objectivist view of management research that denies neither the epistemological dynamics uncovered by historical

[22] This term is first coined by (McKevely 1999: 383).

[23] McKelvy & Baum (1999) and Moldoveany & Baum (2002: 733) discuss the paradigm wars and they point out that these paradigm wars refer to the on going debates between advocates of normal science and postmodernism (positivists and relativists approaches loosely defined), even though philosophers abandoned both decades ago (Suppe, 1977). For an example of the positivist view, refer to (Pfeffer, 1993, 1995) and for a relativist view refer to (van Maanen 1995a; 1995b).

relativists such as (Kuhn, 1962) nor the sociology of knowledge developed by interpretists and social constructionists such as (Burrell & Morgan, 1979). Table 4-1 contrasts the two major paradigms used by researchers.

Table 4-1: Contrasting positivism with social constructionism (Source: Easterby-Smith, Thorpe, & Lowe, 2002: 30)

	Positivism	Social Constructionism
The observer	Must be independent	Is part of what is being observed
Human interests	Should be irrelevant	Are the main drivers of science
Explanations	Must demonstrate causality	Aim to increase general understanding of the situation
Research progress through	Hypotheses and deductions	Gathering rich data from which ideas are induced
Concepts	Need to be operationalised so that can be measured	Should incorporate stakeholder perspectives
Unit of analysis	Should be reduced to simplest terms	May include the complexity of the 'whole' situation
Generalization through	Statistical probability	Theoretical abstraction
Sampling requires	Large numbers selected randomly	Small numbers of cases chosen for specific reasons

Campbellian realism offers a way out of the paradigm wars by folding the positivist and relativist epistemologists into single epistemology that is capable of resolving the following four enduring dilemmas in the philosophy of social science:

1. How to build a postpositivist science that maintains the "goal of objectivity" in science (Campbell, 1974) without forcing metaphysical terms[24] out of theories in favor of operationalist observable terms?
2. How to develop a selectionist evolutionary epistemology that does not steer scientist toward Comtean positivism, instrumentalism, naïve realism,

---

[24] "Operational terms are not the same as observation terms (making up observation language) in logical positivism. An operational term is the actual measure—a "number" coming from a mercury barometer vs. one from an aneroid barometer. An observation term accessible to the human senses could be measured any one of several competing operational terms" (McKelvey, 2002b: 894).

or operationalism at the expense of theory terms less detectable or more metaphysical in nature (1974)?

3. How to build an objectivist postrelativist epistemology that incorporates the dynamic of science changing over time without abandoning the goal of objectivity (1988)?
4. How to develop an objectivist epistemology while remaining sensitive to the differing perceptions, interpretations, and social constructions of individual scientists and scientific communities (1988)?

(McKelevy, 1999: 383, Campbell name is not in the original text)

The resolution[25] of the dilemmas and the resulting Campbell's epistemology involves: (McKelvey, 2002a)

1- Dealing with metaphysical terms.
2- Objectivist empirical investigation
3- Recognition of socially constructed meanings of terms
4- A dynamic process by which a multiparadigm discipline might reduce to fewer but more significant theories.

Campbell defines a *critical, hypothetical, corrigible, scientific realist selectionist evolutionary* epistemology that have the following characteristics: (McKelvey, 1999: 403)

1- A scientific realist postpositivist epistemology that maintains the goal of objectivity in science without excluding metaphysical terms and entities.
2- A selectionist evolutionary epistemology governing the winnowing out of less probable theories, terms, and beliefs in the search for increased verisimilitude[26] that may do so without the danger of systematically replacing metaphysical terms with operationalisms.

---

[25] McKelvey (1999: 384-402) offers a comprehensive discussion of the resolutions of the dilemmas. I opted not to include it in the thesis because it requires lengthy discussion of the backgrounds of various philosophical positions.

[26] Verisimilitude is the same as truthlikeness. Because philosophers moved away from an absolutist view of theoretical statements as either True or False (that is, the logical positivist ' verification of theories), toward Carnap's testability and evolutionary epistemology, Popper (1979) developed the idea of verisimilitude. As poorer theories are winnowed out in selectionist fashion, theories with improved verisimilitude remain. (McKelvey, 2002b: 897).

3- A postrelativist epistemology that incorporate the dynamics of science without abandoning the goal of objectivity.

4- An objectivist selectionist evolutionary epistemology that includes as part of its path toward increased verisimilitude the inclusion of, but also the winnowing out of, the more fallible, individual interpretations and social constructions of the meaning of theory terms comprising theories purporting to explain and objective external reality.

The scientific realist ontological position is the first keystone of the Campbellian realism. It states that a mind-independent reality exists. Bhaskar (1978) describes the stratified realist ontology in the following statement. "...Real structures exist independently of and are often out of phase with actual patterns of events. Indeed it is only because of the latter that we need to perform experiments and only because of the former that we can make sense of our performances of them. Similarly it can be shown to be a condition of the intelligibility of perception that events occur independently of experiences. And experiences are often (epistemically speaking) 'out of phase' with events - e.g. when they are misidentified. It is partly because of this possibility that the scientist needs scientific education or training. Thus I will argue that what I will call the domains of the real, the actual and empirical are distinct. This is represented in Table 4-2.

Table 4-2: Domain of reality (Source: Bhaskar, 1978)

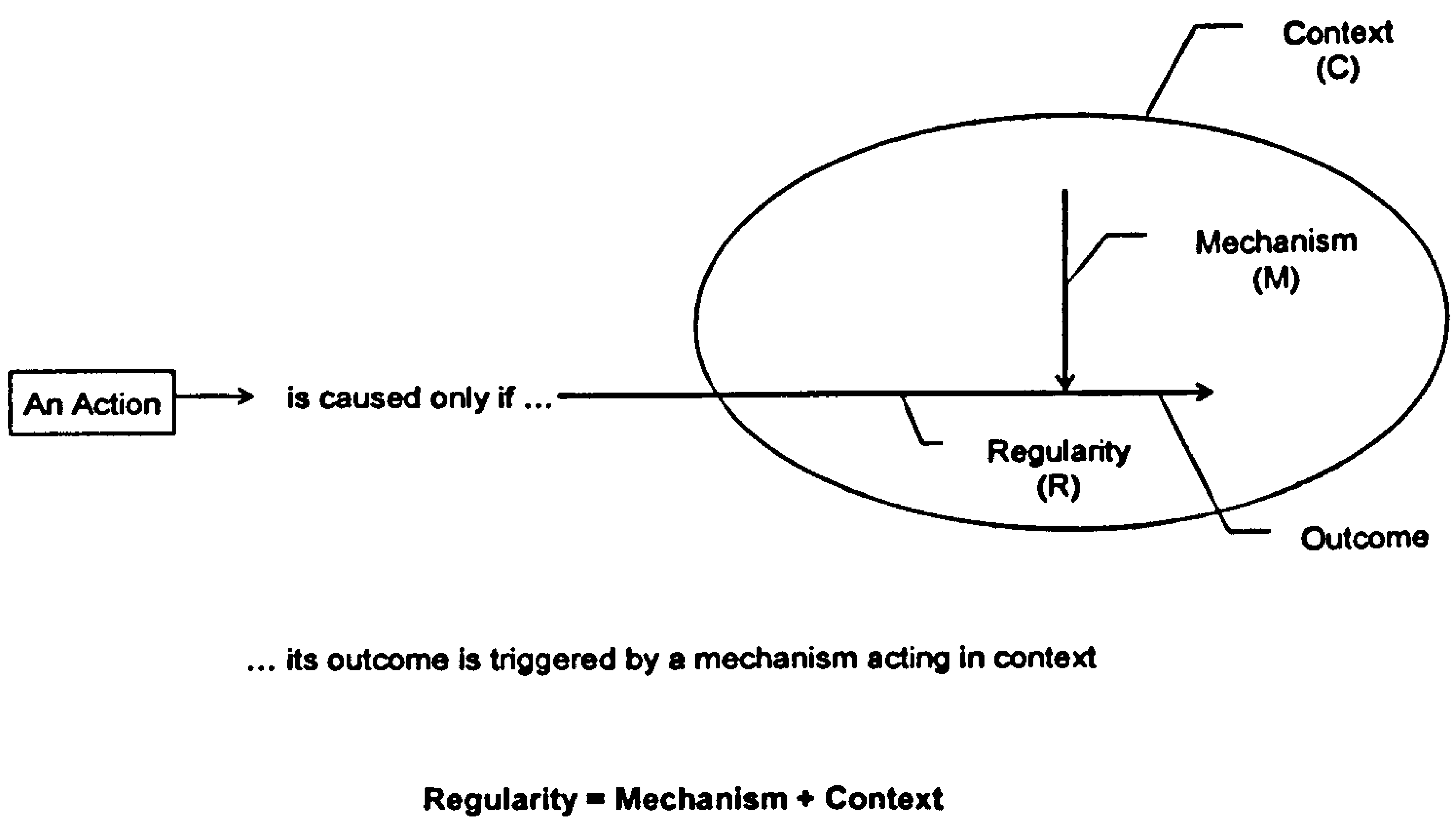
	Domain of Real	Domain of Actual	Domain of Empirical
Mechanisms	√		
Events	√	√	
Experiences	√	√	√

The real basis of causal laws is provided by the generative mechanisms of nature. Such generative mechanisms are, it is argued, nothing other than the ways of acting of things. And causal laws must be analysed as their tendencies.

Tendencies may be regarded as powers or liabilities of a thing which may be exercised without being manifest in any particular outcome."

Furthermore, Pawson & Tilley (1997) posit that "the objective of realist inquiry is to explain social 'regularities', 'rates', 'associations', 'outcomes', 'patterns'. They summarize the basic logic of realist explanation in the following statement: "The basic task of social inquiry is to explain interesting, puzzling, socially significant regularities (R). Explanation takes the form of positing some underlying mechanism (M) which generates the regularity and thus consist of proposition about how the interplay between structure and agency has constituted the regularity. Within realist investigation there is also investigation of how the workings of such mechanisms are contingent and conditional, and thus only fired in particular local, historical, or institutional context (C)." In other words, regularity = mechanism + context as illustrated in Figure 4-5.

Figure 4-5: Ingredients of Realist Social Explanation (Source: adapted from Pawson & Tilley, 1997)



To summarize, scientific realism holds that there are (McKelvey, 2002b)

1. real entities in the world "out there"
2. that exist independently of our perception, experience, or knowledge of them, and

3. that they have properties and relationships that are independent of the concepts or language we use to describe them.
4. Metaphysical and real terms, holding that underlying Generative mechanisms or causes not directly accessible to the human senses are nevertheless real, and not to be relegated to the scientific dustbin.

Evolutionary epistemology is the second keystone of the Campbellian realism. "Evolutionary epistemology ... holds that the dynamics of science are best interpreted as an evolutionary Darwinian selection process in which a less fallible version of truth results as the more fallible individual interpretations of facts and expositions of theory and social construction of facts by scientific communities of real world (causal) processes, are winnowed out over time" (McKelvey, 2002: 890).

Mckelvey (1999: 398) points out that Campbell credited Popper for introducing and developing a Darwinian selectionist evolutionary epistemology. Popper (1979) states "The growth of our knowledge is the result of a process closely resembling what Darwin called "natural selection"; that is, *the natural selection of hypotheses*; our knowledge consists, at ever moment, of those hypotheses which have shown their (comparative) fitness by surviving so far in their struggle for existence; a competitive struggle which eliminates those hypotheses which are unfit. Quoted in (Mckelvey, 1999: 398)

4.3 Research Strategy

Research strategy is the logic of inquiry. Research strategies provide logic or a set of procedures, for answering research questions (Blaikie, 2000: 24). The choice of research strategies determines the choice of methods and techniques that could be used to answer the research questions.

Consistent with my Campbellian realist philosophical position, I use a deductive research strategy to answer my research questions as prescribed by Popper (1972: 32-33)

"From a new idea, put up tentatively, and not yet justified in any way – an anticipation, a hypothesis, a theoretical system, or what you will – conclusions are drawn by means of logical deduction. These conclusions are then compared with one another and with other relevant statements, so as to find what logical relations (such as equivalence, derivability, compatibility, or in compatibility) exist between them.

We may if we like distinguish four different lines along which the testing of a theory could be carried out. First, there is the logical comparison of the conclusions among themselves, by which the internal consistency of the system is tested. Secondly, there is the investigation of the logical form of the theory, with the object of determining whether it has the character of an empirical or scientific theory, or whether it is, for example tautological. Thirdly, there is the comparison with other theories, chiefly with the aim of determining whether the theory would constitute a scientific advance should it survive our various tests. And finally, there is the testing of the theory by way of empirical applications of the conclusions which can be derived from it.

The purpose of this last kind of test is to find out how far the new consequences of the theory – whatever may be new in what it asserts – stand up to the demands of practice, whether raised by purely scientific experiments, or by practical technological applications. Here too the procedure of testing turns out to be deductive. With the help of other statements, previously accepted, certain singular statements – which we may call 'predictions' are deduced from the theory; especially predictions that are easily testable or applicable. From among these statement, those are selected which are not derivable from the current theory, and more specially those which the current theory contradicts. Next we seek a decision as regards these (and other) derived statements by comparing them with the results of practical applications and experiments. If this decision is positive, that is, if the singular conclusions turn out to be acceptable, or *verified*, then the theory has, for the time being, past its test: we have found no reason to discard it. But if the decision is negative, or in other words, if the conclusions have been *falsified*, then their falsification also falsifies the theory from which they were logically deduced.

It should be noticed that a positive decision can only temporarily support the theory, for subsequent negative decisions may always overthrow it. So long as a theory withstands detailed and severe tests and is not superseded by another theory in the course of scientific progress, we may say that it has 'proved its mettle' or that it is '*corroborated*' by past experience."

4.4 Research Rigour

4.4.1 Defining Research Rigour

The criteria for assessing rigour in scholarly management research have been discussed extensively (Varadarjan, 2003). For example, Academy of Management (2006) in its Code of Ethical Conduct defines research rigour as the "careful design, execution, analysis, interpretation of results, and retention of data." Shrivastava (1987) advanced three criteria for the evaluation of the rigour of research projects[27]: conceptual adequacy of the framework that guides the research project; methodological rigour; and accumulated empirical evidence. Varadarajan (2003) offered two criteria: conceptual rigour – quality of conceptual development and methodological rigour – quality of empirical research.

4.4.2 Evaluating Research Rigour

Using the criteria advanced by Shrivastava (1987) and Varadarajan (2003), I construct Table 4-3 to assess the rigour in my research. The third criterion proposed by Shrivastava was dropped because it is more relevant in evaluating research programs[28].

[27] The unit of analysis in (Shrivatava, 1987) is the research program.
[28] Shrivastava applied the three criteria to evaluate the rigour of research programs rather than individual research projects in strategic management.

Table 4-3: Criteria for assessing rigour of research projects (Source: adapted from Shrivastava, 1987 and Varadarajan, 2003)

Property	Description
Conceptual Rigour	Is the research well grounded in a basic discipline?
	Does the research use a conceptual framework consistent with existing theories in the field?
	Does the research review the current literature?
	Does the research pay attention to definitional issues?
	Does the research use evidence to support position – conceptual reasoning underlying conceptual model and hypotheses?
	Does the research objectively treat complementing and competing perspectives?
Methodological Rigour	Does the research use the appropriate research design to answer the research questions?
	Does the research pay attention to the measurement related issues – construct operationalisation, validity, and reliability?
	Does the research use the appropriate data to empirically examine the research questions?
	Does the research use the appropriate methods of analysis/statistical procedures to empirically examine research questions?
	Does the research accurately thoroughly report the results and procedures leading to the results?
	Does the research report the reliability and validity of the empirical findings?

4.4.3 Validity of Research Findings

4.4.3.1 Defining Validity

Consistent with my ontological and epistemological philosophy of evolutionary epistemology and critical realism ontology, I adopt Campbell's position on validity. Campbell's work on validity is documented in Campbell, 1957; Campbell & Stanley, 1966; Cook & Campbell, 1979; Cook, Campbell, & Peracchio, 1990; and Shadish, Cook, & Campbell, 2001.

Validity is the sine qua non for empirical research (Campbell & Stanley, 1966: 5). Shadish et al (2001: 34) uses the term "validity to refer to the approximate[29]

[29] Cook & Campbell (1997: 37) note that "we should always use the modifier 'approximately' when referring to validity, since one can never know what is true. At best one can know what has

truth of inference... Validity is a property of inferences. It is not property of designs or methods, for the same design may contribute to more or less valid inferences under different circumstances."

Assessing validity always entails fallible human judgment because we evaluate evidence supporting our inference not only based on the empirical findings "methodological rigour" but also on the consistency of these empirical findings with the current literature base "conceptual rigour" (Shadish et al, 2001: 34).

Campbell and associates (1957; 1966; 1979; 1990; 2001) elaborated a theory of validity that consisted of typology of validities and threats to validities. The four types of interrelated validities are: construct validity, internal validity, external validity and statistical conclusion validity. The threat of validities are "specific reasons why we can be partly or completely wrong when we make an inference about covariance, about causation, about constructs, or about whether the causal relationship holds over variations in persons, settings, treatments, and outcomes" Shadish et al, 2001: 30). I address the threats of validities that are relevant to the chosen research design throughout the sections that address the design, execution, analysis, interpretation, and retention of data in this thesis.

4.4.3.2 Construct Validity

"Construct validity[30] involves making inferences from the sampling particulars of a study to the higher-order constructs they represent" (Shadish et al, 2001: 65). As noted by Kerlinger & Lee (2000: 40), "The terms "concept" and "construct" have similar meaning, yet there is an important distinction. A *concept* expresses an abstraction formed by generalization from particulars ... A construct is a concept. It has the added meaning, however, of having been deliberately and consciously invented or adopted for a special scientific purpose." Furthermore,

---

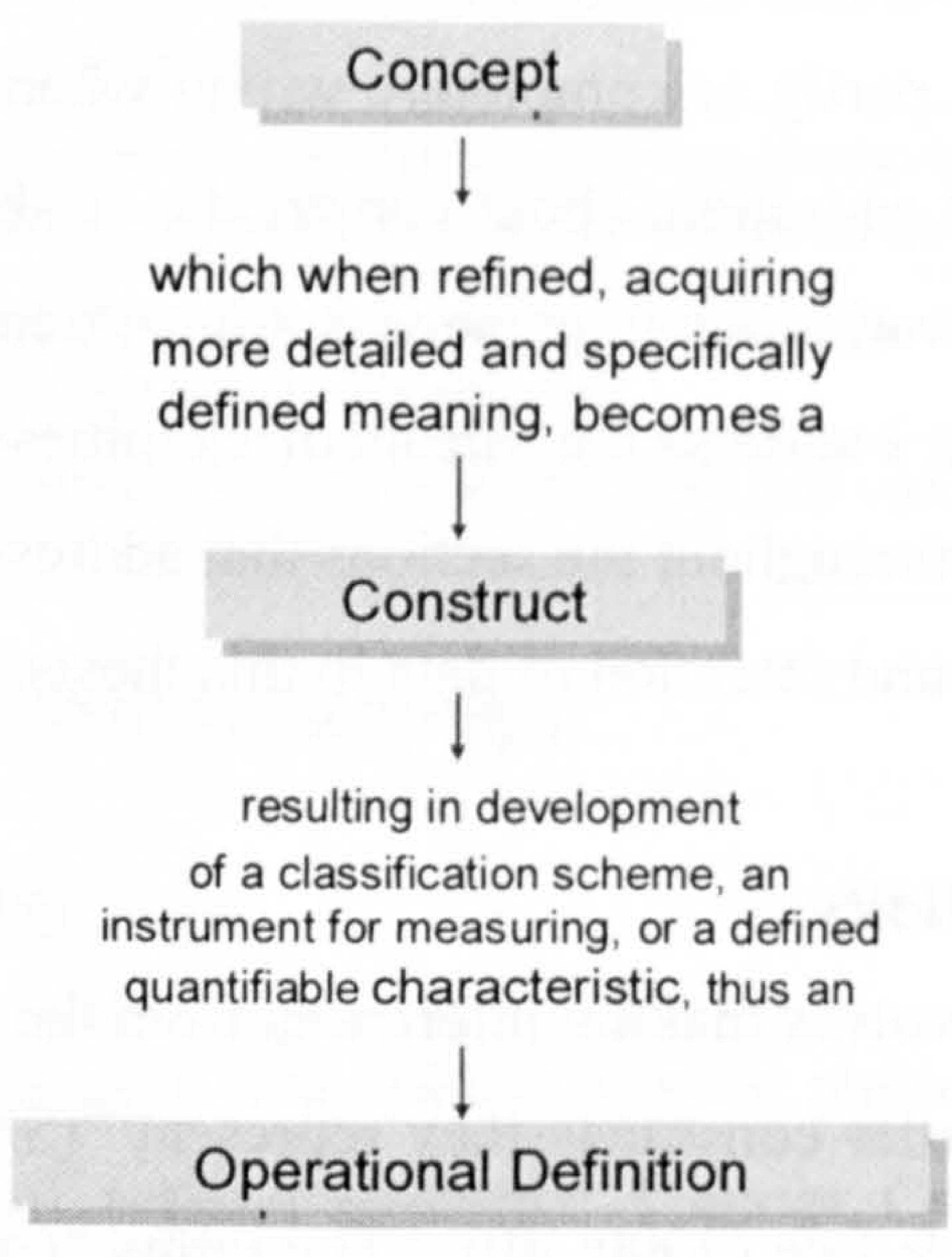
not been ruled out as false. Hence, when I use the terms valid or invalid in this thesis, they should always be understood to be prefaced by the modifiers "approximately" or "tentatively".

[30] In the initial validity topology reported in (Campbell & Stanley, 1966: 5), construct validity was part of external validity. Cook & Campbell (1979: 38) spun off construct validity as a separate type.

they highlight that scientists consciously and systematically use constructs in two ways: (1) it enters into theoretical schemes and is related in various ways to other constructs and (2) it is so defined and specified that it can be measured.

Constructs reside in the theoretical domain and need to operationalised, so we could measure them. Therefore, an "operational definition[31] assigns meaning to a construct or a variable by specifying the activities or "operations" necessary to measure it and evaluate the measurement" (Kerlinger & Lee, 2000: 41). See Figure 4-6 for an illustration of the relationship between concepts, constructs, and operational definitions.

Figure 4-6: Relationship between concepts, constructs, and operational definitions (Source: Black, 1999: 36)

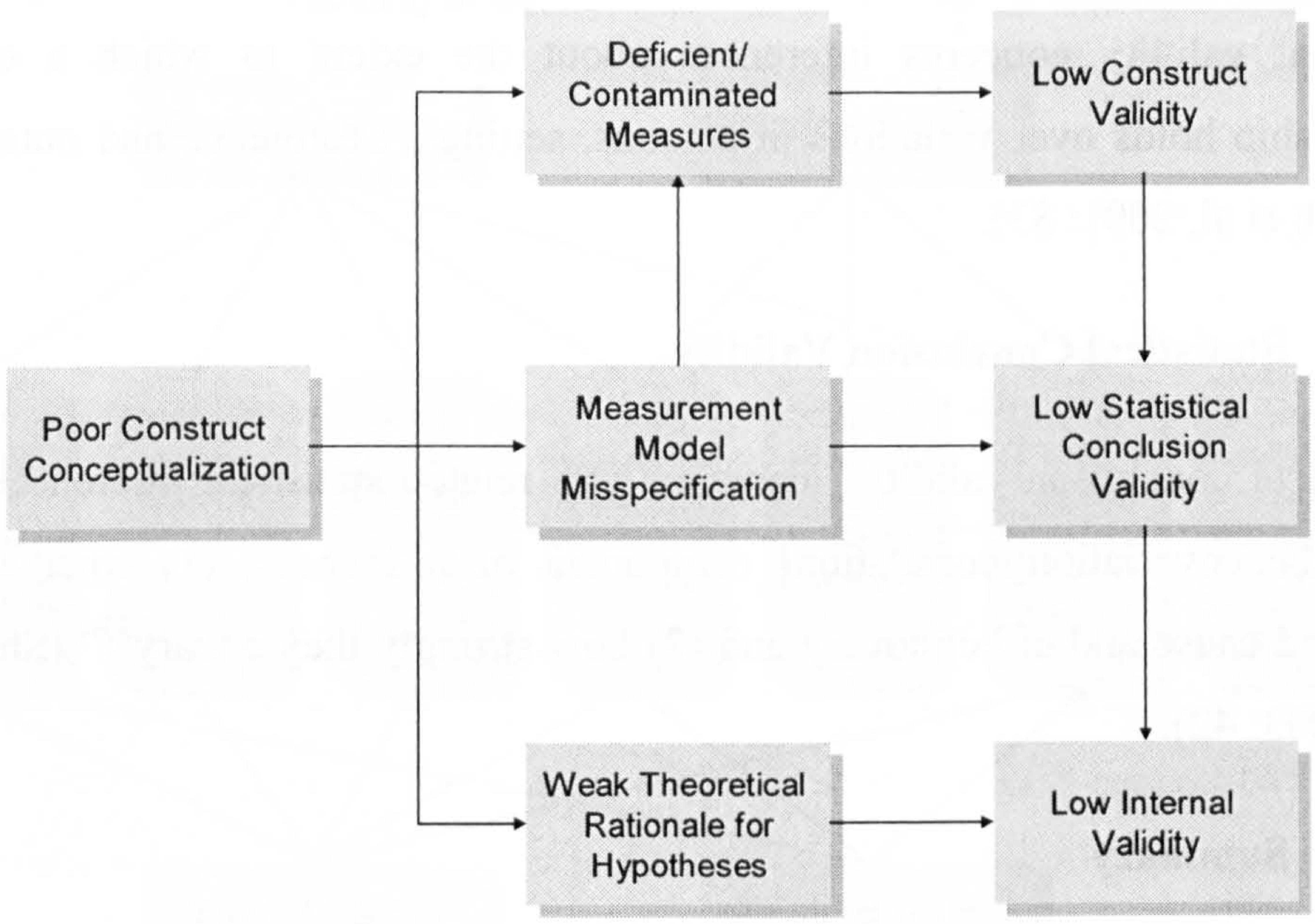


Recently, construct validity has received increased attention in various streams of business research literature because of its importance in drawing valid inferences and its effects on other types of validities. (For example, Boyd, Gove, & Hitt, 2005 in strategic management; O'Leary-Kelly & Vokurka, 1998 in operations management; Jarvis, MacKenzie, & Podsakoff, 2003 in marketing). In Table 4-4, McKenzie (2003) illustrates the devastating impact of poor construct validity.

---

[31] Kerlinger & Lee (2000: 41) offers a second definitional type called constitutive definition in which a construct is defined using other constructs.

Table 4-4: The consequences of poor construct conceptualisation (Source: MacKenzie, 2003)



4.4.3.3 Internal Validity

Shadish et al (2001: 53) use the term internal validity "to refer to the inferences about whether observed covariation between A and B reflects a causal relationship from A to B in the form in which the variables were manipulated or measured."

Internal validity was first discussed by Campbell (1957) making it the oldest between the four types. Moreover, Campbell & Stanley (1966: 5) called it the sine qua none of validities. Campbell et al (1990: 493) note that three conditions must be met for concluding that two variables are causally related and that the direction of causation is from A to B. First, a cause[32] must precede an effect[33] in time. Second, the effect and cause must covary. Third, there are no plausible alternative explanations of B other than A.

[32] A cause is also referred to as a treatment (Cook & Campbell, 1979: 4).
[33] An effect is also referred to as an outcome (Cook & Campbell, 1979: 4).

4.4.3.4 External Validity

"External validity concerns inferences about the extent to which a causal relationship holds over variations in persons, settings, treatments, and outcomes (Shadish et al, 2001: 83).

4.4.3.5 Statistical Conclusion Validity

"Statistical conclusion validity[34] concerns two related statistical inferences that affect the covariation[/correlation] component of inferences: (1) whether the presumed cause and effect covary and (2) how strongly they covary[35]" (Shadish et al, 2003: 42).

4.4.3.6 Summary

Table 4-5 summarizes the four types of validities and the corresponding questions they address.

Table 4-5: Types of Validities (Source: Adapted from Shadish, et al: 38-39)

Type of Validity	Description	Question
Construct Validity	The validity of inferences about the higher order constructs that represent sampling particulars.	Which general constructs are involved in the persons, settings, treatments, and observations used in the experiment?
Internal Validity	The validity of inferences about whether observed covariation between A (the presumed treatment) and B (the presumed outcome) reflects a causal relationship from A to B as those variables were manipulated or measured.	Is the covariation causal, or would the same covariation have been obtained without the treatment?
External Validity	The validity of inferences about whether the cause-effect relationship holds over variation in persons, settings, treatment variables, and measurement variables.	How generalizable is the locally embedded causal relationships over varied persons, treatments, observations, and settings?
Statistical Conclusion Validity	The validity of inferences about the correlation (variation)	How large and reliable is the covariation between the presumed cause and effect?

Figure 4-7 maps the four types of validities into the research phases.

---

[34] In the initial validity topology reported in (Campbell & Stanley, 1966: 5), statistical conclusion validity was part of internal validity. Cook & Campbell (1979: 37) spun off statistical conclusion validity as a separate type.

[35] The emphasis on the magnitude of an effect has been strengthened in Shadish et al's (2001) treatment of external validity. The earlier treatment of statistical conclusion validity (Cook & Campbell, 1979: 39-50) emphasized statistical significance.

Figure 4-7: Types of validities mapped into research design (Source: Black, 1999: 58)

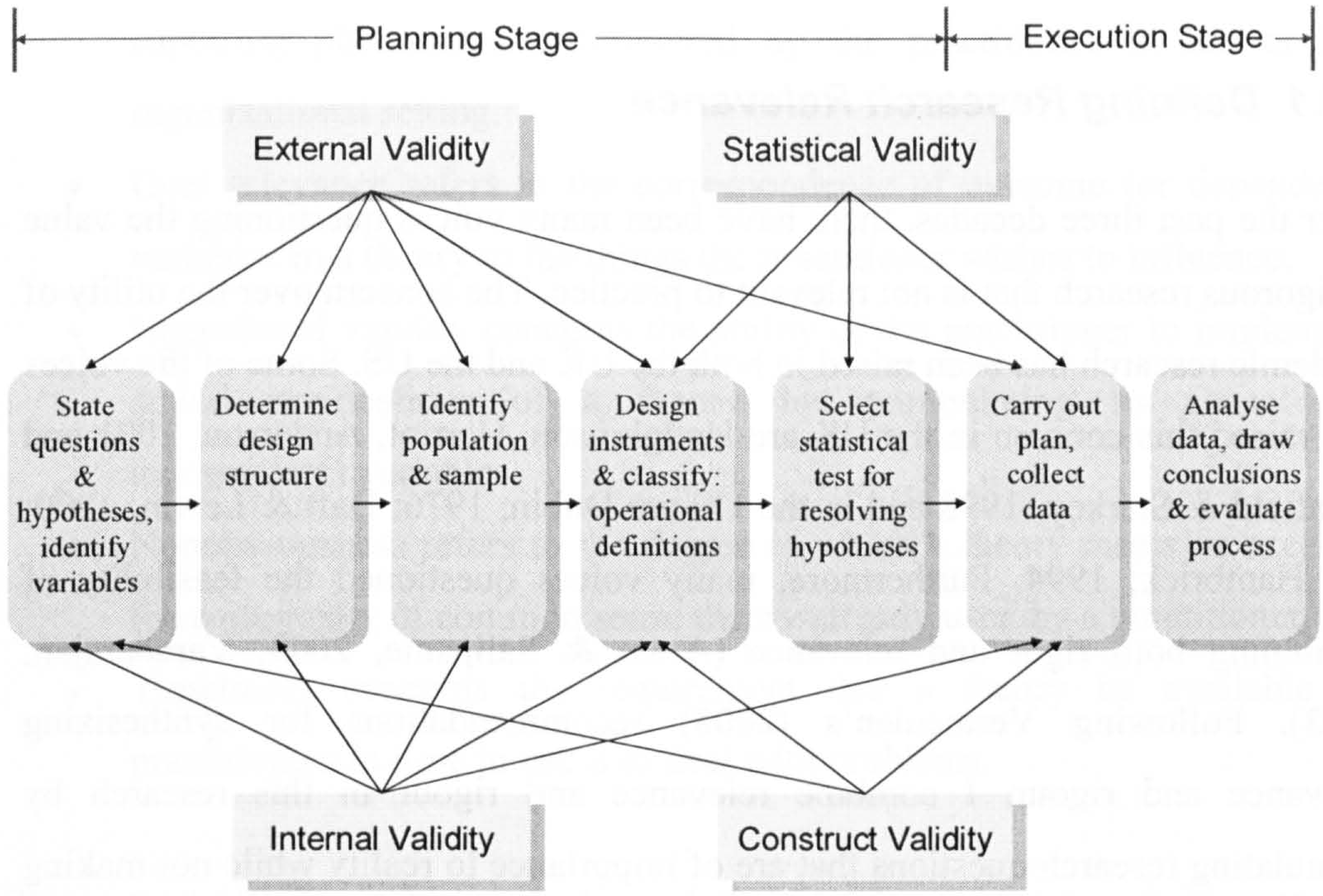
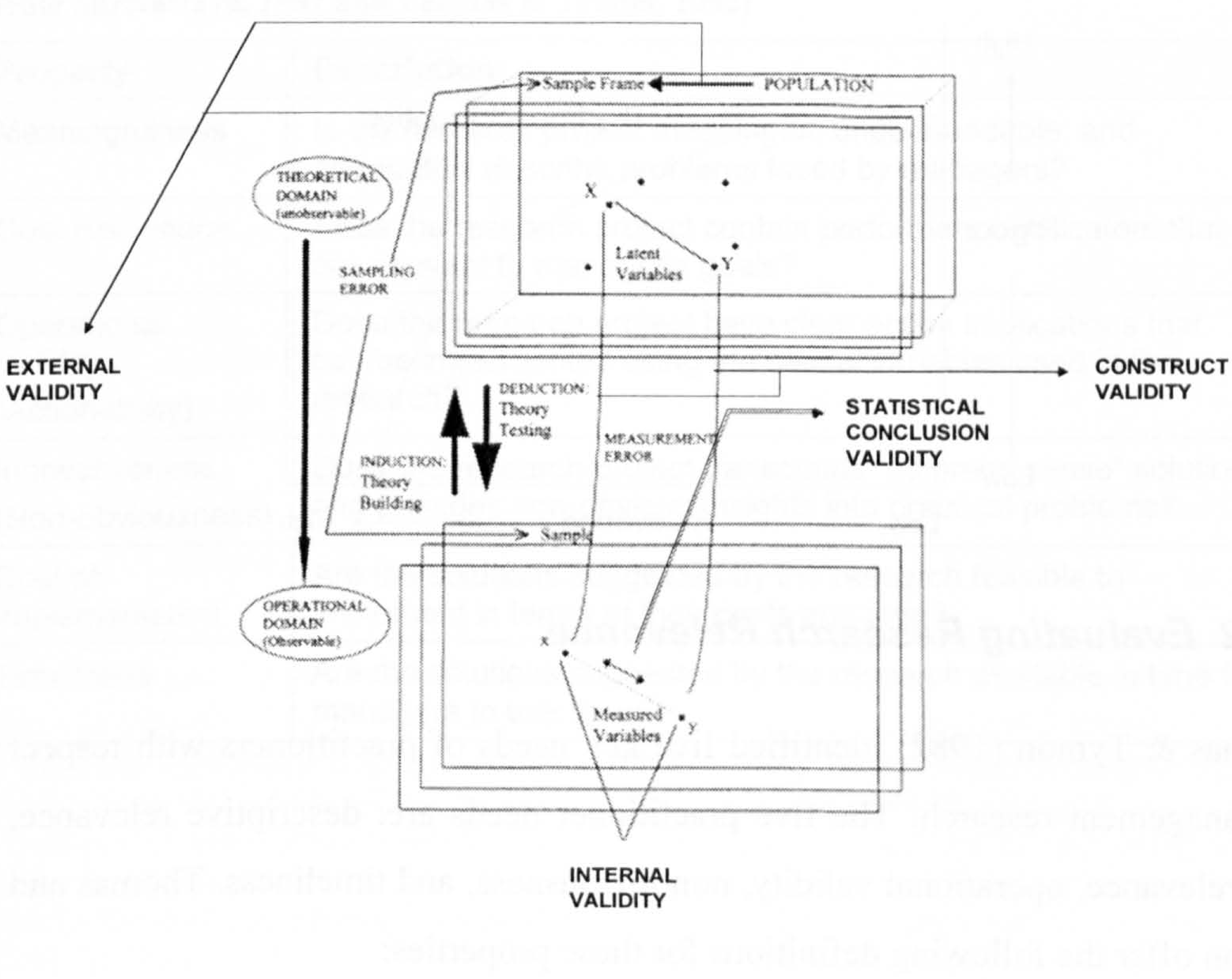


Figure 4-8 maps the four types of validities into the theoretical and operational domains of research projects.

Figure 4-8: Types of validities mapped into theoretical and operational domains (Source: Adapted from Malhotra & Grover, 1998)

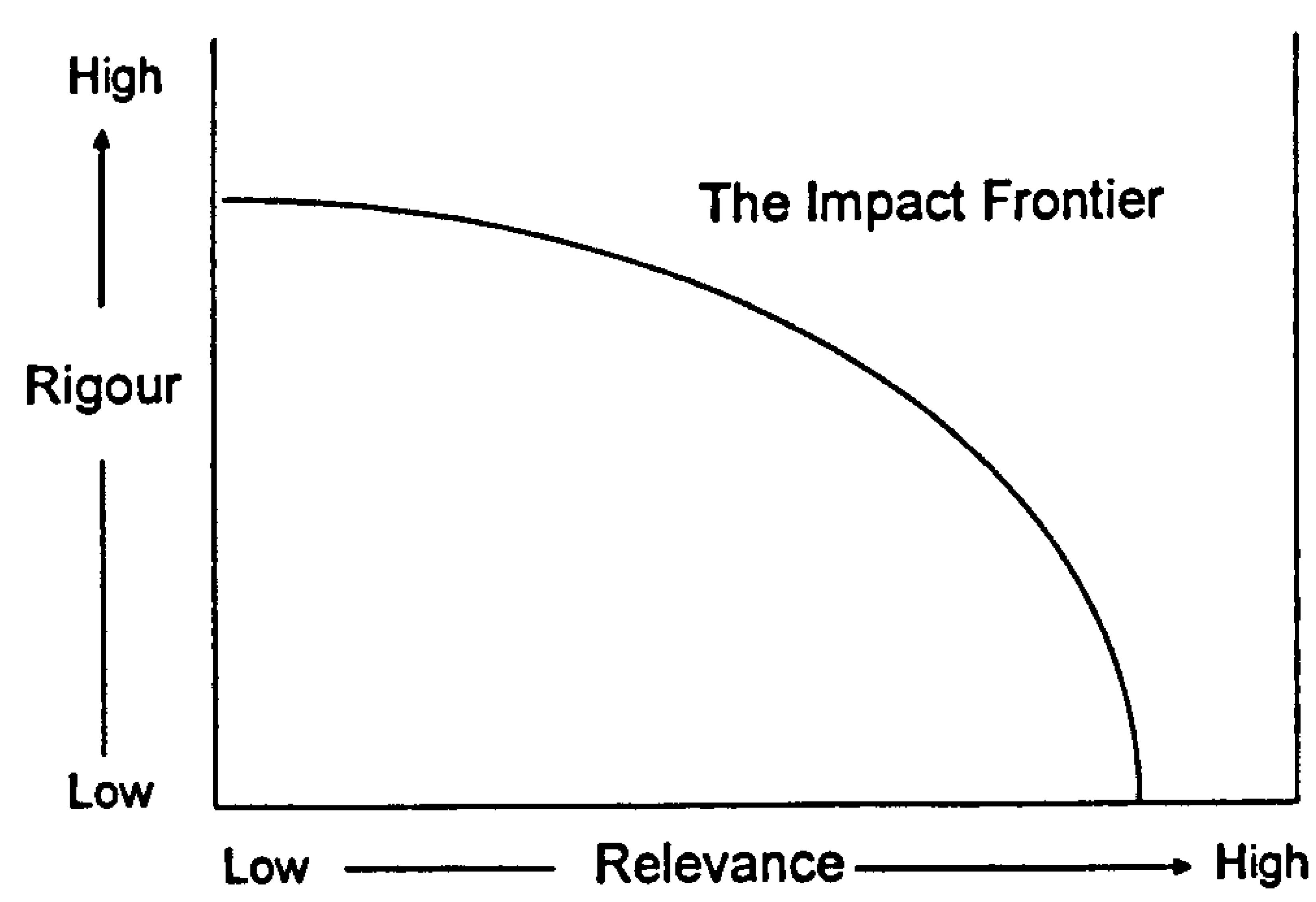


4.5 Research Relevance

4.5.1 Defining Research Relevance

Over the past three decades, there have been many voices questioning the value of rigorous research that is not relevant to practice. The concern over the utility of academic research has been raised in both the UK and the US. Some of the voices that raised this concern in the UK are Hodgkinson, Herriot, Anderson, 2001 and Tranfield & Starkey, 1998 and in the US are Dubin, 1976; Daft & Lewin, 1990; and Hambrick, 1994. Furthermore, many voices questioned the feasibility of combining both rigor and relevance (Aram & Salipante, 2003; Varadarajan, 2003). Following Vermeulen's (2005) recommendations for synthesizing relevance and rigour, I combine relevance and rigour in this research by formulating research questions that are of importance to reality while not making concessions in terms of rigour in developing theory and empirical evidence. Therefore, this research may be placed in the "impact frontier", a term used by Davenport & Markus (1999), as depicted in Figure 4-9.

Figure 4-9: The impact frontier: rigour and relevance (Source: Adapted from Davenport & Markus, 1999)



4.5.2 Evaluating Research Relevance

Thomas & Tymon (1982) identified five key needs of practitioners with respect to management research. The five practitioner needs are: descriptive relevance, goal relevance, operational validity, nonobviousness, and timeliness. Thomas and Tymon offer the following definitions for these properties:

- Descriptive relevance refers to the accuracy of research findings in capturing phenomena encountered by the practitioner in his or her organizational setting.
- Goal relevance refers to the correspondence of outcome (or dependent) variables in a theory to the things the practitioner wishes to influence.
- Operational validity concerns the ability of the practitioner to implement action implications of a theory by manipulating its causal (or independent) variables.
- Nonobviousness refers to the degree to which a theory meets or exceeds the complexity of common sense theory already used by a practitioner.
- Timeliness concerns the requirement that a theory be available to practitioners in time to use it to deal with problems.

Shrivastava (1987), in his evaluation of the usefulness of research in strategic management, converted these properties into five criteria for evaluating the usefulness of research projects as shown in Table 4-6.

Table 4-6: Criteria for assessing practical usefulness of research projects (Source: adapted from Shrivastava, 1987 and Thomas & Tymon, 1982)

Property	Description
Meaningfulness	Is the research project meaningful, understandable, and adequately describe problems faced by managers?
Goal Relevance	Does the research project contain performance indicators that are relevant to manager's goals?
Operational Validity (Actionability)	Does the research project have clear action implications that can be implemented using the causal variables used in the research?
Innovativeness (Non-obviousness)	Does the research project transcends "common sense" solutions and provides non-obvious insights into practical problems?
Cost of Implementation	Are the solutions suggested by the research feasible to implement in terms of their costs and time?
Timeliness	Are the solutions suggested by the research available in time for managers to use them?

4.6 Research Design

4.6.1 Introduction

The primary aim of this research is to assess the impact of organizational performance measurement diversity on product innovation and the moderating role of performance measurement use on this relationship. To test the posited hypotheses, a cross-sectional[36], multi-year, multiple industry, probability sample, survey[37] research design is used.

Taking in consideration the tradeoffs among generalizability, accuracy, simplicity (Weick, 1979) and critical requirements of the research: (1) testing of developed hypotheses; (2) generalizing the findings to a wider population; and (3) measuring the diagnostic and interactive use of performance measurement systems (primary data), the chosen research design is well suited for meeting the objectives of the research. Table 4-7 contrasts the ability of different research methods to meet the critical requirements of the research. The research methods that were evaluated are: survey, experimental, case study, and archival. They are the most used research methods for investigating social phenomena (Singleston & Straits, 2005).

Table 4-7: Evaluating different research methods against critical research requirements

Critical Research Requirement	Survey	Experiment	Case Study	Archival
Testing hypotheses	√	√	X[38]	√
Measurement: Performance Measurement Use	√	√	√	X
Generalizability	√	X	X	√

[36] "Cross-sectional design collects data at one point in time from a sample selected to represent the population of interest at that time" (Pinsonneault & Kraemer, 1993).
[37] Kerlinger & Lee (2000: 599) classifies survey research as field studies with quantitative orientation.
[38] Large multi-case design can be used to test theory but they are prohibitively costly for PhD projects.

Using cross-sectional survey design[39] is justified for having numerous advantages over other research methods. Some of the advantages are:

1- "[P]rimary data[40] developed through a survey methodology have one very important advantage over almost all secondary data. That advantage is that the research design is developed specifically to address the research question ... Survey research is the most appropriate, if not the only, approach for addressing some questions" (Stanley & Atuahene-Gima, 2004: 228).
2- "Survey research has the advantage of wide scope: a great deal of information can be obtained from a large population" (Kerlinger & Lee, 2000:613).
3- Survey research is accurate (Kerlinger & Lee, 2000: 613; Stanley & Atuahene-Gima, 2004: 228). For example, subjective, self-reported performance measures have been found to be highly correlated with objective measures of firm performance (Dess & Robinson, 1984; Ketokivi & Schroeder, 2004; Robinson & Pearce, 1988; Venkatraman & Ramanujam, 1986, 1987; Zahra, 1993a). Also, Slater & Atuahene-Gima (2004) point out that research has shown congruence between archival measures and managers assessments of market conditions (Keats & Hitt, 1988) and product-market strategy (Shortell & Zajac, 1990).
4- Knoke, Marsden, & Kalleberg (2000: 799) indicate that "survey research methods ... are especially valuable for organizational studies when representativeness and generalization are central study objectives."
5- Giving the amount of data that could be collected using survey research, it is economical. (Kerlinger & Lee, 2002:613).

---

[39] This is a non-experimental research design. As noted by Cook & Campbell (1979: 295) and Stone-Romero (2002: 85), non-experimental research has been often referred to inappropriately as correlation research. Correlation is a statistical technique and not a research strategy. Furthermore, correlation techniques could be used with different research strategies.

[40] Primary data are collected directly from organizations and secondary data are collected from available records (Venkatraman & Ramanujam, 1986). Primary data is collected specifically to address the research questions whereas secondary data was collected for a different purpose at hand (Malhotra, 1999).

6- Pinsonneault & Kraemer (1993) argue that survey research method is appropriate for theory testing and explanation.

Moreover, the cross-sectional survey designs are the most widely used field study designs in operations management and innovation research. In operations management, Flynn, Sakakibara, Schroeder, Bates, & Flynn (1990) state that "survey is undoubtedly the most commonly used research design in operations management". In innovation research, a recent meta-analysis study by Vincent, Bharadwai, & Challagalla (2005)[41] investigating the determinants and consequence of organizational innovation found that ninety five out of 134 studies examined employed cross-sectional survey designs. Also, twenty six of these studies aggregated innovation across industries.

The research design used multiple industries to increase the external validity and generalizability of the study (Scandura & Williams, 2000).

Table 4-8 summarizes some of the arguments made for the justification of the research design.

Table 4-8: Justification of the research design

Critical Research Requirement	Research Design Element	Comments
Testing hypotheses	Cross-sectional Survey	Testing hypotheses require large dataset. Case studies were ruled out because they do not meet this criterion. Although longitudinal survey designs would increase the internal validity of the research, they were ruled out given the time constraint of PhD projects.
Measurement: Performance	Survey	The research calls for operationalising the performance measurement use and

[41] The studies were reported in the following highly ranked scholarly journals: Economics (American Economic Review, Journal of Technology Transfer, R&D Management, RAND Journal of Economics, Research Policy, Quarterly Journal of Economics); Management (Administrative Science Quarterly, Academy of Management Journal, Journal of Management, Management Science, Organization Science, Strategic Management Journal); Marketing (Journal of the Academy of Marketing Science, Journal of Marketing, Journal of Marketing Research, Journal of Product Innovation management).

Critical Research Requirement	Research Design Element	Comments
Measurement Use and Organizational performance measurement diversity		diversity constructs. This requirement calls for a field research design to obtain primary data, therefore excluding archival designs for two reasons. First, there are no publicly available sources containing secondary data on performance measurement use in organizations. Second, even if public records of organizational performance measurement diversity are available, they may be problematic (Georgiou, 1973) because they may differ from the actual measures in the organization. Perrow (1961) differentiates between official or formal organizational goals that are communicated to the public and operative goals that are used internally. Similarly, Etzioni (1964; 1975) differentiates between formal goals and real goals.
Generalizability	Multiple industry	The use of multiple industries research design increase the generalizability of the research findings. McKee, Varadarjan, & Pride (1989) reminds us that single industry studies increase the internal validity of the research design but reduces the extent to which the findings can be generalized (external validity). Experimental and small multi-case research designs are ruled based on this criterion. Black (1999: 48) notes that "[case studies] do not allow one to generalize either to larger populations or general situations". Although experimental designs increase the internal validity of the study, they decrease the external validity and make it difficult for generalizing the findings beyond the experiment setting (Campbell and associates, 1966, 1979, 1990, 2001). Furthermore, they are very difficult to set up in real organizational settings.

Research designs entails making trade-off between research objectives of generalizability, accuracy, and simplicity (Weick, 1979). Therefore, all research designs are subject to limitations. This research is no exception. It has several limitations and they will be discussed in the research limitations section in chapter 7.

4.6.2 Survey Research Method

4.6.2.1 Definition

As noted by Pinsonneault & Kraemer (1993), there is an important difference between survey and survey research. They define a survey as a means of gathering information about the characteristics, actions, or opinions of a large group of people, referred to as a population, which may include marketing, or opinion surveys, but survey research is conducted for the purpose of advancing scientific knowledge.

Survey research has three characteristics (Malhotra & Grover, 1998; Pinsonneault & Kraemer, 1993):

1- It involves collection of information by *asking people* structured and predefined questions. Their answers, which might refer to themselves or to some other unit of analysis, constitute the data to be analysed. In management research, there are five common units of analysis: the individual participant, the subunit or department, the establishment (e.g. strategic business unit), the multi-establishment firm, and the interorganizational network (Knoke et al., 2002).

2- Information is gathered *via a sample*, which is a fraction of the population, with the need to be able to generalize findings from the sample to the population.

3- Survey research is a quantitative field study method. The purpose of the survey is to produce *quantitative descriptions* of some aspects of the studied population. Therefore, it requires standardized information in order to define or describe variables, or to study relationships between variables.

4.6.2.2 Types of Survey Research

Survey research can be used for exploration, description, or explanation purposes (See Table 4-9). In this research project, I use survey research for explanation purpose, more specifically to test theory. This type of survey research is also called confirmatory, theory testing or relational. It answers the question "Does the hypothesized causal relationship exist, and does it exist for the reasons posited?"

Table 4-9: Purposes of survey research (Source: Adapted from Forza, 2002; Pinsonneault & Kraemer, 1993)

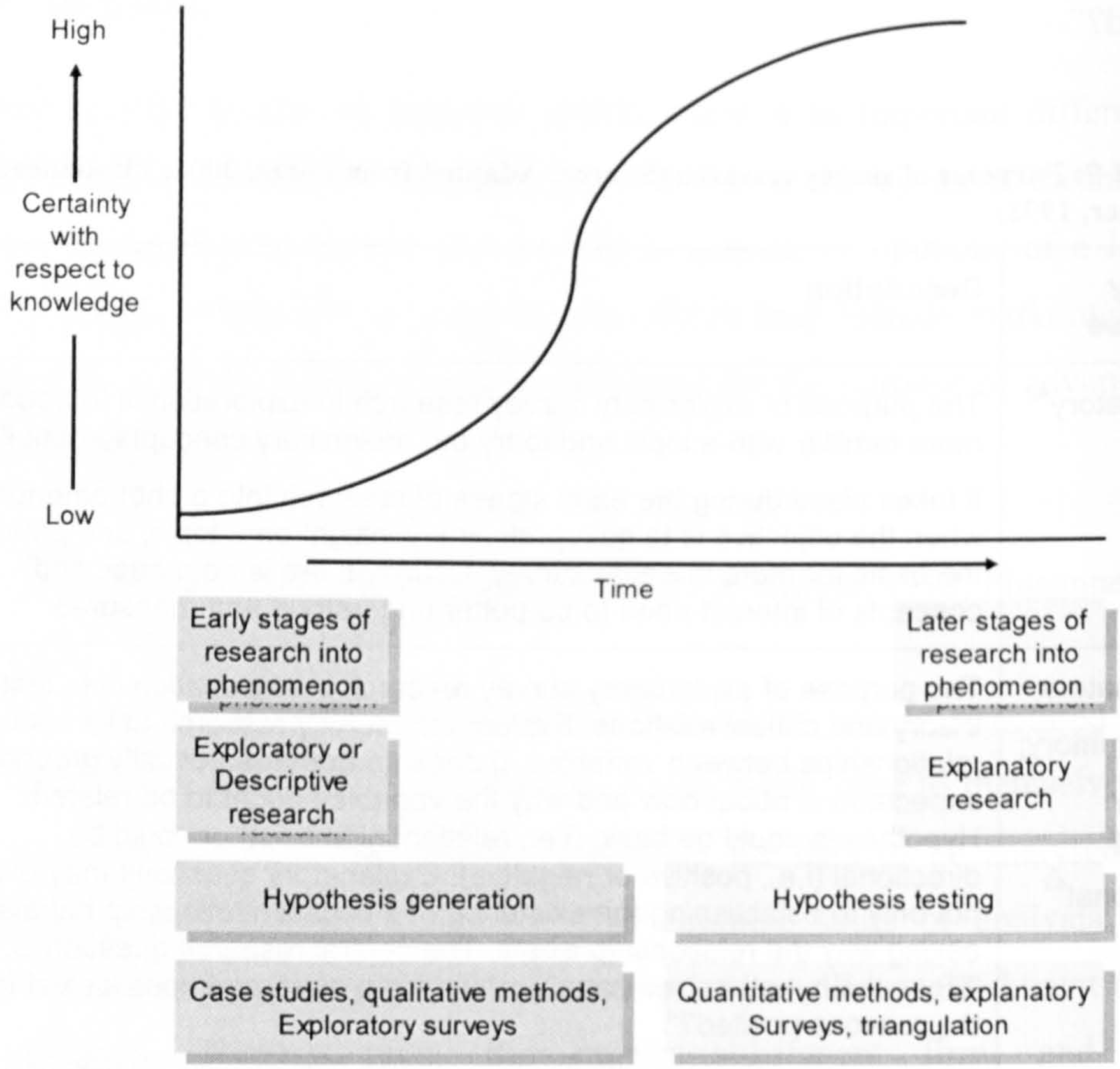
Survey Purpose	Description
Exploratory[42]	The purpose of exploratory survey research in exploration is to become more familiar with a topic and to try out preliminary concepts about it. It takes place during the early stages of research into a phenomenon, when the objective is to gain preliminary insight on a topic, and provides the basis for more in-depth survey. Usually there is no model, and concepts of interest need to be better understood and measured.
Explanatory; Confirmatory; Theory Testing; Relational[43]	The purpose of explanatory survey research in explanation is to test theory and causal relations. Explanatory survey research asks about the relationships between variables. It does so from theoretically grounded expectations about how and why the variables ought to be related. Hypotheses could be basic (i.e., relationships exist) or could be directional (i.e., positive or negative). Explanatory questions may extend not only to establishing the existence of a causal relationship but also to asking why the relationship exists. The central research question is: "Does the hypothesized causal relationship exist, and does it exist for the reasons posited?" It takes place when knowledge of a phenomenon has been articulated in a theoretical form using well-defined concepts, models and propositions.
Description	The purpose of descriptive survey research is to understand the relevance of a certain phenomenon and describe the distribution of the phenomenon in a population. Its primary aim is not theory development, even though through the facts described it can provide useful hints both for theory building and for theory refinement.

The different purposes of survey research are associated with different stages of the maturity cycle of research (See Figure 4-10). Explanatory research is associated with the hypotheses testing phase that follows from hypotheses generation phase. Therefore, it corresponds to the later stages of the maturity cycle of research.

[42] Malhotra & Grover (1998) identified two purposes for survey research: explorative and explanatory. Subsequently, they classified exploratory research into exploratory and descriptive.

[43] Rungtusanatham, Choi, Hollingworth, Wu, & Forza (2003) define studies that empirically examine relationships among two or more variables in either confirmatory or exploratory manner as relational studies.

Figure 4-10: The maturity cycle of research (Source: Malhotra & Grover, 1998)



4.6.2.3 Requirements for Survey Research

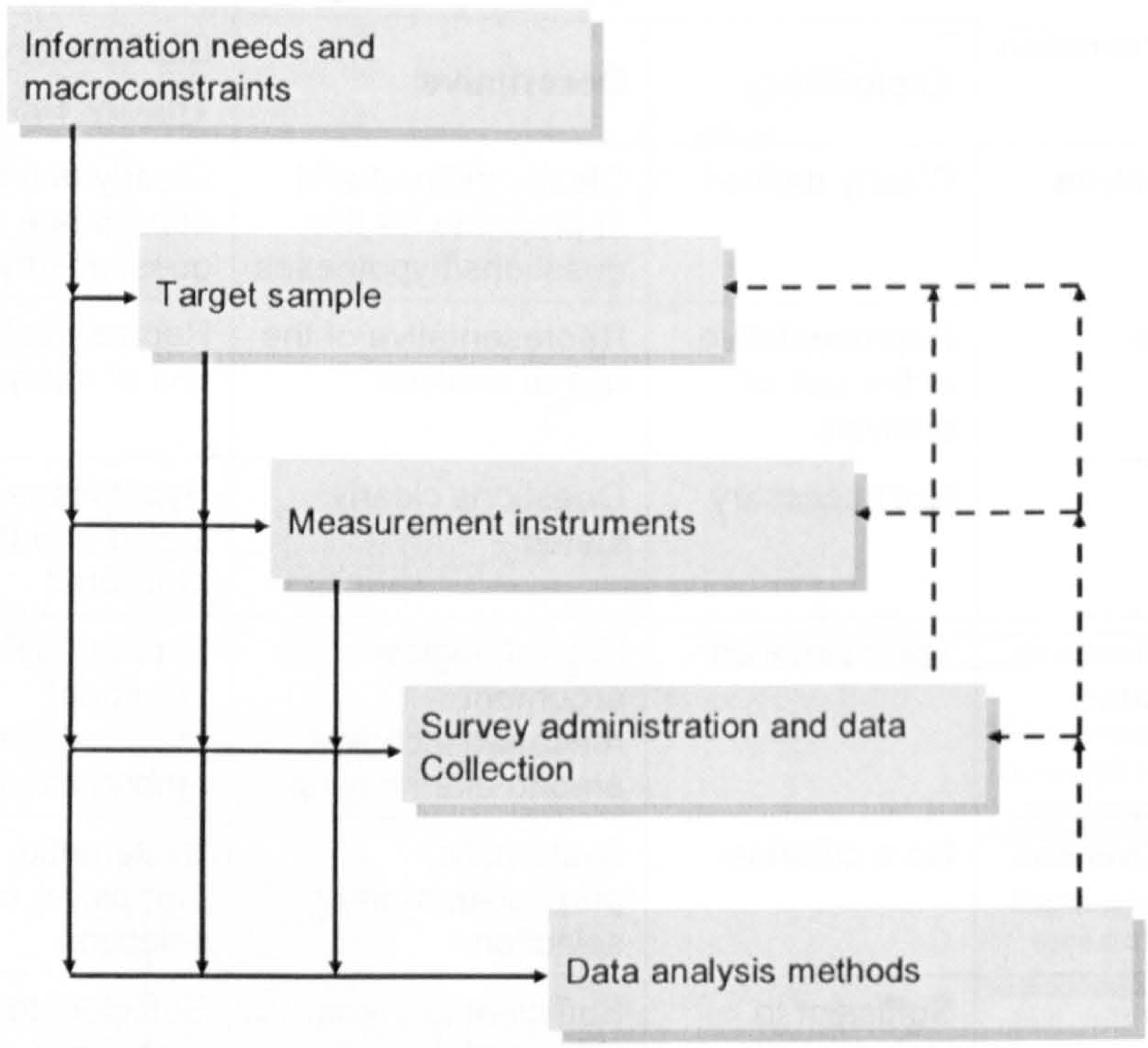
Although the three different types of survey research share the same dimensions, they differ in their requirements in some of these dimensions. Table 4-10 contrasts the requirement for the three types of survey research against the following design elements of survey research: unit of analysis, choice of respondents, representativeness of sample frame, representativeness of sample, sample size, Pre-test of questionnaires, Response rate, and Mixed of data collection methods.

Table 4-10: Requirements for different survey research purposes (Source: adapted from Forza, 2002 based on Pinsonneault & Kraemer, 1993)

Element/Dimension	Purpose of Survey Research		
	Exploratory	Descriptive	Explanatory Theory Testing
Unit(s) of analysis	Clearly defined	Clearly defined and appropriate for the questions/hypotheses	Clearly defined and appropriate for the questions/hypotheses
Respondents	Representative of the unit of analysis	Representative of the unit of analysis	Representative of the unit of analysis
Research hypotheses	Not necessary	Questions clearly stated	Hypotheses clearly stated and theoretical motivated
Representativeness of sample frame	Approximation	Explicit, logical argument; reasonable choice among alternatives	Explicit, logical argument; reasonable choice among alternatives
Representativeness of sample	No a criterion	Systematic, purposive, random selection	Systematic, purposive, random selection
Sample size	Sufficient to include the range of the interest phenomenon	Sufficient to present the population of interest and perform statistical tests	Sufficient to test categories in the theoretical framework with statistical power
Pre-test of questionnaires	With subsample of sample	With subsample of sample	With subsample of sample
Response rate	No minimum	Greater than 50% of targeted population and study of bias	Greater than 50% of targeted population and study of bias
Mixed of data collection methods	Multiple methods	Not necessary	Multiple methods

Research is an iterative process and survey research is no exception. Figure 4-11 highlights the iterative nature of survey research and the decision points in survey research planning.

Figure 4-11: Linkages between decisions in survey planning (Source: Adapted in Alreck & Settle, 1985 as reported in Forza, 2002)



4.6.3 Empirical Setting: UK Manufacturing

In this section the empirical setting of the study is introduced. I will briefly discuss the UK manufacturing sector, show why the sector is a good empirical setting for the study, and give examples of prior studies on UK manufacturing.

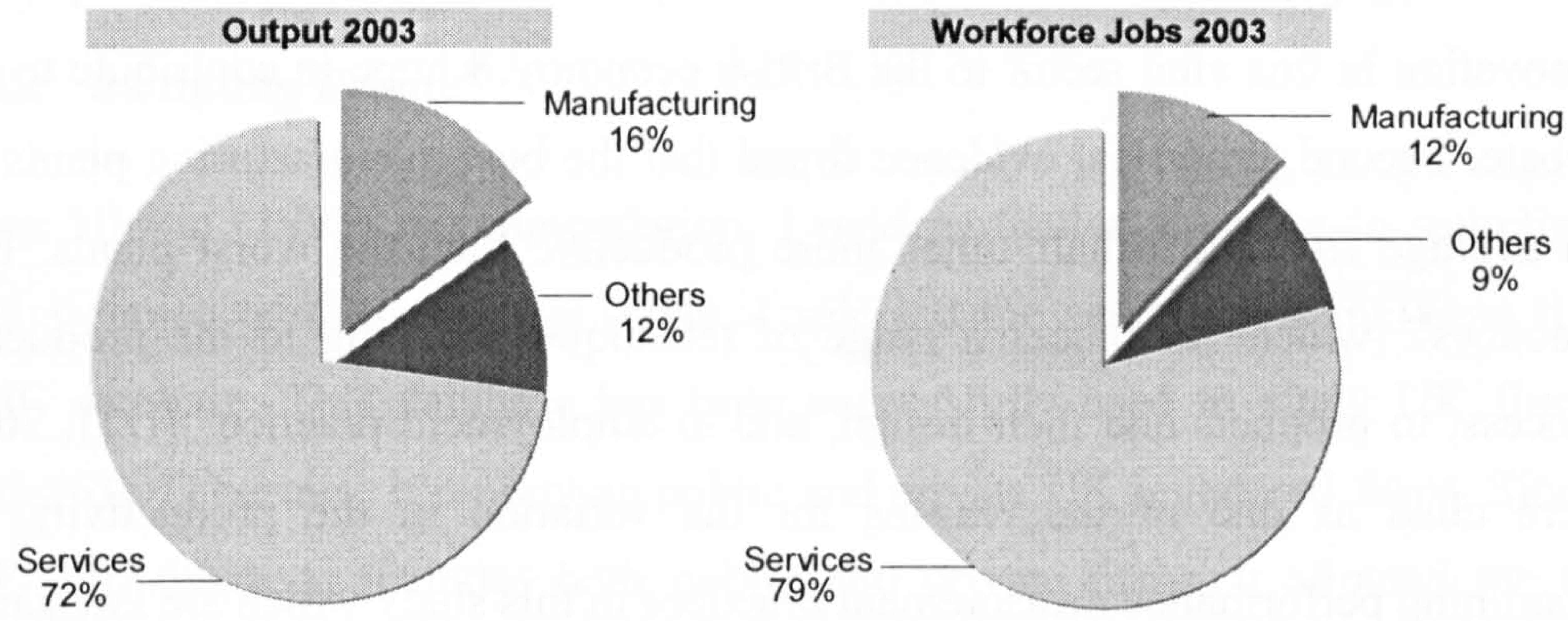
4.6.3.1 Industrial Context and Background

DTI (2002; 2006a) declared that the success of the United Kingdom manufacturing sector is crucial for the UK prosperity and highlights the following characteristics of the UK manufacturing industry:

- Manufacturing accounts for sixth of the economy as illustrated in Figure 4-12.
- It is vital for the UK's trading position - being responsible for around two-thirds of all UK exports
- It generates almost 3.1 million jobs directly - and supports a further 2 million in the service sector as illustrated in Figure 4-12.

- It is responsible for around 75% of the spending on business research & development.
- It is a key generator of productivity in the wider economy, through introducing new products and processes.

Figure 4-12: Composition of the UK economy



Source: DTI

As noted by DTI (2002), although the UK has world-class companies in every manufacturing industry and productivity in the manufacturing sector is 25% higher than in the rest of the UK economy, the UK manufacturing has a substantial productivity gap when compared to its overseas competitors. DTI (2002) reports that "[a]nalysis by the National Institute of Economic and Social Research shows that manufacturing productivity is 55 per cent higher in the US than in the UK, 32 per cent higher in France and 29 per cent higher in Germany. As a result, manufacturing contributes disproportionately to the overall gap. Although the manufacturing sector comprises only 20 per cent of output, it accounts for between 30 per cent and 40 per cent of the total shortfall in productivity between the UK and our competitors."

DTI (2002) focused on seven drivers of success in UK manufacturing: macroeconomic stability, investment, science and innovation, best practice, skills and education, modern infrastructure, and the right market framework.

4.6.3.2 Motivation for Studying UK Manufacturing

At least three issues motivated the study on product innovation in UK manufacturing. First, although the UK science base is considered one of the best in the world, UK manufacturing firms are less innovative than their counterparts in the US. DTI (2005) notes that innovation is one of the reasons for the major productivity gap between the UK and US. By studying determinants of product innovation in this vital sector to the British economy, I hope to contribute to this debate. Second, empirical evidence found that the best manufacturing plants are on average five and a half times more productive than the worst plants. Best practices[44] which "embraces a range of techniques, relating to the production process, to products and their design, and to employment practice" (DTI, 2002) were cited as one of the reasons for the variation in the productivity. By examining performance measurement practices in this study which are key part of management practices (Bloom & Van Reenen, 2006), I hope to contribute to the performance measurement and management practitioner literature. Third, the manufacturing sector has been the object of several previous empirical studies that explored the impact of performance measurement systems on product innovation, thus providing this study with the appropriate empirical benchmarks (e.g., Bisbe & Otley, 2004).

4.6.3.3 Prior Studies on Innovation in UK Manufacturing

The UK manufacturing has been studied extensively. Examples of these studies are as follows: Craggs & Jones, 1998; Stockdale, 2002; Laursen & Salter, 2006; and Robson & Ortmans, 2006. However, none of the studies examined the impact of performance measurement systems on product innovations in the UK manufacturing sector or the moderating role of performance measurement use on this relationship.

---

[44] A recent systematic literature review of the adoption of best practices by AIM scholars (Leseure, Bauer, Birdi, Neely, & Denyer, 2005) highlighted the other terms that are currently in use such as promising, appropriate, and integrated frameworks. Some authors argue that best denotes universality, which is usually not the case with management practices (Bessant, Caffyn, & Gilbert, 1996; Bessant, Kaplinsky, & Lamming, 2003).

4.6.4 Sample

4.6.4.1 Population[45]

The population that this study aims to generalize its findings to is the medium and large manufacturing firms in UK. Therefore, the sampling frame is selected to reflect this choice of generalization.

4.6.4.2 Sampling Frame[46]

As per Hinkin (1995) recommendation, I paid particular attention to sampling and the choice of the sampling frame. I selected the sampling frame from the FAME database. This database has been successfully used to study UK firms (Harris, 2001) because it reports on public and private UK registered firms. Since the FAME database includes both public and private firms, it allowed me to include in my sample frame both types of firms, thus allowing me to avoid selection bias associated with sampling frames chosen from databases that contain only public traded firms. A probability sample of 913 firms was selected that had more than 150 employees because these firms are more likely to have formal performance measurement systems (Bisbe & Otley, 2004).

4.6.4.3 Sampling Method

Probability sampling[47] is used to select the sample. Specifically, systematic sampling was used in which the tenth firm was selected.

Table 4-11: Sampling Approaches (Source: Forza, 2002)

Representativeness	Purpose is mainly	Types of sampling
Essential for the study	Generalizability	Simple random sampling
		Systematic sampling

[45] "A population is the entire set of elements about which the survey researcher wishes to make generalizations" (Van der Stede, Young, & Chen, 2005).
[46] The sampling frame is "the list or quasi list of elements from which a probability sample is selected" (Babbie, 2001; 194).
[47] Babbie (2001: 176) states that probability sampling is one of key aspects to generalizing from the sample to the larger population. Random selection is the cornerstone of probability sampling and in which "each element has an equal chance of selection independent of any other event in the selection process" (Babbie, 186). Random sampling ensures that the sample of elements contain essentially the same variation that exists in the population (Babbie, 2001, 182).

Representativeness	Purpose is mainly	Types of sampling
Probabilistic sampling	Assessing differential parameters in subgroups of population	Proportionate stratified sampling (for subgroups with an equal number of elements)
		Disproportionate stratified sampling (for subgroups with an different number of elements)
	Collecting information in localized areas	Area sampling
	Gathering information from a subset of the sample	Double (or multi stage) sampling
Not essential for the study Non-probabilistic sampling	Obtain quick, even if unreliable, information	Convenience sampling
	Obtain information relevant to available from certain groups	Judgment sampling (when looking for information that only few experts can provide)
		Quota sampling (when the responses of special interest minority group are needed)

4.6.4.4 Unit of Analysis[48]

The unit of analysis in this study is the firm or the strategic business units (SBU). I adopt Hax & Majluf's (1996) definition of SBU. SBU is an operating unit or planning focus that groups a distinct set of product or services, which are sold to a uniform set of customers, facing well-defined set of competitors. SBU is the appropriate unit of analysis for research investigating strategic implications of performance measurement systems because different SBUs employ different strategies that may require different performance measurement systems (Bruggeman & Van der Stede, 1993; Hoque, 2004).

4.6.4.5 Key Informants

The key informants in this study are defined as members of the top management team. A full treatment of key informants is offered in the data collection chapter.

[48]Forza (2002) states that the "unit of analysis refers to the level of data aggregation during subsequent analysis [and it is the] empirical parallel of the level of reference of the theory.

4.7 Operationalisation of Constructs – Measures

The choice of measures is a critical element of survey research (Malhotra & Grover, 1998), therefore I undertook an extensive review of the existing literature on existing measurement scales (See appendix E) to ensure adherence to the methodological guidelines on developing and selecting measures (See for example, Hensley, 1999; Hinkin, 1995; Malhotra & Grover, 1998; O'Leary-Kelly & Vokurka, 1998).

The following are general characteristics of the measures used in this study:

1- For four of the variables, subjective, primary measures are used and for the other two variables, objective, primary measures are used.
2- All perceptual measures employ multi-item[49] scales with four or more items.
3- All measures are validated scales that have been used in studies published in high-ranking scholarly journals.
4- All Likert-type scales have at least 6 points[50].
5- Reverse coding[51] is not used.

4.7.1 Dependent Variables

4.7.1.1 Product Innovation

I captured ***Product Innovation*** by first presenting the top management team a definition of product innovation based on DTI's community innovation survey (CIS)[52]. The definition given was "A product innovation is the market

[49] "The theoretical construct should be measured using as few items as possible but at least three items Cronbach and Meehl, 1955" (Hensley, 1999).
[50] "Reliability increases as the number of scale points increases to five and continues to increase at a much smaller rate for additional points above five" (Hensley, 1999).
[51] "There appear to be two schools of thought regarding the advisability of using reverse-scored items. The use of reverse-scored items is supported on the basis that response bias will be reduced because the respondent will be more alert in completing the survey Spector, 1992. Other researchers have opposed the use of reverse-scored items suggesting that validity may be lowered and the possibility of systematic error may be increased if reverse-scored items are used Jackson et al., 1993; . Schriesheim and Hill, 1981; Hinkin, 1995" (Hensley, 1999).
[52] CIS is performed every four years by a core set of countries, including the UK. The process is overseen by Eurostat, who are responsible for coordinating the development of a harmonized

introduction of a *new* good or service or a *significantly* improved good or service. The innovation must be new to your enterprise, but it does not need to be new to your market. It does not matter if the innovation was originally developed by your enterprise or by other enterprises."

Second, for the first three questions, I asked the top management team to specify to what extent their organisations introduced product innovations in the past two years on six-point Likert scale anchored with "To a very great extent" and "Not at all" on the following items:

1. *Significantly* improved goods or services
2. *New* goods or services onto your market before your competitors *(New to your market)*
3. *New* goods or services that was essentially the same as a product already available from your competitors in your market *(Only new to your company)*

For the fourth question, I asked the top management team to specify to what extent their organisations in the past two years on six-point Likert scale anchored with "To a very great extent" and "Not at all" increased the following item:

4. The percentage of new products or services significantly in your portfolio

The instrument used is adapted from the instrument developed by DTI (2005a) and is based on the guidelines set in the "Oslo Manual[53]" by OCED (1997). The last item was adapted from the instrument developed by Bisbe & Otley (2004). This CIS based instrument is widely used by innovation researchers[54] (see for, example, Cassiman & Veugelers, 2002; DTI, 2005a; Laursen & Salter, 2006; Mairesse & Mohnen, 2002)

In order to capture product innovation, I had to answer two questions:

1- What type of innovation indicator(s) should I use?

---

methodology to allow international comparison of results. For a general critique of innovation surveys and CIS, refer to (Salzar & Holbrook, 2004).

[53] The Oslo manual offers comprehensive procedures for collecting information on technological innovations through surveys.

[54] Laursen & Salter (2006) note that over than 60 recent academic studies used CIS data.

2- What method should I use to collect data on the chosen innovation indicator(s)?

Choosing an innovation indicator[55]:

In evaluating the multitude of different innovation indicators, I answered the following two questions:

1- Should I use input, intermediate or output indicator?

2- Should I choose single or multiple indicators?

Type of innovation indicator

In Table 4-12, I contrast the various advantages and disadvantages of using different types of indicators. I choose an output-based innovation indicator because it better reflects my theoretical argument that product innovation are very risky solutions when compared to patenting or increasing R&D spending. Greve, 2003b argues that "patents ... correspond better with the success of a development process than with a decision to launch, since firms have a reason to seek patent protection even for ***innovations*** judged to be too risky to launch as products: a firm can earn license fees by patenting such ***innovations***. Patenting can also slow the depreciation of an ***innovation*** kept in the buffer by extending the time it takes other firms to invent around it (Levin et al., 1987)." [Emphases are in the original text]. Figure 4-13 clarifies the relationship between patenting, invention, and innovation.

[55] For a comprehensive historical review of the development of innovation indicators, refer to work of Godin (2003; 2005a)

Table 4-12: Measures of innovation (Source: adapted from Brouwer & Kleinknecht, 1999; Coombs, Goffin & Mitchell, 2005: 45; Narandren, & Richards, 1996)

	Measure	Advantages of the Measure	Limitations of the Measure
Input	**R&D[56]-Based** • R&D expenditure • R&D intensity (R&D/Sales) • R&D employees	1. Data investment levels are normally published in company annual reports.	1. This is 'input' measure, rather than a measure of R&D output. 2. R&D measures reflect innovative activities taken by formal R&D organizations. However, many innovation activities are taken outside the formal R&D organizations (Kleinknecht, 1987; Kleinknecht & Verspagen, 1989; Kleinknecht & Reijnen, 1991)

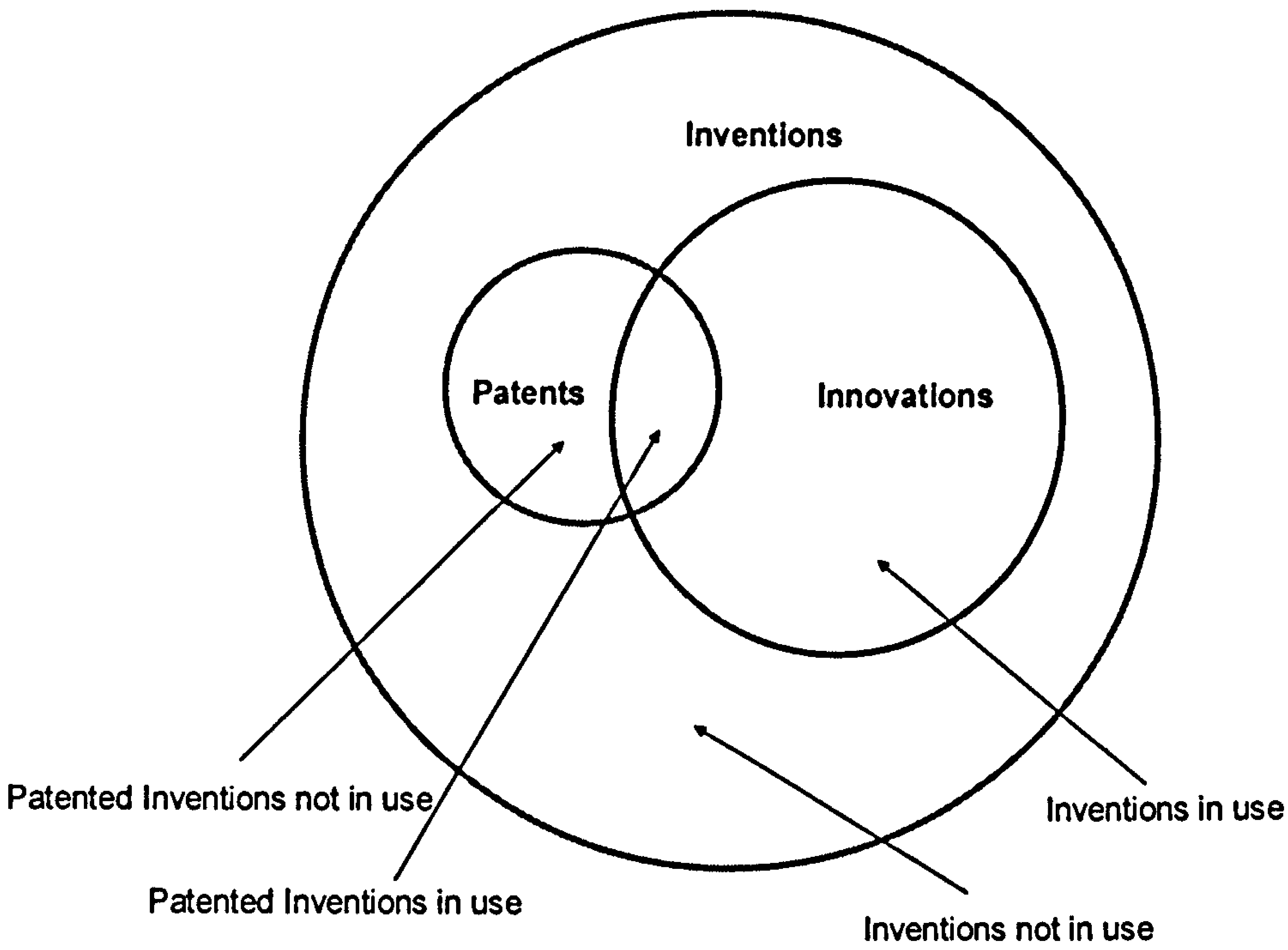
[56] The standard definition of R&D given in the "Frascati Manual" by OECD (2002) is "Research and experimental development (R&D) comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications." Furthermore, "[t]he term R&D covers three activities: basic research, applied research and experimental development .... Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts, without any particular application or use in view. Applied research is also original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific practical aim or objective. Experimental development is systematic work, drawing on existing knowledge gained from research and/or practical experience, that is directed to producing new materials, products or devices, to installing new processes, systems and services, or to improving substantially those already produced or installed. R&D covers both formal R&D in R&D units and informal or occasional R&D in other units."

Measure		Advantages of the Measure	Limitations of the Measure
Intermediate	**Patent-based** • Number of patents[57] • Number of patents per employee	1. Data on the number of patents are readily available and can easily be analysed by industrial sector, country, etc.	1. Patents are more a measure of invention than of innovation (Basberg, 1987; Griliches, 1990) 2. Some firms may chose not to apply for patents, as it is a time-consuming process and it does not always offer good protection to small organizations 3. Larger firms have a greater probability to seek patent protections. 4. There is a significant difference among sectors. Higher technology sectors have a higher tendency to patent than low technology sectors[58]. 5. Firms that have R&D collaboration agreements, have a higher propensity to patent. 6. Patents measure codified knowledge and most of innovation knowledge is tacit (Patel & Pavitt, 1994; Pavitt, 1985; 1988). 7. The tendency to bundle claims together in one or more patents varies widely among countries (Santarelli & Piergiovanni, 1996).
Output	**New Products-Based** • Number of new products • New product scales • Sales from new products	1. It is a measure of the output of R&D (but not strictly of innovation, unless product success is considered)	1. The meaning of 'new product' is equivocal and this can lead to measurement problems. 2. Figures on the numbers of new products developed by companies are not easily accessible.

[57] Smith (2005) defines patent as "a public contract between an inventor and a government that grants time-limited monopoly rights to the applicant for the use of a technical invention". For a comprehensive review of patent-based measure, refer to (Iversen, 1998).

[58] Pavitt (1984) advanced a classification based on (1) sectoral sources of technology and (2) sources and nature of the technology produced and characteristic of innovating firms. The classification groups the sectors into five technological trajectories: supplier dominated, science-based, scale intensive, information intensive and specialized supplier.

Figure 4-13: The relationship between inventions, patents, and innovations (Source: Basberg, 1987)



Single versus multiple composite index

Although a number of researchers called for the use of composite measure of innovation based on a number of innovation indicators Hagedoorn & Cloodt (2003) and Mohnen & Dagenais (2002), I chose a single indicator of innovation because empirical research by Grupp & Mogee (2004) highlighted the sensitivity of such composite measures. The difficulty in constructing a composite measure of innovation arises from two decisions that need to be made: what indicator should go into the composite and what weight should be assigned to each indicator in the composite.

Choosing a collection approach:

Kleinknecht (1993) detailed four approaches used in measuring innovation. According to Kleinknecht, studies use one or more of the following approaches:

1- Identification of major innovations from the historical literature.
2- Identification of innovations by consulting experts.

3- Survey-based - Managers' assessment through postal surveys of numbers of innovations

4- Literature-based - Counting of innovations in 'new product announcements' sections of trade and technical journals.

Table 4-13 contrasts the two widely used approaches in innovation research: survey and literature based approaches.

I choose the survey-based approach because the literature (journal) based approach is not suitable for studying multiple industries for the following two reasons:

a. The quality of journals covering the different industries differs significantly from one industry to another.

b. Classifying product innovations in multiple industries requires a significant investment in coding and classifying innovations because of the requirement for coders who understand the technical specifications of products in the different industries under study.

Table 4-13: Approaches to measuring innovation (Source: adapted from Coombs, Narandren, & Richards, 1996; Kleinknecht, 1993; 1996)

Approach	Advantages of the Approach	Limitations of the Approach
Survey based	1- There are established standards for data collection on innovations such as the "Oslo Manual". 2- It captures in-house innovations. 3- It is easy to administer and simple to score and code. 4- It is replicable. 5- If standard classification are used (e.g. CIS), the method can be applied in different countries and country comparison made. 6- It requires little knowledge of the industries that are sampled because the coding of innovation is done by the respondents.	1- There is a danger that the indicator will be distorted by companies that which to 'inflate" the perceived rate of new product introduction for their own company to gain public relation (PR) or market benefits. 2- The collection and use of data may be hampered by secrecy. 3- The dataset may not be extendable to the past, hence comparisons overtime become impossible. Panel designs are not very common. 4- The approach may be susceptible to non-response bias where innovative firms are more likely to respond.
Literature (Journal) based	1- The announcements times are close to the date of commercialisation. 2- The data collection can be performed without contacting the firms, hence there is no non-response problem and firms are not burdened. 3- The method may well capture innovations from small firms better than other indicators. 4- If standard classification are used, the method can be applied in different countries and country comparison made. 5- The collection and use of data is not hampered by secrecy. 6- The dataset can be extended to the past, hence comparisons overtime become possible.	1- The method does not capture in-house process innovation. 2- There is a danger that the indicator will be distorted by companies that which to 'inflate" the perceived rate of new product introduction for their own company to gain public relation (PR) or market benefits. 3- There are problems of judgment involved in the selection of relevant journals. 4- There are problems in the coding of innovations. 5- There are problems in the coding the type of innovations. 6- Journals may differ in their editorial control of new product announcements. 7- Different industries and countries are covered by different journals. The quality of coverage of new product introductions in these journals may differ substantially.

4.7.2 Independent Variables

4.7.2.1 Organizational Performance Measurement Diversity

I measure ***Organizational Performance Measurement Diversity*** by asking top management team to rate the extent they use the following 10 performance measurement categories on six-point Likert scale anchored with "To a very great extent" and "Not at all".

1. *Short term financial results* - e.g. operating income, sales growth, etc
2. *Customer relations* - e.g. market share, customer satisfaction, etc.
3. *Employee relations* - e.g. employee satisfaction, safety, etc
4. *Operational performance* - e.g. productivity, lead times, etc.
5. *Quality* – e.g. quality performance, defect rates, etc.
6. *Innovation and learning* - e.g. number of new products and /or services launched training, etc.
7. *Supplier relations* - e.g. on-time delivery, suppliers' integration etc
8. *Alliances* - e.g. joint marketing, joint product designs, etc.
9. *Environnemental performance* - environnemental compliances, etc.
10. *Community* - public image, community involvement, etc.

The organizational performance measurement diversity is measured with an adapted version of the instrument used by Ittner, Larcker, & Randall (2003). The instrument was adapted to fit the purpose of the research. Specifically, since performance measurement uses were hypothesized as moderator of organizational performance measurement diversity, they were removed from the original scale.

Ittner and colleagues (2003) constructed their ten performance categories from a number of literature streams: balanced scorecard; intangible assets; intellectual capital; and value based management (for example, Edvinsson & Malone, 1997; Kaplan & Norton, 2001; Shiemann & Lingle, 1999) making it a comprehensive measure of diversity. This instrument was selected because it was suitable for my multiple industry research design. Specifically, the instrument did not hard-wire specific measures in each of the ten performance categories. This was desirable because firms competing in different industries and firms following different strategies will emphasize different performance categories and measures. Second, the instrument had a comprehensive coverage of the performance categories because it was not only based on the balanced scorecard literature. The other widely used instrument in performance measurement research developed by Hoque and associates (Henri, 2006a; Hoque, 2004; 2005; Hoque & James, 2000) did not meet the above two criteria.

4.7.3 Mediating Variables

4.7.3.1 Innovation Search

I measure ***Innovation Activities*** by asking top management team to rate the extent during the past two years their organizations engaged in the following innovation activities on six-point Likert scale anchored with "To a very great extent" and "Not at all".

1. Intramural (in-house) R&D - Creative work undertaken within your enterprise on an occasional or regular basis to increase the stock of knowledge and its use to devise new and improved goods, services and processes.
2. Acquisition of R&D (extramural R&D) - Same activities as above, but purchased by your enterprise and performed by others.
3. Acquisition of machinery, equipment and software - Acquisition of advanced machinery, equipment and computer hardware or software to produce new or significantly improved goods and services.
4. Acquisition of external knowledge - Purchase or licensing of patents and non-patented inventions, know-how, and other types of knowledge from other enterprises or organisations
5. Training - Internal or external training for your personnel specifically for the development and/or introduction of innovations
6. All forms of Design - Expenditure on design functions for the development or implementation of new or improved goods, services and

processes. Expenditure on design in the R&D phase of product development should be excluded.

7. Market introduction of innovations - Activities for the market preparation and introduction of new or significantly improved goods and services, including market research and launch advertising

Innovation search is a subset of the innovation activities. By using factor analysis, the subscale of innovation search will be factored out and used in the study. The innovation activities scale is developed by DTI (2005a) and is based on the guidelines set in the "Oslo Manual" by OCED (1997).

4.7.3.2 Risk

I measure ***Risk*** by asking top management team to rate the extent their organizations have an operating top management philosophy of the following qualities on six-point Likert scale anchored with "To a very great extent" and "Not at all".

1. Strong emphasis on research and development, technological leadership, and innovations.
2. Strong proclivity to high risk, high return investments.
3. Growth strategy primarily through external financing (borrowings, capital issues, etc.)
4. Very competitive, "undo-the-competitors" philosophy.

Khandwalla (1977) developed this risk scale that was subsequently used by Singh (1986).

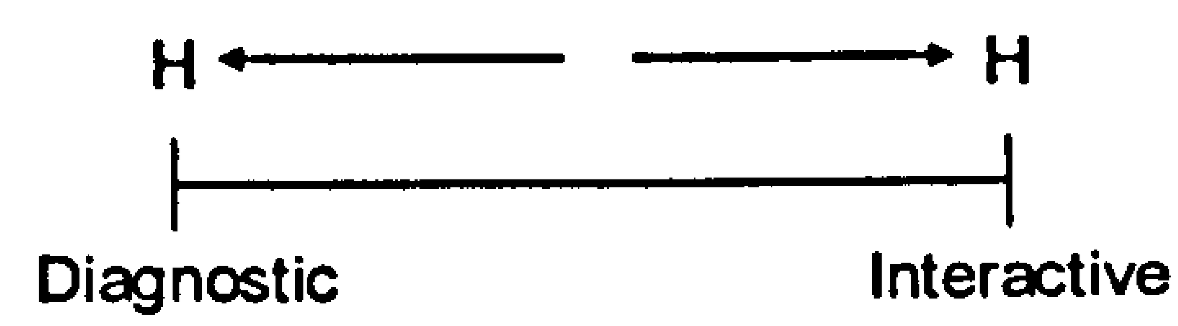
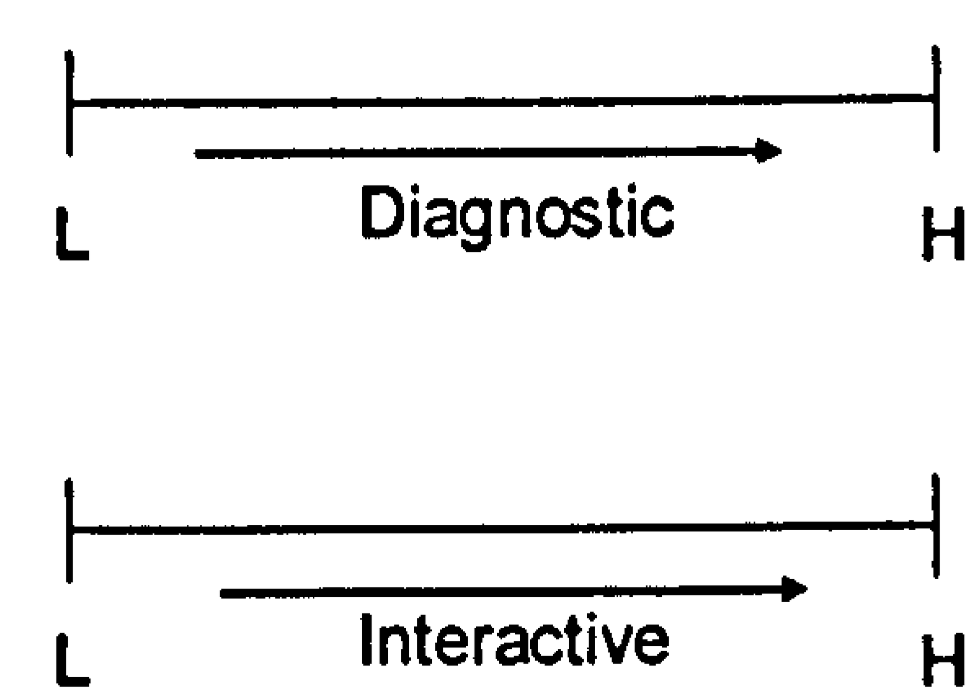
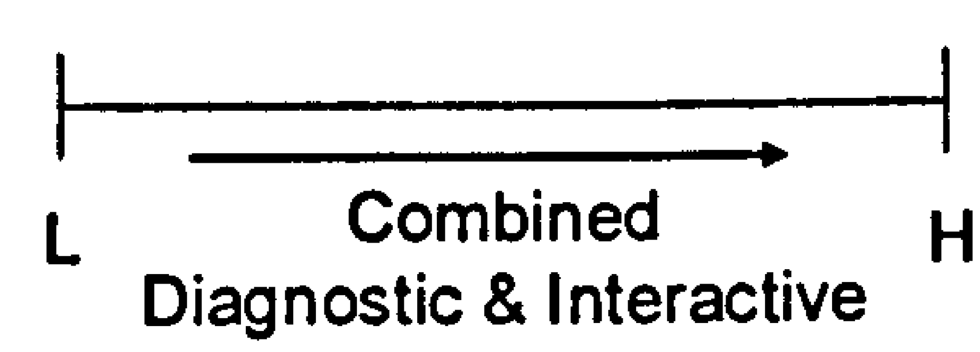
4.7.4 Moderator Variables

Survey based studies (variance research[59]) that have used Simons' (1995) diagnostic and interactive uses operationalised the constructs in different ways

[59] Mohr (1982) distinguishes between two types of research – variance and process. Variance also known as outcome research (Aldrich, 2001) involves explaining a dependent variable with a set of independent variables that statistically explain the variation in the dependent variable whereas

(See Table 10-5). Schwab (1982) notes that it is common for researchers conducting social research to adopt different operationalisations of the same construct as long as these operationalisations represent the underlying construct. However, this may not be the case here as illustrated in Figure 4-14.

Figure 4-14: Different operationalisations of diagnostic and interactive uses

Operationalisation	Studies	Illustration
Opposite Poles	Abernethy & Bowens, 1991; Davila, 2000; Abernethy, Bouwens, & Van Lent, 2006; Bisbe & Otley, 2004	
Separate	Gill et al., 2003; Henri, 2006b; Widener, 2005	
Combined	Henri, 2006b	

In this research, I conceptualise diagnostic and interactive uses as two separate constructs, which is consistent with Simons' conceptualisation.

"process" or "event-driven" involves explaining the temporal order and sequence in which a discrete set of events occur based on a story or historical narrative (Van de Ven, 2007)

4.7.4.1 Diagnostic Use

Diagnostic Use by asking top management team to what extent they use performance measures to do the following using 6-point Likert scales anchored with "To a very great extent" and "Not at all".

1- Track progress toward goals
2- Review key measures
3- Monitor results
4- Compare outcomes to expectations

4.7.4.2 Interactive Use

I capture ***Interactive Use*** by asking top management team to what extent they use performance measures to do the following using 6-point Likert scales anchored with "To a very great extent" and "Not at all".

1. Integrate the organization – i.e. tie the organization together
2. Enable the organization to focus on common issues
3. Enable the organization to focus on your critical success factors
4. Develop a common vocabulary in the organization
5. Provide a common view of the organization
6. Enable discussion in meetings of superiors, subordinates and peers
7. Enable continual challenge and debate underlying results, assumptions and action plans

The scales for diagnostic and interactive uses were originally developed by Vanenbosch (1999) to measure the intensity of use of executive support systems (ESS) and subsequently adapted by Henri (2006a; 2006b). Following Henri (2006a; 2006b), the choice of scales are justified by its development was based on theories of control systems use. More specifically, it was based on the works of Simon & colleagues (1954) and Burchell & colleagues (1980). Furthermore, ESS is used as a surrogate for accounting and management information and is restricted to the accounting, management, and control information provided. Thus, performance measurement systems and ESS have a common base that allows the adaptation of the scale to my specific context. In developing the scales, Vendenbosch (1999) used the guidelines prescribed by Churchil (1979)[60] and Q-sorting techniques. For descriptions of the various instruments for diagnostic and interactive use of performance measurement systems, please refer to appendix E.

4.7.5 Control Variables

Given that my research design is non-experimental design, I could not use design elements such as random assignment[61], control groups, and pretest/posttest to construct counterfactual[62] inference. Therefore, I had to measure alternative explanations individually and then statistically control for them (Shadish et al., 2001: 18).

I identified a number of environmental and organizational determinants of product innovation in the innovation literature stream. These determinants represent sources of potential extraneous variance. To reduce the possibility of spurious results caused by correlations among these variables and my constructs of interest, I included firm level and industry level controls in my statistical model and analysis. Furthermore, for each control variable I have included, I answered the following four questions:

1. How do I measure the control variable?
2. What is the rationale for choosing the measure?
3. Who else have used the same measure?
4. Why do I need to control for the variable?

---

[60] Churchill's (1979) scale development method is the dominant paradigm in management research. However, other researchers proposed variations and extensions to the Chrchill's method. See for example, (DeVellis, 1991; Gerbing & Anderson, 1988; Hinkin, 1995; Schwab, 1980; Spector, 1992). For an alternative approach to Churchill's paradigm, see C-OARS-E method proposed by Rossiter (2002).

[61] "Random assignment is the great ceteris paribus - that is, other things being equal – of causal inference (Cook & Campbell, 1979: 5).

[62] "A counterfactual is something that is contrary to fact. In an experiment, we observe what *did happen* when people received treatment. The counterfactual knowledge of what would have happened to those same people if they simultaneously had not received treatment. An *effect* is the difference between what did happen and what would have happened" Shadish et al, 2001: 5).

4.7.5.1 Firm Level Controls

I used two objective measures: firm size and firm age to control for firm-level effects.

4.7.5.1.1 Firm Size

Firm size is measured as the number of employees in the firm. I ask the respondents to indicate how many employees their organization employ. This is an ordinal variable with 7-levels: fewer than 500; 500-999; 1,000-1,499; 1,500-1,999; 3,000-9,999 & more than 10,000. This measure is widely used in performance measurement and innovation research (e.g., Bisbe & Otley, 2004; Henri, 2006a; 2006b).

I controlled for firm size effects for the following reasons:

1. Firm size has been shown to positively influence innovation (Balkin, Markman, & Gomez-Mejia, 2000; Camisón-Zornoza, Lapiedra-Alcami, Segarra-Cipres, & Boronat-Navarro, 2004; Chaney & Devinney, 1992; Chaney, Devinney, & Winter, 1991; Damanpour, 1991; 1992; Greve, 2003b). Other studies have pointed to a negative relationship between firm size and innovation (Aldrich & Ellen, 1986; Wade, 1996)[63].
2. Firms size has been shown to influence R&D (Balkin et al., 2000; Greve, 2003b; Lee & Sung 2005)
3. Controlling for firm size helps avoid omitted variable bias[64] (Abernethy & Brownell, 1999; Govindarajan, 1988).
4. George (2005) argues that controlling for firm size acts as a control for slack[65]. Slack has been found to increase firms' propensity to innovate (Greve, 2003b; 2004; Singh, 1986) and has been linked to

---

[63] Camisón-Zornoza et al (2004) conducted a meta-analysis on the link between firm size and innovation to update the earlier meta-analysis conducted by Damanpour (1992). Consistent with the earlier finding of Damanpour, they found a positive link.

[64] Omitting a relevant variable would constitute a specification error resulting in bias in the coefficient estimates of the included independent variables (Studenmund, 2001: 394).

[65] Cyert & March (1963: 42) defines slack as "the difference between total resources and total necessary payments". Singh (1986) and Bourgeois & Singh (1983) classify slack as available (unabsorbed), recoverable (absorbed), and potential. For a recent and comprehensive review of slack literature, refer to Daniel et al. (2004).

organizational performance (Cyert & March, 1963; Daniel, Lohrke, Fornaciari, & Turner, 2004)

4.7.5.1.2 Firm Age

Firm age is measured as the number of years the organization is in business. This measure is widely used in management research (Zahra, 1996). Few innovation studies control for firm age (for an exception see, Zahra, 1996).

I control for firm age effects for the following reasons:

1. Firm age has been shown to influence innovation. Some studies found a positive relationship (Damanpour, 1991; Sorensen & Stuart, 2000) and others found negative relationship (Boeker, 1997; Wade, 1996)
2. The age of the firm may influence its product innovation strategy. For example, Rosen (1991) found that young companies may pursue radical innovations.

4.7.5.2 Industry Level Controls

I control for industry effects on product innovations by incorporating the industry comparison in the questions except for the last question, which is internal to the firm. This technique is widely used in performance measurement and innovation research and effectively control for industry effects on product innovation (Bisbe & Otley, 2004; Li & Atuahene-Gima, 2001, 2002; Miller & Friesen, 1983).

I control for industry effects on product innovations for the following reasons:

1- Industries may determine the availability of opportunities for innovation for firms (Covin & Slevin, 1991; Zahra, 1993b, 1996).

2- Competitive intensity in the industry is determinant of product innovation (Vincent et al., 2005; Weerawardena, O'Cass, & Julian, 2006).

3- By controlling for industry effects on product innovations, I control for the number of the innovations in the industry, which was found in prior studies to influence the firms' propensity to innovate (Greve &

Taylor, 2000). Greve (2003b) observes that "[i]nnovations observed by managers facilitate discovery of market and technological opportunities, which increases the innovation rate of a firm". Furthermore, Greve (2003b) argues that by controlling for innovations in the industry you control for environmental sources of innovation solution.

4- Ebben & Johnson (2005) state that controlling for industry helps to control for environmental effects, such as hostility, complexity, and dynamism (Naman & Slevin, 1993)

5- Ebben & Johnson (2005) state that controlling for industry helps to control for firm goals[66] (Bromiley, 1991; Fiegnbaum & Thomas, 1988). Therefore, by controlling for firm goals, I implicitly control for past performance, which has been shown to influence firm's strategic decisions (Greve, 1998; 2003a; 2003c; Boeker & Goodstein, 1991). Also, past performance has been shown to influence R&D expenditure (Antonelli, 1989; Greve, 2003b; Kamien & Schwartz, 1982) and firm's propensity to innovate (Bolton, 1993; Greve, 2003a; 2003b, 2004).

4.7.5.3 Country Level Controls

I control for country effects by selecting a single country, UK, to conduct the study.

4.7.6 Summary

Table 4-14 summarizes the scales used in this study. It shows that all the scales are existing validated scales.

---

[66] Fiegnbaum & Thomas (1988) state that "Lev (1969: 290) suggested that firms adjust their performance to the industry average. He emphasized "the desirability of adjusting the firm's financial rations to predetermined targets which are usually based on industry wide averages." Lev also performed an empirical study on 900 major U.S. firms, in which he confirmed the hypothesis that firms periodically adjusted their financial rations to their industry means"

Table 4-14: Summary of survey scales

Construct	Variable	Level	Type / No. of Items	Source	Used By
Product Innovation	Product Innovation	Interval	6-point Likert Scale / 4	DTI, 2005a; Bisbe & Otley, 2004	Cassiman & Veugelers, 2002 ; DTI, 2005a; Laursen & Salter, 2006; Mairesse & Mohnen, 2002
Organizational Performance Measurement Diversity	Organizational Performance Measurement Diversity	Interval	6-point Likert Scale / 10	Ittner, Larcker, & Randall, 2003	Ittner, Larcker, & Randall, 2003
Performance Measurement Use	Diagnostic Use	Interval	6-point Likert Scale / 4	Vendenbosch, 1999	Vendenbosch, 1999; Henri, 2006a, 2006b; Widener, 2005
	Interactive Use	Interval	6-point Likert Scale / 7	Vendenbosch, 1999	Vendenbosch, 1999; Henri, 2006a, 2006b; Widener, 2005
Innovation Search	Innovation Search	Interval	6-point Likert Scale / 7	DTI, 2005a	DTI, 2005a; Laursen & Salter, 2006
Risk	Risk	Interval	6-point Likert Scale / 4	Khandwalla, 1977	Singh, 1986
Firm Size	Firm Size	Ordinal	Ordinal Check-box / 7		
Firm Age	Firm Age	Ratio	Blank space / 1	Zahra, 1996	Zahra, 1996

4.8 Model Estimation

The study consists of three sets of statistical analyses corresponding to the three types of the hypotheses: moderation analysis, mediation analysis, and direct effect analysis.

4.8.1 Moderation Analysis

Hypotheses, H2, H3, H4, H5, H6, and H7, postulate moderating role for performance measurement use (diagnostic and interactive). Therefore, moderated regression analysis (MRA) is used to test the following hypotheses: H2a, H3a, H4a, H5a H6a, and H7a and partial derivative of the moderated regression equation is used to test the following hypotheses: H2b, H2c, H3b, H4b, H4c H5b, H6b, and H7b.

Moderated regression analysis is used to test contingency relationships (Schoonhoven, 1981) and it is a specific application of multiple linear regression analysis, in which the regression equation contains an 'interaction term' (Hartmann & Moers, 1999). The typical format of a multiple regression equation containing two independent variables is given by equation (1):

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon \qquad (1)$$

However, when Z is hypothesized to act as a moderator[67], the multiple regression equation containing the moderating effect is given by equation (2).

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon \qquad (2)$$

where Y is the dependent or outcome variable; X is the independent or predictor variable; Z is the moderator variable; XZ[68] is the moderating effect of Z on the relationship between X and Y; and ε is the error term.

The meaning of the moderating effect could be illustrated by taking the partial derivative of equation (2) with respect to X, which has the format expressed by equation (3):

$$\partial Y/\partial X = \beta_1 + \beta_3 Z \qquad (3)$$

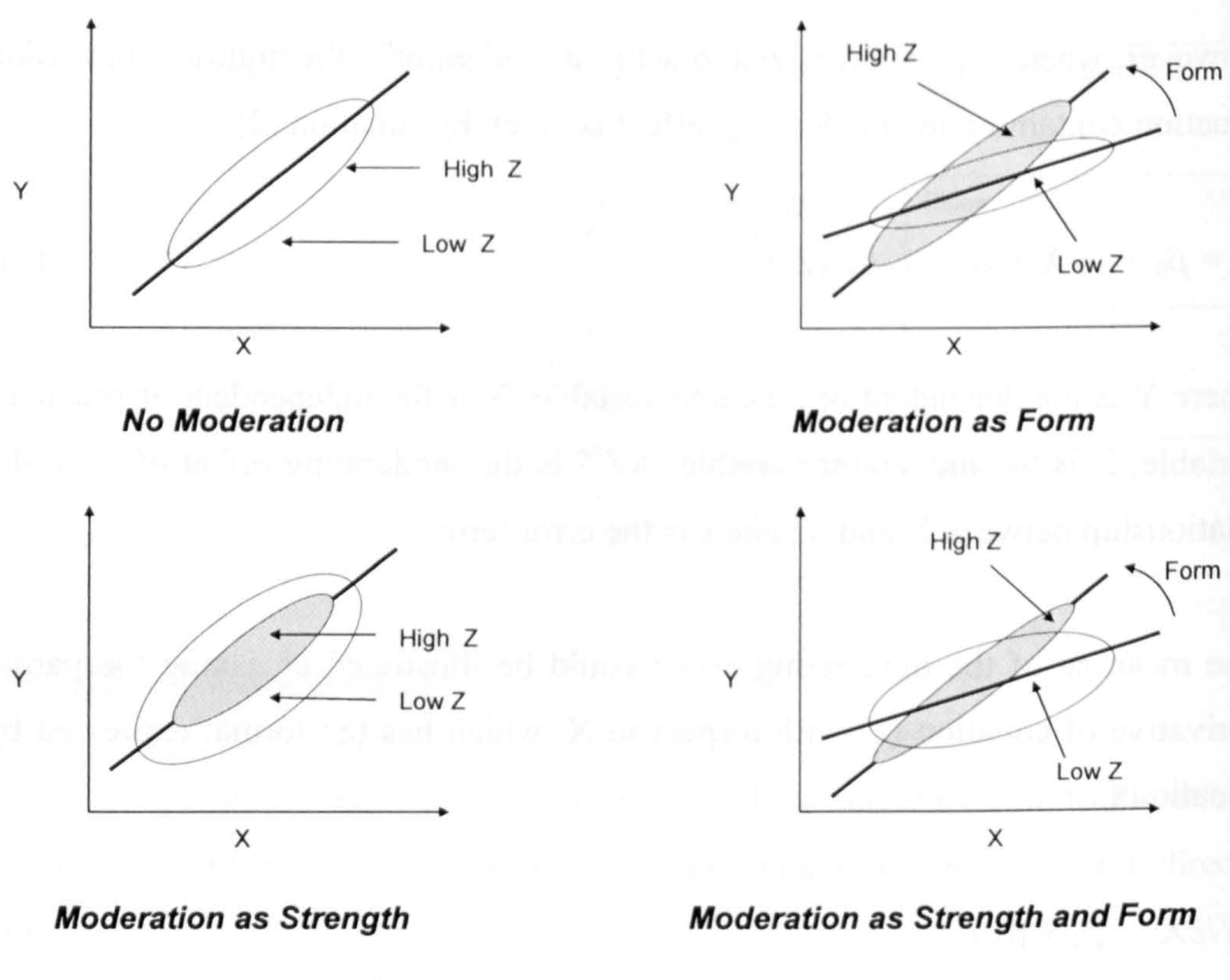
As illustrated in equation (3), the slope between Y and X is dependent on Z. For every one unit increase in Z, the slope of Y on X increases by the value of β_3. It

[67] The XZ interaction is symmetrical. It could also be hypothesized that X is the moderator variable. Therefore, the selection of the moderator variable should be theory driven (Aiken & West, 1991)

[68] Jaccard & Turrisi (2003) note that the simple product term is only one of many possible functional forms of the interaction effect and it is called bilinear interaction. It indicates that the slope between Y and X changes as a linear function of scores on Z, which is the case in this study.

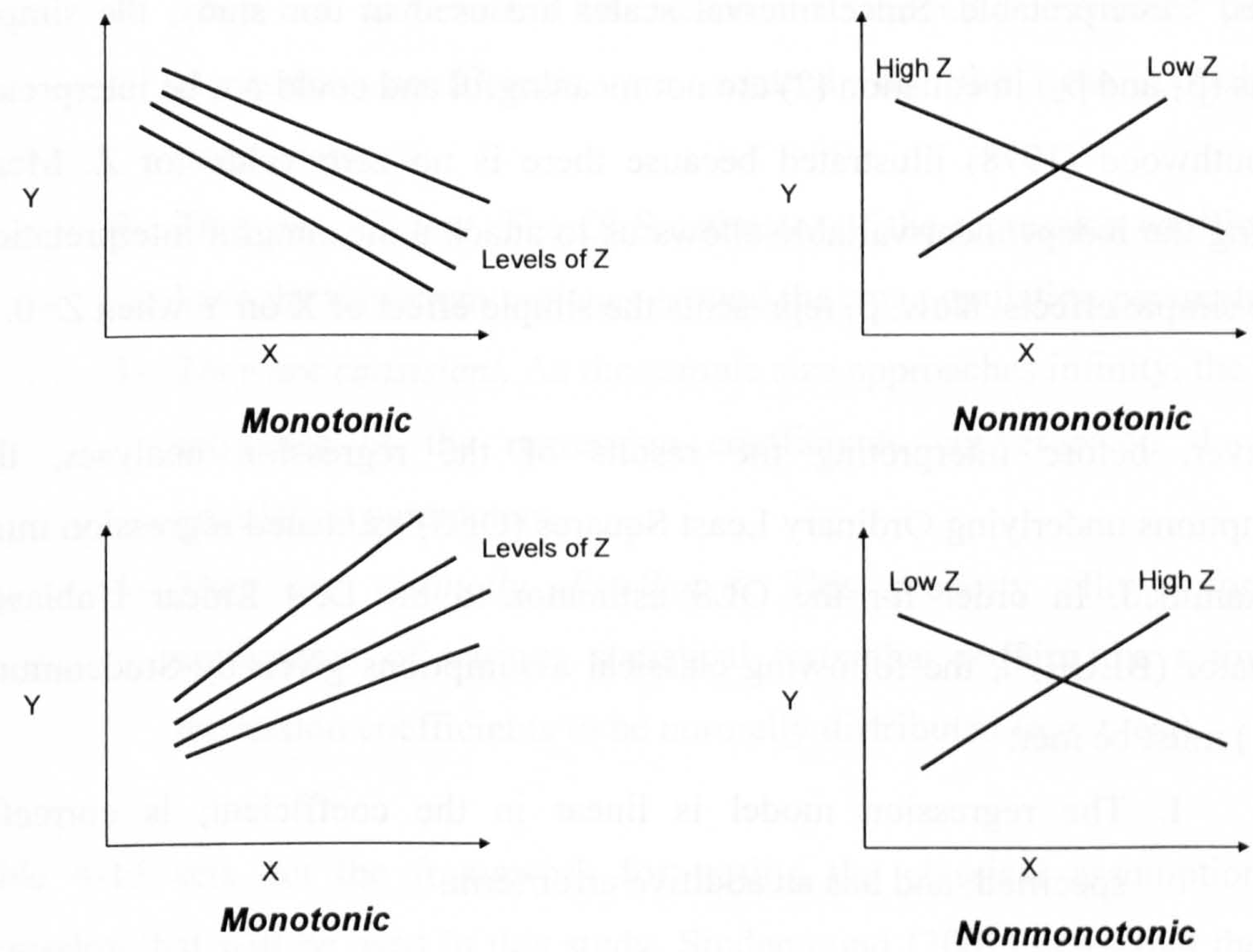
could be said the Z moderates the relationship between X and Y or alternatively, the relationship between Y and X is contingent on X. It is important to note here that the statistical significance of β_3 indicates that Z moderates the form (slope) of the relationship instead of the strength (predictive efficacy) of the relationship (Sharma, Durand, & Gur-Arie, 1981; Venkatraman, 1989) as illustrated in Figure 4-15. Hence, testing the statistical significance of the regression coefficient of the interaction term (β_3) allow us to test the following hypotheses: H2a, H3a, H4a, H5a, and H6a.

Figure 4-15: The two types of moderations: moderation as form and moderation as strength (Source: Hartmann & Moers, 1999)



However, for us to test the other moderation hypotheses that call for specific forms of the moderation relationship, we need to analyse the partial derivative (Shoonhoven, 1981; Southwood, 1977). As Shoonhoven (1981: 362) notes "[m]erely inspecting the signs and magnitudes of regression coefficients is insufficient analysis of contingency hypotheses. Graphing a partial derivative from the larger regression equation will reveal nonmonotonic effects not readily apparent in the tabled coefficients. If a nonmonotonic effect is present, as revealed by the graph and its calculations, it will add substantially to our knowledge to know where in the range of ... (moderator variable)... a change in the direction of slope occurs" (Emphases added). A monotonic[69] relationship exists when the partial derivative does not cross the horizontal line and it indicates that the moderating effect of Z changes the slope between Y and X within positive values or negative values only whereas nonmonotonic[70] relationship exists when the partial derivative crosses the horizontal line and it indicates that the moderating effect of Z changes the slope between Y and X within both positive values and negative values; within one range of values the relationship between Y and X will be positive and within the other range it will be negative (Hartmann & Moers, 1999; Schoonhoven, 1981). Figure 4-16 illustrates the concept of monotonicity of the moderated relationships.

Figure 4-16: Monotonic and nonmonotonic forms of moderation relationships (Source: Adapted from Gerdin, 2005; Hartmann & Moers, 1999)



---

[69] In some literatures, monotonic relationships are called ordinal or noncrossover and non monotonic relationships are called disordinal or crossover relationships (Aiken & West, 1991).

[70] Jaccard & Turrisi (2003) note that in theory all interactions are nonmonotonic because for any given pair of nonparallel regression lines, there is always a point where the lines intersect. So what determines monotonicity if the intersection point is present within the range of scores being studied.

In using the moderated regression analysis, two issues must be addressed: (1) the mean centring of independent variables and the interpretation of the coefficients.

The independent variables (Diversity, Diagnostic, and Interactive) were mean centred (the mean of the variable is deducted from the value of the variable) before they were entered into MRA and the interaction terms (Diversity X Diagnostic and Diversity X Interactive) were formed by multiplying the mean-centred independent variables. Given that the mean centring of the independent variables does not change the value and the statistical significance of the regression coefficient of interaction term for the unstandardized regression solution (Southwood, 1978), it is recommended for two reasons: (1) it minimizes the correlation between the independent variables and their products, thus reducing the potential problems caused by multicollinearity (Aiken & West, 1991; Jaccard Turrisi, 2003) and (2) it makes the simple effects (conditioning effects)[71] interpretable. Since interval scales are used in this study, the simple effects (β_1 and β_2) in equation (2) are not meaningful and could not be interpreted as Southwood (1978) illustrated because there is no zero value for Z. Mean centring the independent variable allows us to attach a meaningful interpretation of the simple effects. Now, β_1 represents the simple effect of X on Y when Z=0.

However, before interpreting the results of the regression analyses, the assumptions underlying Ordinary Least Squares (OLS) estimated regression must be examined. In order for the OLS estimator to the Best Linear Unbiased Estimator (BLUE)[72], the following classical assumptions given by Studenmund (2001) must be met:

I. The regression model is linear in the coefficient; is correctly specified; and has an additive error term.

---

[71] Simple effects are different than main effects. As noted by Aiken & West (1991: 37), "[c]onditional effects describe the effects of one predictor on the criterion variable under the condition in which the other predictor equals a specified value". Contrast this with the interpretation of main effects in equation (1) where the regression coefficient of X estimates the effect X on Y for each level of Z and the regression coefficient of Z estimates the effect of Z on Y for every level of X (Jaccard & Turrisi, 2003)

[72] The proof of this is given by the Gauss-Markov theorem.

II. The error term has zero population mean.

III. All explanatory variables[73] are uncorrelated with the error term.

IV. Observations of the error term are uncorrelated with each other (no serial correlation/no autocorrelation).

V. The error term has a constant variance (no heteroscedasticity)

VI. No explanatory variable is a perfect linear function of any other explanatory(s) (no perfect multicollinearity).

VII. The error term is normally distributed (this assumption is optional). Meeting this option in addition to the other 6 makes OLS the Best Unbiased Estimator (BUE) between all linear and non-linear estimators.

Studenmund (2001) notes that if the seven classical assumptions are met, the OLS coefficient estimators will have the following properties:

1- *They are unbiased.* This means that the OLS estimates of the regression coefficients are centred on the true population parameters.

2- *They are efficient.* The OLS estimates of the regression coefficients have the minimum variance around the true population parameters.

3- *They are consistent.* As the sample size approaches infinity, the OLS estimates of the regression coefficients converge on the true population parameters.

4- *They are normally distributed.* This property allows for the application of various statistical tests that require the estimated regression coefficients to be normally distributed (e.g. t test).

Table 4-15 sets out the framework for testing the classical assumptions of regression that will be used in this study. Studenmund (2001) notes that there is no need to test if the error term has zero population mean because it is always

[73] The terms explanatory variables and independent variables are used interchangeably in this study.

forced to zero in the regression equation by the existence of the constant term. Hence, it is not included in the proposed framework.

Table 4-15: Framework for testing classical regression assumptions (Source: Hair et al., 2006; Studenmund, 2001)

Assumption	Test Procedure
Linearity	▪ Plot of standardized residuals versus the predicted values
Homoscedasticity (No heteroscedasticity)	▪ Plot of standardized residuals versus the predicted values
No serial correlation/No auto correlation	▪ Plot of standardized residuals versus sequencing variable (case ID that is sequenced based on time of arrival) ▪ Perform Durbin Watson test
Normality	▪ Graph histogram of standardized residuals. ▪ Plot Normal P-P plot of standardized residuals
Multicollinearity	▪ Assess correlation matrix ▪ Calculate tolerance and VIF values
Explanatory variables are uncorrelated with error term	▪ Plot of standardized residuals versus each of the explanatory variables.

The equation for the hypothesized moderating effects of performance measurement use (diagnostic and interactive) on the relationship between performance measurement diversity and innovation search is given by the following equation:

$$Search_i = \beta_0 + \beta_1\, Diversity_i + \beta_2\, Diagnostic_i + \beta_3\, Interactive_i + \beta_4(Diversity_i\, X\, Diagnostic_i) + \beta_5(Diversity_i\, X\, Interactive_i) + \beta_6\, Size_i + \beta_7\, Age_i$$

Where,

$Diversity_i$ = Performance measurement diversity for i^{th} firm

$Daignostic_i$ = Daignostic Use for i^{th} firm

$Interactive_i$ = Interactive Use for i^{th} firm

$Diversity_i$ X $Diagnostic_i$ = Interaction term between Diversity and Daignostic Use for i^{th} firm

$Diversity_i$ X $Interactive_i$ = Interaction term between Diversity and Interactive Use for i^{th} firm

$Search_i$ = Innovation Search for i^{th} firm

$Size_i$ = Organizational Size for i^{th} firm

Age_i = Organizational Age for i^{th} firm

β_n = Parameters to be estimated

i = Firm = 1, 2, 3... n where n = total number of firms (observations)

The equation for the hypothesized moderating effects of performance measurement use (diagnostic and interactive) on the relationship between performance measurement diversity and risk is given by the following equation:

$$Risk_i = \beta_0 + \beta_1\, Diversity_i + \beta_2\, Diagnostic_i + \beta_3\, Interactive_i + \beta_4\, (Diversity_i\, X\, Diagnostic_i) + \beta_5\, (Diversity_i\, X\, Interactive_i) + \beta_6\, Size_i + \beta_7\, Age_i$$

Where,

$Risk_i$ = Risk taking for i^{th} firm

The equation for the hypothesized moderating effects of performance measurement use (diagnostic and interactive) on the relationship between performance measurement diversity and product innovation is given by the following equation:

$$Product\_Innovation_i = \beta_0 + \beta_1\, Diversity_i + \beta_2\, Diagnostic_i + \beta_3\, Interactive_i + \beta_4\, (Diversity_i\, X\, Diagnostic_i) + \beta_5\, (Diversity_i\, X\, Interactive_i) + \beta_6\, Search_i + \beta_7\, Risk_i + \beta_8\, Size_i + \beta_9\, Age_i$$

Where,

$Product\_Innovation_i$ = Product innovation for i^{th} firm

I used the statistical software SPSS version 14.0 to run multiple regression analysis with the Ordinary Least Squares (OLS) estimation method to test the hypothesized moderation effects.

4.8.2 Mediation Analysis

Hypotheses, H9 and H11, postulate that innovation search and risk taking mediate the moderating effects of performance measurement use on product innovation. Innovation search and risk are called mediators and they represent the generative mechanism through which the moderating effects are able to influence the product innovation (Baron & Kenny, 1986). As noted by Baron & Kenny (1986), a given variable may be said to function as a mediator if it accounts for the relation between the predictor (moderating effect in this study) and the criterion (product innovation).

To test the mediation hypotheses, I adopted the approach used by Singh, Goolsby, & Rhoads (1994) and Tippins & Sohi (2003). I checked for the presence of a mediating effect, by performing a competing model analysis (i.e., two substantive models are estimated and evaluated for significant differences). As shown in Figure 4-17, the first model (direct effects) examines the direct relationship between the moderating effects of performance measurement use on product innovation, while a second model (partial mediation) shown in Figure 4-18 represents the posited hypotheses examining the same relationship with innovation search and risk as mediators.

The mediating effects of innovation search and risk on the relationship between the moderating effects of performance measurement use are said to be supported when:

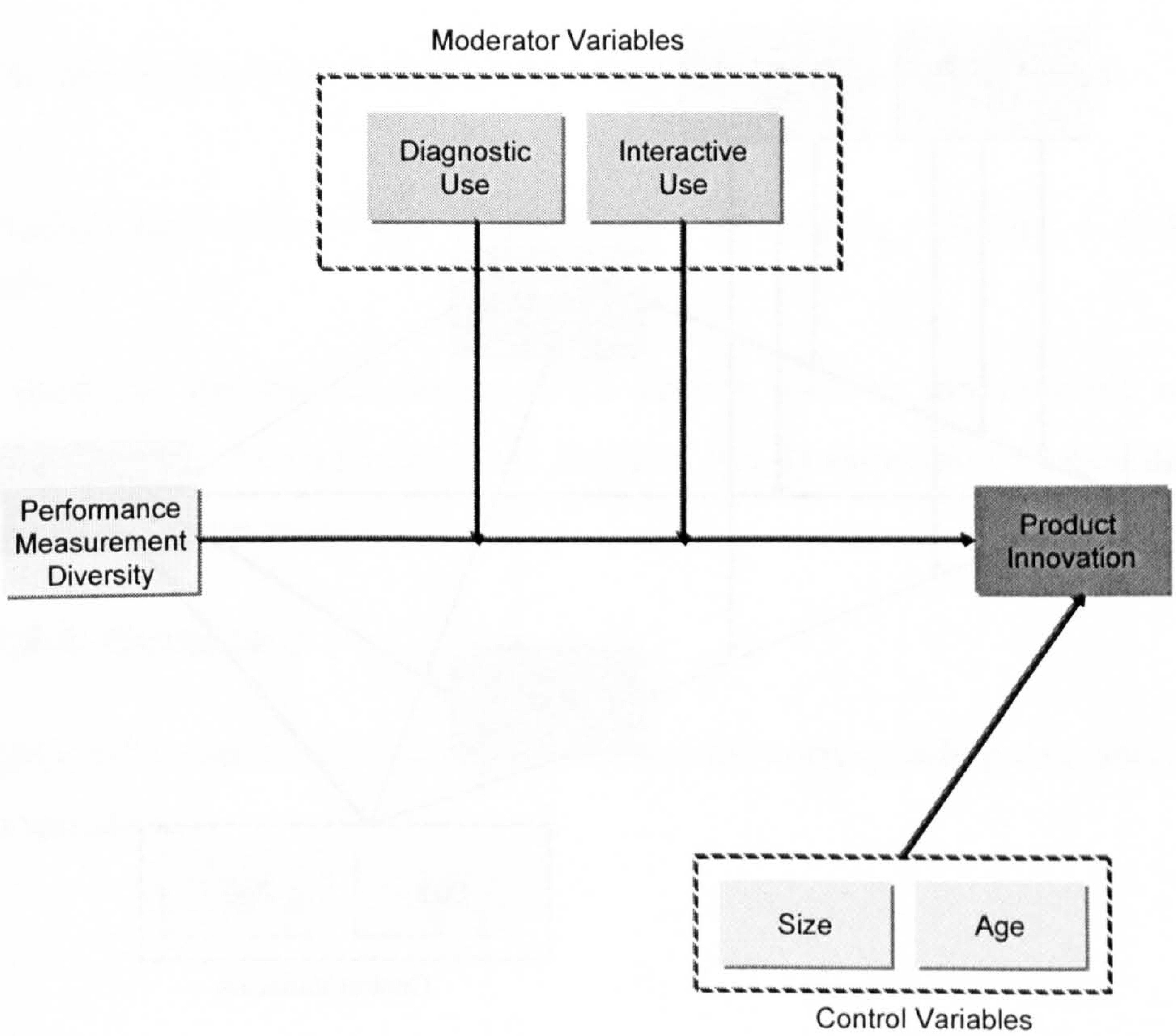
1) The partial mediation model explains more variance (i.e., R^2) in product innovation than the direct model;
2) There is a significant relationship between moderating effects of performance measurement use and innovation search and risk;

3) A significant relationship between moderating effects of performance measurement use and product innovation (as observed in the direct model) is greatly diminished or eliminated in the partial mediation model; and
4) There is a significant relationship between innovation search and risk and product innovation.

The first model (direct effects) shown in Figure 4-17 is represented by the following equation and illustrated in Figure 4-17.

$$Product\_Innovation_i = \beta_0 + \beta_1\ Diversity_i + \beta_2\ Diagnostic_i + \beta_3\ Interactive_i + \beta_4\ (Diversity_i\ X\ Diagnostic_i) + \beta_5\ (Diversity_i\ X\ Interactive_i) + \beta_6\ Size_i + \beta_7\ Age_i$$

Figure 4-17: Direct effects model



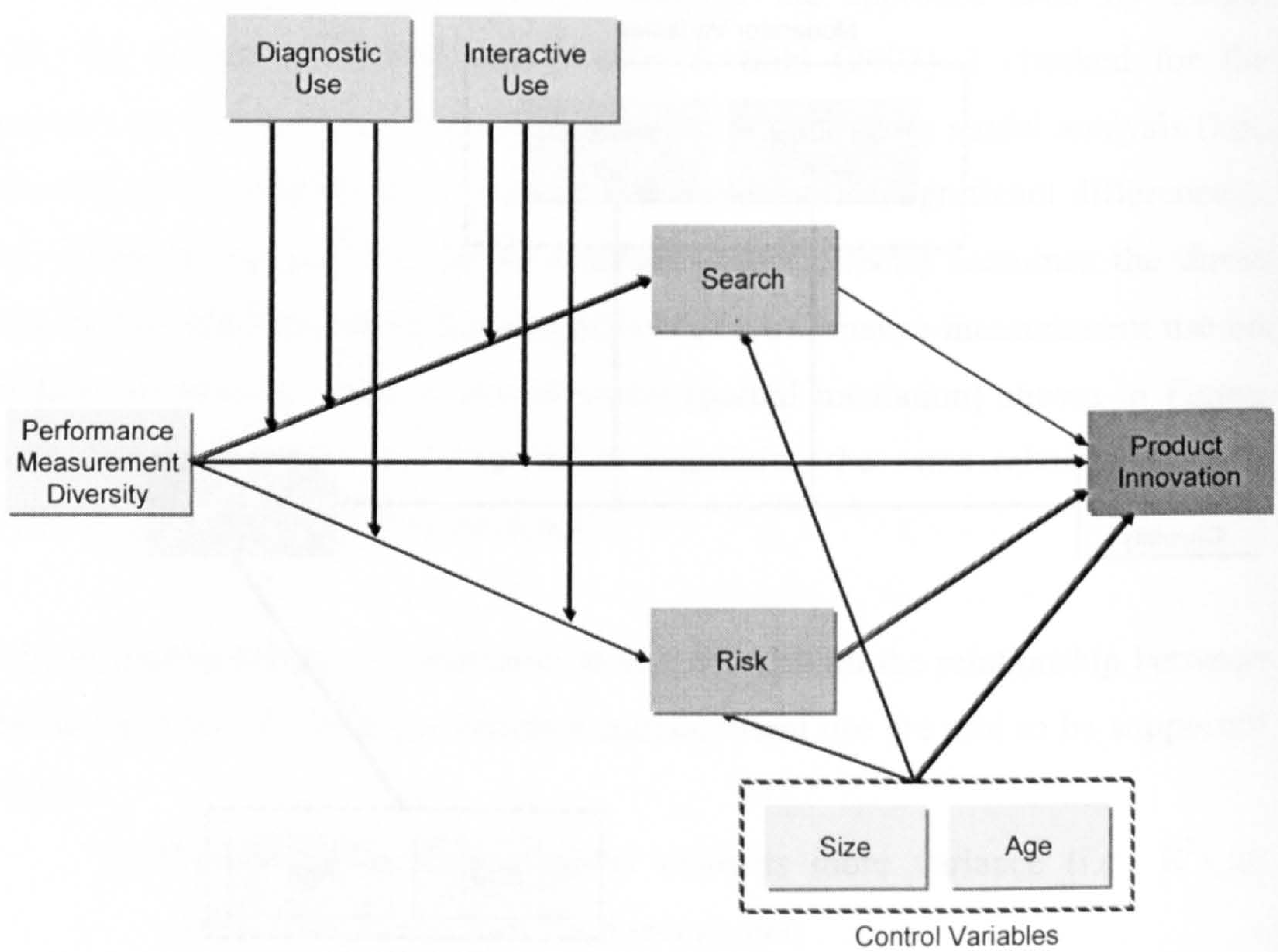
The second model (partial mediation) shown in Figure 4-18 is represented by the following equations and illustrated in Figure 4-18.

$Search_i = \beta_0 + \beta_1\, Diversity_i + \beta_2\, Diagnostic_i + \beta_3\, Interactive_i + \beta_4\, (Diversity_i \times Diagnostic_i) + \beta_5 (Diversity_i \times Interactive_i) + \beta_6\, Size_i + \beta_7\, Age_i$

$Risk_i = \beta_0 + \beta_1\, Diversity_i + \beta_2\, Diagnostic_i + \beta_3\, Interactive_i + \beta_4\, (Diversity_i \times Diagnostic_i) + \beta_5\, (Diversity_i \times Interactive_i) + \beta_6\, Size_i + \beta_7\, Age_i$

$Product\_Innovation_i = \beta_0 + \beta_1\, Diversity_i + \beta_2\, Diagnostic_i + \beta_3\, Interactive_i + \beta_4\, (Diversity_i \times Diagnostic_i) + \beta_5\, (Diversity_i \times Interactive_i) + \beta_6\, Search_i + \beta_7\, Risk_i + \beta_8\, Size_i + \beta_9\, Age_i$

Figure 4-18: Partially mediated model



I used the statistical software AMOS version 6.0 (Arbuckle, 2005) to run path analysis (structural equations modelling technique) with the maximum likelihood (ML) estimation method to test the mediation hypotheses. Although I could have also used multiple regression analysis to test the mediation hypotheses, I selected path analysis because it has two major advantages: (1) there are numerous statistical indexes of overall fit that can be used to assess the model fit and (2) most of the results that have to be computed manually in multiple regression analysis are provided by the AMOS V.6 structural equation modelling software (Maruyama, 1997).

4.8.3 Direct Effects Analysis

Hypotheses, H1, H8 and H10, postulate that performance measurement diversity, innovation search, and organizational risk taking have direct positive effect on product innovation.

The direct effect hypotheses can be tested by the following equations:

$$Product\_Innovation_i = \beta_0 + \beta_1\ Diversity_i + \beta_2\ Search_i + \beta_3\ Risk_i + \beta_4\ Size_i + \beta_5\ Age_i$$

I used the statistical software SPSS version 14.0 to run multiple regression analysis with the Ordinary Least Squares (OLS) estimation method to test the direct effect hypotheses.

4.8.4 Summary

Table 4-16 summarizes the hypotheses and the corresponding data analytics used to test them.

Table 4-16: Summary of the hypotheses and the data analytics used to test them

No.	Hypothesis	Statistical Model(s)	Data Analytic	Test Parameter	Predicted Sign
Direct Effect Hypotheses					
H1	There is a positive relationship between performance measurement diversity and product innovation.	$Product\_Innovation_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Search_i + \beta_3 Risk_i + \beta_4 Size_i + \beta_5 Age_i$	OLS Regression	β_1	+
H8	There is a positive relationship between innovation search and product innovation.			β_2	+
H10	There is a positive relationship between organizational risk taking and product innovation.			β_3	+
Moderating Hypotheses					
H2a	Diagnostic use moderates the form of the relationship between performance measurement diversity and product innovation.	$Product\_Innovation_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Search_i + \beta_7 Risk_i + \beta_8 Size_i + \beta_9 Age_i$	Moderated regression	β_4	+
H2b	When levels of diagnostic use are high, performance measurement diversity will be positively related to product innovation.	*The impact of performance measurement diversity on product innovation is non-monotonic over the range of diagnostic use. For higher levels of diagnostic use, the relationship between performance measurement diversity and product innovation is positive and for lowers levels of diagnostic use, the relationship is negative.*	Partial derivative	Partial derivative crosses the diagnostic use (horizontal) line over the range of diagnostic use.	
H2c	When levels of diagnostic use are low, performance measurement diversity will be negatively related to product innovation.				
H3a	Interactive use moderates the form of the relationship between performance measurement diversity and product innovation.	$Product\_Innovation_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Search_i + \beta_7 Risk_i + \beta_8 Size_i + \beta_9 Age_i$	Moderated regression	β_5	-
H3b	When levels of interactive use are high, performance measurement diversity will be more negatively related to product innovation.		Partial derivative	∂ (Product_Innovation)/ ∂(Diversity) has a more negative value when interactive use is high than when it is low.	
H4a	Diagnostic use moderates the form of the relationship between performance measurement diversity and innovation search.	$Search_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$	Moderated regression	β_4	+
H4b	When levels of diagnostic use are high, performance measurement diversity will be positively related to innovation search.	*The impact of performance measurement diversity on innovation search is non-monotonic over the range of diagnostic use. For higher levels of diagnostic use, the relationship between performance measurement diversity and innovation search is positive and for lowers levels of diagnostic use, the relationship is negative.*	Partial derivative	Partial derivative crosses the diagnostic use (horizontal) line over the range of diagnostic use.	
H4c	When levels of diagnostic use are low, performance measurement diversity will be negatively related to innovation search.				
H5a	Interactive use moderates the form of the relationship between performance measurement diversity and innovation search.	$Search_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$	Moderated regression	β_5	-

No.	Hypothesis	Statistical Model(s)	Data Analytic	Test Parameter	Predicted Sign
H5b	When levels of interactive use are high, performance measurement diversity will be more negatively related to innovation search.		Partial derivative	∂ (Search)/ ∂(Diversity) has a more negative value when interactive use is high than when it is low.	
H6a	Diagnostic use moderates the form of the relationship between performance measurement diversity and organizational risk taking.	$Risk_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$	Moderated regression	β_4	+
H6b	When levels of diagnostic use are high, performance measurement diversity will be positively related to organizational risk taking.		Partial derivative	∂ (Risk)/ ∂(Diversity) has a more positive value when diagnostic use is high than when it is low.	
H7a	Interactive use moderates the form of the relationship between performance measurement diversity and organizational risk taking.	$Risk_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$	Moderated regression	β_5	-
H7b	When levels of interactive use are high, performance measurement diversity will be more negatively related to organizational risk taking.		Partial derivative	∂ (Risk)/ ∂(Diversity) has a more negative value when interactive use is high than when it is low.	
Mediating Hypotheses					
H9	Innovation search mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.	*Two competing models are compared for significant difference: direct effect model and partial mediation model.* Direct Effect Model: $Product\_Innovation_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i X Diagnostic_i) + \beta_5 (Diversity_i X Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$ Partial Mediation Model. $Search_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i x Diagnostic_i) + \beta_5 (Diversity_i x Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$ $Risk_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i x Diagnostic_i) + \beta_5 (Diversity_i x Interactive_i) + \beta_6 Size_i + \beta_7 Age_i$ $Product\_Innovation_i = \beta_0 + \beta_1 Diversity_i + \beta_2 Diagnostic_i + \beta_3 Interactive_i + \beta_4 (Diversity_i x Diagnostic_i) + \beta_5 (Diversity_i x Interactive_i) + \beta_6 Search_i + \beta_7 Risk_i + \beta_8 Size_i + \beta_9 Age_i$	Path Analysis	ΔR^2	
H11	Organizational risk taking mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.				

4.9 Summary

Table 4-17 shows how the research design complies with all the survey research design criteria advanced by Malhotra & Grover (1998).

Table 4-17: Assessing Survey Research Design (Source: Adapted from Malhotra & Grover, 1998)

Area	Ideal Survey Attributes	Met? (Y/N)	Comments
General	Is the type of survey research identified?	Y	The research design is defined as explanatory survey. Hypotheses are stated based on theoretical model.
	Is the research identified as cross-sectional or longitudinal?	Y	The research design is defined as cross-sectional survey.
	Is the unit of analysis clearly defined for the study?	Y	A formal statement defining the unit of analysis was offered.
	Is the unit of analysis justified?	Y	
	Does the instrumentation consistently reflect that unit of analysis?	Y	
	Is the respondent(s) chosen appropriate for the research question?	Y	I followed the guidelines set by Huber & Power (1985).
	Is any form of triangulation used to cross validate results?	Y	The survey was sent to more than one respondent in organization.
Measurement Error	Are multi-item variables used?	Y	All scales use multi-item. The number of the items per scale ranged from 4-10 items.
	Is content validity assessed?	Y	
	Is field-based pretesting of measures performed?	Y	The measures were tested on academics.
	Is reliability assessed?	Y	
	Is construct validity assessed?	Y	
	Is pilot data used for purifying measures or are existing validated measures adapted?	Y	All the measures used in the study are existing measures that have been validated and extensively used.
	Are confirmatory methods used	Y	
Sampling Error	Is the sample frame defined and justified?	Y	
	Is random sampling used from the sample frame?	Y	Systematic probability sampling was used.
	Is the response rate over 20%?	Y	
	Is non-response bias estimated?	Y	
Internal Validity Error	Are attempts made to establish internal validity of the findings?	Y	I included firm level and industry level controls in my statistical model to establish internal validity of the findings.
Statistical Conclusion Error	Is there sufficient statistical power to reduce statistical conclusion error?	Y	Sample size analysis was undertaken.

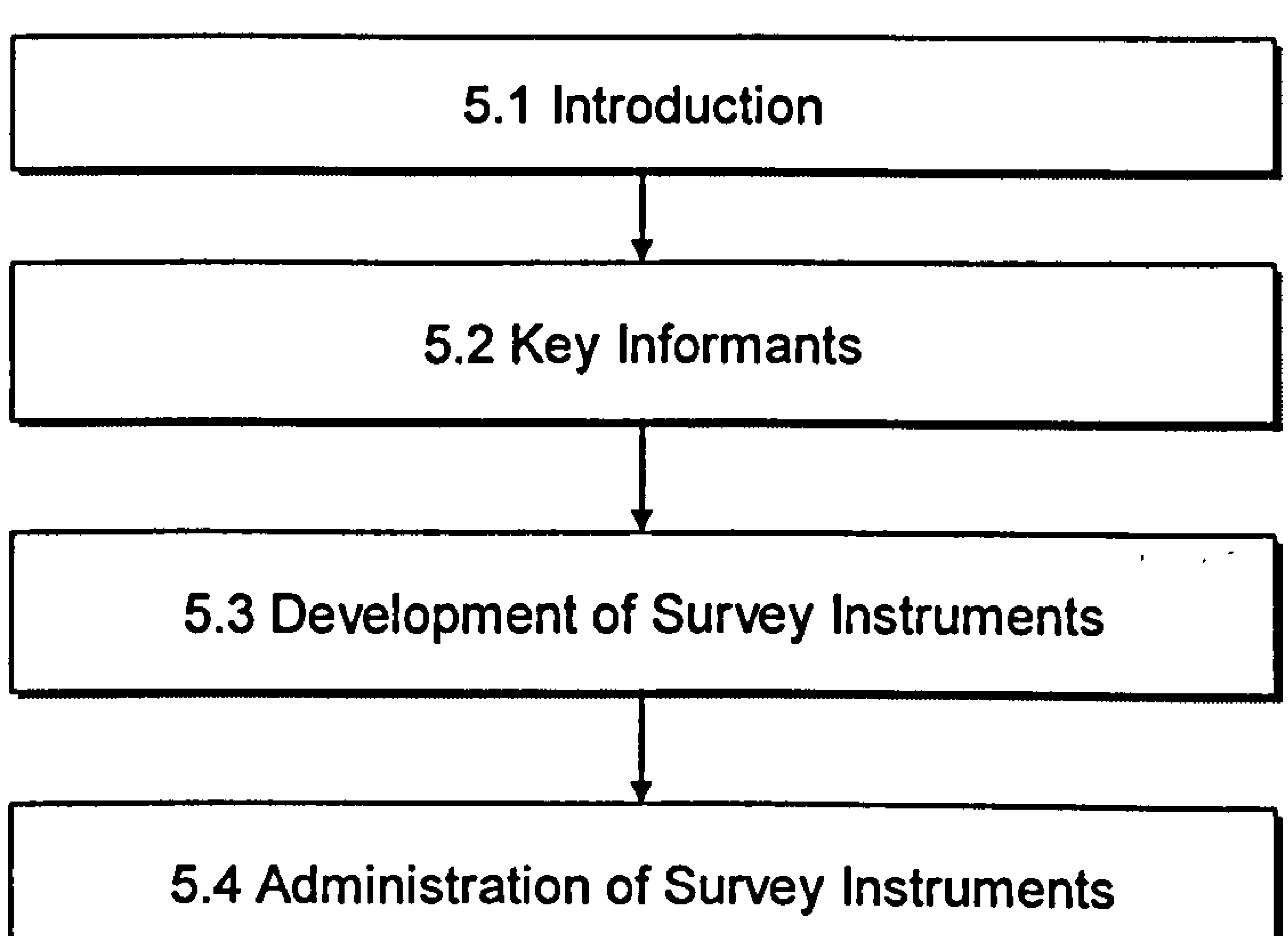
5 Data Collection

5.1 Introduction

The previous chapter examined the methodology of the study to ensure that its design is appropriate to provide answers to the research questions and to test the research hypotheses. More specifically, a cross-sectional probability sample, survey research design was chosen and justified and the study constructs were operationalised using existing validated scales. This chapter explains how the survey instruments are developed incorporating the chosen validated scales in the previous chapter and how they are administered to collect data on the variables under study.

There are four sections in this chapter, as illustrated in Figure 5-1. The first section introduces the chapter. In the second section, I define the key informants in this study and I detail the guidelines used to minimize the inaccuracy and bias of their retrospective accounts. In the third section, I explain the development of the mail-based and web-based versions of the survey using Dillman's (1978) total design method. In the fourth section, I explain the administration of the mail- and web-based surveys using Dillman's TDM procedure and I report the results.

Figure 5-1: Outline of chapter 5



5.2 Key Informants

The key informants in this study are defined as members of the top management team. Snow & Hrebiniak (1980) posit, "top managers have the best vantage point for viewing the entire organizational system". In defining what constitutes top management team, referred to as TMT[74], I adopt the "strategic issue processing view" advocated by Dutton, Fahey, & Narayanan (1983) and Jackson (1992). Finkelstein & Hambrick (1996) explain that such perspective "assumes that the top decision-making body is not constant and implies that the appropriate definition of a TMT is the set of executives who are most active on a particular issue." This perspective is also compatible with the dominant coalition theory advanced by March (1962) and Cyert & March (1963).

I have made every effort to contact the organizations in the sample to identify the appropriate key informants (Huber & Power, 1985; Mitchell, 1994) who could provide accurate information regarding the constructs of interest. I accomplished this using several approaches:

1. I called the firms in the sample and I asked them to name the executive who should receive the survey;
2. I used several databases and directories to identify the key executives (e.g., FAME, Cranfield Database, Thompson/Gale (Goliath), Proquest Industry News, DTI industry directories;
3. I consulted the firms' web sites, press releases and annual reports if available; and
4. I consulted trade journals and web portals such as TheManufacturer.com.

According to Huber and Power's (1985), "[t]here are four primary reasons that informants provide inaccurate or biased data.

1. They are motivated to do so.
2. Their perceptual and cognitive limitation result in inadvertent errors.
3. They lack crucial information about the event of interest.

[74] Finkelstein & Hambrick (1996) offer five different operationalisation of TMT. For a recent comprehensive review of the TMT literature, refer to (Carpenter, Geletkanycz, & Sanders, 2004).

4. They have been questioned with an appropriate data elicitation procedure."

Table 5-1 sets how I followed the guidelines that have been advanced by Huber & Brown (1985) to minimize the inaccuracy and bias of retrospective reports. The first column divides the guidelines into three areas: selecting key informants; motivating key informants to participate; and ensuring data accuracy. The three areas address the above reasons numbered 3, 1, and 4, respectively. Although guidelines set in the three areas may also act on the second reason, they can not completely eliminate the errors. The second column states the guideline to be followed. The third column specifies if the guideline has been implemented. Where it was not possible to conform, or amendments were made, the reasons have been given under the 'comments' column.

Table 5-1: Key informants strategies (Source: Huber & Power, 1985)

Issue	Guideline	Met (Y/N)	How It is met OR Why It is not met
Selecting key informants	If only one informant per organization is to be questioned, attempt to identify the person most knowledgeable about the issue of interest.	Y	Top management team was chosen because they are the most knowledgeable of the organizational-level variables under investigation.
	If there is more than one informant per unit of analysis, choose informants whose unique biases or lack of knowledge are likely to offset those of other informants.	Partial	This was partially met because the questionnaire was sent to more than one member of the top management team (Bowman & Ambrosini, 1997). However, given that the data collection method was carried through self- administered survey and the limitation of information on top management team, it was not possible to select respondents based on their potential biases and emotional involvement. However, this problem has been minimized by asking respondents to refer to their organizations in their responses.
	When choosing key informants, recognize that the person's emotional involvement may either increase or decrease the accuracy of the responses. 1. Choose informants with moderate levels of emotional involvement. 2. Seek factual data from informants with higher emotional involvement, as their ability to recall is probably greater, and seek judgmental data from those with lower involvement, as their responses are less likely to be distorted by their motives.	Partial	
Motivating participation	Attempt to motivate the informants to co-operate. 1. Remove as many disincentives to responding as possible. a. Ensure the anonymity and confidentiality of the responses. b. Minimize respondent effort. 2. Explain how the research results can be useful to the manager, organization, the field of study, and the researcher.	Y	I reduced respondents' disincentives to respond by offering them anonymity and confidentiality. I also minimized their effort to complete the survey by reducing the time for survey completion to 15 minutes. I informed the respondents of the importance of the research to them, their organization, performance management discipline and UK industry in general. This assertion was strengthened by informing them that this research is funded by the Department of Trade and Industry (DTI) and the Engineering and Physical Science Research Council (EPSRC) in the mail-based survey. By participating in the survey, the respondents were promised a free copy of the study report. Furthermore, 100 free tickets to attend the conference in which the results of the study would be presented were promised the first 100 responding organizations in the mail-based survey.
Enhancing the accuracy of data	Minimize the elapsed time between the events of interest and the collection of data.	Y	To minimize respondent recall problems, none of the questions in the survey required respondents to recall data over than two years old.
	Consider how the framing of questions will affect the informant's response.	Y	Questions were framed using the guidelines set by Dillman (2000) and by using existing and validated scales (Malhotra & Grover, 1999; Forza, 2002).
	Use questions that are pretested, structured and that impart an image of being rich in information content without being complex.	Y	The questions were pretested with practitioners and academics. Furthermore, all the scales used in the survey have been published in high-ranking academic journals.

5.3 Development of Survey Instruments

In order to increase the response rate, two versions of the survey were developed and subsequently administered: a mail-based version and a web-based version. Dillman's (1978) total design method, referred to as TDM, which is a comprehensive approach to the development and application of mail surveys was used to develop and administer the survey. It represents best practice in survey research and it has been recommended by many methodological researchers (see for example, Slater & Atuahene-Gima, 2004; Forza, 2002; Young, 1996).

Table 5-2 is based on the TDM recommendations and was tabulated by Ranaweera (2002). I updated the original table contents to incorporate the new recommendations in Dillman's (2000) tailored design method[75]. Table 5-2 sets how the mail-based survey development and administration procedure that was followed conform to the recommendations set by the TDM method. The first column groups the recommendations into nine areas. The second column states the recommendation. The third column specifies if the recommendation has been followed. Where it was not possible to conform, or amendments were made, the reasons have been given under the 'comments' column.

Table 5-2: Dillman's (1978; 2000) TDM recommendations (Source: adapted from (Ranaweera, 2002: 150-152)

TDM Recommendation		Met? (Y/N)	Comments
Appearance	Print questionnaire as a booklet – concise, professional, easy to read	Y	
	No questions on front or back cover	Y	
	Questionnaire reduced to booklet size	Y	
	Reproduction on good quality white or off white paper	Y	The cover page was printed in colour to increase the attractiveness of the survey. The rest of the survey was printed on yellow paper again to make it more distinctive. Furthermore, to reduce

[75] Note that Dillman in 1978 referred to the method as "Total Design Method" and in 2000 as "Tailored Design Method". Since both names have the same acronym 'TDM', I use TDM hence after to refer to the updated method.

TDM Recommendation		Met? (Y/N)	Comments
			the problems referred to in the TDM during the reproduction phase, a professional firm was contracted to produce the survey.
Order of questions	Order them according to importance	Partial	Some of the organizational demographics were placed in the beginning of the survey.
	Keep questions that are of similar content together	Y	This was followed though out the survey. Questions measuring the same construct were grouped together.
	Keep a logical flow. Dillman argued that there is no evidence to say that validity of measures is compromised in anyway by this ordering	Y	
	Keep objectionable questions till the end	Y	No questions in the survey could be called 'objectionable'. All the questions referred to the organization under study.
Question	Be clearly related to the survey topic	N	After several consultations on this issue with subject experts, it was deemed necessary to start the survey by asking respondents to the type of organization they work for because I am interested in firm-level constructs.
	Easy to answer	Y	
	Convey a sense of neutrality (Avoid agree/disagree type)	Y	None of the questions used agree/disagree anchors. Furthermore, all the questions in the survey conveyed a sense of neutrality by using anchors such as To a very great extent/Not at all.
	Be clearly applicable to everyone	Y	The survey was applicable to the targeted respondents.
	Be interesting to everyone	Y	The survey is interesting to the targeted respondents.
Page design	Clearly distinguish questions from answers	Y	All the question formats used in the survey allowed the separation of questions and answers to facilitate answering questions and reduce completion time.
	Number the questions	Y	
	Establish vertical flow so that answers are registered in a vertical line	Y	The tabulated format ensured that.
	Provide directions for how to answer	Y	Whenever directions were warranted, they were used. For example before asking the questions on product innovations, I advanced a definition for product innovation. However, most of the questions were Likert- type with clear anchors that did not require any explanation.
	Make questions fit each page	Y	None of the items related to the same question were in separate pages.
Front cover design	Create a positive first impression	Y	A large Cranfield University logo was used at the top. Also, the

TDM Recommendation		Met? (Y/N)	Comments
			Centre for Business Performance logo was included. The logos portray integrity, the importance of the study, and academic need.
	Study title	Y	The study title is the "Impact of performance measurement & management systems." The title was preceded by "Cranfield Survey" label to indicate the academic nature of the survey and to increase its credibility.
	A graphic illustration	Y	To increase the appeal of the first page, a graphic illustration was incorporated.
	Any needed directions	N	The directions were given in the cover letter.
	Name and address of the sponsor	N	The sponsor information was given in the cover page. The sponsor of this research is EPSRC.
Back cover design	Keep it simple not to compete with front cover	N	Questions were included in the back cover making it difficult to reduce complexity in the page design.
	Invitation to make additional comments – respondents appreciate this opportunity (in addition to value per se)	N	This was implemented in the cover page.
	A thank you	N	
	Plenty of space	N	I did not have plenty of space because I included questions in the last page.
	Absolutely no questions on back cover since the last questions are bound to be the most objectionable	N	
	Promise of a summary of the results	N	This was implemented in the cover page.
Cover letter	It should comfortably fit on one side of A4	Y	
	Explain study and describe it as useful	Y	I informed the respondents of the importance of the research to them, their organization, performance management discipline and UK industry in general. This assertion was strengthened by informing them that this research is funded by the Department of Trade and Industry (DTI) and the Engineering and Physical Science Research Council (EPSRC).
	State that respondent is important to success of study	Y	
	Ensure confidentiality	Y	
	Other important messages	Y	It was mentioned that this is the first survey of its kind in the UK.
	Select the date on the letter appropriately	Y	
	Individually address the respondent	Y	See the key informant strategies table for further details on how this has been achieved.
	Reproduce on quality letterheads	Y	The cover letter was produced in

TDM Recommendation		Met? (Y/N)	Comments
			the official colour letter head of Cranfield School of Management.
	Put signature on each letter	Y	Each letter was hand signed.
The envelope	The aim should be to ensure that the envelope is opened	Y	The envelope carried the Cranfield School of Management logo. This would have made most respondents inquisitive.
	Use first class stamps	Y	
	Individually type the names and addresses on to the envelope itself	Y	
	Include individual identification numbers		
	Include a pre-addressed , stamped envelope (results have shown that business reply envelopes result in lower response)	Y	
	Select a suitable mail out date	Y	
Administration	A brief letter that is sent to the respondent a few days prior to the questionnaire.	Y	
	A questionnaire mailing that includes a detailed cover letter explaining why a response is important.	Y	
	A thank you post card that is sent a few days to a week after the questionnaire.	N	
	A replacement questionnaire that is sent to nonrespondents 2-4 weeks after the previous questionnaire mailing.	Y	
	A final contact that may be made by telephone a week or so after the fourth contact (if telephone numbers are available). It may also be made by special delivery mail services.	N	

Before administering the survey, the survey was piloted with 4 academics. The pilot resulted in minor modifications to the wording of some questions. The final mail and web-based surveys along with the accompanying cover letters are shown in Appendices C and D, respectively.

5.4 Administration of Survey Instruments

After the mail- and web-based survey instruments were developed and piloted, they were administered to the key informants. The mail-administered survey used in this study was incorporated into a large performance measurement survey[76] undertaken by the Centre for Business Performance (CBP) at Cranfield School of Management.

The web–based survey was adapted from the mail administered survey and it contained only the scales that were used in this study, so it was shorter than the mail-based survey. It was hosted in a domain that was specifically bought and designed for this study. The domain name was www.performancemeasurementsurvey.co.uk.

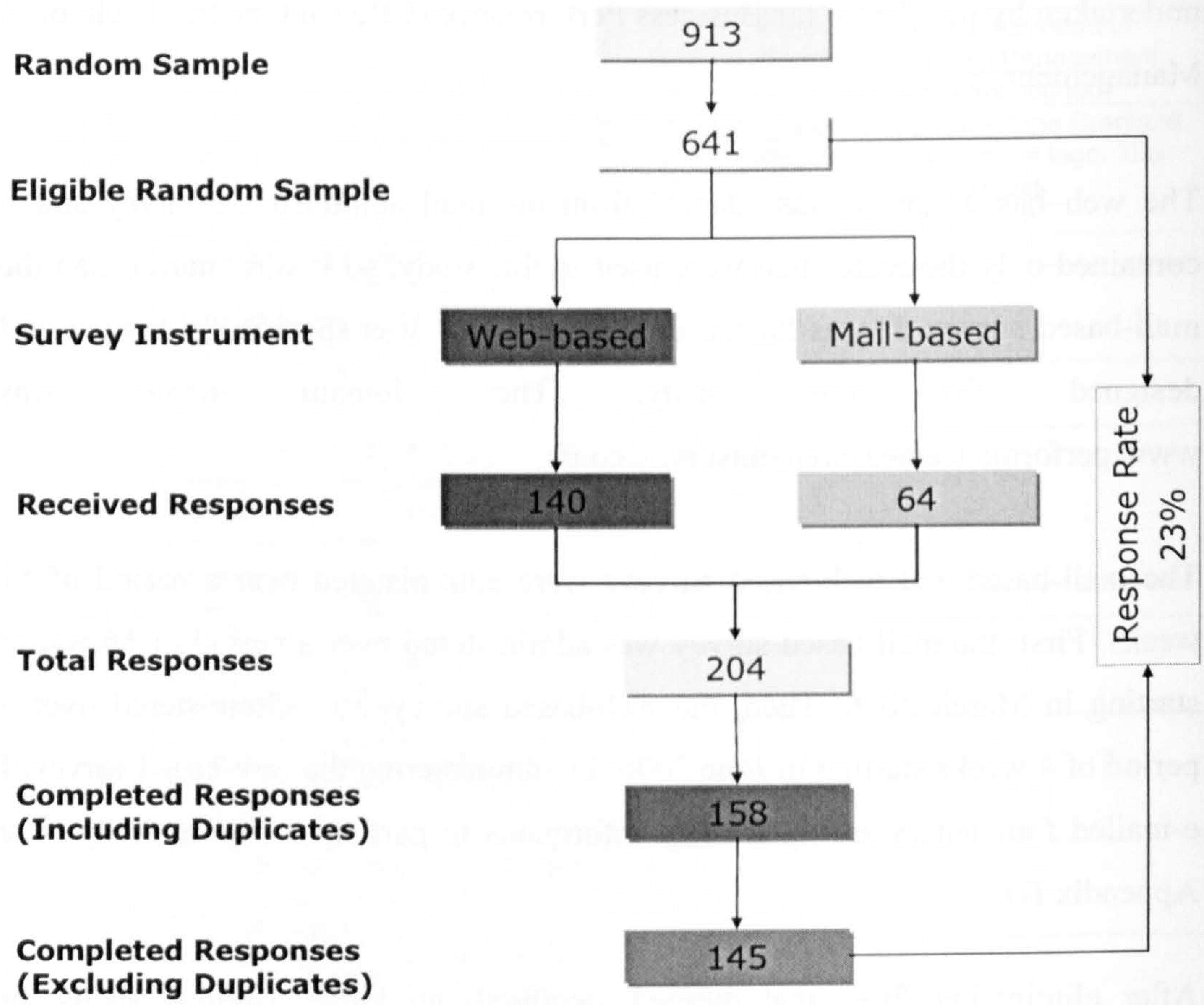
The mail-based and web-based surveys were administered over a period of 17 weeks. First, the mail-based survey was administered over a period of 10 weeks starting in March 2006. Then, the web-based survey was administered over a period of 4 weeks starting in June 2006. In administering the web-based survey, I e-mailed four letters asking the key informants to participate in the study (See Appendix D).

After eliminating firms that merged, acquired, no longer existed, shells, or holding, the initial random sample of 913 was reduced to 641. The screening of the firms was accomplished using the same different approaches used in identifying the key informants discussed in the previous section. The mail-based survey resulted in 64 responses. The web-based survey resulted in 140 responses. The total responses received were 204 responses. After eliminating the non-completed surveys, the total number of completed surveys was 158. Since there were 13 duplicates, the final number of survey responses was reduced to 145. As illustrated in Figure 5-2. The response rate in this study is 23%, which meets Malhotra and Grover's (1998) 20% response rate hurdle.

---

[76] This study was called the Impact of Performance Measurement and Management Systems and it was supported by the Engineering and Physical Science Research Council (EPSRC) [under the research grant: GR/S28846]. It was led by Dr Mike Kennerly and Dr Veronica Martinez. I would like to thank them for incorporating my survey within their study.

Figure 5-2: Responses from the surveys



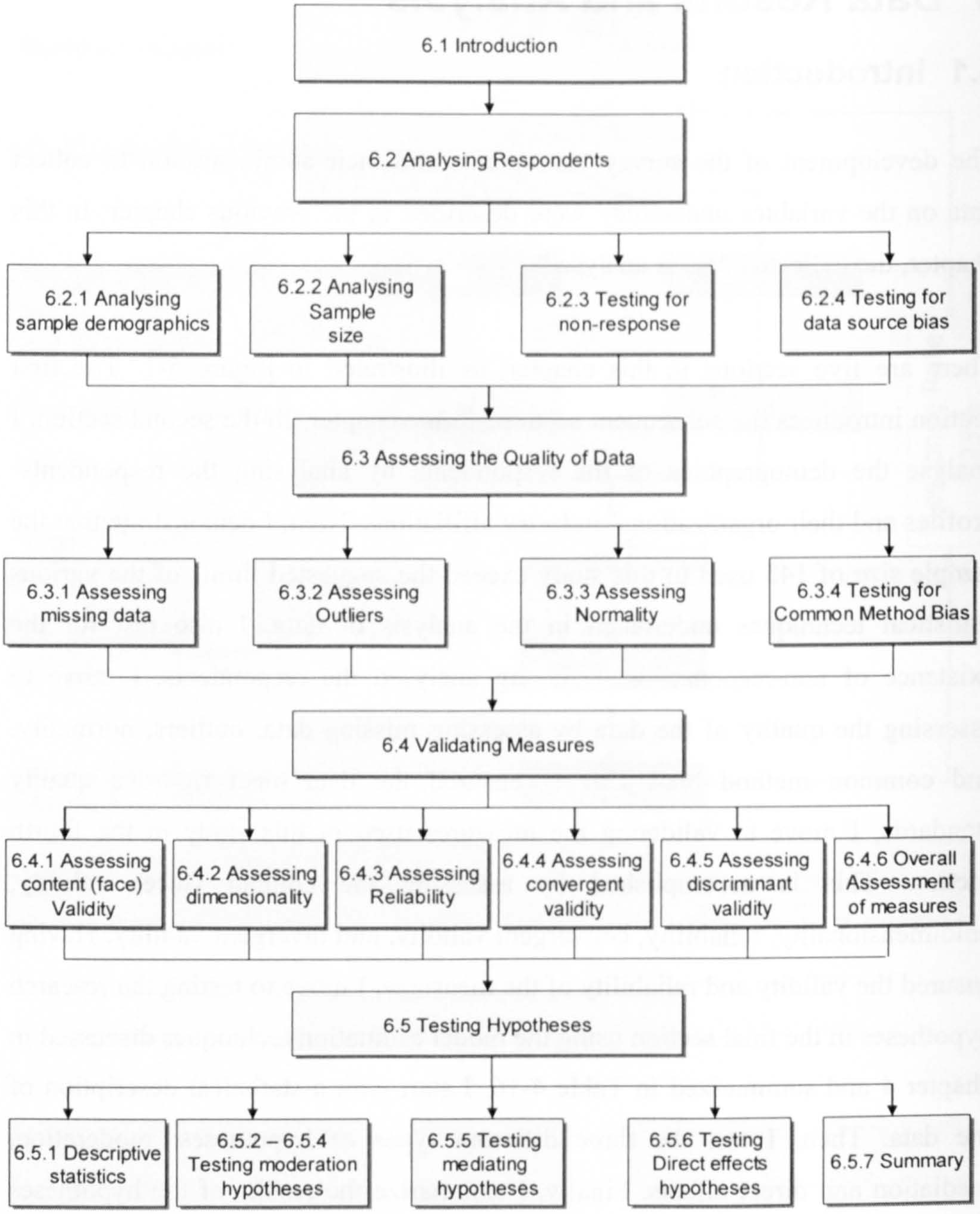
6 Data Results and Analysis

6.1 Introduction

The development of the survey instruments and their administration to collect data on the variables understudy were described in the previous chapter. In this chapter, the collected data is analysed.

There are five sections in this chapter, as illustrated in Figure 5-1. The first section introduces the subsequent sections of the chapter. In the second section, I analyse the demographics of the respondents by analysing the respondents' profiles and their organizations' industry affiliations. Next, I demonstrate that the sample size of 145 used in this study exceed the suggested limits of the various statistical techniques undertaken in the analysis of data. I also test for the existence of non-response bias. Having analysed the respondents, I move to assessing the quality of the data by assessing missing data, outliers, normality, and common method bias. Having ensured the data meet rigorous quality standards, I move to validating the measures used in this study in the fourth section. This is accomplished by assessing the content (face) validity, unidimensionality, reliability, convergent validity, and divergent validity. Having ensured the validity and reliability of the measures, I move to testing the research hypotheses in the final section using the model estimation techniques discussed in chapter 4 and summarized in Table 4-16. I start with a statistical description of the data. Then, I test the three different types of hypotheses: moderation, mediation and direct effects. Finally, I summarize the results of the hypotheses testing.

Figure 6-1: Outline of chapter 6



6.2 Analysing respondents

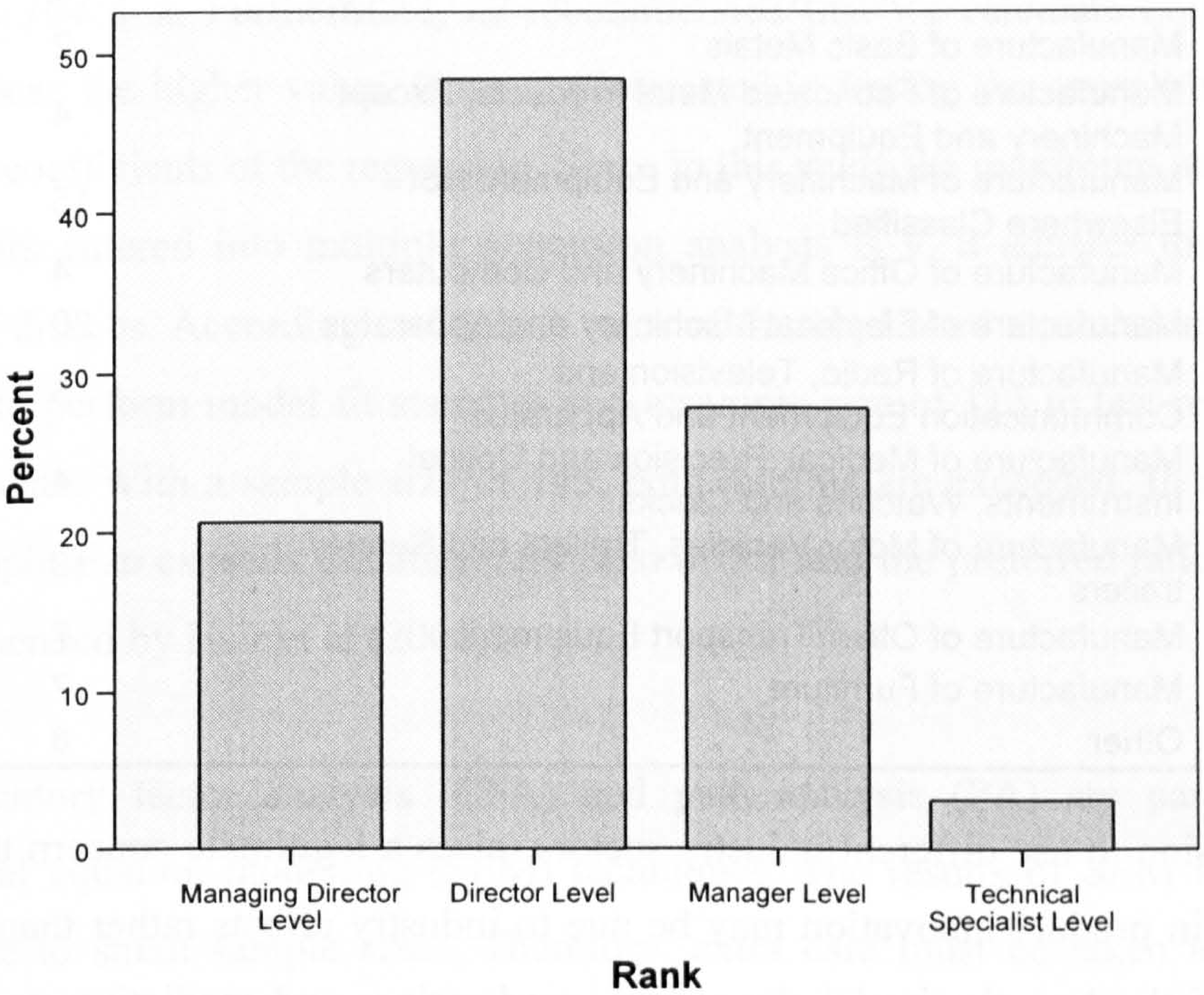
6.2.1 Analysing sample demographics

As shown in Table 6-1 and Figure 6-2, the majority of respondents filled positions with titles such as managing director, president, vice president, director, and manger. The profile of the survey's respondents was as follows: 69% were members of the top management team, 29% were managers, and only 2% were specialists. The profile of the survey's respondents suggests that respondents possess the knowledge required to answer the questions appropriately.

Table 6-1: Respondents profile

Management Rank	Frequency	Percent	Cumulative Percent
Managing Director, CEO, President, General Manager	19	19.6	19.6
Director, Vice President	48	49.5	69.1
Manager	28	28.9	97.9
Specialist	2	2.1	100.0

Figure 6-2: Respondents profile



Although the organizations that make up the sample represent various manufacturing industries (SIC[2003][77] 2-digit codes 15–36), two industries accounted for 30% of the organizations in the sample. The electrical machinery and apparatus industry (SIC 31) accounted for 15.6% and the automotive industry (SIC 34) accounted for 14.1%. These percentages are not surprising given the UK strength in these sectors. For example, the automotive industry accounts for

[77] The 2003 standard industry code classification is used in this study.

12.4% of the UK exports of goods and contributes some £9.8 billion value-added to the UK (DTI, 2006b). The cross-industry sample as discussed earlier is appropriate for this study because it allow us to generalize the findings of the research to a wider population. Table 6-2 provides information on the standard industry codes (SIC-2003) for the responding firms.

Table 6-2: Standard industry code (SIC) of the responding organizations

SIC	Description	Frequency	Percent
15	Manufacture of Food Products and Beverages	5	3.9
21	Manufacture of Pulp, Paperand Paper Products	2	1.6
23	Manufacture of Coke, Refined Petroleum Products and Nuclear Fuel	2	1.6
24	Manufacture of Chemicals and Chemical Products	10	7.8
25	Manufacture of Rubber and Plastic Products	11	8.6
26	Manufacture of Other Non-metallic Mineral Products	4	3.1
27	Manufacture of Basic Metals	2	1.6
28	Manufacture of Fabricated Metal Products, Except Machinery and Equipment	4	3.1
29	Manufacture of Machinery and Equipment Not Elsewhere Classified	15	11.7
30	Manufacture of Office Machinery and Computers	4	3.1
31	Manufacture of Electrical Machinery and Apparatus	20	15.6
32	Manufacture of Radio, Television and Communication Equipment and Apparatus	5	3.9
33	Manufacture of Medical, Precision and Optical Instruments, Watches and Clocks	4	3.1
34	Manufacture of Motor Vehicles, Trailers and Semi-trailers	18	14.1
35	Manufacture of Other Transport Equipment	7	5.5
36	Manufacture of Furniture	7	5.5
Mixed	Other	8	6.2

The decision to use different industry sectors raises a legitimate concern that the variance in product innovation may be due to industry effects rather than to the effects of the hypothesized independent, moderating, and mediating variables. However, this is unlikely in this study because I controlled for industry effects on product innovation by incorporating the industry comparison in the questions except for the last question, which is internal to the firm. This technique is widely used in performance measurement and innovation research and effectively control for industry effects on product innovation (Bisbe & Otley, 2004; Li & Atuahene-Gima, 2001, 2002; Miller & Friesen, 1983).

6.2.2 Analysing sample size

There are three primary analytical techniques used in this study to test the hypotheses: multiple regression analysis, confirmatory factor analysis, and path analysis. In order to achieve stability of research findings, the sample size should be adequate to meet the requirements of the different statistical techniques.

For multiple regression analysis, Green (1991) advanced two rules of thumb for the minimum acceptable sample size. The first is based on whether we are interested in testing the overall model fit (R^2) and is given by the formula $50 + 8k$ where k is the number of predictors and the second is based on whether we are interested in testing the individual regression coefficients and is given by the formula $104 + k$. Furthermore, he recommended that we calculate both values and choose the higher value, if we are interested in testing the overall model fit and the coefficients of the regression. Since in this study the maximum number of predictors entered into multiple regression analysis is 9, it adheres to Green's rules of thumbs. According to Green's rules of thumb, I need a minimum sample of 122 to perform model fit statistics and a sample size of 113 to test regression coefficients. With a sample size of 145, both minima are exceeded. In addition, the sample size exceeds the minimum ratio of 5:1 and the preferred ratio of 15:1 recommended by Hair et al (2006).

Confirmatory factor analysis (CFA) and path analysis (PA) are part of the structural equation modelling (SEM) techniques. The results of SEM are more sensitive to small sample sizes. Therefore, extra care must be taken to ensure adequate sample size. Although SEM methodologists do not agree on a formula for calculating the proper sample size, there is some consensus on having 5 cases per free parameter estimated (Kline, 2003). Furthermore, Hair et al (2006) note that Maximum Likelihood (ML) estimation in structural equation modelling yields stable solutions when the sample size is 100 to 150. The sample size of 145 used in the ML estimated confirmatory factor analysis and path analysis exceeds the suggested lower limit of 5:1 and is within the recommended range.

6.2.3 *Testing for Non-response Bias*

Over a period of 4 months, 158 eligible surveys were received out of 641 organizations in the target sample. 13 of the received surveys were from second respondents. Excluding the 13 duplicates brings the number of organizations that have responded 145. Thus, the response rate is 23%, which meets Malhotra and Grover's (1998) 20% response rate hurdle.

One key issue is commonly raised in survey research is non-response bias (Fowler, 1993). One method for testing for non-response bias is to test for significant differences between the early and late returned surveys (Armstrong & Overton, 1977; Lambert & Harrington, 1990). The logic underlying this test is that late responders are similar to non-responders (Armstrong & Overton, 1977).

For the present study, all the variables used in this study are compared across the early and late responders. First, the two groups were constructed from the mail-administered survey. Second, the mail respondents were compared to the web respondents. This served two purposes: (1) to test for differences between early and late responders and (2) to test for differences between the responders to the mail administered survey and responders to the web administered survey.

The results of the two tests are show in Table 6-3 and Table 6-4. The *t*-tests for both tests yielded no statistically significant differences among the 11 survey variables tested. Although these results do not rule out the possibility of non-response bias, they suggest that non-response may not be a problem to the extent that late responders represent the opinions of non-respondents. Also, the results suggest that there is no difference between responders to the mail administered survey and web administered survey.

Table 6-3: Tests of mean equality between early versus late mail- survey respondents

Variable	t	Sig. (2-tailed)
Age	-.662	.511
Diagnostic	-.245	.808

Variable	t	Sig. (2-tailed)
Interactive	-.308	.760
Diversity	-1.203	.236
Search	-.673	.504
Risk	-.124	.902
Innovation	-.785	.437
Size	-.087	.931

Table 6-4: Tests of mean equality between early mail- survey and web-survey respondents

Variable	t	Sig. (2-tailed)
Age	.397	.692
Diagnostic	1.550	.124
Interactive	-.193	.847
Diversity	-.402	.688
Search	-.176	.861
Risk	-.105	.917
Innovation	-1.262	.209
Size	2.105	.091

6.2.4 Testing for Data Source Bias (Multiple rater agreement)

Following the recommendation of Malhotra & Grover (1988), I collected surveys from multiple data sources within the organization. Boyer & Verma (2000) reminds us that research that relies on single respondents may suffer from single rater bias. Single respondent bias or data source bias may result from the respondent's subjective bias that is due to the respondent's unique views and limited access to information. Several procedures have been taken to minimize this bias as discussed in the early section addressing the selection of respondents.

Unfortunately, I received only 13 responses from second sources with two of these responses coming from the same organization. So, effectively, I had only second responses from 12 organizations. However, this should not be a concern in the current study because as highlighted in the respondent profile section, most of the respondents were part of the top management team. Although, there are a number of measures designed to test for interrater agreement (see Boyer & Verma, 2002 for full discussion), they are not appropriate for this study given the small sample and the number of raters.

6.3 Assessing the Quality of the Data

6.3.1 Assessing Missing data

Missing data are a very common problem in empirical research and especially in survey research because surveys usually involve a larger number of responses and a larger number of respondents (Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2004). Missing data have two negative effects (1) they reduce statistical power and (2) they may result in biased estimates (Tsikriktsis, 2005). Thus, the first step in assessing the quality of survey data is to identify the cases with missing data and determine why the data is missing, so remedies could be applied (Hair et al., 2006).

In identifying missing data and applying remedies, the 4-step approach advanced by Hair et al (2006) was followed. The first step involved determining of the missing data is ignorable. In this study, the missing data could not be ignored. In the second step, following Hair et al's (2006) recommendation two deletion criteria were used: (1) cases that had over than 70% missing data were deleted and (2) cases that were missing the dependent variable (product innovation) were deleted. This action resulted in the deletion of 46 cases bringing the total responses from 204 to 158. Third, the missing data was diagnosed for its randomness. To find out if the missing data is missing completely at random (MCAR), I performed the Missing Value Analysis procedure in SPSS V.14. The resulting Little's MCAR test was statistically significant (Chi-Square = 1090.332, DF = 970, Sig. = .004). Therefore, we could conclude that the missing data are missing completely at random. The final step entailed choosing an imputation method to replace the missing data. The SPSS EM approach was chosen because research has shown that it works very effectively (Hair et al., 2006; Scahfer & Graham, 2002). EM[78] is Expectation-Maximization algorithm that uses maximum

[78] EM works also for missing data that is not completely missing at random (i.e. missing at random (MAR))

likelihood (ML) estimation, which makes it also appropriate for use with structural equation modelling[79] as well.

6.3.2 Assessing Outliers

Hair et al (2006) defines outliers as "observations with a *unique combination of characteristics identifiable as distinctly different* from the other observations" (Emphasis is in the original). In this study, I follow the procedure advocated by Hair et al (2006). This procedure involves classifying observations as potential outliers when the absolute values of their standard scores are greater than 3 or 4 for samples sizes greater than 80. In this study, I flagged out all the absolute standard values that are higher than 3. Table 6-5 shows that 5 cases were flagged as potential outliers. The first 4 cases were 19, 37, 47, 49, and 83 and they all have high standard values on the age of their organizations. To confirm if this was a respondent entry error, I checked secondary sources (organizations' web sites and FAME database) and I confirmed that the information is correct. For case 83, three items of the diagnostic score had a low standard value. I reviewed the remaining scale item measuring diagnostic use and I found it low. Therefore, the low values were consistently low across the items measuring the diagnostic use of performance measurement. Byrne (1995) notes that five cases are usually flagged out as potential outliers regardless if they are not truly outliers when this procedure is used. Hence, I could tentatively conclude that none of the five cases is an outlier case. Moreover, the assessment of the normality next will provide us with more information on the effect of these cases and if I need to take action on them.

Table 6-5: Potential outliers

Case No.	Item	Survey Question	Standard Value
19	Age	My organisation's age is: ------ years	3.2837
37	Age	My organisation's age is: ------ years	3.2837
47	Age	My organisation's age is: ------ years	3.8795

[79] Confirmatory factor analysis (CFA) and path analysis are two techniques that are part of the family of structural equation modelling and they will be used in this study. Both of these techniques use ML estimators.

49	Age	My organisation's age is: —— years	3.6873
83	D1	Track progress towards goals	-3.0136
	D2	Review key measures	-4.1849
	D3	Monitor results	-3.8638

6.3.3 Assessing Normality

The data was assessed for normality because it is one of the basic assumptions of multiple regression and path analysis. Hair, et al. (2006) define normality as the shape of the data distribution of individual variables and its correspondence to the normal distribution. They note that if the variation from the normal distribution is sufficiently large, all the resulting statistical tests are invalid, because normality is required to use the F and t statistics.

Two measures can be used to determine the normality of the variable: kurtosis and skewness. Kurtosis refers to the "peakedness" or "flatness" of the distribution compared to the normal distribution. Positive kurtosis values indicate leptokurtic (peaked) distribution, while negative values indicate platykurtic (flatter) distribution (Hair et al., 2006). Excessive kurtosis occurs when its values are outside the range of -7 to +7 (Curran, West, & Finch, 1996). Skewness, on the other hand, refers to the symmetry of the distribution compared to the normal distribution. If the distribution is asymmetrical, it is unbalanced and therefore skewed. If the distribution is shifted to the left, it is positively skewed and if it is shifted to the right, it is negatively skewed. Excessive skewness occurs when its values are outside the range of -2 to +2 (Curran et al., 1996).

I examine the univariate kurtosis and skewness of all the items in the survey. As seen in Table 6-6, all univariate kurtosis and skewness values are well within their respective rule-of-thumb ranges, which provide support for univariate normality.

Table 6-6: Assessment of the normality of the items

Scale	Item	Mean	Standard Deviation	Skewness	Kurtosis
Age	Age	58.14	52.031	1.645	2.851
Diagnostic	D1	4.87	.952	-1.252	2.887
	D2	4.84	.918	-1.208	2.550
	D3	5.06	.793	-1.045	2.209
	D4	4.66	1.030	-1.083	1.306
Diversity	Div1	5.39	.689	-1.325	3.510
	Div2	4.44	1.020	-.718	.621
	Div3	3.90	1.153	-.457	-.133
	Div4	4.88	.942	-.709	.398
	Div5	4.84	1.012	-1.267	2.402
	Div6	3.58	1.147	-.379	-.230
	Div7	4.28	1.177	-.646	.045
	Div8	2.71	1.280	.256	-.697
	Div9	4.08	1.275	-.490	-.253
	Div10	3.14	1.328	-.053	-.840
Interactive	I1	3.99	1.106	-.758	.486
	I2	4.35	.961	-1.050	1.999
	I3	4.57	.998	-.745	.560
	I4	4.15	1.147	-1.042	.974
	I5	4.20	1.079	-.761	.482
	I6	4.22	1.135	-.870	.694
	I7	4.29	1.178	-.996	.702
Risk	R1	4.33	1.162	-.584	.097
	R2	2.78	1.162	.077	-.755
	R3	2.67	1.372	.311	-.996
	R4	3.68	1.240	-.493	-.250
Search	S1	4.43	1.172	-.741	.634
	S2	2.94	1.295	-.051	-.937
	S3	4.21	1.066	-.360	.229
	S4	2.84	1.243	.131	-.740
	S5	4.02	1.072	-.363	-.124
	S6	3.94	1.286	-.334	-.477
	S7	3.85	1.239	-.421	-.033
Size	Size	2.61	2.018	1.050	-.327
Product	V1	4.39	.907	-.505	1.238

Scale	Item	Mean	Standard Deviation	Skewness	Kurtosis
Innovation	V2	4.09	1.125	-.666	.510
	V3	3.48	1.195	-.426	-.337
	V4	3.88	1.098	-.468	.058

Univariate normality is a necessary but not insufficient condition for multivariate normality (West, Finch, & Curran, 1995). Multivariate normality will be examined later in regression diagnostics. Therefore, the results of our univariate analysis strengthen our confidence, but it does not guarantee multivariate normality. Hair et al (1996) points out as the sample size grows, departure from normality may have negligible impact. This also strengthens our confidence in the results because the sample size is 145.

6.3.4 Testing for common method variance

Since this research uses the same respondent to provide information on all of the variables under study, it is susceptible to common method variance. Podsakoff, MacKenzie, & Lee (2003) define common method variance as the "variance that is attributable to the measurement method rather than to the constructs the measures represent". Common method variance is present when the correlations between measures are not due to substantive relationships between them but because they were provided by the same respondent (Podsakoff & Organ, 1986).

On approach for assessing the existence of common method variance is Harman's one factor test (or single-factor) test (Podsakoff & Organ, 1986; Podsakoff et al., 2003). Podsakoff & Organ (1986) explain the procedure as follows "all of the variables of interest are entered into factor analysis. Following this, the results of the unrotated factor solution are examined to determine the number of factors that are necessary to account for the variance in the variables. Basic assumption of this technique is that if a substantial amount of common method variance is present, either (a) a single factor will emerge from the factor analysis, or (b) one "general" factor will account for the majority of the covariance in the independent and criterion variables."

I conducted an exploratory factory analysis using principle component analysis[80]. The unrotated factor solution resulted in three factors with the first factor accounting for 31.377% of the total variance (See Table 6-7). While not conclusive proof that common method variance does not exist in the data, the results suggests that any common method variance that does exist is too small to be problematic. Podsakoff & Organ (1986) note that the Harman's one factor test is sensitive to the number of variables in the analysis. The probability of finding more than one factor increases as the number of variables increase. However, this is not an issue in this research given that only 8 variables entered in the analysis and the low variance accounted by the first factor.

Table 6-7: Results of the unrotated exploratory factor analysis (EFA) testing for common method variance

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.510	31.377	31.377	2.510	31.377	31.377
2	1.548	19.345	50.722	1.548	19.345	50.722
3	1.124	14.055	64.778	1.124	14.055	64.778
4	.803	10.035	74.812			
5	.746	9.331	84.143			
6	.539	6.736	90.879			
7	.470	5.879	96.758			
8	.259	3.242	100.000			

Extraction Method: Principal Component Analysis.

6.4 Validating Measures

Construct validity is the extent to which a scale or a set of measures accurately represent the construct of interest (Carmines & Zeller, 1979; Hair et al., 2006). Before testing the proposed hypotheses (structural relationships), it is important

[80] I also conducted exploratory factor analysis using two other extraction techniques: principal axis factoring and maximum likelihood. Since the results were similar to that of the principle component analysis, only the latter is reported here.

to validate the measures to ensure construct validity. The process used to assess measure validity is shown in Figure 6-3.

The process consisted of 5 steps performed on each scale. First, content validity was established. Once content validity was established, unidimensionality was assessed because it is a necessary condition for assessing the reliability and convergent and discriminant validities. Reliability is required before convergent validity could be assessed and convergent validity is required before discriminant validity could be assessed.

Figure 6-3: A framework for measure validation (Sources: Bagozzi, Yi, Philips, 1991; Hair et al., 2006; Kaynak & Hartley, 2005; Ping, 2004; O'Leary-Kelly & Vokurka, 1998)

Establish Content Validity

1- Use of existing and validated scales
2- Pre-test of the questionnaire with academics

↓

Establish Unidimensionality

1- Use Exploratory Factor Analysis (EFA) to determine the multiple dimensions (factors) of the innovation activities and diversity scales.
2- Use Confirmatory Factor Analysis (CFA) to evaluate the fit of all scales.

↓

Establish Reliability

1- Estimate reliability coefficient Cronbach's α
2- Estimate construct reliabilities

↓

Establish Convergent Validity

1- Evaluate statistical significance of factor loadings.
2- Evaluate the standardized factor loading.
3- Estimate Average Variance Extracted (AVE).
4- Evaluate construct reliabilities estimated earlier

↓

Establish Discriminant Validity

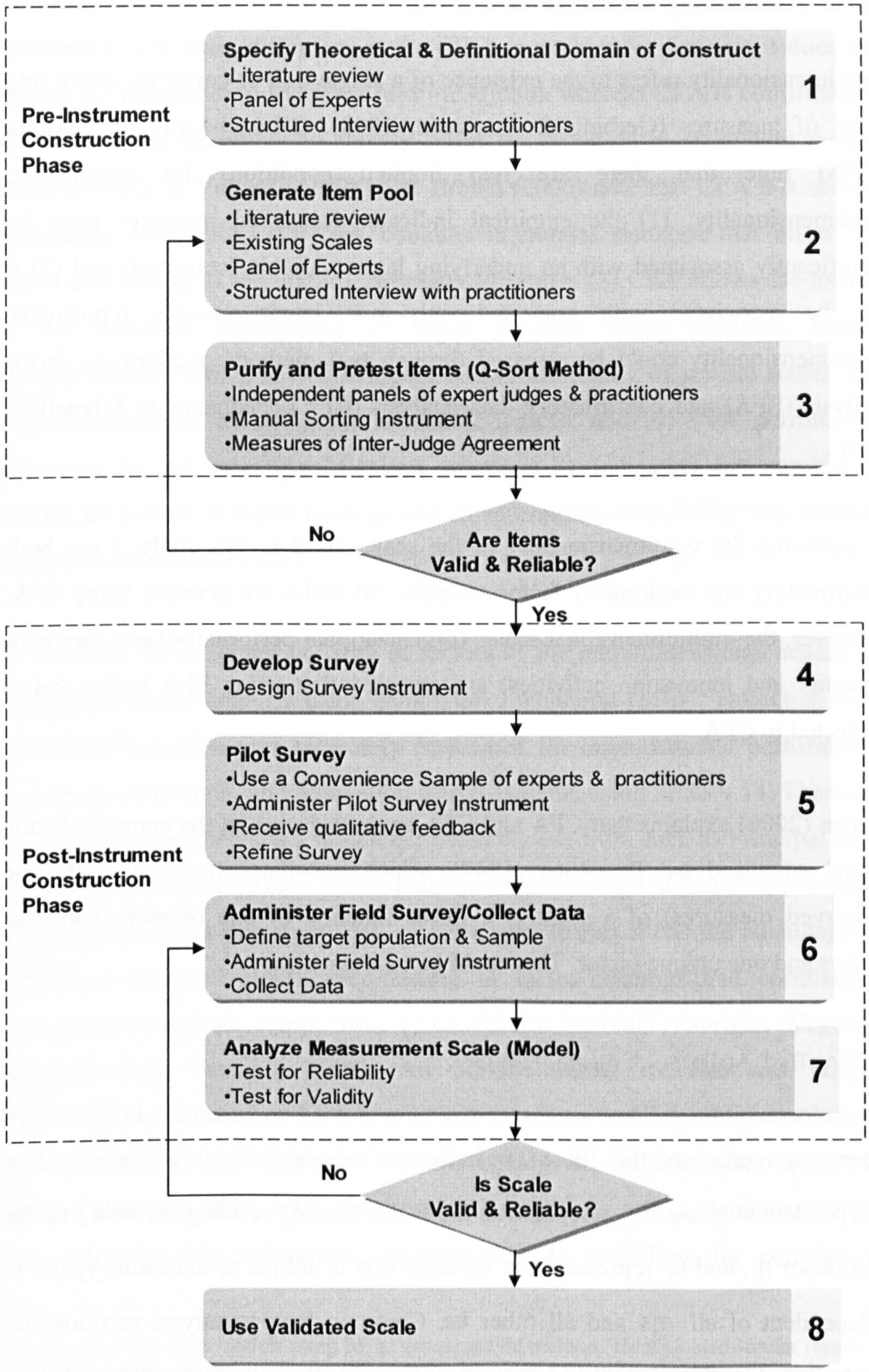
1- Compare the squared correlation of each pair of factors to AVE for each factor.
2- Compare the correlation of each pair of factors to reliability for each factor.
3- construct confidence intervals (±2 standard errors) around the correlation estimate between scales.
4- Compare the Chi Squared difference between the constrained and unconstrained CFAs of all pair-wise combinations of the scales

6.4.1 Assessing Content (Face) Validity

Content validity relates to the subjective judgment of content based on the procedures used to generate and select items to represent the domain of the construct (Viswanathan, 2005). It is assessed on theoretical and methodological grounds rather than statistical testing (Nunnally & Bernstein, 1994). Content validity involves assessing the correspondence between items selected to constitute a scale and its conceptual definition (Hair et al., 2006).

In this study, I used existing and validated scales. To assess the content validity of the scales, I used the scale development process illustrated in Figure 6-4 to assess the methodological rigour of the development process. I also piloted the questionnaire with academics to ensure content validity before I administered it to the target sample. Based on the pilot results, minor adjustments were made in terms of wording and presentation. Therefore, it is argued that the scales satisfy the conditions for content validity.

Figure 6-4: Scale development process (Source: Adapted from (Churhill, 1979; Froehle & Roth, 2004; Li, Rao, Ragu-Nathan, & Ragu-Nathan, 2005; Lori, Malhorta, Sharma, 2004; Menor & Roth, 2004; Moore & Benbasa, 1991; Stratman & Roth, 2002)


Specify Theoretical & Definitional Domain of Construct
•Literature review
•Panel of Experts
•Structured Interview with practitioners
1
Pre-Instrument
Construction
Phase
Generate Item Pool
•Literature review
•Existing Scales
•Panel of Experts
•Structured Interview with practitioners
2
Purify and Pretest Items (Q-Sort Method)
•Independent panels of expert judges & practitioners
•Manual Sorting Instrument
•Measures of Inter-Judge Agreement
3
No
Are Items
Valid & Reliable?
Yes
Develop Survey
•Design Survey Instrument
4
Pilot Survey
•Use a Convenience Sample of experts & practitioners
•Administer Pilot Survey Instrument
•Receive qualitative feedback
•Refine Survey
5
Post-Instrument
Construction
Phase
Administer Field Survey/Collect Data
•Define target population & Sample
•Administer Field Survey Instrument
•Collect Data
6
Analyze Measurement Scale (Model)
•Test for Reliability
•Test for Validity
7
No
Is Scale
Valid & Reliable?
Yes
Use Validated Scale
8

6.4.2 Assessing Unidimensionality

Unidimensionality refers to the existence of a single trait or construct underlying a set of measures (Gerbing & Anderson, 1988). O'Leary-Kelly & Vokurka (1998) note that there are two implicit conditions for establishing unidimensionality: (1) the empirical indicator (observed measure) must be significantly associated with an underlying latent variable (construct) and (2) it can be associated with one and only one latent variable (construct). Unidimensionality could be assessed through two methods: exploratory factor analysis (EFA) and confirmatory data analysis (CFA) (Pedhazur & Schmelkin, 1991).

In assessing the unidimensionality of the scales used in this study, I use both confirmatory and exploratory factor analysis. All scales are assessed using CFA. However, the multidimensional scales (organizational performance measurement diversity and innovation activities) are assessed first using EFA before being subjected to CFA.

Brown (2006) explains that CFA and EFA are both based on the common factor model advanced by Thurstone (1947), which postulates that each indicator (observed measures) of a scale is a linear function of one or more common factors and one unique factor. This could be expressed as,

$$y_j = \lambda_{j1}\eta_1 + \lambda_{j2}\eta_2 + \ldots + \lambda_{jm}\eta_m + \varepsilon_j \,,$$

where y_j represents the jth of p indicators obtained from a sample of n independent subjects, λ_{jm} represents the factor loading relating variable j to the mth factor η , and ε_j represents the variance that is unique to indicator y_j and is independent of all ηs and all other εs. Common factor analysis partition the variance of the indicator into common variance that is share with other indicators and unique variance that is not shared with other indicators. The unique variance is a combination of specific variance (systematic) and error variance (random).

Although CFA and EFA are based on the same concept, they have some differences. The main difference is that EFA is exploratory because it does not require the researcher to specify a priori restriction whereas CFA is confirmatory and require the researcher to specify priori the proposed measurement model. O'Leary-Kelly & Vokurka (1998) and Brown (2006) note that CFA has several advantages over EFA: (1) CFA contains inferential statistics that allow for hypothesis testing of the unidimentionality of scales; (2) CFA allows the testing of the statistical significance of factor loadings using t-staistics (factor loading/standard error); (3) CFA model allows the testing of several scales that may be correlated by using the model fit indices; and (4) CFA permits the researcher to test more parsimonious solutions by indicating the number of factors, the pattern of factor loadings and an appropriate error theory (e.g. random or correlated indicator error).

To establish the unidimensionality of factors of the multidimensional scales, an exploratory factor analysis using Maximum Likelihood (ML)[81] extraction with PROMAX[82] rotation was separately performed for organizational performance measurement diversity and innovation activities items using SPSS v.14. There are several extraction methods a researcher could choose from such as Principal Axis Factoring (PF), Generalized Least Square (GLS), Unweighted Least Squares, and Alpha Factoring. In this study, ML was chosen because it has the advantage of permitting statistical significance testing of factor loadings and correlations among factors and the computation of confidence intervals (Fabrigar, Wegener, MacCallum, & Strahan, 1999). An oblique factor rotation was chosen (PROMAX) because when EFA is used as a precursor to CFA, oblique solutions are more likely to generalize to CFA than orthogonal solutions (Brown, 2006). Oblique solutions allow for correlation between factors making it more realistic and preferable than orthogonal solutions that do not allow for correlations

[81] I also ran all the EFA models using PF as the extraction method. They yielded similar results to ML. So ML results are only reported in this study. Note that PF has the advantage of being free of distributional assumptions whereas ML requires the assumption of multivariate normal distribution.

[82] I also ran all the EFA models using the Direct Oblimin Method as the rotation method. They yielded similar results to PROMAX. Therefore, PROMAX results are only reported in this study.

between factors. PROMAX was chosen over other methods such as Direct Oblimin Method because it has the advantage of being fast and conceptually simple.

Maximum Likelihood estimated confirmatory factor analysis was performed on all the scales using AMOS software version 6.0 (Arbuckle, 2005) to evaluate the fit of the measurement model of the scales. Jöreskog (1993) suggests that the measurement model for each scale should be estimated separately, then, combining the scales into pairs, estimate each pair separately, and finally estimate the measurement model for all scales. Because items are omitted as required at each step to obtain adequate measurement model fit, Ping (2004) questions the efficiency of such approach. As per Jöreskog (1993) recommendation, I assess each scale individually and I assess only the related scales in the combined measurement models because of the small sample size. CFA requires at least 5 cases per estimated parameter (Kline, 2004). The related scales are that of diagnostic and interactive performance measurement use and organizational performance measurement diversity's subscales of financial, operational, and stakeholder.

Schumacker & Lomax (2004) remind us of the difficulty of choosing one fit index to evaluate SEM models because a good fit index must: (1) be independent of sample size, (2) accurately reflects difference in fit, (3) imposes penalty for inclusion of additional parameters, and (4) supports the choice of the true model when it is known. They conclude, "*No model fit criterion can actually meet all of these criteria*". So as recommended by the structural equation methodologists (such as, Hoyle & Panter, 1995; Hu & Bentler, 1999; Kline, 2003; Hair et al., 2006), the fit of the CFA measurement models was assessed using a number of fit indices that belonged to two different types: absolute fit measures and incremental fit measures. Absolute fit measures are direct measures of how well the model specified by the researcher reproduces the observed data whereas incremental fit measures assess how well a specified model fits relative to some alternative base line, which is usually the null model that assumes there are no correlations between the observed measures (Hair et al., 2006). In this study, the absolute fit indices used are ratio Chi-square χ^2 to degree of freedom, Goodness-of-Fit Index (GFI), and Root Mean Square Error of Approximation (RMSEA) and the incremental fit indices used are the Comparative Fit Index (CFI), Tucker Lewis Index (TLI) which is also known as Bentler-Bonett non-normed fit index (NNFI), and Incremental fit Index (IFI Delta2). These indices were chosen because they reflect the current state of practice and recommendations about what should be reported (Kline, 2003). Table 6-8 provides a brief description of the fit along with the recommended critical value.

Table 6-8: Description of fit indices and their recommended critical values (Compiled from different sources: (Byrne, 2001; Hoyle & Panter, 1995; Hu & Bentler, 1999; Kline, 2003)

Index	Reference	Description	Critical Value
Ratio χ^2 to degree of freedom	Bollen, (1989b)	Statistical test of the lack of fit resulting from over identifying restrictions placed on a model	<3
Goodness-of-Fit Index (GFI)	Joreskog & Sorbom (1981)	Indexes the relative amount of the observed variances and covariances accounted for by a model Analogous to R2	>.90
Root Mean Square Error of Approximation (RMSEA)	Browne & Cudeck (1993)	It is a parsimony-adjusted index in that its formula includes a built-in correction for model complexity.	<.10
Comparative Fit Index (CFI)	Bentler (1989, 1990)	Indexes the relative reduction in lack of fit as estimated by the noncentral chi square of a target model versus a baseline model. Varies between 0 and 1.	>.90
Tucker-Lewis Index (TLI)/Nonnormed Fit Index (NNFI)	Bentler & Bonett (1980) Tucker & Lewis (1973)	Compares the lack of fit of a target model to the lack of fit of a baseline model, usually the independence model. Value estimates the relative improvement per df of the target model over a baseline model. Not recommended for very small samples (<150) or with GLS estimation.	>.90
Incremental fit Index (IFI Delta2)	Bollen (1989a)	Same interpretation as TLI/NNFI. Less variable than TLI/NNFI in small samples and more consistent across estimators than TLI/NNFI.	>.90

6.4.2.1 Diagnostic Use Scale

Table 6-9 provides information on the items composing the diagnostic scale and their descriptive statistics. The kurtosis statistics are within the acceptable range of normality (-7 to +7) and they ranged from 1.306 to 2.887. The skewness statistics are also within the acceptable range of normality (-2 to +2) and they ranged from -1.045 to -1.252.

Table 6-9: Diagnostic scale: Descriptive statistics

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
D1	Track progress towards goals	4.87	.952	-1.252	2.887
D2	Review key measures	4.84	.918	-1.208	2.550
D3	Monitor results	5.06	.793	-1.045	2.209
D4	Compare outcomes to expectations	4.66	1.030	-1.083	1.306

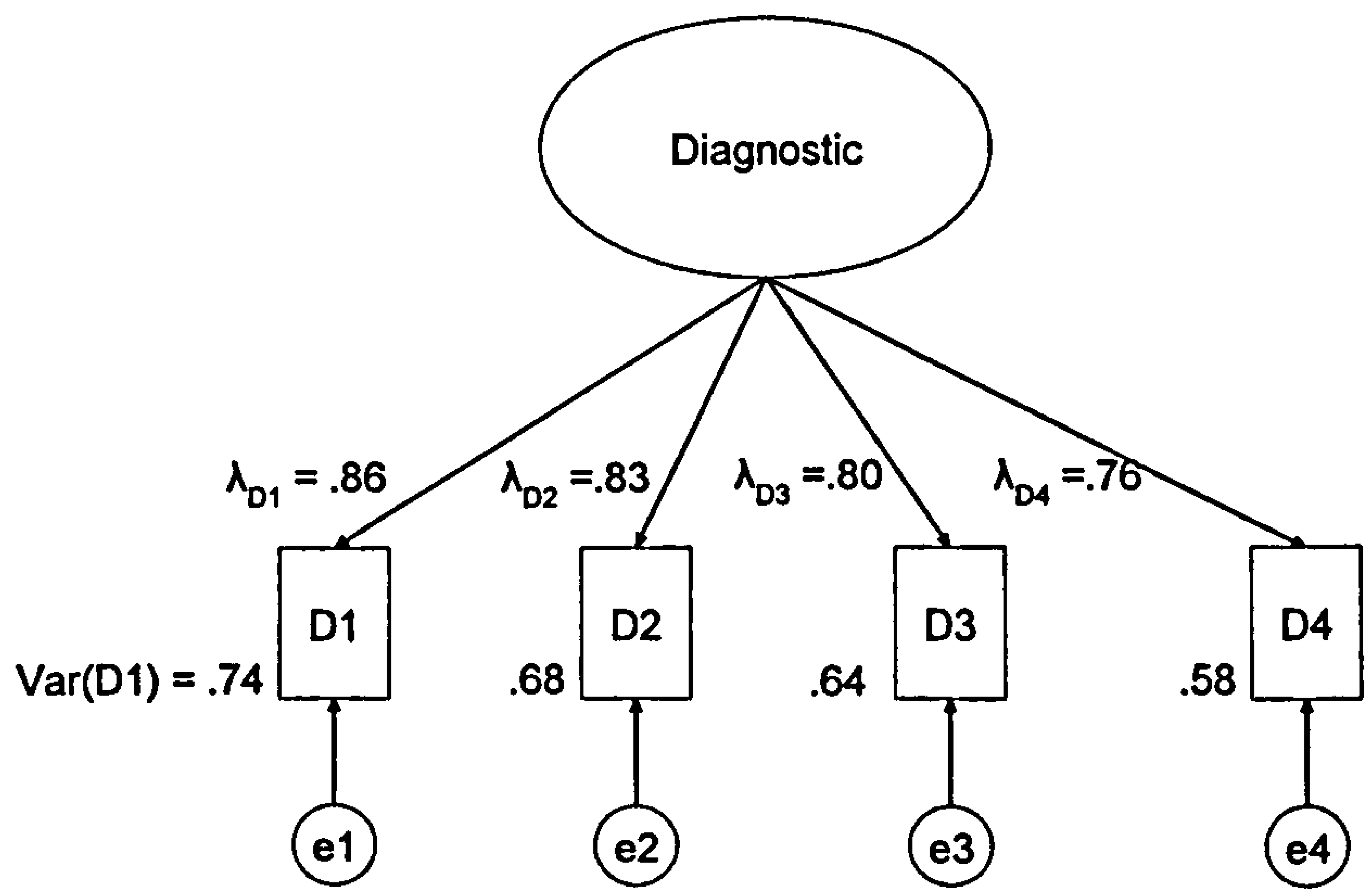
As discussed, unidimensionality of the diagnostic scale is assessed with ML estimated CFA using AMOS 6.0 software. The scale items for the diagnostic use of performance measurement exhibits unidimensionality, as shown by the results from the confirmatory factor analysis (see Table 6-10 and Figure 6-5). The factor loading of the four items composing the scale were statistically significant at $p<.001$ and their standardized values ranged from 0.76 to 0.86 exceeding the recommended lower value of 0.5 (Hair et al., 2006). The model fit indices suggest good fit although RMSEA is higher than the critical value of 0.1. However, this is not a concern because RMSEA tends to improve when the model complexity increase.

Table 6-10: Confirmatory factor analysis of diagnostic use scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
D1	1.00	.86	--	
D2	.91 (.08)	.83	11.68	***
D3	.79 (.07)	.80	11.19	***

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
D4	.96 (.09)	.76	10.49	***
Convergent Validity				
Average Variance Extracted (AVE)	.66			
Reliability				
Composite Reliability	.89	Cronbach's α	.89	
Model Fit				
χ^2 (df) → P; χ^2/df	7.80 (2) → .02; 3.9	CFI	.98	
GFI	.97	TLI (NNFI)	.95	
RMSEA	.14	IFI Delta2	.98	

Figure 6-5: Diagnostic use measurement model



Standardized factor loading are reported

6.4.2.2 Interactive Use Scale

Table 6-11 provides information on the items composing the interactive scale and their descriptive statistics. The kurtosis statistics are within the acceptable range of normality (-7 to +7) and they ranged from 0.482 to 1.999. The skewness statistics are also within the acceptable range of normality (-2 to +2) and they ranged from -1.050 to -0.745.

Table 6-11 : Interactive scale: Descriptive statistics

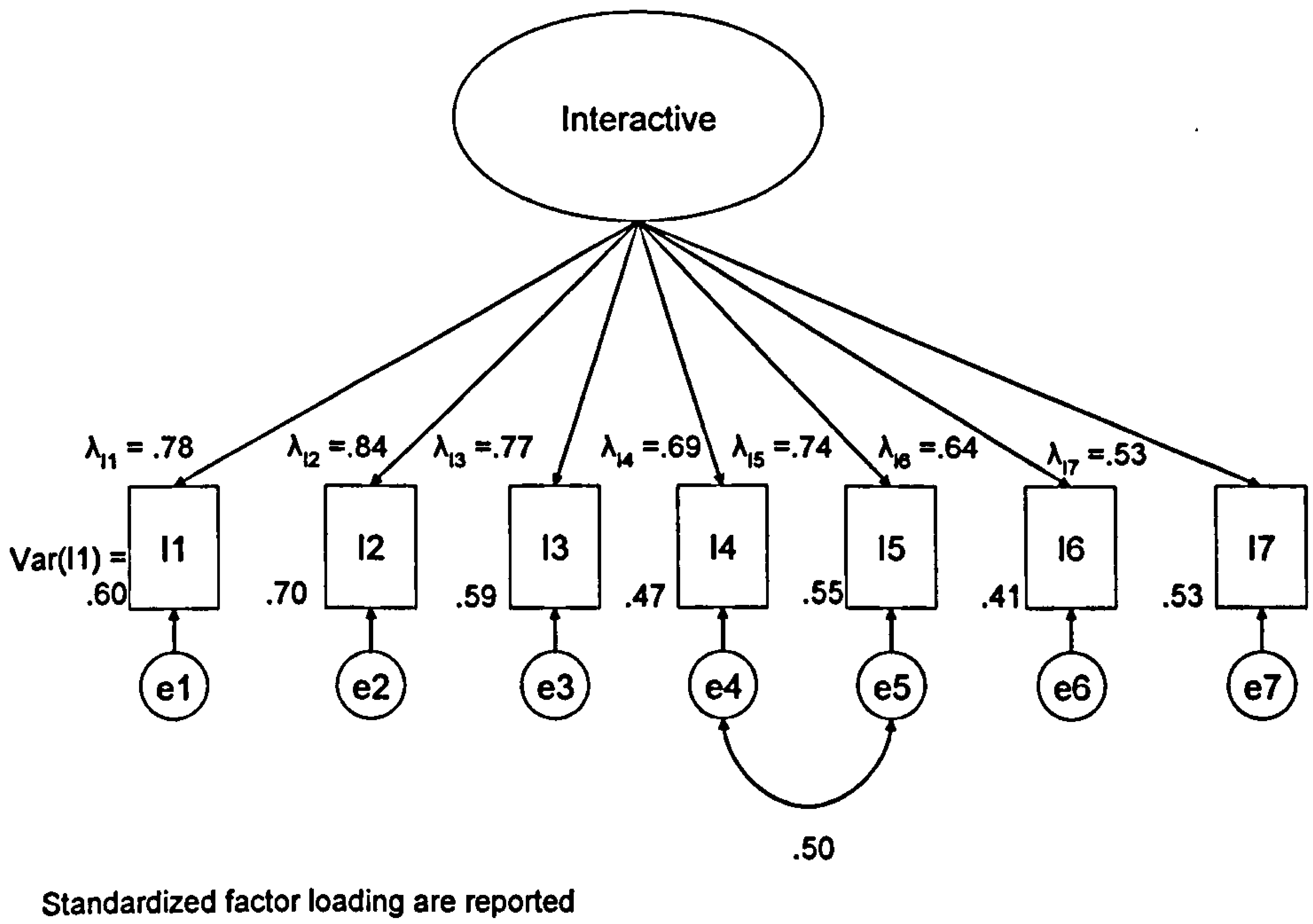
Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
I1	Integrate the organization- i.e. tie the organization together	3.99	1.106	-.758	.486
I2	Enable the organization to focus on common issues	4.35	.961	-1.050	1.999
I3	Enable the organization to focus on your critical success factors	4.57	.998	-.745	.560
I4	Develop a common vocabulary in the organisation	4.15	1.147	-1.042	.974
I5	Provide a common view of the organisation	4.20	1.079	-.761	.482
I6	Enable discussion in meetings of superiors, subordinates and peers	4.22	1.135	-.870	.694
I7	Enable continual challenge and debate underlying results, assumptions and action plans	4.29	1.178	-.996	.702

Unidimensionality of the interactive scale is assessed with ML estimated CFA. The scale items for the interactive use of performance measurement exhibit unidimensionality, as shown by the results from the confirmatory factor analysis (see Table 6-12 and Figure 6-6). The factor loading of the seven items composing the scale were statistically significant at $p<.001$ and their standardized values ranged from 0.64 to 0.84 exceeding the recommended lower value of 0.5. The model fit indices suggest good fit. Two of the measurement errors (disturbances), e4 and e5, were correlated. There are two reasons why that may happen: (1) the two indicators may share a substantive latent (construct) other than interactive use or (2) the measurement may be due to method effect due to the measurement approach, which may result because of similar wording of the items (Brown, 2006). A close examination of the two indicators (I4, I5) associated with the measurement errors suggest that the correlated measurement errors are not due to substantive reasons but due to measurement method. Both items are closely worded as seen in Table 6-11. Hence, unidimensionality is not affected.

Table 6-12: Confirmatory factor analysis of interactive use scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
I1	1.00	.78	–	
I2	.93 (.09)	.84	10.51	***
I3	.89 (.09)	.77	9.51	***
I4	.92 (.11)	.69	8.37	***
I5	.93 (.10)	.74	9.16	***
I6	.84 (.11)	.64	7.73	***
I7	.99 (.11)	.72	8.92	***
Convergent Validity				
Average Variance Extracted (AVE)	.54			
Reliability				
Composite Reliability	.89	**Cronbach's α**	.93	
Model Fit				
χ^2 (df) → P; χ^2/df	38.16 (13) → .00; 3.9	**CFI**	.96	
GFI	.93	**TLI (NNFI)**	.93	
RMSEA	.12	**IFI Delta2**	.96	

Figure 6-6: Interactive use scale measurement model

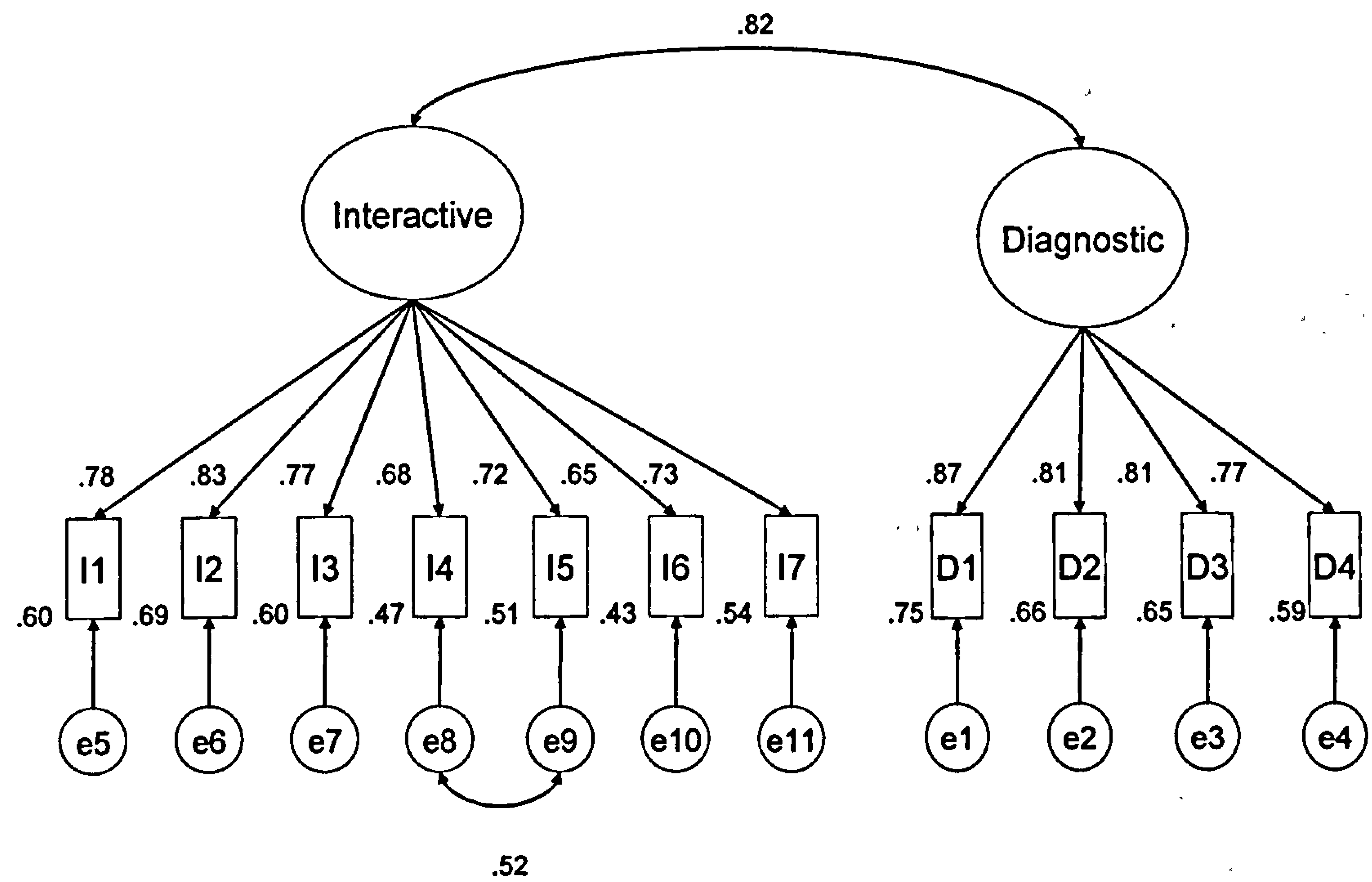


As discussed earlier, the diagnostic and interactive scales were combined into a single CFA measurement model to determine their unidimensionality. The results were consistent with earlier results for the individually assessed CFA models. These results further strengthen our confidence in the earlier findings. The scale items for both scales exhibit unidimensionality, as shown by the results from the confirmatory factor analysis (see Table 6-13 and Figure 6-7). The factor loading of the eleven items composing the scale were statistically significant at $p<.001$ and their standardized values ranged from 0.66 to 0.87 exceeding the recommended lower value of 0.5. All model fit indices respected their critical values suggesting a good fit.

Table 6-13: Confirmatory factor analysis of performance measurement use (diagnostic + interactive)

Item	ML Unstd. Λ(Std. Err)	ML Std. Λ	Critical Ratio	P
Diagnostic use scale				
D1	1.00	.87	--	
D2	.91 (.08)	.81	11.94	***
D3	.79 (.07)	.81	11.80	***
D4	.96 (.09)	.77	10.87	***
Interactive use scale				
I1	1.00	.78	--	
I2	.93 (,09)	.83	10.65	***
I3	.90 (.09)	.77	9.75	***
I4	.91 (.11)	.68	8.41	***
I5	.90 (.10)	.72	8.91	***
I6	.86 (.11)	.65	8.01	***
I7	1.00 (.11)	.73	9.17	***
Model Fit				
χ^2 (df) → P; χ^2/df	95.44 (42) → .00; 2.3	CFI	.95	
GFI	.90	TLI (NNFI)	.93	
RMSEA	.09	IFI Delta2	.95	

Figure 6-7: Performance measurement use (diagnostic + interactive) measurement model



Standardized factor loading are reported

6.4.2.3 Organizational Performance Measurement Diversity Scale

Table 6-14 provides information on the items composing the performance measurement scale and their descriptive statistics. The kurtosis statistics are within the acceptable range of normality (-7 to +7) and they ranged from -0.625 to 3.510. The skewness statistics are also within the acceptable range of normality (-2 to +2) and they ranged from -1.325 to 0.256.

Table 6-14: Organizational Performance measurement diversity scale : Descriptive statistics

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
Div1	Short term financial results- e.g. operating income, sales growth,etc	5.39	.689	-1.325	3.510
Div2	Customer relations- e.g. market share, customer satisfaction, etc.	4.44	1.020	-.718	.621
Div3	Employee relations- e.g. employee satisfaction, safety, etc	3.90	1.153	-.457	-.133
Div4	Operational performance- e.g. productivity, lead times, etc.	4.88	.942	-.709	.398

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
Div5	Quality – e.g. quality performance, defect rates, etc.	4.84	1.012	-1.267	2.402
Div6	Innovation and learning- e.g. number of new products and /or services launched, training, etc.	3.58	1.147	-.379	-.230
Div7	Supplier relations- e.g. on-time delivery, suppliers' integration etc	4.28	1.177	-.646	.045
Div8	Alliances- e.g. joint marketing, joint product designs, etc.	2.71	1.280	.256	-.697
Div9	Environmental performance- environmental compliances, etc.	4.08	1.275	-.490	-.253
Div10	Community- public image, community involvement, etc.	3.14	1.328	-.053	-.840

After excluding Div6 and Div8 from analysis in this study because they may introduce circularity in the findings since they measure innovation, unidimensionality of the performance measurement subscales are first assessed using EFA with ML extraction and PROMAX oblique rotation in SPSS Version 14.0. First, two tests were conducted to ensure EFA was appropriate: Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's Test of Sphericity. The results of the tests as shown in Table 6-15 indicate that EFA was appropriate because KMO was above 0.6 and the Bartlett's test was statistically significant (Tabachnick & Fidell, 1996).

Table 6-15: KMO and Bartlett's test for organizational performance measurement diversity items

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.778
Bartlett's Test of Sphericity	Approx. Chi-Square	350.807
	df	21
	Sig.	.000

In EFA solutions, the pattern matrix is interpreted because it contains the unique contribution of each item to the extracted factors (Hair et al., 2006). The pattern matrix of the rotated solution shown in Table 6-16 confirmed the existence of two dimensions (subscales) in addition to the financial performance subscale. They were labelled stakeholder and operations. None of the items had a significant loading on more than one factor (cross loading). These findings suggest we have unidimensional subscales. The next step is to use CFA to assess the unidimensionality of the subscales.

Table 6-16: Organizational performance measurement diversity pattern matrix

	Factor	
Item	**Diversity - Operations**	**Diversity - Stakeholder**
Div2		.425
Div3		.769
Div4	.829	
Div5	.863	
Div7	.614	
Div9		.568
Div10		.769

Extraction Method: Maximum Likelihood.
Rotation Method: Promax with Kaiser Normalization.
Factors loading less than 0.3 have not been printed

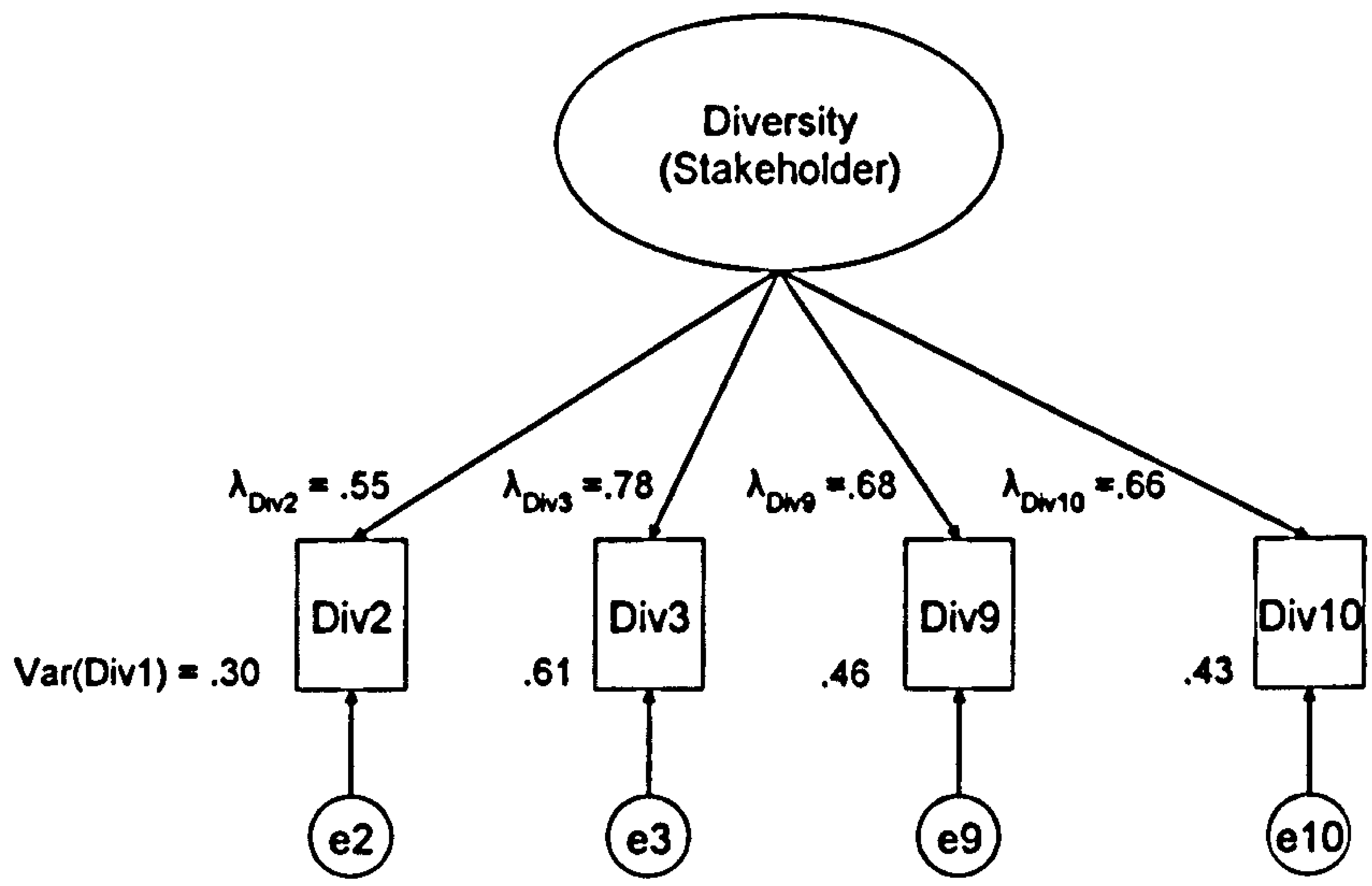
Unidimensionality of the stakeholder subscale is assessed with ML estimated CFA. The scale items exhibit unidimensionality, as shown by the results in Table 6-17 and Figure 6-8. The factor loading of the four items composing the scale were statistically significant at $p<.001$ and their standardized values ranged from 0.55 to 0.78 exceeding the recommended lower value of 0.5. All model fit indices respected their critical values except for RMSEA, which was slightly high at .11 suggesting a good fit.

Table 6-17: Confirmatory factor analysis of organizational performance measurement diversity (stakeholder) scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
Div2	1.00	.55	--	
Div3	1.60 (.28)	.78	5.67	***
Div9	1.53 (.28)	.68	5.44	***
Div10	1.55 (.29)	.66	5.35	***
Convergent Validity				
Average Variance Extracted (AVE)	.46			
Reliability				

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
Composite Reliability	.76		Cronbach's α	.76
Model Fit				
χ^2 (df) → P; χ^2/df	5.53 (2) → .06; 2.77		CFI	.97
GFI	.98		TLI (NNFI)	.92
RMSEA	.11		IFI Delta2	.98

Figure 6-8: Organizational performance measurement diversity (Stakeholder) scale measurement model



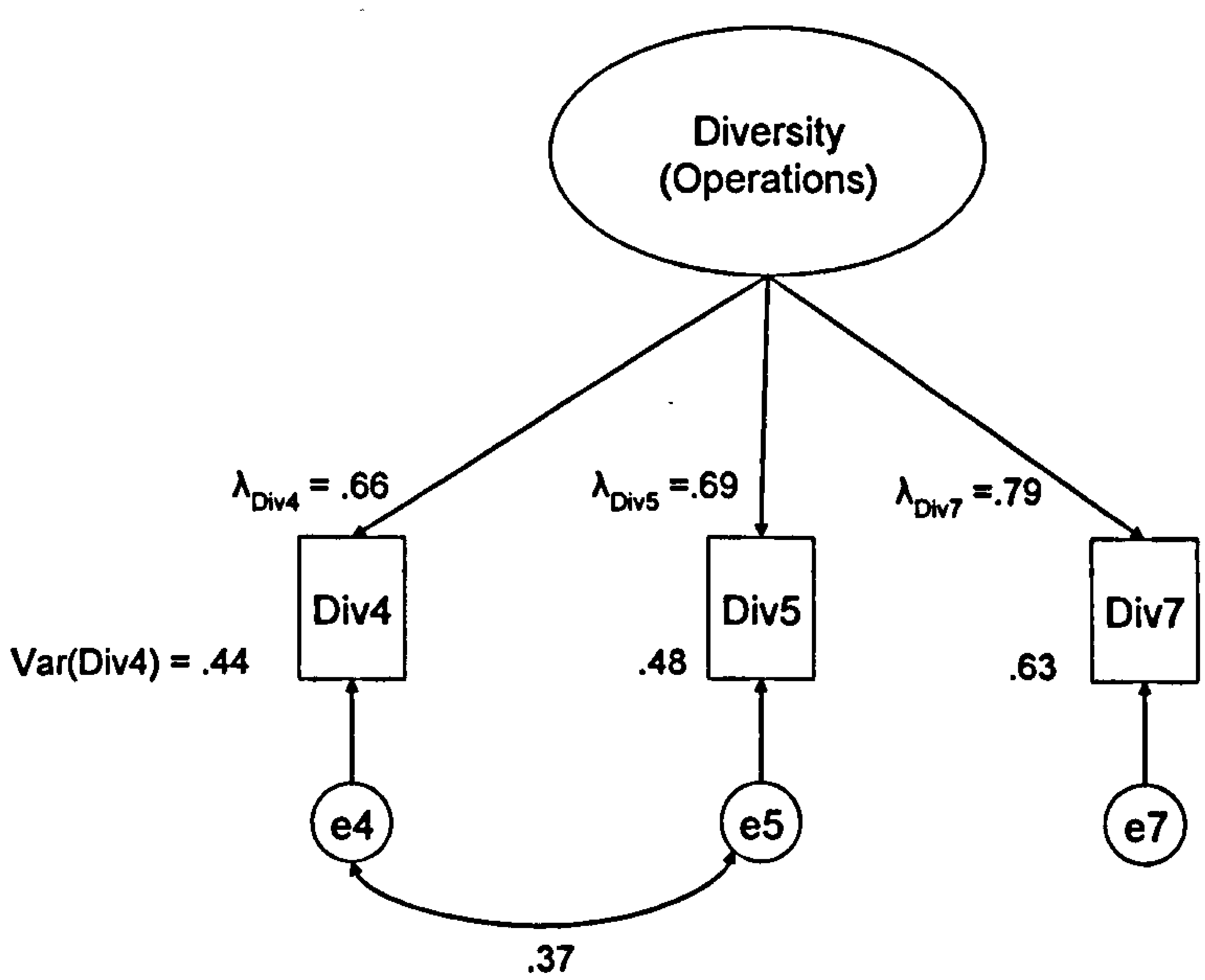
Standardized factor loading are reported

Unidimensionality of the operations subscale is assessed with ML estimated CFA. The scale items exhibit unidimensionality, as shown by the results in Table 6-18 and Figure 6-9. The factor loading of the three items composing the scale were statistically significant at p<.001 and their standardized values ranged from 0.66 to 0.79 exceeding the recommended lower value of 0.5. All model fit indices respected their critical values suggesting a good fit. Two of the measurement errors (disturbances), e4 and e5, were correlated. A close examination of the two indicators (Div4, Div5) associated with the measurement errors suggest that the correlated measurement errors are not due to substantive reasons but due to measurement method. Both items are closely worded as seen in Table 6-14. Hence, unidimensionality is not affected.

Table 6-18: Confirmatory factor analysis of organizational performance measurement diversity (operations) scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
Div4	1.00	.66	–	
Div5	1.08 (.13)	.69	8.50	***
Div7	1.45 (.19)	.79	7.50	***
Convergent Validity				
Average Variance Extracted (AVE)	.53			
Reliability				
Composite Reliability	.77	Cronbach's α	.80	
Model Fit				
χ^2 (df) → P; χ^2/df	1.33 (1) → .25; 1.33	CFI	1.00	
GFI	.99	TLI (NNFI)	.99	
RMSEA	.05	IFI Delta2	1.00	

Figure 6-9: Organizational performance measurement diversity (Operations) scale measurement model



Standardized factor loading are reported

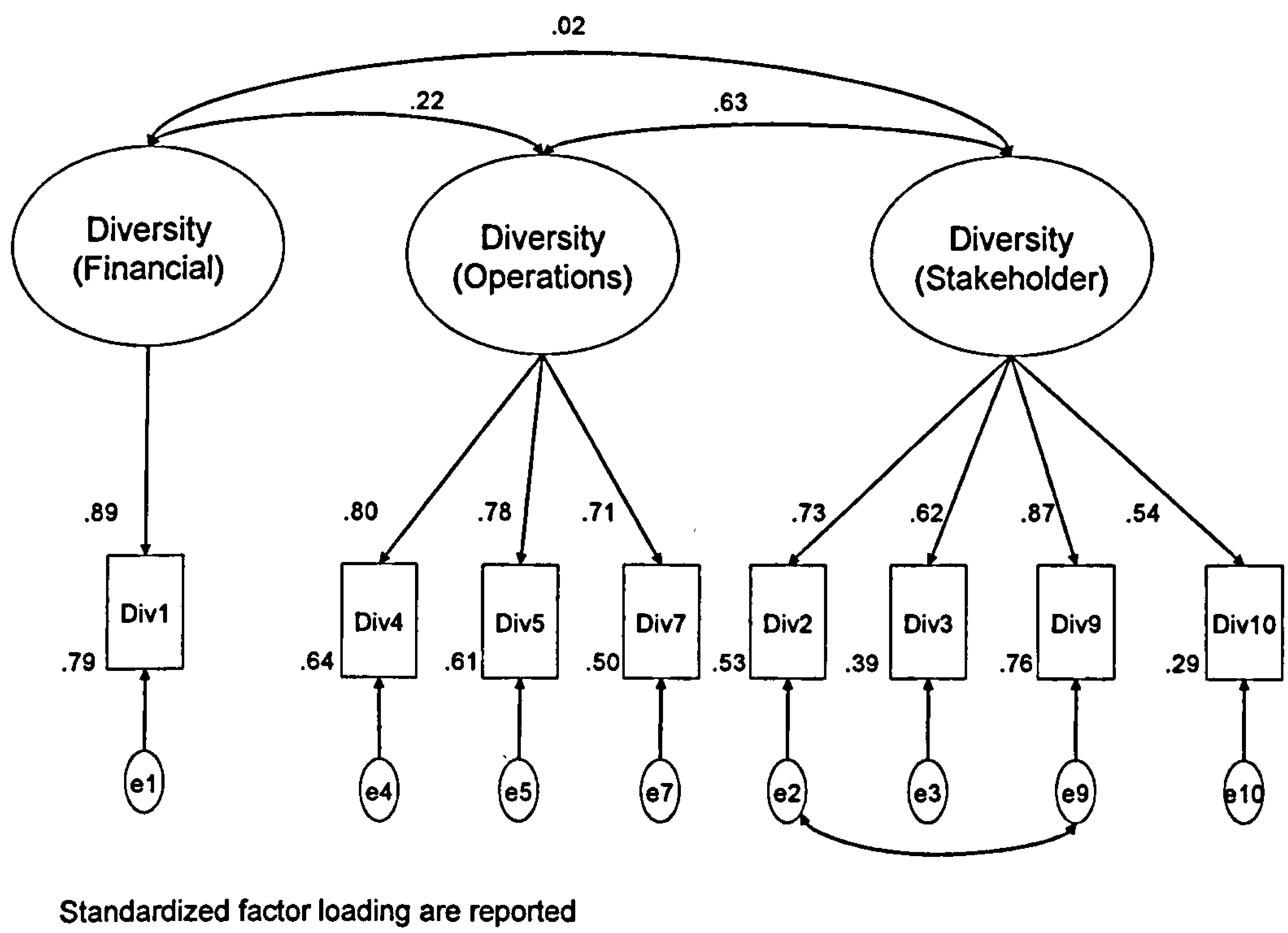
As discussed earlier, the financial, stakeholder, and operations subscales of the organizational performance measurement diversity construct were combined into a single CFA measurement model to determine their unidimensionality. The results were consistent with earlier results for the individually assessed CFA models. These results further strengthen our confidence in the earlier findings. The scale items for the three subscales exhibit unidimensionality, as shown in Table 6-19 and Figure 6-10. The factor loading of the eight items composing the scale were statistically significant at p<.001 and their standardized values ranged from 0.54 to 0.89 exceeding the recommended lower value of 0.5. All model fit indices respected their critical values suggesting a good fit.

Table 6-19: Confirmatory factor analysis of organizational performance measurement diversity (financial + stakeholder + operations) scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
Organizational Performance Measurement Diversity (Financial)				
Div1	1.00	.89	--	
Organizational Performance Measurement Diversity (Stakeholder)				
Div2	1.00	.73	--	
Div3	.97 (.17)	.62	5.80	***
Div9	1.50 (.24)	.87	6.29	***
Div10	.96 (.18)	.54	5.27	***
Organizational Performance Measurement Diversity (Operations)				
Div4	1.00	.80	--	
Div5	1.05 (.12)	.78	8.81	***
Div7	1.10 (.14)	.71	8.14	***
Model Fit				

χ^2 (df) → P; χ^2/df	1.33 (1) → .25; 1.33	CFI	1.00
GFI	.99	TLI (NNFI)	.99
RMSEA	.05	IFI Delta2	1.00

Figure 6-10: Organizational performance measurement diversity (Financial + Stakeholder + Operations) scale measurement model



6.4.2.4 Search Scale

Table 6-20 provides information on the items composing the innovation activities scale and their descriptive statistics. The kurtosis statistics are within the acceptable range of normality (-7 to +7) and they ranged from -0.937 to 0.634. The skewness statistics are also within the acceptable range of normality (-2 to +2) and they ranged from -0.741 to 0.131.

Table 6-20: Search scale: Descriptive statistics

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
S1	Intramural (in-house) R&D - Creative work undertaken within your enterprise on an occasional or regular basis to increase the stock of knowledge and its use to devise new and improved goods, services and processes	4.43	1.172	-.741	.634
S2	Acquisition of R&D (extramural R&D) - Same activities as above, but purchased by your enterprise and performed by others.	2.94	1.295	-.051	-.937
S3	Acquisition of machinery, equipment and software -Acquisition of advanced machinery,	4.21	1.066	-.360	.229

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
	equipment and computer hardware or software to produce new or significantly improved goods and services.				
S4	Acquisition of external knowledge - Purchase or licensing of patents and non-patented inventions, know-how, and other types of knowledge from other enterprises or organisations	2.84	1.243	.131	-.740
S5	Training - Internal or external training for your personnel specifically for the development and/or introduction of innovations	4.02	1.072	-.363	-.124
S6	All forms of Design - Expenditure on design functions for the development or implementation of new or improved goods, services and processes. Expenditure on design in the R&D phase of product development should be excluded.	3.94	1.286	-.334	-.477
S7	Market introduction of innovations - Activities for the market preparation and introduction of new or significantly improved goods and services, including market research and launch advertising	3.85	1.239	-.421	-.033

Unidimensionality of the innovation search subscale is first assessed using EFA with ML extraction and PROMAX oblique rotation. First, two tests were conducted to ensure EFA was appropriate: Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's Test of Sphericity. The results of the tests as shown in Table 6-21 indicate that EFA was appropriate because KMO was above 0.6 and the Bartlett's test was statistically significant (Tabachnick & Fidell, 1996).

Table 6-21: KMO and Bartlett's test for innovation activities performance items

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.783
Bartlett's Test of Sphericity	Approx. Chi-Square	241.852
	df	21
	Sig.	.000

The pattern matrix of the rotated solution shown in Table 6-22 confirmed the existence of three dimensions (subscales) for the items comprising innovation activities questionnaire. They were labelled innovation search, technology acquisition, and outsourcing. None of the items comprising the innovation search subscale had a significant loading on the other two factors (cross loading). These findings suggest we have unidimensional innovation search scale subscale. The next step is to use CFA to assess the unidimensionality of the innovation search subscale.

Table 6-22: Innovation activities pattern matrix

	Factor		
Item	Innovation Search	Technology Acquisition	Outsourcing
S1	.520		
S2			1.027
S3		.756	
S4		.360	.326
S5		.640	
S6	.994		
S7	.421		

Extraction Method: Maximum Likelihood.
Rotation Method: Promax with Kaiser Normalization.
Factors loading less than 0.3 have not been printed

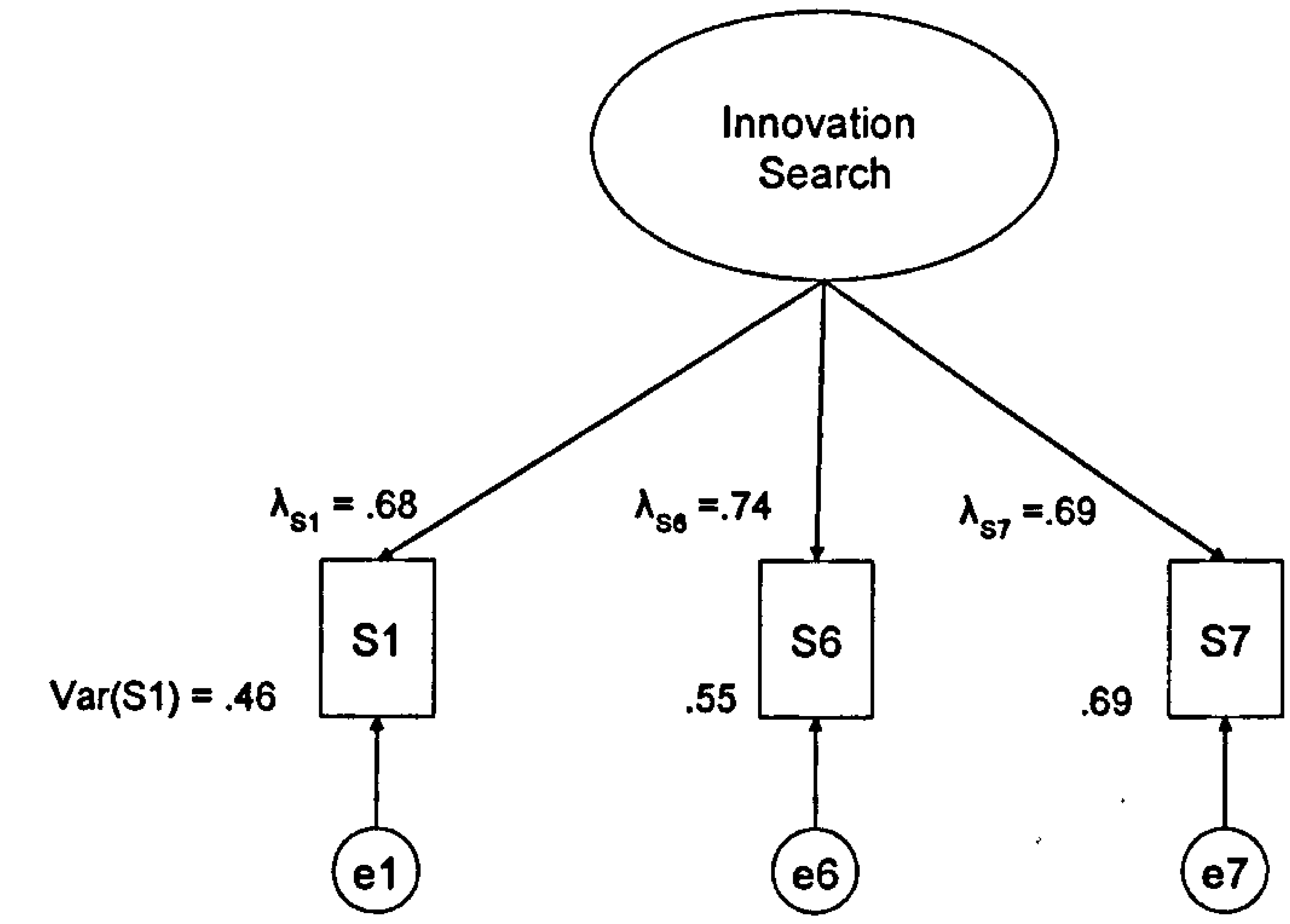
Unidimensionality of the innovation search scale is assessed with ML estimated CFA. The scale items exhibit unidimensionality, as shown by the results in Table 6-23 and Figure 6-11. The factor loading of the three items composing the scale were statistically significant at p<.001 and their standardized values ranged from 0.69 to 0.74 exceeding the recommended lower value of 0.5. All model fit indices respected their critical values suggesting a good fit.

Table 6-23: Confirmatory factor analysis of innovation search scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
S1	1.00	.68	--	
S6	1.19(.16)	.74	7.44	***
S7	1.03(.15)	.69	6.96	***
Convergent Validity				
Average Variance Extracted (AVE)	.50			
Reliability				
Composite Reliability	.75	**Cronbach's α**	.75	
Model Fit				
χ^2 (df) → P; χ^2/df	3.23 (2) → .20; 1.62	CFI	.99	

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
GFI	.99		TLI (NNFI)	.98
RMSEA	.07		IFI Delta2	.99

Figure 6-11: Innovation search scale measurement model



Standardized factor loading are reported

6.4.2.5 Risk Taking Scale

Table 6-24 provides information on the items composing the risk scale and their descriptive statistics. The kurtosis statistics are within the acceptable range of normality (-7 to +7) and they ranged from -0.755 to 0.097. The skewness statistics are also within the acceptable range of normality (-2 to +2) and they ranged from -0.584 to 0.311.

Table 6-24: Risk scale : Descriptive statistics

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
R1	Strong emphasis on research and development, technological leadership, and innovations	4.33	4.33	-.584	.097
R2	Strong proclivity to high risk, high return investments	2.78	2.78	.077	-.755
R3	Growth strategy primarily through external financing (borrowings, capital issues, etc.)	2.67	2.67	.311	-.996
R4	Very competitive, "undo-the-competitors" philosophy	3.68	3.68	-.493	-.250

After excluding R1 from analysis in this study because it is measured as part of the innovation search scale and including it may introduce circularity in the findings, unidimensionality of the risk scale is assessed with ML estimated CFA. The scale items did not exhibit unidimensionality (see Table 6-25 and Figure 6-12). The factor loading of the three items composing the scale were 0.42, 0.57, and 0.46, which were less than the recommended lower value of 0.5. Also, both measures of reliability indicated that the scale is not reliable at 0.48. Furthermore, the average variance extracted from the scale was very low at 0.24. These results are disappointing given that the scale was validated in previous research. Moreover, no concerns were raised when it was piloted with academics. However, in hindsight, the results may be due to the fact that this scale was not validated in the UK. It was validated in an American context. The wordings of the scale may have worked for the American context and not for the British context. For example, statements such as "undo-the-competitors" may be part of everyday American business vocabulary but not the British. Academics who participated in the piloting of the questionnaire may have not picked on that because they are fluent with both contexts. Subsequently, I employed EFA and CFA with two item measures to maintain multi-item scale for risk but none of the EFA or CFA solutions were acceptable. Therefore, the risk scale is turned into a single item scale. After extensive review of the risk taking literature (e.g. Bromiley, Miller, & Rau, 2001; Miller & Bromiley, 1990) to ensure content validity and minimum measurement error, item R2 was chosen. Item R2 stated "Strong proclivity to high risk, high return investments."

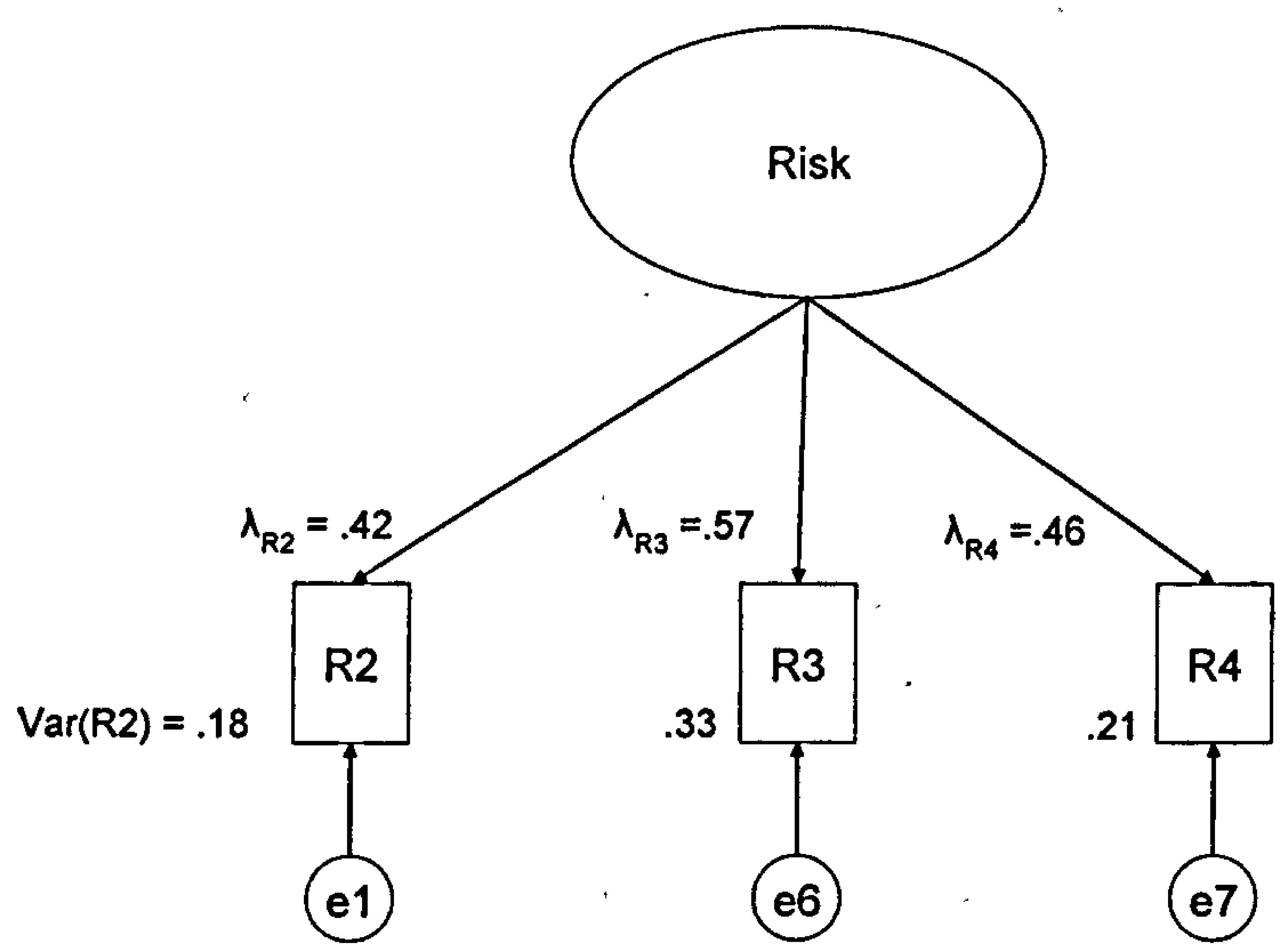
Table 6-25: Confirmatory factor analysis of risk scale

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
R2	1.00	.42	--	
R3	1.51 (.47)	.57	3.23	***
R4	1.12 (.39)	.46	2.9	**
Convergent Validity				
Average Variance Extracted (AVE)		.24		
Reliability				

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P

Composite Reliability	.48	Cronbach's α	.48
Model Fit			
χ^2 (df) → P; χ^2/df	4.284 (2) → .12; 2.14	CFI	.89
GFI	.98	TLI (NNFI)	.83
RMSEA	.09	IFI Delta2	.89

Figure 6-12: Risk scale measurement model



Standardized factor loading are reported

6.4.2.6 Product Innovation Scale

Table 6-26 provides information on the items composing the product innovation scale and their descriptive statistics. The kurtosis statistics are within the acceptable range of normality (-7 to +7) and they ranged from -0.337 to 1.238. The skewness statistics are also within the acceptable range of normality (-2 to +2) and they ranged from -0.666 to -0.426.

Table 6-26: Product innovation scale: Descriptive statistics

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
V1	Significantly improved goods or services	4.39	.907	-.505	1.238

Item	Survey Question	Mean	Std. Dev.	Skew	Kurt
V2	New goods or services onto your market before your competitors (New to your market)	4.09	1.125	-.666	.510
V3	New goods or services that was essentially the same as a product already available from your competitors in your market (Only new to your company)	3.48	1.195	-.426	-.337
V4	The percentage of new products or services significantly in its portfolio	3.88	1.098	-.468	.058

Unidimensionality of the product innovation scale is assessed with ML estimated CFA. As shown in Table 6-27, one of the items of the scales (V3) was statistically insignificant at $p<0.05$ and the model indices indicated a poor fit. Item V3 was dropped from the purified scale, which now included three items. The new purified is assessed again with ML estimated CFA. All the scale items exhibit unidimensionality (see Table 6-28 and Figure 6-13). The factor loading of the three items composing the scale were statistically significant at $p<.001$ and their standardized values ranged from 0.64 to 0.76 exceeding the recommended lower value of 0.5. All model fit indices improved drastically from the initial model and they respected their critical values suggesting a good fit. The construct reliability also improved at 0.77 from the unacceptable level of 0.67 for the initial model. A major concern in deleting an item from a scale is the danger of throwing content validity; this is not the case here since the other items contained the relevant information.

Table 6-27: Confirmatory factor analysis of product innovation scale

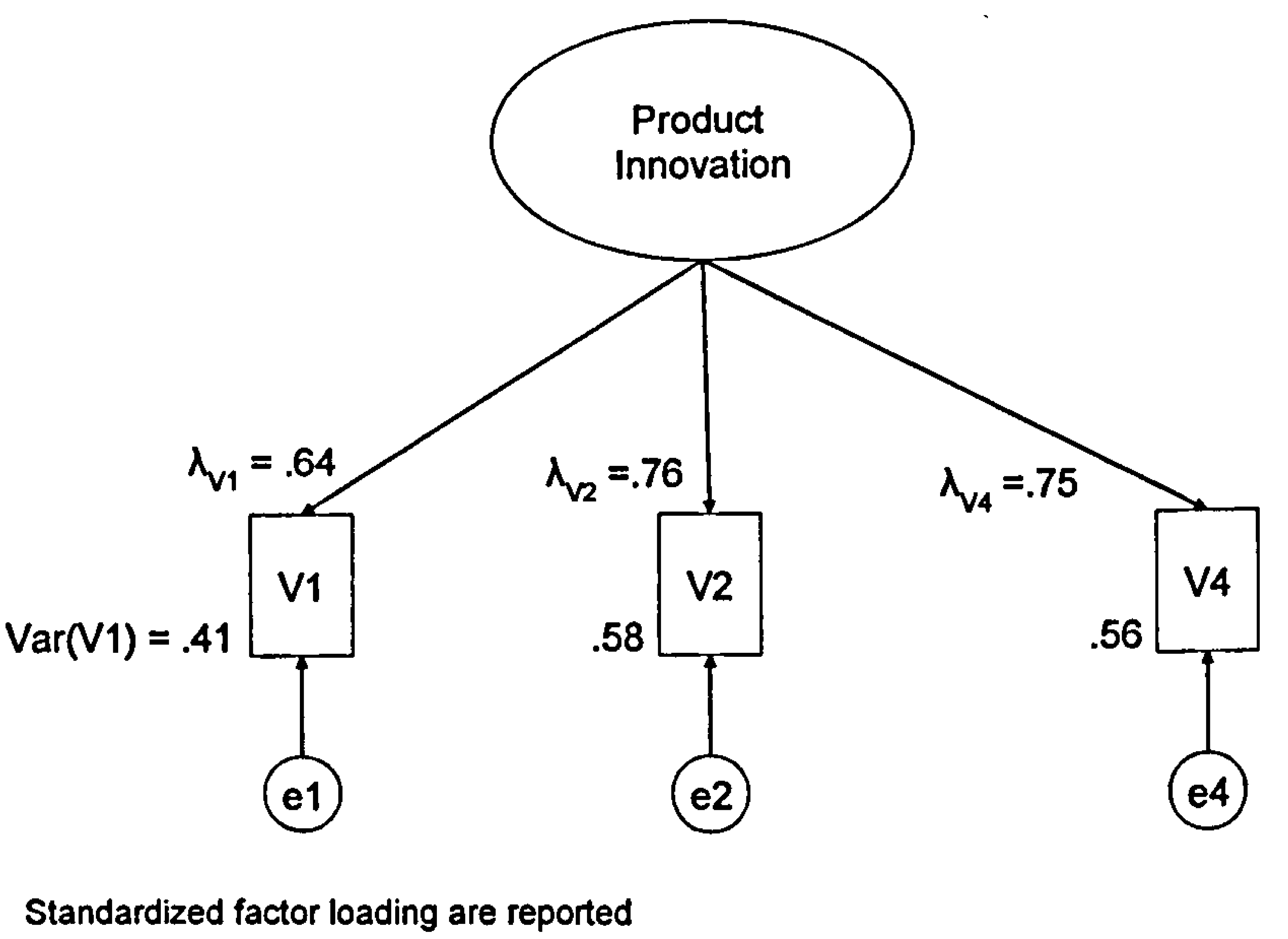
Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
V1	1	.64	--	
V2	1.21 (.19)	.76	6.30	***
V3	.33 (.18)	.76	1.86	Not Significant P > .05
V4	1.34 (.21)	.75	6.30	***
Convergent Validity				
Average Variance Extracted (AVE)	.37			
Reliability				

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
Composite Reliability	.67		Cronbach's α	.93
Model Fit				
χ^2 (df) → P; χ^2/df	6.426 (2) → .04; 3.21		CFI	.96
GFI	.98		TLI (NNFI)	.88
RMSEA	.124		IFI Delta2	.96

Table 6-28: Confirmatory factor analysis of product innovation scale after deleting item 3 (V3)

Item	ML Unstd. Λ (Std. Err)	ML Std. Λ	Critical Ratio	P
V1	1	.64	--	
V2	1.37 (.19)	.76	7.09	***
V4	1.40 (.20)	.75	7.15	***
Convergent Validity				
Average Variance Extracted (AVE)		.53		
Reliability				
Composite Reliability	.77		Cronbach's α	.76
Model Fit				
χ^2 (df) → P; χ^2/df	2.31 (2) → .32; 1.16		CFI	1.00
GFI	.99		TLI (NNFI)	1.00
RMSEA	.03		IFI Delta2	1.00

Figure 6-13: Product innovation scale measurement model



6.4.3 Assessing Reliability

Reliability is an assessment of the degree of consistency between multiple measurements of the variable (Hair et al., 2006). It indicates the amount of random error present in a measurement scale and the greater error, the greater the risk of drawing incorrect conclusions when using the scales to explore relationships among constructs (Carmine & Zeller, 1979; O'Leary-Kelly and Vokurka, 1998). One form of reliability examined in this research is internal consistency. The scale is internally consistent if the items forming the scale are highly intercorrelated since the items are all measuring the same underlying construct. Two tests are used to measure reliability in this study: (1) reliability coefficient Cronbach's α (Cronbach, 1951) and (2) construct reliability (Werts, Linn, & Jöreskog, 1974).

Cronbach's α is given by the following formula (Allen & Yen, 2002):

$$\alpha = \frac{N}{N-1} \; \frac{(\sigma_X)^2 - \sum_{i=1}^{N} (\sigma_{Yi})^2}{(\sigma_X)^2},$$

Where X = the observed score for a test formed by combining N components,

$X = \sum_{i=1}^{N} Y_i$,

$(\sigma_X)^2$ = the population variance of scale X,

$(\sigma_{Yi})^2$ = the population variance of the ith item composing the scale, Y_i, and

N = number of items that are combined to form scale X.

Ping (2004) explains that there have been several proposals for computing construct reliability of items measured with error (i.e. CFA is performed on the measures). In this research, I compute construct reliability using the most frequently used formula based on the work of Werts et al (1974). Ping (2004) explains that the construct reliability of a measure X, with indicators (items) x_1, $x_2, \dots, x_n$, is given by,

$$\rho_X = \frac{(\Sigma\lambda_i)^2\,\mathrm{Var}(X)}{(\Sigma\lambda_i)^2\,\mathrm{Var}(X) + \Sigma\mathrm{Var}(e_i)},$$

where λ_i is the loading of x_i on X, e_i is the error term for x_i, Var(X) is the disattenuated (measurement error free) variance of X (i.e., available in a CFA model), and Σ denotes a sum. He further notes that Gerbing and Anderson (1988) pointed out that for unidimensional measures there is little practical difference between coefficient alpha and construct reliability, which was indeed the case in this study.

As shown in Table 6-29, both measures of reliability exceeded the recommended lower limit of 0.7 (Nunnally & Bernstein, 1994) for all of the scales. The reliabilities of the scales ranged from .75 to .94.

Table 6-29: Reliability measures for multi-tem scales

	Diagnostic	Interactive	Diversity Stake.	Diversity Ops	Search	Prod. Innov.
Cronbach's α	0.89	0.89	0.76	0.77	0.75	0.77
Construct Reliability (WLJ)	0.89	0.93	0.76	0.80	0.75	0.76

6.4.4 Assessing Convergent Validity

Convergent validity was originally advanced by Campbell & Fiske (1959) and it involved assessing the correspondence (correlation) between different measures of the construct using different methods. However, since research studies typically employ only one measurement method, different methods of the construct are substituted for with multiple items of the scale. Each item may be considered a different method for measuring the construct. Thus, convergent validity involves assessing the correspondence between the items of the scale.

The items that are indicators of a specific construct should share a high proportion of their variance with each other (Hair et al., 2006). In this research, I assess convergent validity using four approaches.

The first approach of assessing convergent validity involves reviewing the *t*-tests for the factor loadings. If all of the *t*-tests are significant, then the indicators are effectively measuring the same construct (Anderson & Gerbing, 1988). As shown in Table 6-30, all the factor loadings are significant at $p<0.001$.

The second approach of assessing convergent validity involves reviewing the standardized factor loadings to see if they are greater than 0.5 (Hair et al., 2006). As shown in Table 6-30, all the standardized factor loadings are greater than the minimum threshold of 0.5.

Table 6-30: Standardized factor loadings and their statistical significance

	Diagnostic	Interactive	Diversity Stake.	Diversity Ops	Search	Prod. Innov.
Standardized Factor Loadings (Range)	0.76 to 0.86	0.64 to 0.84	0.55 to 0.78	0.66 to 0.79	0.69 to 0.74	0.64 To 0.75
Statistical Significance	***	***	***	***	***	***

The third approach is suggested by Fornell & Larker (1981) and involves calculating a statistic involving the percentage error variance in a measure, Average Variance Extracted (AVE). Ping (2004) explains that the variance of a measure can be expressed as,

$$\mathrm{Var}(x_1+\ldots+x_n) = \mathrm{Var}(\lambda_1 X+e_1+\ldots+\lambda_n X+e_n) = (\Sigma\lambda_i^2)\mathrm{Var}(X)+\Sigma\mathrm{Var}(e_i),$$

if X and e are independent, where λ_i is the loading of the indicator x_i on the latent variable X., Var(X) is the disattenuated (error free) variance of X, and e_i is the measurement error of x_i. AVE is given by,

$$\text{AVE}_X = \frac{(\Sigma\lambda_i^2)\text{Var}(X)}{(\Sigma\lambda_i^2)\text{Var}(X) + \Sigma\text{Var}(e_i)},$$

where Σ indicates a sum. The result is the percentage of the total variance of a measure represented or extracted by the variance due to the construct, $\lambda_1^2\text{Var}(X) + \ldots + \lambda_n^2\text{Var}(X) = (\Sigma\lambda_i^2)\text{Var}(X)$. AVE ranges from 0 to 1.

Fornell & Larker (1981) suggest adequately convergent valid measures should contain less than 50% error variance (i.e., AVE should be .5 or above). As shown in Table 6-31, all the AVEs are 0.5 or greater except the diversity (stakeholder) scale, which is slightly less than the recommended lower limit at 0.46.

Table 6-31: Average Variance Extracted (AVE) from multi-tem scales

	Diagnostic	Interactive	Diversity Stake.	Diversity Ops	Search	Prod. Innov.
AVE	0.66	0.54	0.46	0.53	0.50	0.53

The fourth approach of assessing convergent validity involves estimating construct reliabilities of the scales using information from CFA. Scales that have construct reliabilities that are 0.7 and higher are considered convergent (Hair et al., 2006). As shown in Table 6-29, all scales have construct reliabilities higher than 0.7.

6.4.5 Assessing Discrminant Validity

Discriminant validity represents the degree to which measures of different constructs are unique (Bagozzi, Yi, & Phillips, 1982; O'Leary-Kelly & Vokurka, 1998). Thus, high discriminant validity provides evidence that the construct is unique and captures some phenomenon other measures do not (Hair et al., 2006). I assess the discriminant validity of the scales using four different tests.

In the first test, I demonstrate discriminant validity by using a technique suggested by Fornell & Larcker (1981) in which I show that the average variance extracted (AVE) for each construct (within construct variance) is greater than the squared correlations between constructs (between construct variance). The logic here is based on the idea that the latent construct should explain its item measures (indicators) better than it explains other constructs (Hair et al., 2006). Put differently, constructs should each have more error-free (extracted) variance than variance shared with other constructs (R^2) (Ping, 2004).

Table 6-32 reports the discriminant validity of the scales. By comparing each entry in the diagonal of the table containing the average extracted variance (AVE) of the scale with the entries to the top and to the right of it containing the squared correlations between it and other scales, I could confirm that all the AVEs are higher suggesting discriminant validity is demonstrated.

Table 6-32: Results of discriminant validity: Correlations exhibited left of the diagonal, squared correlations (R^2) between measures exhibited right of the diagonal and extracted variances (AVE) exhibited at the diagonal

	1	2	3	4	5	6	Cronbach's α
Innovation	0.53	0.20	0.00	0.01	0.01	0.01	.76
Search	0.45	0.50	0.02	0.05	0.05	0.04	.75
Diagnostic	0.07	0.14	0.66	0.51	0.16	0.15	.89
Interactive	0.11	0.22	0.72	0.54	0.27	0.23	.93
Diversity - Operations	0.10	0.22	0.40	0.52	0.53	0.22	.80
Diversity - Stakeholder	0.09	0.20	0.39	0.48	0.47	0.46	.76

In the second test, I demonstrate discriminant validity by showing that the correlations between scales of different constructs using the same method of measurement are lower than the reliability coefficients (Crocker and Algina, 1986; Kaynak, 2003). The bivariate correlations between the measures and the reliability coefficients (Cronbach's α) are presented in Table 6-32. Significant correlations are expected because of the theoretical relation between them. Nevertheless, the correlation coefficients were lower than the reliability coefficients, suggesting that measures have discriminant validity.

The third test involves constructing confidence intervals (±2 standard errors) around the correlation estimate between scales. If the confidence interval does not include 1.0, discriminant validity is achieved (Anderson and Gerbing, 1988). None of the confidence intervals (±2 standard errors) for each bivariate correlation of factors included 1.0.

The last test involves comparing two CFAs models on each pair of scales. One of the two models is constrained with the correlation between the two scales is set to 1 and the other is free (Bagozzi & Phillips, 1982). If the resulting single degree of freedom difference in the chi-squares of the two models are significant then the two scales are different. Hence, discriminant validity is established. Since I have 8 scales, I estimated 30[83] models (15 constrained, 15 unconstrained) and conducted 30 Chi square (χ^2) tests. Table 6-33 reports the results of the pair-wise tests. All the differences between the models were statistically ($p<0.001$) significant suggesting discriminant validity.

Table 6-33: Results of discriminant validity pair-wise tests

Test (i with j)	ML Estimate Correlation	Constrained Model ML X^2	Unconstrained Model ML X^2	ML X^2 Difference
Diagnostic with...				
Innovation	0.07	99.2 (14)	22 (13)	77.2***
Search	0.15	73.4 (14)	24.2 (13)	49.2***
Interactive	0.82	106.98 (43)	95.44 (42)	11.54***
Diversity (Operations)	0.53	66.2 (14)	37.9 (13)	28.3***
Diversity (Stakeholder)	0.46	80 (20)	36.3 (19)	43.7***
Interactive with...				
Innovation	0.12	140.4 (34)	73.3 (33)	67.1***
Search	0.28	95.5 (34)	58.9 (33)	36.6***

83 The formula for calculating the number of pairs of scales to test is C(m,2)=m!/[(m-2)!*2!] (Ahire, Golhar, Walker, 1996).

Test (i with j)	ML Estimate Correlation	Constrained Model ML X^2	Unconstrained Model ML X^2	ML X^2 Difference
Diagnostic	0.82	106.98 (43)	95.44 (42)	11.54***
Diversity_Operations	0.72	97.1 (33)	78.4(32)	18.9***
Diversity (Stakeholder)	0.56	112.2 (43)	79.2(42)	33***
Innovation with...				
Search	0.56	34 (9)	2.4 (8)	31.6***
Diagnostic	0.07	99.2 (14)	22 (13)	77.2***
Interactive	0.12	140.4 (34)	73.3 (33)	67.1***
Diversity (Operations)	0.12	89.1 (8)	13 (7)	76.1***
Diversity (Stakeholder)	0.06	113.7 (14)	30.5 (13)	83.2***
Search with...				
Innovation	0.56	34 (9)	2.4 (8)	31.6***
Diagnostic	0.15	73.4 (14)	24.2 (13)	49.2***
Interactive	0.28	95.5 (34)	58.9 (33)	36.6***
Diversity (Operations)	0.31	53.9 (8)	7.9 (7)	46.0***
Diversity (Stakeholder)	0.19	96.4 (14)	37.6 (13)	58.8***
Diversity (Operations) with...				
Innovation	0.12	89.1 (8)	13 (7)	76.1***
Search	0.31	53.9 (8)	7.9 (7)	46.0***
Diagnostic	0.53	66.2 (14)	37.9 (13)	28.3***
Interactive	0.72	97.1 (33)	78.4 (32)	18.9***
Diversity (Stakeholder)	0.64	82.0 (13)	41.4 (12)	40.6***
Diversity (Stakeholder) with...				
Innovation	0.06	113.7 (14)	30.5 (13)	83.2***
Search	0.19	96.4 (14)	37.6 (13)	58.8***
Diagnostic	0.46	80 (20)	36.3 (19)	43.7***
Interactive	0.56	112.2 (43)	79.2 (42)	33.0***
Diversity (Operations)	0.64	82.0 (13)	41.4 (12)	40.6***

6.4.6 Overall Assessment of the Measures

As shown in Table 6-34, all the scales except the risk scale passed all the procedures and tests in the measure validity framework. These results allow us now to proceed to hypotheses testing were the validated measures would be used to test the structural relationships between them.

Table 6-34: Summary results of construct validation

Step	Procedure/Test	Results
Establish Content Validity	1- Use of existing and validated scales 2- Pre-test of the questionnaire with academics	Minor adjustments were made in terms of wording and presentation.
Establish Unidimensionality	1- Use Exploratory Factor Analysis (EFA) to determine the multiple dimensions (factors) of the innovation activities and diversity scales. 2- Use Confirmatory Factor Analysis (CFA) to evaluate the fit of all scales.	Unidimensionality was achieved for all the multi-item scales except the risk scale. The risk scale was converted into single item scale and was dropped from subsequent steps.
Establish Reliability	1- Estimate reliability coefficient Cronbach's α. 2- Estimate construct reliabilities	All scales passed the tests.
Establish Convergent Validity	1- Evaluate statistical significance of factor loadings. 2- Evaluate the standardized factor loading. 3- Estimate Average Variance Extracted (AVE). 4- Evaluate construct reliabilities estimated earlier	All scales passed the tests.
Establish Discriminant Validity	1- Compare the squared correlation of each pair of factors to AVE for each factor. 2- Compare the correlation of each pair of factors to reliability for each factor. 3- Construct confidence intervals (±2 standard errors) around the correlation estimate between scales. 4- Compare the Chi Squared difference between the constrained and unconstrained CFAs of all pair-wise combinations of the scales	All scales passed the tests.

Having ensured that the scales exhibited content validity, unidimensionality, reliability, convergent validity, and discriminant validity, summated scales were developed for the scales of diagnostic, interactive, innovation search, and product innovation and for the subscales of diversity (stakeholders), and diversity (operations). The composite measure was formed by averaging the items in the scale (Hair et al., 2006). Organizational performance measurement diversity was measured using an index of the three subscales: financial, stakeholder, and operations. The averages of each of the subscales were summed and averaged to form the index (Hair et al., 2006). These composite measures most likely have reduced measurement error associated with single indicators of constructs. The risk construct was measured with one item, which may have undoubtedly introduced measurement error. However, since it has strong content validity, the measurement error may have had little impact on the results. The employee size and the age of the organization were both objective single indicator measures. For ease of interpretation of results, Size was converted into an interval measure because it exhibited a positively skewed normal distribution.

6.5 Testing hypotheses

To test the hypotheses advanced in this study, the statistical techniques shown in Table 4-16 will be undertaken.

6.5.1 Descriptive Statistics

Table 6-35 provides descriptive statistics on the variable used to test the proposed hypotheses. It also includes information on the correlation between the variables and their corresponding 2-tailed statistical significance.

Table 6-35: Statistical descriptives and correlations of the variables used in hypotheses testing

Variable	Mean	Std. Dev.	1	2	3	4	5	6	7	8
Diagnostic	4.8569	.79644	1	.716**	.511**	.136	.147	.066	.078	.036
Interactive	4.2518	.85510	.716**	1	.590**	.218**	.267**	.105	.041	-.031
Diversity	4.6470	.58048	.511**	.590**	1	.217**	.245**	.073	.107	.127
Search	4.0699	1.00374	.136	.218**	.217**	1	.273**	.446**	-.024	-.012
Risk	2.7848	1.16171	.147	.267**	.245**	.273**	1	.234**	-.056	.002
Product Innovation	4.1188	.86236	.066	.105	.073	.446**	.234**	1	-.165*	-.101
Age[84]	58.14	52.031	.078	.041	.107	-.024	-.056	-.165*	1	.234**
Size[85]	2.61	2.018	.036	-.031	.127	-.012	.002	-.101	.234**	1

** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

[84] In the initial analyses, Age was Log transformed to correct for its slightly skewed normal distribution. However, the log transformation of Age did not change the regression and path results, so for maintaining simplicity of interpretation, Age was entered into the analyses untransformed.

[85] As with Age, the same analysis was carried out for Size yielding similar results. Hence, Size was entered into the analysis untransformed.

6.5.2 Testing moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation (Hypotheses 2a, 2b, 2c, 3a, and 3b)

The moderated regression analysis was carried out using OLS hierarchical multiple regression method, which calls for entering variables into the analysis cumulatively in several steps. Table 6-37 displays the results of the OLS hierarchical regression of the moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation. First, only the control variables (Age, Size, Search, and Risk) were entered into the regression analysis forming model 1. Second, the main effects (Diversity) and moderators (Diagnostic and Interactive) were added forming model 2. Third, the moderating effects (Diversity X Diagnostic and Diversity X Interactive) were added forming model 3. Model 3 will be used to test the moderated hypotheses because it contains the moderated effects.

To ensure the robustness of the findings of the analysis, I asked SPSS to produce casewise diagnostics, so I could identify potential outliers and assess their impact on the stability of the regression estimates. Table 6-36 shows the results of the SPSS procedure. I ran the regression models after excluding case 62 but I found there were no changes in the estimates, so case 62 was included back into the reported analysis, as there is no substantive reason for excluding it.

Table 6-36: Casewise diagnostics

Case Number	Standard Residual	Innovation	Predicted Value	Residual
62	3.088	5.67	3.3479	2.31875

Table 6-37: OLS Regression results of the moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation

	Product Innovation		
	Model 1	Model 2	Model 3
Intercept	2.649***	2.615***	2.849***
	(.298)	(.318)	(.320)
Control variables			
Age	-.002+	-.002+	-.002+
	(.001)	(.001)	(.001)
Size	-.028	-.027	-.031
	(.032)	(.033)	(.032)
Search	.353***	.357***	.326***
	(.066)	(.068)	(.067)
Risk	.085+	.090+	.072
	(.057)	(.060)	(.059)
Main effects			
PM Diversity		-.043	-.195+
		(.143)	(.149)
Moderators			
Diagnostic		.039	.114
		(.118)	(.118)
Interactive		-.023	-.114
		(.119)	(.125)
Moderated effects			
Diversity x Diagnostic			.453**
			(.189)
Diversity x Interactive			-.552***
			(.179)
Model statistics			
R^2	.238	.239	.289
ΔR^2		.001	.05**
F-value	.043	1.895+	2.212*
N	145	145	145

Unstandardized regression coefficients are reported. Standard errors are in parentheses. Two-tailed tests for controls, one-tailed tests for hypothesized variables.

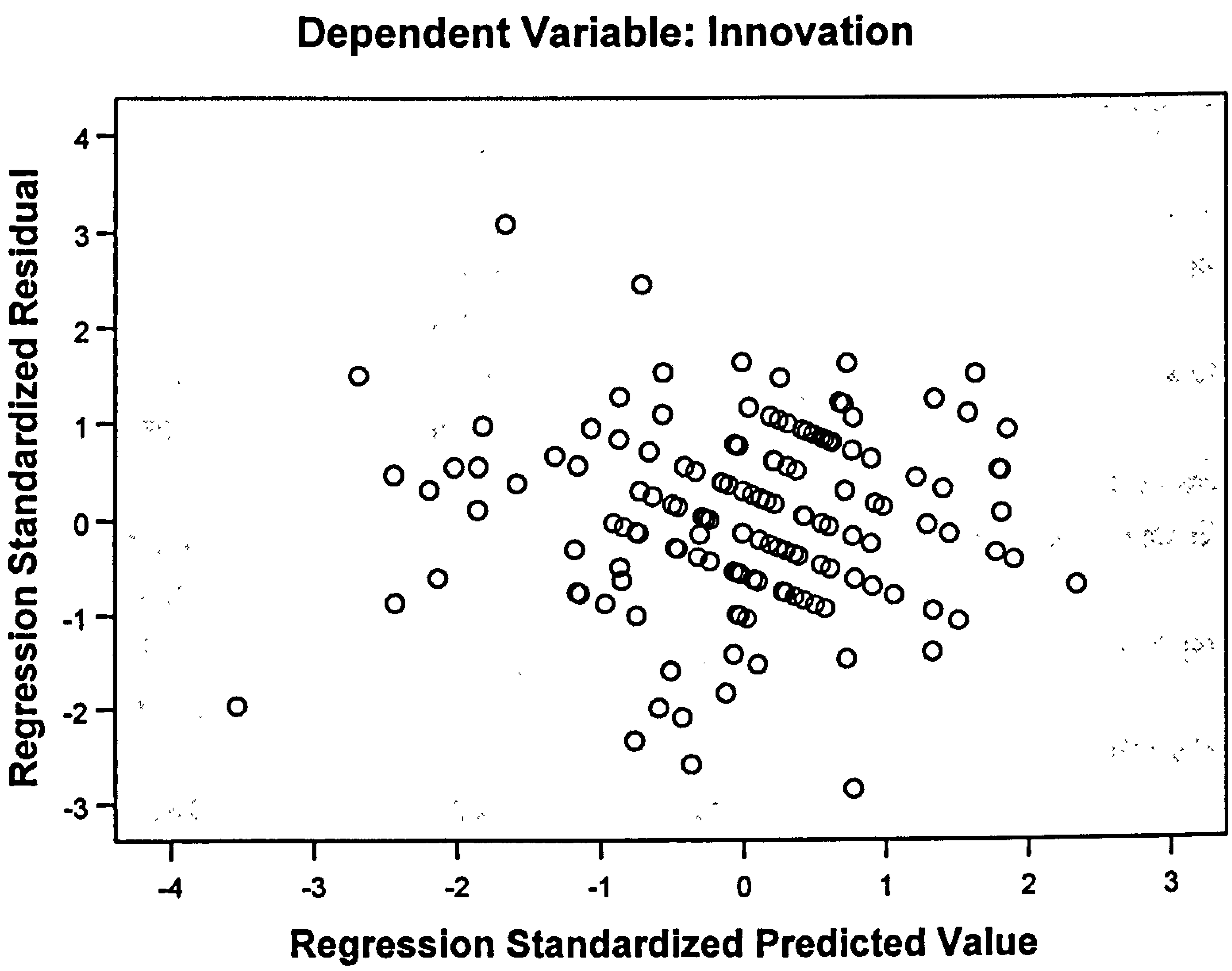
+ p ≤ .10; * p ≤ .05; ** p ≤ .01; *** p ≤ .001

6.5.2.1 Testing multiple regression assumptions

Before interpreting the results of the regression, I used the framework shown in Table 4-15 to test the underlying assumption of the OLS regression[86].

First, linearity is assessed through the analysis of the standardized residuals[87]. More specifically, it is assessed through the analysis of the plot of standardized residuals versus the predicted values of the dependent variable (Product Innovation). Figure 6-14 does not exhibit any nonlinear pattern to the residuals, thus ensuring that the overall equation (multivariate) is linear.

Figure 6-14: Plot of standardized residuals versus predicted values of product innovation

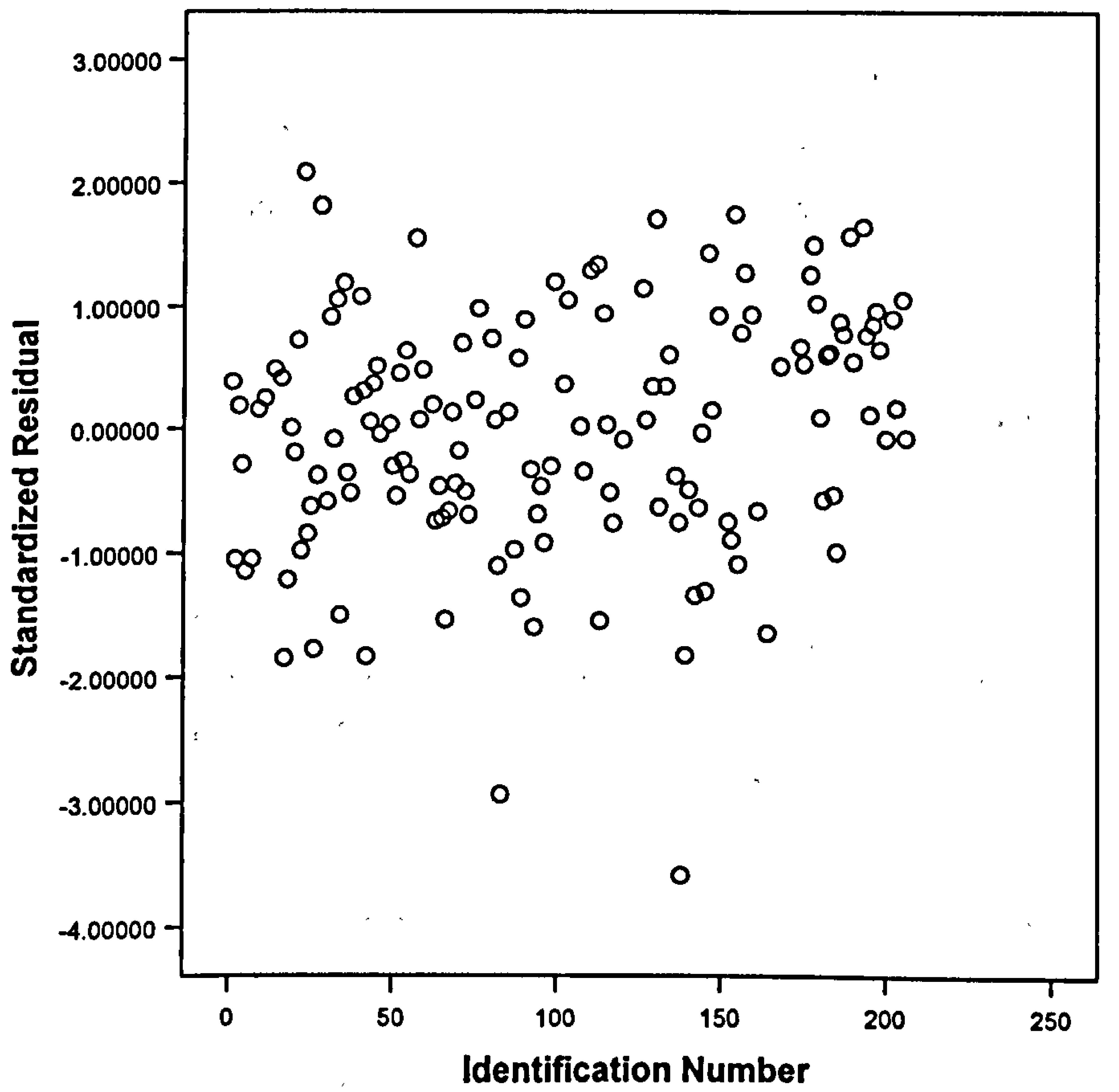


[86] The testing of assumptions underlying OLS regression was carried out for all of the multiple regression analyses in this study. However, I report only the test for the first regression analysis to conserve space.
[87] Standardized residuals are the standard values of the residuals which are the difference between the predicted value of the dependent variable and its observed values.

Figure 6-14 is also used to examine the constancy of the residuals across values of the explanatory variables (Homoscedasticity). The figure does not show a pattern of increasing or decreasing residuals. This finding indicates homoscedasticity in the multivariate (the set of explanatory variables).

Third, I examined the assumption of the independence of the residuals. This is assessed through the analysis of the plot of standardized residuals versus time sequenced variable (Case ID). Since the received survey responses were assigned identification ID based on their time of arrival, this created a time sequenced variable. Figure 6-15 does not show any consistent pattern of the standardized residuals. This finding indicates that the observations of the error term are uncorrelated with each other (no serial correlation/no autocorrelation). Also, I conducted the Durbin Watson test (Durbin & Watson, 1951) to test for the presence of serial correlation. The result of Durbin Watson test was 1.920, which indicates the absence of serial correlation.

Figure 6-15: Plot of residuals versus time sequence case ID



Fourth, the assumption of the absence of perfect multicollinearity was assessed. Severe multicollinearity makes it difficult to separate the effect of individual independent variables because it increases the shared variance between explanatory variables and lower their individual unique variances (Hair et al., 2006). I examined the correlation matrix Table 6-35 to assess collinearity. None of the variable had zero-order (bivariate) correlations greater than 0.9 indicating the absence of collinearity.

However, examining only the correlation matrix for collinearity is not enough because it could also be due to the combined effect of two or more independent variables (multicollinearity). To assess multicollinearity, I evaluated two measures: tolerance and variance inflation factor (VIF). Tolerance is the amount of variability of the selected independent variable not explained by other independent variables and VIF is the inverse of the tolerance value (Hair et al., 2006). Tolerance is calculated for the independent variable by running an OLS regression with the independent variable as a function of the other independent variables and calculating the variance of the independent variable that is explained by the other independent variables (R^2) . The tolerance of the dependent variable is 1- R^2 and the VIF is the inverse of 1- R^2 (Hair et al., 2006). Hair et al (2006) notes that a common cutoff threshold is a tolerance value of 0.10, which corresponds to a VIF value of 10. Table 6-12 shows that the tolerance and VIF values of all the independent variables are not close to the recommended cut-off thresholds. These findings indicate that multicollinearity is not an issue in this research.

Table 6-38: Collinearity statistics

	Age	Size	Search	Risk	Diversity	Diagnostic	Interactive	Diversity X Diagnostic	Diversity X Interactive
Tolerance	.924	.915	.861	.846	.523	.447	.342	.363	.300
VIF	1.08	1.09	1.16	1.18	1.91	2.24	2.93	2.75	3.33

Fifth, the assumption of the normality of the residuals is examined through a visual check of the residuals' histogram and the normal P-P plot of the residuals.

Figure 6-16 illustrates how the residuals falls approximately in a normal distribution. Next, the normal probability plot is examined which plots the standardized residuals against a diagonal line representing normal distribution. As shown, in Figure 6-17, the values fall along the diagonal with no substantial or systematic departures; thus, the residuals are considered to represent normal distribution.

Figure 6-16: Histogram of the residuals

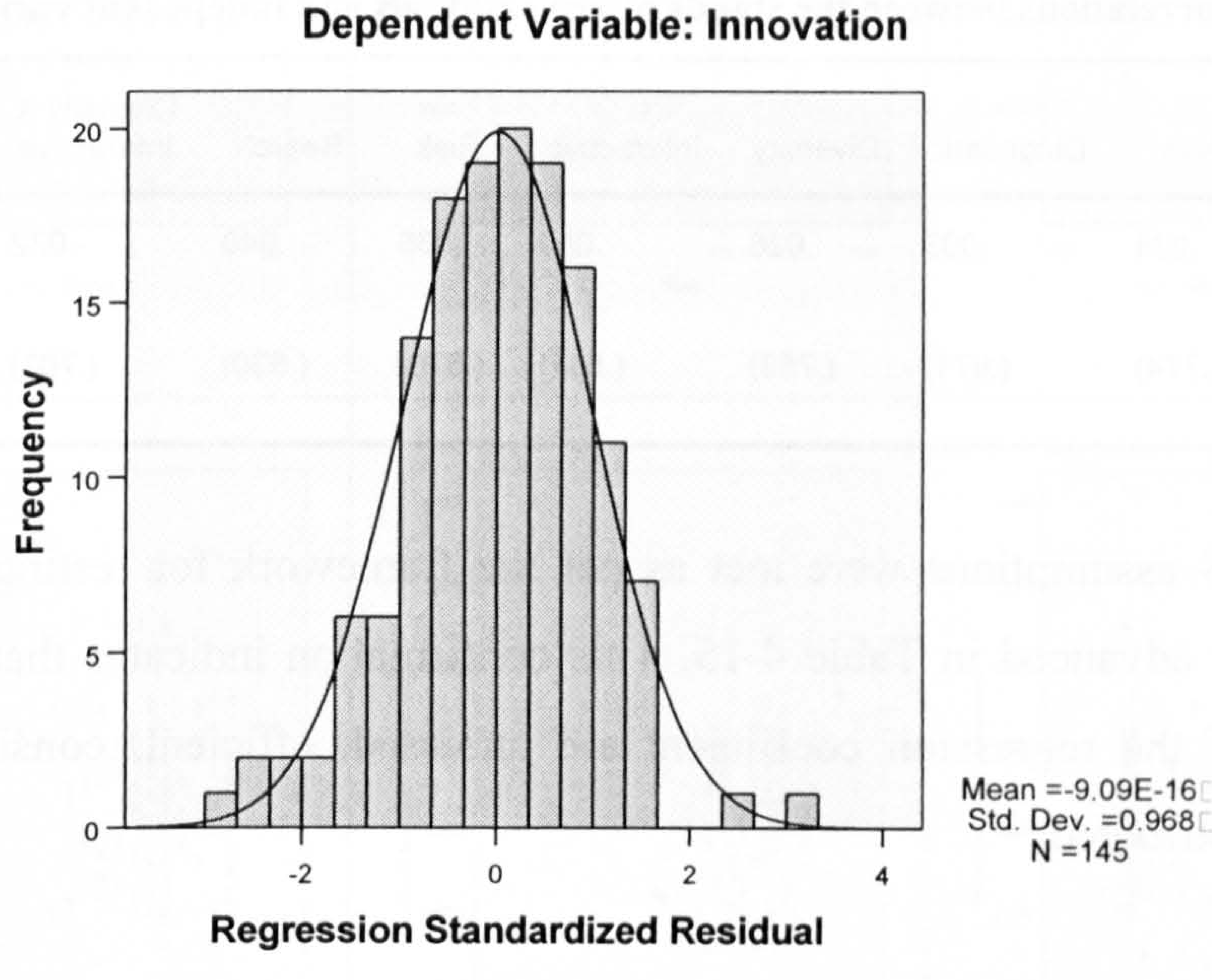
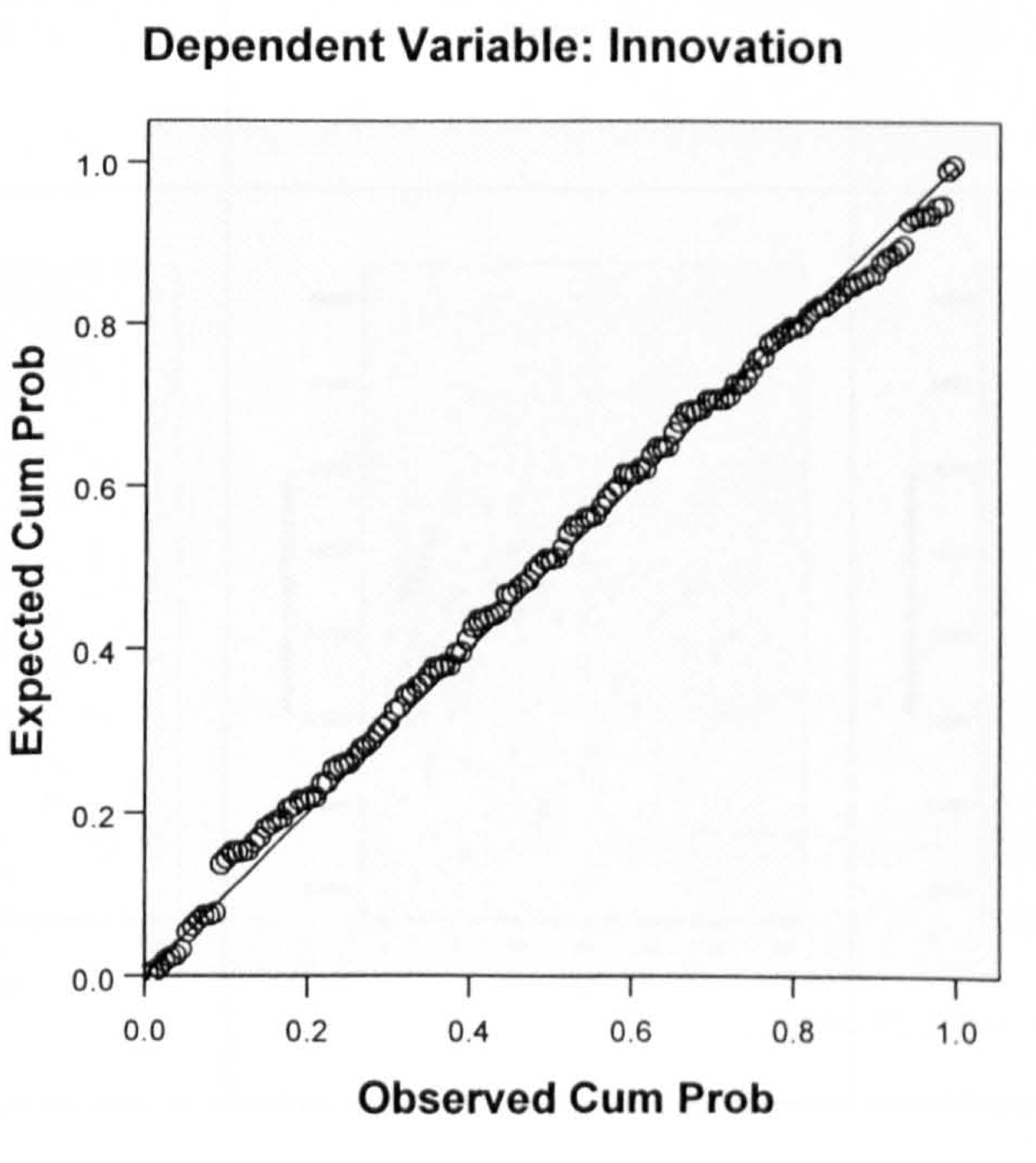


Figure 6-17: Normal P-P plot of regression standardized residual



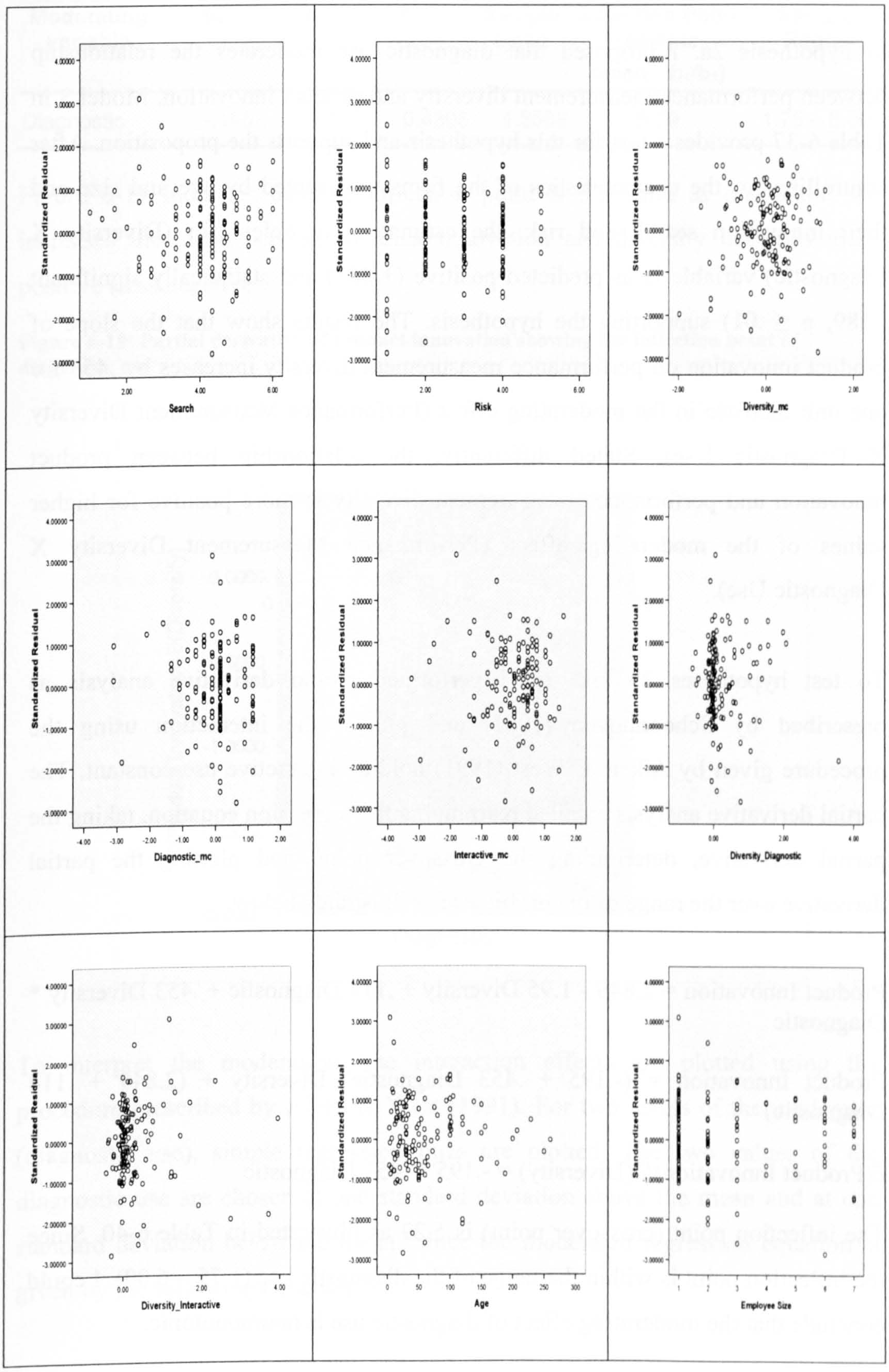
Six, the final assumption calls for the lack of correlation between the independent variables and the residuals. This assumption is examined through the scatter plots of standardized residuals versus each of the independent variables. As shown, in Table 6-39 and Figure 6-18, none of the independent variables is correlated with the standardized residuals; thus, the residuals are considered uncorrelated with the independent variables.

Table 6-39: Correlations between the standardized residuals and indepedent variables

	Age	Diagnostic	Diversity	Interactive	Risk	Search	Diversity X Interactive	Diversity X Diagnostic
Residuals	.024	.003	.026	.050	.036	.040	-.032	.007
Sig. (2-tailed)	(.774)	(.971)	(.758)	(.551)	(.670)	(.630)	(.703)	(.931)

All the OLS assumptions were met as per the framework for testing the OLS assumptions advanced in Table 4-15. This confirmation indicates that the OLS estimates of the regression coefficient are unbiased, efficient, consistent, and normally distributed.

Figure 6-18: Standardized residuals versus the explanatory variables



6.5.2.2 The moderating role of diagnostic use

In hypothesis 2a, I proposed that diagnostic use moderates the relationship between performance measurement diversity and product innovation. Model 3 in Table 6-37 provides a test for this hypothesis and supports the proposition. After controlling for the characteristics of the firms represented by age and size and their innovation search and risk, the estimated coefficient for (Diversity X Diagnostic) variable is as predicted positive (+.453) and statistically significant (.189, $p \leq .01$) supporting the hypothesis. The results show that the slope of product innovation on performance measurement diversity increases by .453 for one unit increase in the moderating effect (Performance Measurement Diversity X Diagnostic Use). Stated differently, the relationship between product innovation and performance measurement diversity is more positive for higher values of the moderating effect (Performance Measurement Diversity X Diagnostic Use).

To test hypotheses 2b and 2c, I performed partial derivative analysis as prescribed by Schoonhoven (1981) and plotted the interaction using the procedure given by Aiken & West (1991) holding interactive use constant. The partial derivative analysis entailed rearranging the regression equation, taking the partial derivative, determining the crossover point, and plotting the partial derivative over the range of diagnostic use as illustrated below.

Product Innovation = 2.849 - 1.95 Diversity + .114 Diagnostic + .453 Diversity * Diagnostic

Product Innovation = (-.195 + .453 Diagnostic) Diversity + (2.849 + .114 Diagnostic)

∂(Product Innovation)/∂(Diversity) = -.195 + .453 Diagnostic

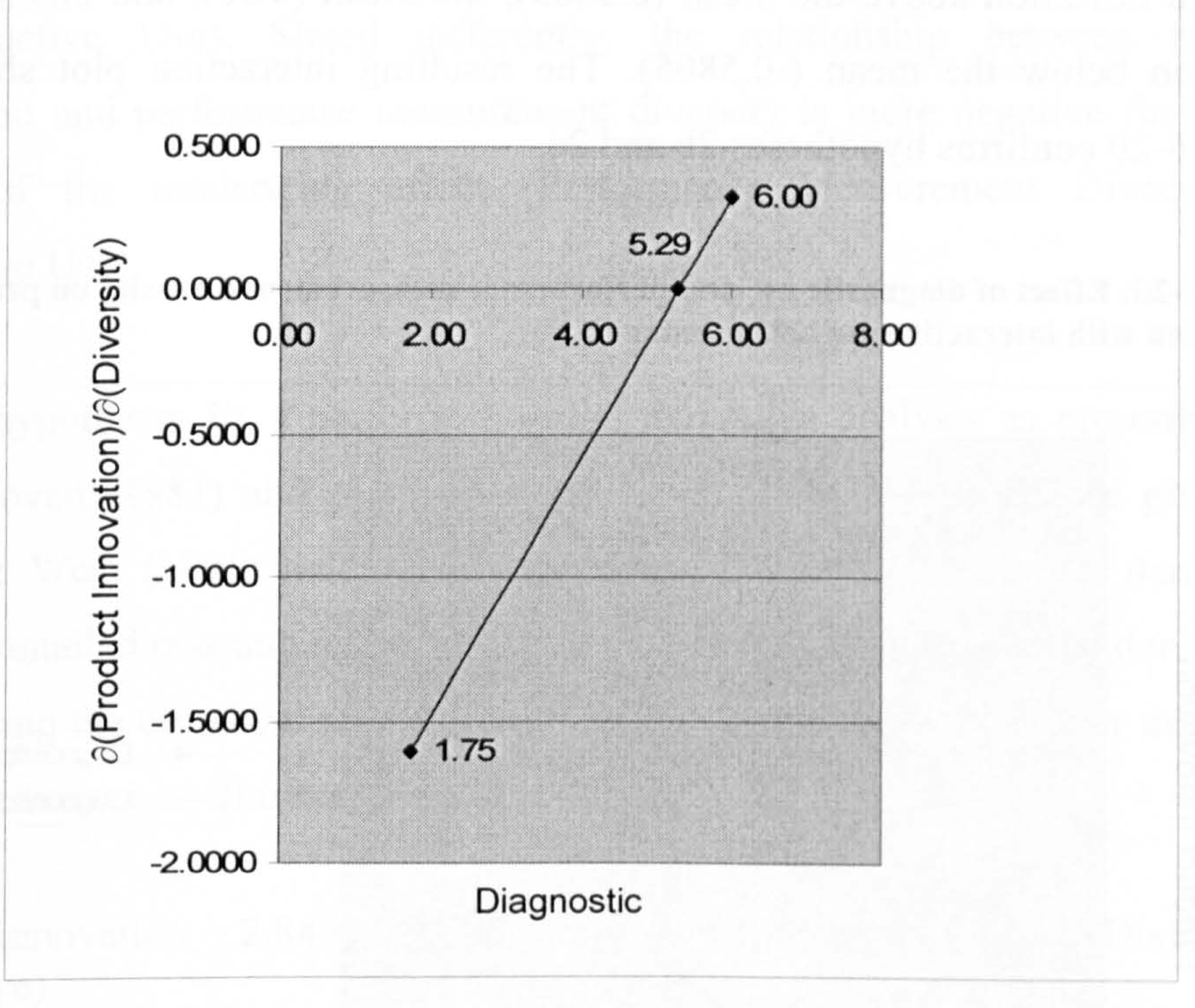
The inflection point (crossover point) is 5.29 as illustrated in Table 6-40. Since the inflection point is within the range of the diagnostic use (1.75 – 6.00). I could conclude that the moderating effect of diagnostic use is nonmonotonic.

Table 6-40: Location of inflection point for diagnostic use in moderating the relationship between performance measurement diversity and product innovation

Moderating variable	b_1	b_3	$-b_1/b_3$	Sample Mean	Inflection Point (sample mean - b_1/b_3)	Sample Range
Diagnostic	-.195	.453	0.4305	4.8569	5.29	1.75 – 6.00

Figure 6-19 indicates that the inflection point is 5.29 and as diagnostic use increases the slope between product innovation and diversity becomes more positive (less negative).

Figure 6-19: Partial derivative of product innovation showing the inflection point of diagnostic use



To interpret the moderation, the interaction effects are plotted using the procedure prescribed by Aiken & West (1991). For two levels of the moderator (diagnostic use), simple regression lines are plotted. The two values of the diagnostic use are chosen at one standard deviation above the mean and at one standard deviation below the mean. Since the moderated regression equation is given by the following equation:

Product Innovation = (-.195 + .453 Diagnostic) Diversity + (2.849 + .114 Diagnostic);

It follows that when diagnostic use = 1 standard deviation above the mean (0.7964), the simple regression equation is given by:

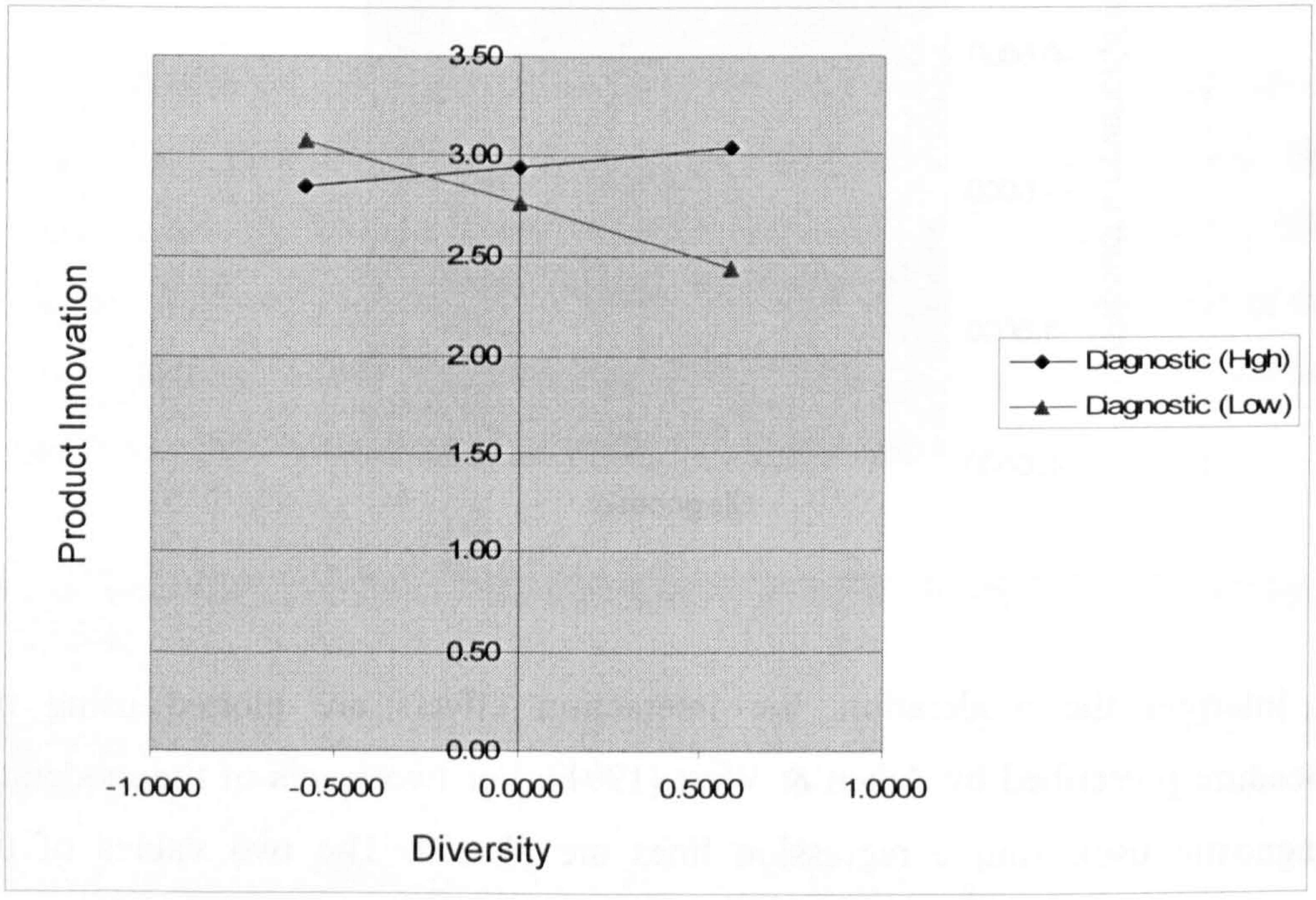
Product Innovation = 2.9398 + .1658 Diversity;

and when diagnostic use = 1 standard deviation below the mean (-0.7964), the simple regression equation is given by:

Product Innovation = 2.7582 -0.5558 Diversity

Now, for each of the two simple regression equations, we substitute three values for Diversity to plot the interaction. The values of Diversity correspond to the one standard deviation above the mean (0.5805), the mean (0.00), and one standard deviation below the mean (-0.5805). The resulting interaction plot shown in Figure 6-20 confirms hypotheses 2b and 2c.

Figure 6-20: Effect of diagnostic use and performance measurement diversity on product innovation with interactive use held constant



6.5.2.3 The moderating role of interactive use

In hypothesis 3a, I proposed that interactive use moderates the relationship between performance measurement diversity and product innovation. Model 3 in Table 6-37 provides a test for this hypothesis and supports the proposition. After controlling for the characteristics of the firms represented by age and size and their innovation search and risk, the estimated coefficient for (Diversity X Interactive) variable is as predicted negative (-.552) and statistically significant (.179, $p \leq .001$) supporting the hypothesis. The results show that the slope of product innovation on performance measurement diversity decreases by .552 for one unit increase in the moderating effect (Performance Measurement Diversity X Interactive Use). Stated differently, the relationship between product innovation and performance measurement diversity is more negative for higher values of the moderating effect (Performance Measurement Diversity X Interactive Use).

To test hypothesis 3b, I performed partial derivative analysis as prescribed by Schoonhoven (1981) and plotted the interaction using the procedure given by Aiken & West (1991) holding diagnostic use constant. The partial derivative analysis entailed rearranging the regression equation, taking the partial derivative, determining the crossover point, and plotting the partial derivative over the range of interactive use as illustrated below.

Product Innovation = 2.849 -1.95 Diversity - .114 Interactive - .552 (Diversity * Interactive)

Product Innovation = (-.195 - .552 Interactive) Diversity + (2.849 - .114 Interactive)

∂ (Product Innovation)/∂(Diversity) = -.195 - .552 Interactive

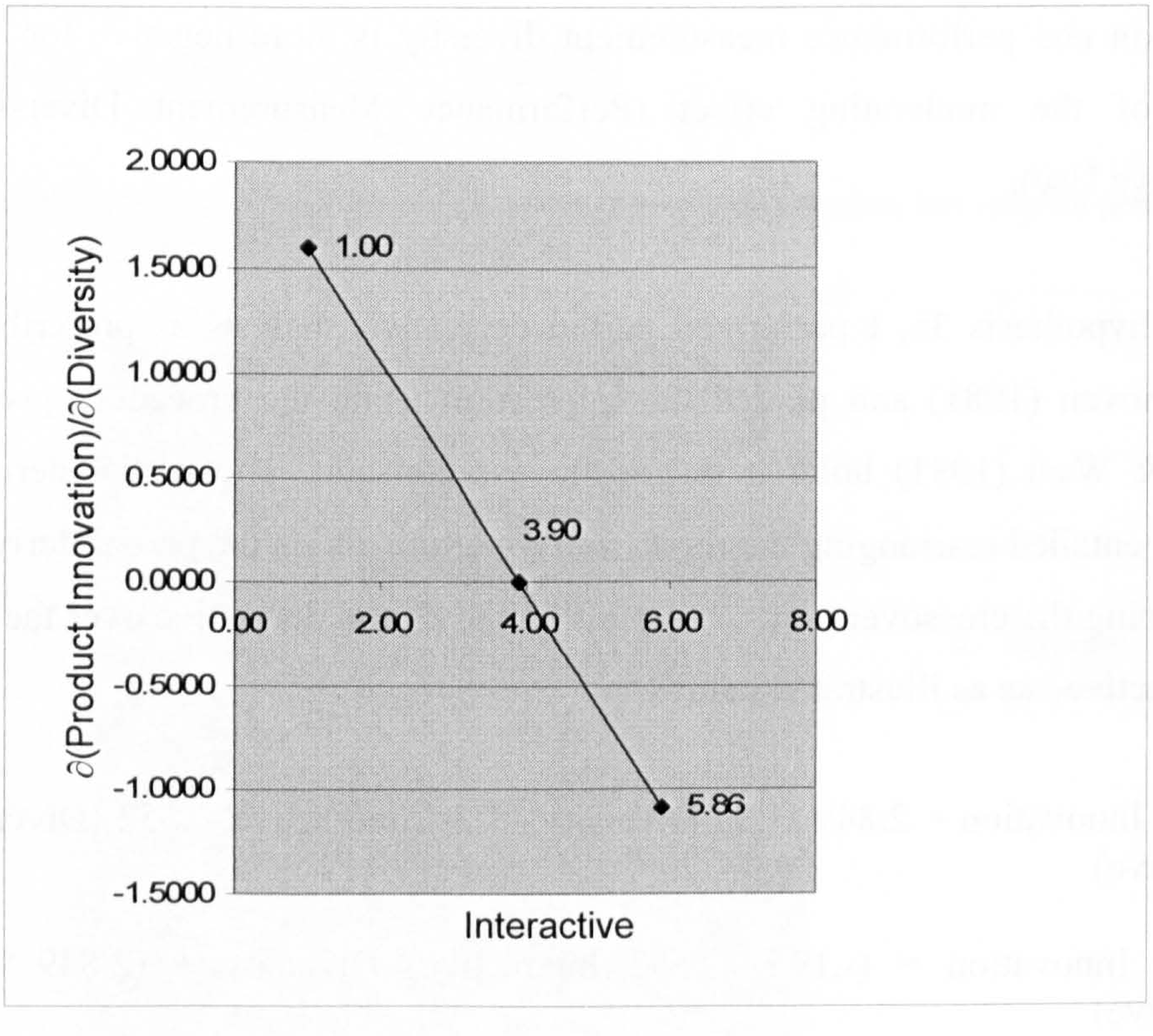
The inflection point (crossover point) is 3.90 as illustrated in Table 6-41 . Since the inflection point is within the range of the interactive use (1.00 – 5.86). I could conclude that the moderating effect of diagnostic use is nonmonotonic.

Table 6-41: Location of inflection point for interactive use in moderating the relationship between performance measurement diversity and product innovation

Moderating variable	b_1	b_3	$-b_1/b_3$	Sample Mean	Inflection Point (sample mean - b_1/b_3)	Sample Range
Interactive	-.195	-.552	-0.3533	4.2518	3.90	1.00 – 5.86

Figure 6-21 indicates that the inflection point is 3.90 and as interactive use increases the slope between product innovation and diversity becomes more negative (less positive).

Figure 6-21: Partial derivative of product innovation showing the inflection point of interactive use



To interpret the moderation, the interaction effects are plotted using the procedure prescribed by Aiken & West (1991). For two levels of the moderator (interactive use), simple regression lines are plotted. The two values of the interactive use are chosen at one standard deviation above the mean and at one standard deviation below the mean. Since the moderated regression equation is given by the following equation:

Product Innovation = (-.195 - .552 Interactive) Diversity + (2.849 - .114 Interactive)

It follows that when interactive use = 1 standard deviation above the mean (0.8551), the simple regression equation is given by:

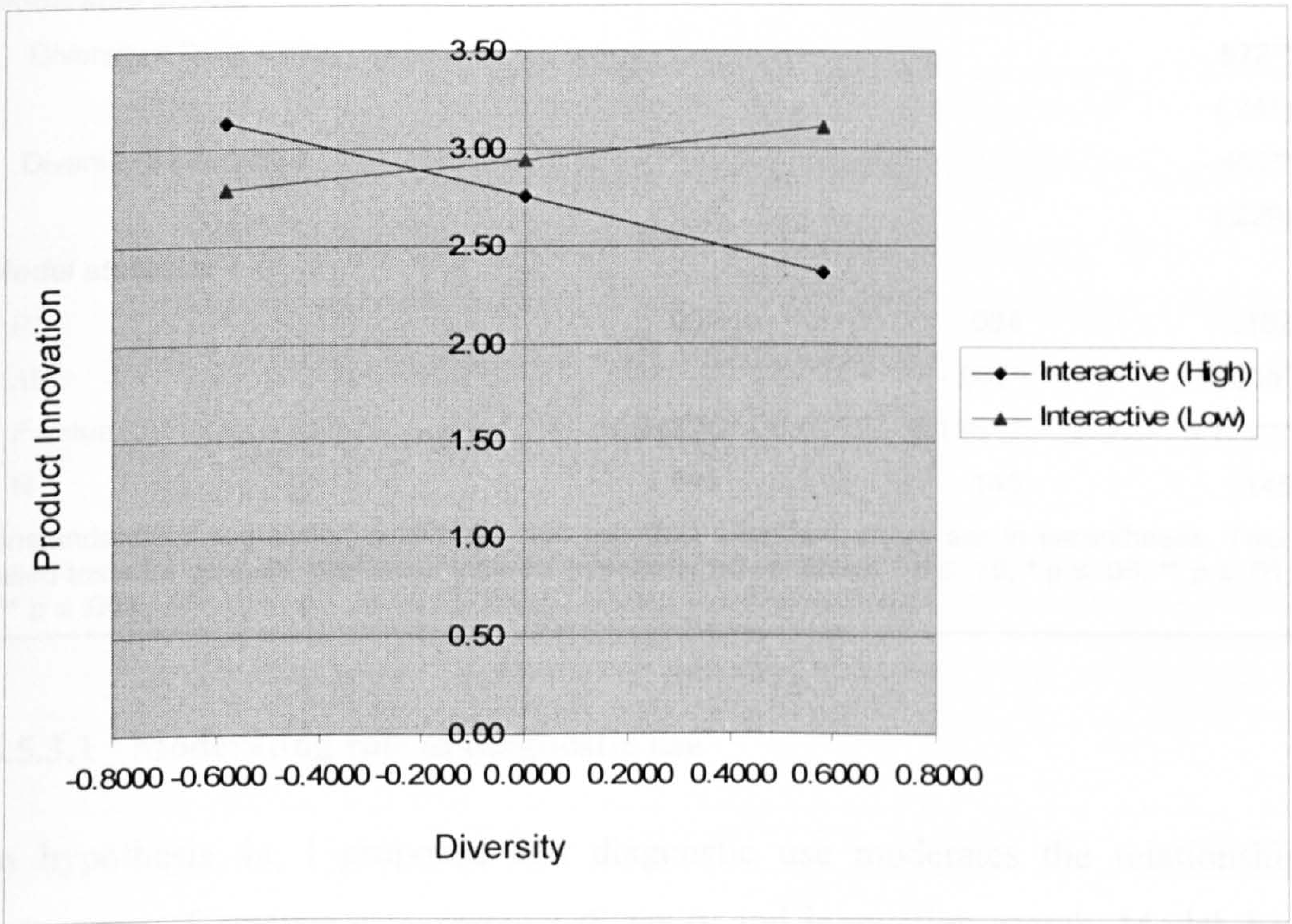
Product Innovation = 2.7515 - 0.6670 Diversity

and when interactive use = 1 standard deviation below the mean (-.8551), the simple regression equation is given by:

Product Innovation = 2.9465 -0.2770 Diversity

Now, for each of the two simple regression equations, we substitute three values for Diversity to plot the interaction. The values of Diversity correspond to the one standard deviation above the mean (0.5805), the mean (0.00), and one standard deviation below the mean (-0.5805). The resulting interaction plot shown in Figure 6-22 confirms hypotheses 3b.

Figure 6-22: Effect of interactive use and performance measurement diversity on product innovation with diagnostic use held constant



6.5.3 Testing moderating effects of performance measurement use on the relationship between performance measurement diversity and innovation search (Hypotheses 4a, 4b, 4c, 5a, and 5b)

Table 6-42 displays the results of the OLS hierarchical regression of the moderating effects of performance measurement use on the relationship between performance measurement diversity and innovation search. First, only the control variables (Age, Size) were entered into the regression analysis forming model 1. Second, the main effects (Diversity) and moderators (Diagnostic and Interactive) were added forming model 2. Third, the moderating effects (Diversity X Diagnostic and Diversity X Interactive) were added forming model 3. Model 3 will be used to test the moderated hypotheses because it contains the moderated effects.

Table 6-42: OLS regression results of the moderating effects of performance measurement use on the relationship between performance measurement diversity and innovation search

	Innovation Search		
	Model 1	**Model 2**	**Model 3**
Intercept	4.104***	4.132***	4.120***
	(.154)	(.153)	(.160)
Control variables			
Age	.000	-.001	.000
	(.002)	(.002)	(.001)
Size	-.003	-.007	-.010
	(.043)	(.043)	(.042)
Main effects			
PM Diversity		-.043⁺	.092
		(.143)	(.193)
Moderators			
Diagnostic		-.081	.000
		(.150)	(.152)
Interactive		.205⁺	.176
		(.150)	(.161)
Moderated effects			
Diversity x Diagnostic			.572**
			(.241)
Diversity x Interactive			-.450**
			(.229)
Model statistics			
R^2	.001	.064	.102
ΔR^2		.063*	.038⁺
F-value	10.942***	6.156***	6.104***
N	145	145	145

Unstandardized regression coefficients are reported. Standard errors are in parentheses. Two-tailed tests for controls, one-tailed tests for hypothesized variables. ⁺ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

6.5.3.1 Moderating role of diagnostic use

In hypothesis 4a, I proposed that diagnostic use moderates the relationship between performance measurement diversity and innovation search. Model 3 in Table 6-42 provides a test for this hypothesis and supports the proposition. After controlling for the characteristics of the firms represented by age and size, the estimated coefficient for (Diversity X Diagnostic) variable is as predicted positive (+.572) and statistically significant (.241, $p \leq .01$) supporting the hypothesis. The results show that the slope of innovation search on performance measurement diversity increases by .572 for one unit increase in the moderating effect (Performance Measurement Diversity X Diagnostic Use). Stated differently, the relationship between innovation search and performance measurement diversity is more positive for higher values of the moderating effect (Performance Measurement Diversity X Diagnostic Use).

To test hypotheses 4b and 4c, I performed partial derivative analysis and plotted the interaction holding interactive use constant. The partial derivative analysis entailed rearranging the regression equation, taking the partial derivative, determining the crossover point, and plotting the partial derivative over the range of diagnostic use as illustrated below.

Innovation Search = 4.120 +.092 Diversity + .000 Diagnostic + .572 Diversity * Diagnostic

Innovation Search = (.092 + .572 Diagnostic) Diversity + 4.120

∂(Innovation Search)/∂(Diversity) = .092 + .572 Diagnostic

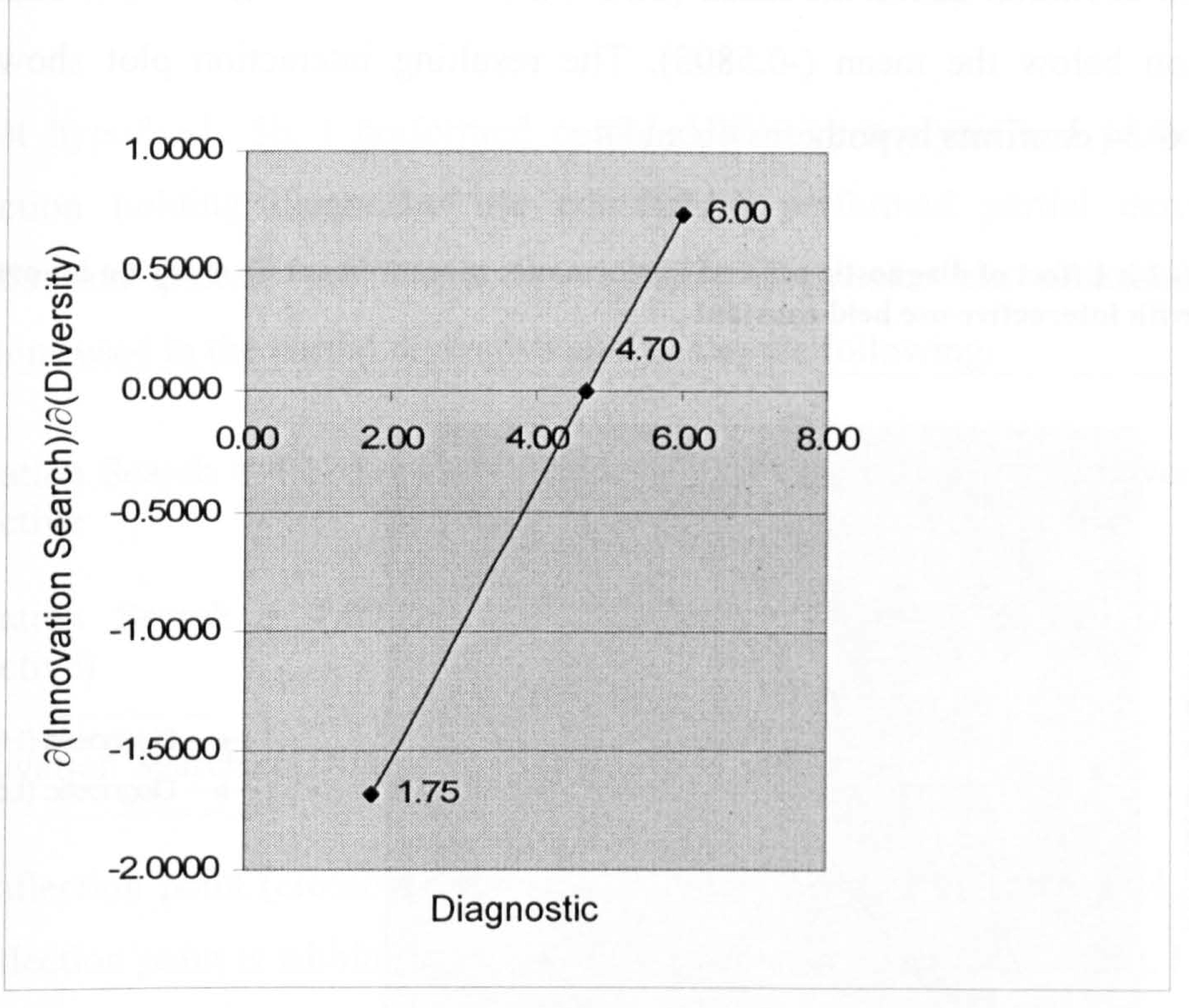
The inflection point (crossover point) is 4.70 as illustrated in Table 6-43. Since the inflection point is within the range of the diagnostic use (1.75 – 6.00). I could conclude that the moderating effect of diagnostic use is nonmonotonic.

Table 6-43: Location of inflection point for diagnostic use in moderating the relationship between performance measurement diversity and innovation search

Moderating variable	b_1	b_3	$-b_1/b_3$	Sample Mean	Inflection Point (sample mean - b_1/b_3)	Sample Range
Diagnostic	.092	.572	-0.1608	4.8569	4.70	1.75 – 6.00

Figure 6-23 indicates that the inflection point is 4.70 and as diagnostic use increases the slope between innovation search and diversity becomes more positive (less negative).

Figure 6-23: Partial derivative of innovation search showing the inflection point of diagnostic use



To interpret the moderation, the interaction effects are plotted using the procedure prescribed by Aiken & West (1991). For two levels of the moderator (diagnostic use), simple regression lines are plotted. The two values of the diagnostic use are chosen at one standard deviation above the mean and at one standard deviation below the mean. Since the moderated regression equation is given by the following equation:

Innovation Search = 4.120 +.092 Diversity + .000 Diagnostic + .572 Diversity * Diagnostic;

It follows that when diagnostic use = 1 standard deviation above the mean (0.7964), the simple regression equation is given by:

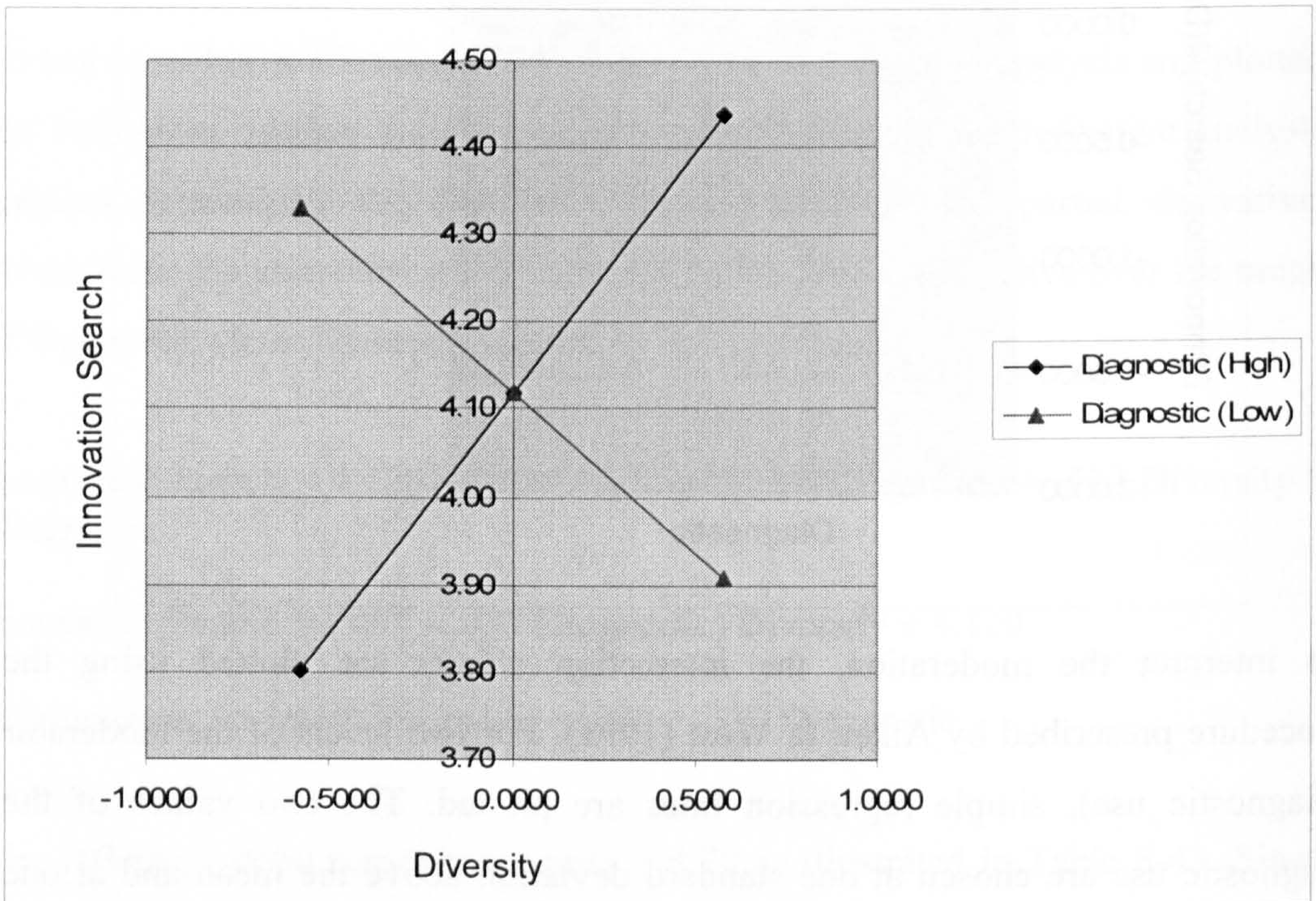
Innovation Search = 4.120 + .5475 Diversity;

and when diagnostic use = 1 standard deviation below the mean (-0.7964), the simple regression equation is given by:

Innovation Search = 4.120 - 0.3635 (Diversity)

Now, for each of the two simple regression equations, we substitute three values for Diversity to plot the interaction. The values of Diversity correspond to the one standard deviation above the mean (0.5805), the mean (0.00), and one standard deviation below the mean (-0.5805). The resulting interaction plot shown in Figure 6-24 confirms hypotheses 4b and 4c.

Figure 6-24: Effect of diagnostic use and performance measurement diversity on innovation search with interactive use held constant



6.5.3.2 Moderating role of interactive use

In hypothesis 5a, I proposed that interactive use moderates the relationship between performance measurement diversity and innovation search. Model 3 in Table 6-42 provides a test for this hypothesis and supports the proposition. After controlling for the characteristics of the firms represented by Age and Size, the estimated coefficient for (Diversity X Interactive) variable is as predicted negative (-.450) and statistically significant (.229, $p \leq .01$) supporting the hypothesis. The results show that the slope of innovation search on performance measurement diversity decreases by .450 for one unit increase in the moderating effect (Performance Measurement Diversity X Interactive Use). Stated differently, the relationship between innovation search and performance measurement diversity is more negative for higher values of the moderating effect (Performance Measurement Diversity X Interactive Use).

To test hypothesis 5b, I performed partial derivative analysis and plotted the interaction holding diagnostic use constant. I performed partial derivative analysis and plotted the interaction holding diagnostic use constant. The equations used in the partial derivative analysis as the following:

Innovation Search = 4.120 +.092 Diversity + .176 Interactive - .450 Diversity * Interactive

Innovation Search = (.092 - .450 Interactive) Diversity + (4.120 +.176 Interactive)

∂(Innovation Search)/∂(Diversity) = .092 - .450 Interactive

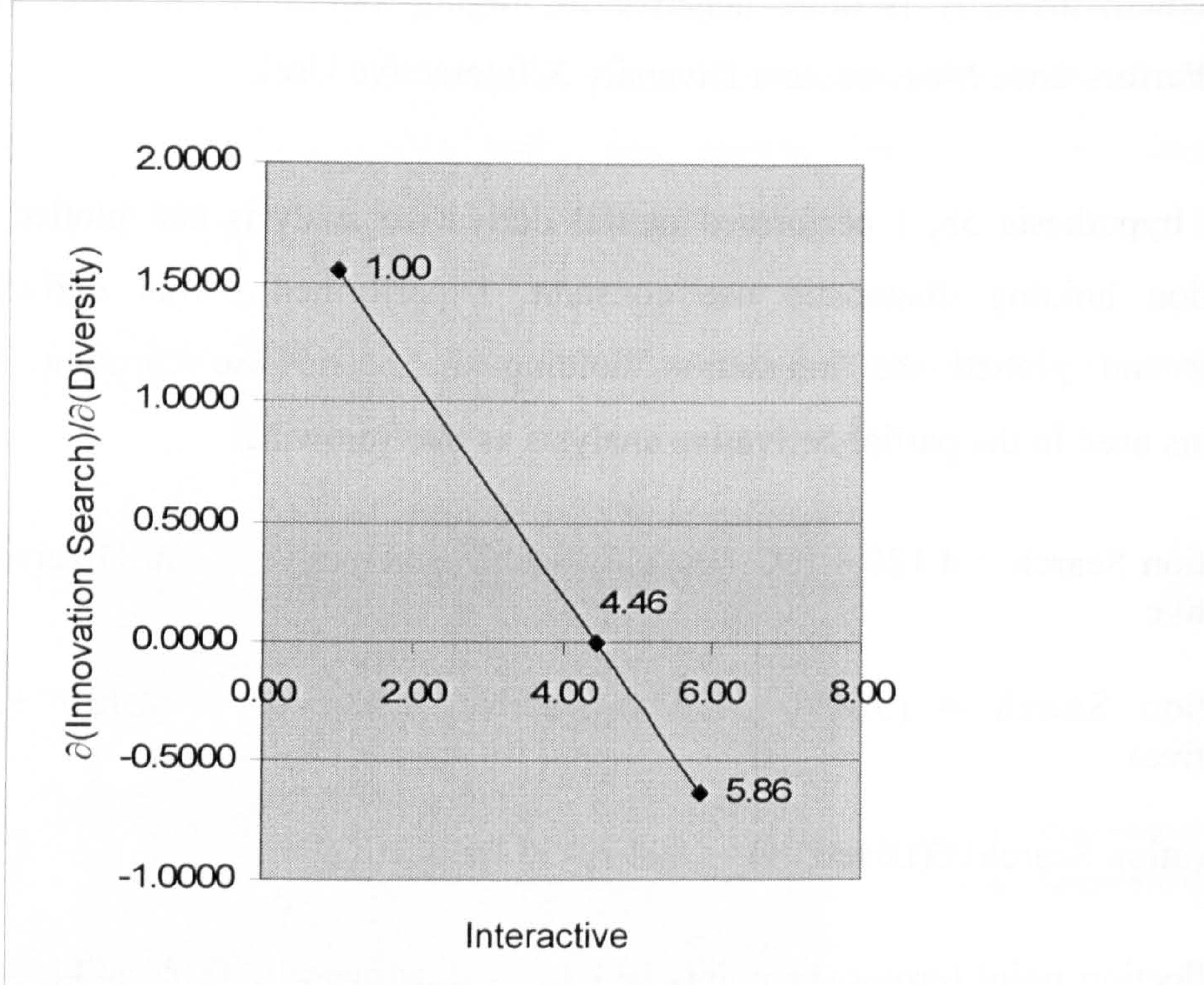
The inflection point (crossover point) is 4.46 as illustrated in Table 6-44. Since the inflection point is within the range of the interactive use (1.00 – 5.86). I could conclude that the moderating effect of diagnostic use is nonmonotonic.

Table 6-44: Location of inflection point for interactive use in moderating the relationship between performance measurement diversity and innovation search

Moderating variable	b_1	b_3	$-b_1/b_3$	Sample Mean	Inflection Point (sample mean - b_1/b_3)	Sample Range
Interactive	-.195	-.552	0.2044	4.2518	4.46	1.00 – 5.86

Figure 6-25 indicates that the inflection point is 4.46 and as interactive use increases the slope between product innovation and diversity becomes more negative (less positive).

Figure 6-25: Partial derivative of innovation search showing the inflection point of interactive use



To interpret the moderation, the interaction effects are plotted using the procedure prescribed by Aiken & West (1991). For two levels of the moderator (interactive use), simple regression lines are plotted. The two values of the interactive use are chosen at one standard deviation above the mean and at one standard deviation below the mean. Since the moderated regression equation is given by the following equation:

Innovation Search = (.092 - .450 Interactive) Diversity + (4.120 +.176 Interactive)

It follows that when interactive use = 1 standard deviation above the mean (0.8551), the simple regression equation is given by:

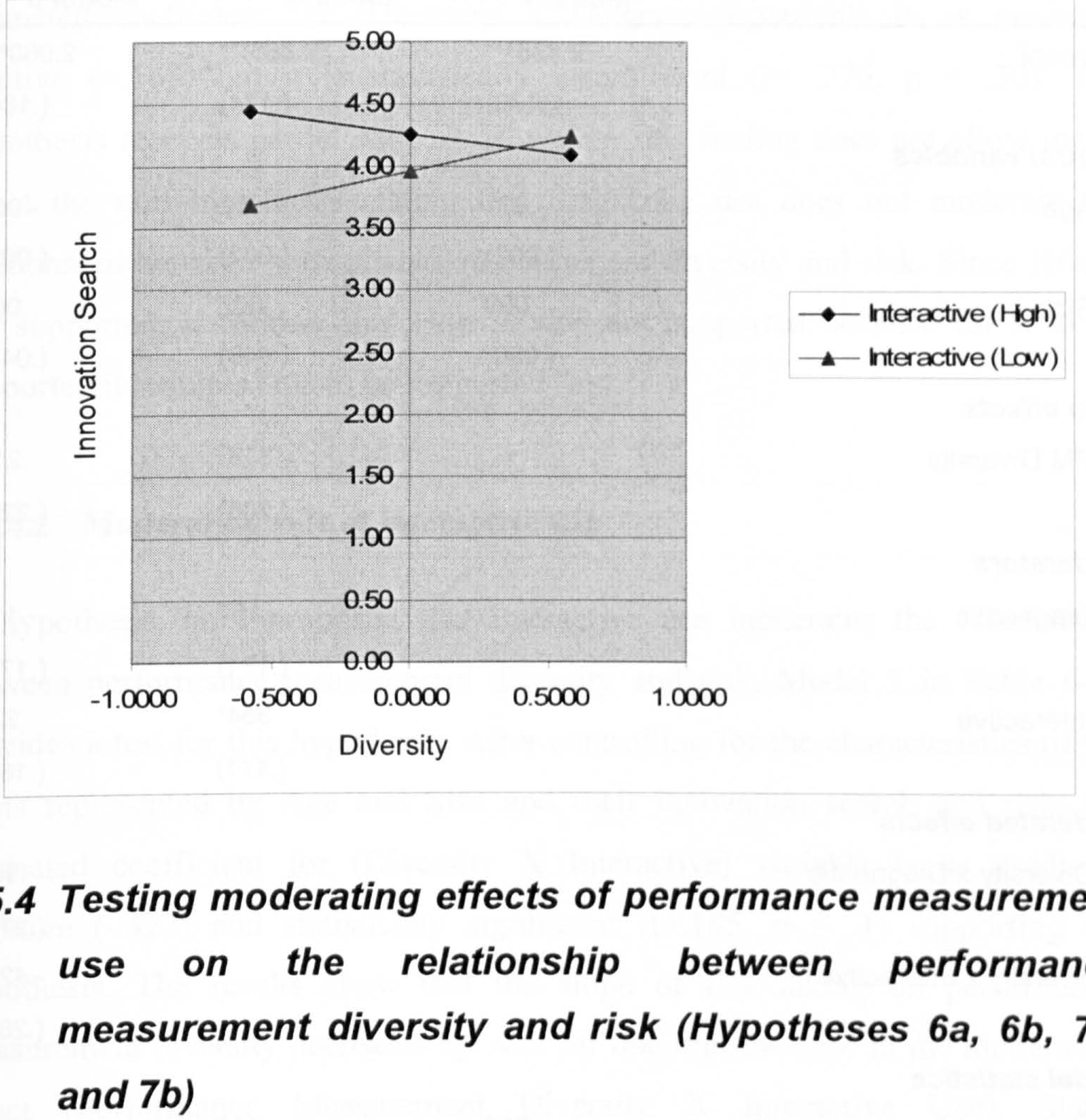
Innovation Search = 4.2705 - 0.2928 Diversity;

and when interactive use = 1 standard deviation below the mean (-.8551), the simple regression equation is given by:

Innovation Search = 3.9712 + 0.4718 (Diversity)

Now, for each of the two simple regression equations, we substitute three values for Diversity to plot the interaction. The values of Diversity correspond to the one standard deviation above the mean (0.5805), the mean (0.00), and one standard deviation below the mean (-0.5805). The resulting interaction plot shown in Figure 6-26 confirms hypothesis 5b.

Figure 6-26: Effect of interactive use and performance measurement diversity on innovation search with diagnostic use held constant



6.5.4 *Testing moderating effects of performance measurement use on the relationship between performance measurement diversity and risk (Hypotheses 6a, 6b, 7a, and 7b)*

Table 6-45 displays the results of the OLS hierarchical regression of the moderating effects of performance measurement use on the relationship between performance measurement diversity and risk. First, only the control variables (Age, Size) were entered into the regression analysis forming model 1. Second, the main effects (Diversity) and moderators (Diagnostic and Interactive) were added forming model 2. Third, the moderating effects (Diversity X Diagnostic and Diversity X Interactive) were added forming model 3. Model 3 will be used to test the moderated hypotheses because it contains the moderated effects.

Table 6-45: OLS regression results of the moderating effects of performance measurement use on the relationship between performance measurement diversity and risk.

	Risk		
	Model 1	**Model 2**	**Model 3**
Intercept	2.838***	2.866***	2.960***
	(.179)	(.174)	(.184)
Control variables			
Age	-.001	-.002	-.002
	(.002)	(.002)	(.002)
Size	.009	.007	.004
	(.050)	(.048)	(.048)
Main effects			
PM Diversity		.313⁺	.219
		(.206)	(.221)
Moderators			
Diagnostic		-.167	-.117
		(.171)	(.175)
Interactive		.354*	.227
		(.171)	(.185)
Moderated effects			
Diversity x Diagnostic			.168
			(.276)
Diversity x Interactive			-.422⁺
			(.262)
Model statistics			
R^2	.003	.096	..117
ΔR^2		.092**	.021
F-value	.242	2.935*	2.588*
N	145	145	145

Unstandardized regression coefficients are reported. Standard errors are in parentheses. Two-tailed tests for controls, one-tailed tests for hypothesized variables.
⁺ p ≤ .10; * p ≤ .05; ** p ≤ .01; *** p ≤ .001

6.5.4.1 Moderating role of diagnostic use

In hypothesis 6a, I proposed that diagnostic use moderates the relationship between performance measurement diversity and risk. Model 3 in Table 6-45 provides a test for this hypothesis. After controlling for the characteristics of the firms represented by Age and Size and their innovation search and risk, the estimated coefficient for (Diversity X Diagnostic) variable is as predicted positive (+.168) but it is statistically insignificant ($t= .276$, $p = .30$). The hypothesis receives partial support. However, this finding does not allow me to reject the null hypothesis stating that diagnostic use does not moderate the relationship between performance measurement diversity and risk. Since H6a is not supported, it follows that H6b is also not supported because for it to be supported it requires H6a to be supported first.

6.5.4.2 Moderating role of interactive use

In hypothesis 7a, I proposed that interactive use moderates the relationship between performance measurement diversity and risk. Model 3 in Table 6-45 provides a test for this hypothesis. After controlling for the characteristics of the firms represented by Age and Size and their innovation search and risk, the estimated coefficient for (Diversity X Interactive) variable is as predicted negative (-.422) and statistically significant ($t=.185$, $p \leq .1$) supporting the hypothesis. The results show that the slope of risk taking on performance measurement diversity decreases by .422 for one unit increase in the moderating effect (Performance Measurement Diversity X Interactive Use). Stated differently, the relationship between risk taking and performance measurement diversity is more negative for higher values of the moderating effect (Performance Measurement Diversity X Interactive Use).

To test hypothesis 7b, I performed partial derivative analysis and plotted the interaction holding diagnostic use constant. The equations used in the partial derivative analysis as the following:

Risk = 2.960 +.219 Diversity + .227 Interactive - .422 Diversity * Interactive

Risk = (.219 - .450 Interactive) Diversity + (4.120 +.227 Interactive)

∂(Risk)/∂(Diversity) = .219 - .422 Interactive

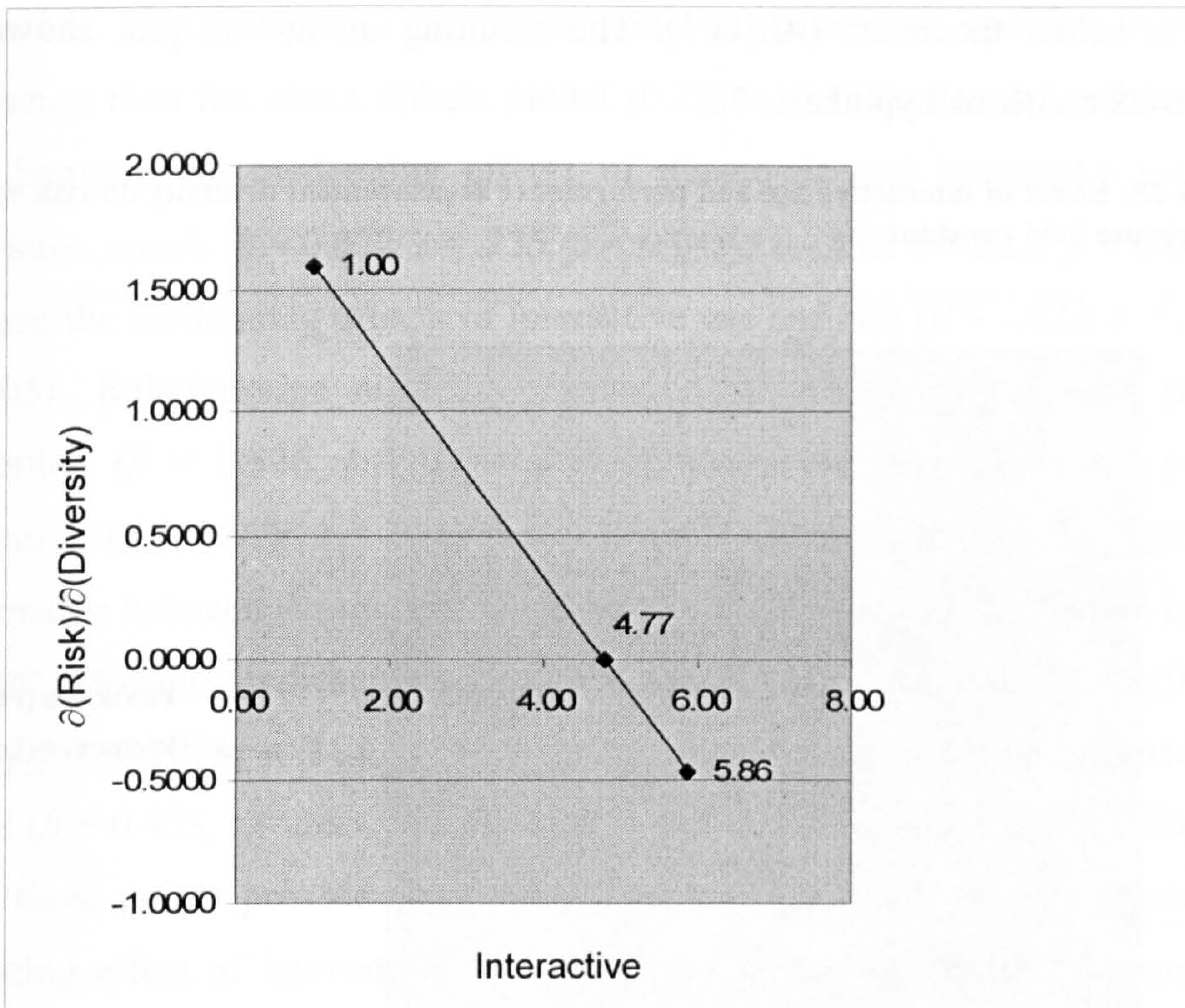
The inflection point (crossover point) is 4.77 as illustrated in Table 6-46. Since the inflection point is within the range of the interactive use (1.00 – 5.86). I could conclude that the moderating effect of diagnostic use is nonmonotonic.

Table 6-46: Location of inflection point for interactive use in moderating the relationship between performance measurement diversity and product innovation

Moderating variable	b_1	b_3	$-b_1/b_3$	Sample Mean	Inflection Point (sample mean - b_1/b_3)	Sample Range
Interactive	.219	-.422	.5190	4.2518	4.77	1.00 – 5.86

Figure 6-27 indicates that the inflection point is 4.77 and as interactive use increases the slope between product innovation and diversity becomes more negative (less positive).

Figure 6-27: Partial derivative of risk showing the inflection point of interactive use



To interpret the moderation, the interaction effects are plotted using the procedure prescribed by Aiken & West (1991). For two levels of the moderator (interactive use), simple regression lines are plotted. The two values of the interactive use are chosen at one standard deviation above the mean and at one standard deviation below the mean. Since the moderated regression equation is given by the following equation:

Risk = (.219 - .450 Interactive) Diversity + (4.120 +.227 Interactive)

Interactive)

It follows that when interactive use = 1 standard deviation above the mean (0.8551), the simple regression equation is given by:

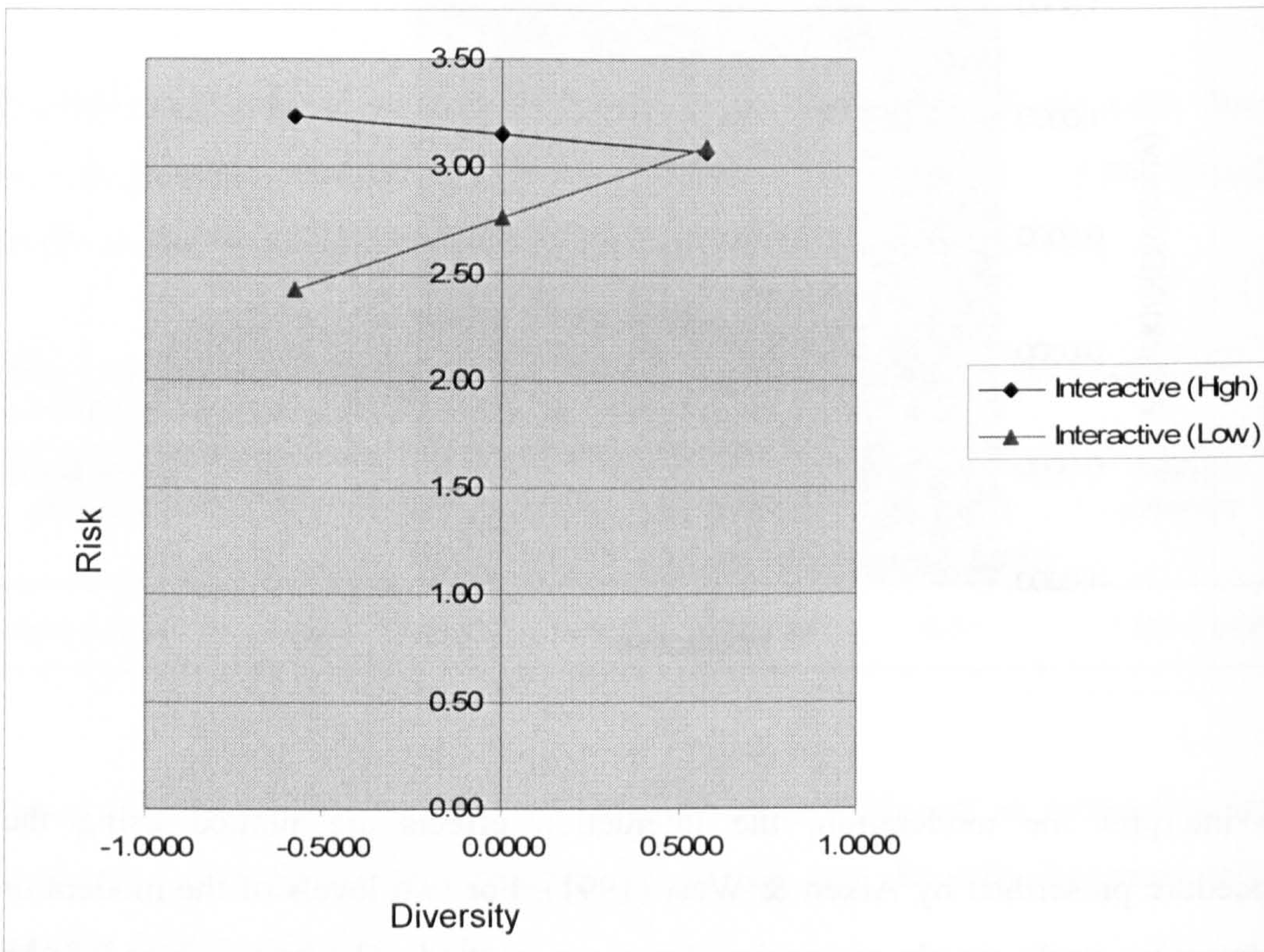
Risk = 3.1541 - 0.1419 Diversity;

and when interactive use = 1 standard deviation below the mean (-.8551), the simple regression equation is given by:

Risk = 2.7622 + .5799 Diversity

Now, for each of the two simple regression equations, we substitute three values for Diversity to plot the interaction. The values of Diversity correspond to the one standard deviation above the mean (0.5805), the mean (0.00), and one standard deviation below the mean (-0.5805). The resulting interaction plot shown in Figure 6-28 confirms hypothesis 7b.

Figure 6-28: Effect of interactive use and performance measurement diversity on risk with diagnostic use held constant



6.5.5 *Testing mediating effects innovation search and risk (Hypotheses 9 and 11)*

In testing the mediation hypotheses, I adopted the approach used by Singh, Goolsby, & Rhoads (1994) and Tippins & Sohi (2003). I checked for the presence of a mediating effect, by performing a competing model analysis (i.e., two substantive models are estimated and evaluated for significant differences).

Table 6-47 shows the results of the competing model analysis. The results confirm the postulated hypothesis that innovation search and risk mediate the relationship between the moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation. First, the partial mediation model explains more variance in product innovation than the direct effects model (0.28 vs. 0.138). Second, relationships exist between the moderating effects of diagnostic and interactive uses and innovation search ($\beta = 0.572$, $t = .235$, $p < 0.01$; $\beta = -.450$, $t = .223$, $p < 0.05$) and between the moderating effects of interactive use and risk ($\beta = -.422$, $t = .256$, $p < 0.05$). Relationships also exist between innovation search and product innovation ($\beta = 0.326$, $t = .064$, $p < 0.001$) and between risk and product innovation ($\beta = 0.072$, $t = .055$, $p < 0.1$). Third, the strength of the significant relationship between the moderating effects of diagnostic and interactive use and product innovation indicated in the direct effects model ($\beta = 0.651$, $t = .197$, $p < 0.001$; $\beta = -0.729$, $t = .188$, $p < 0.001$) becomes weaker in the partial mediation model ($\beta = 0.453$, $t = .183$, $p < 0.001$; $\beta = -0.552$, $t = .174$, $p < 0.001$). Together these three points provide compelling evidence that there exists a discernible mediating effect of innovation search and risk on the relationship between the moderating effects of performance measurement use on the relationship between performance measurement diversity and product innovation.

Table 6-47: Test results of the mediating effects of innovation search and risk

Parameter	Direct Effect Model	Partial Mediation Model
Control variables		
Age → Product Innovation	-.002^{+}	-.002^{+}
	(.001)	(.001)
Size → Product Innovation	-.034	-.031
	(.035)	(.031)
Age → Innovation Search		.000
		(.002)
Size → Innovation Search		-.010
		(.041)
Age → Risk		-.002
		(.002)
Size → Risk		.004
		(.047)
Main effects		
PM Diversity → Product Innovation	-.149	-.195^{+}
	(.158)	(.144)
PM Diversity → Innovation Search		.092
		(.188)
PM Diversity → Risk		.219
		(.216)
Mediators		
Search → Product Innovation		.326***
		(.064)
Risk → Product Innovation		.072^{+}
		(.055)
Moderators		
Diagnostic → Product Innovation	.105	.114
	(.125)	(.114)
Interactive → Product Innovation	-.040	-.114
	(.132)	(.121)
Diagnostic → Innovation Search		.000
		(.149)
Interactive → Innovation Search		.176
		(.157)
Diagnostic → Risk		-.117
		(.171)
Interactive → Risk		.227

Parameter	Direct Effect Model	Partial Mediation Model
		(.180)
Moderated effects		
Diversity x Diagnostic → Product Innovation	.651***	.453**
	(.197)	(.183)
Diversity x Interactive → Product Innovation	-.729***	-.552***
	(.188)	(.174)
Diversity x Diagnostic → Innovation Search		.572**
		(.235)
Diversity x Interactive → Innovation Search		-.450*
		(.223)
Diversity x Diagnostic → Risk		.168
		(.269)
Diversity x Interactive → Risk		-.422*
		(.256)
Model statistics		
R^2 (Product Innovation)	.138	.280
ΔR^2 (Product Innovation)	.142	
N	145	145

ML unstandardized path coefficients are reported. Standard errors are in parentheses. Two-tailed tests for controls, one-tailed tests for hypothesized variables. $^{+}$ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

6.5.6 Testing direct effects of performance measurement diversity, innovation search and risk on product innovation (Hypotheses 1, 8 and 10)

Table 6-48 displays the results of the OLS hierarchical regression of the direct effects of performance measurement diversity, innovation search, and risk taking on product innovation. First, only the control variables (Age, Size) were entered into the regression analysis forming model 1. Second, the main effects (performance measurement diversity, innovation search, and risk taking) were added forming model 2. Model 2 will be used to test the hypotheses.

Table 6-48: OLS regression results for the impact of performance measurement diversity, innovation search, and risk taking on product innovation

	Product Innovation	
Variable	**Model 1**	**Model 2**
Intercept	4.336***	2.786***
	(.131)	(.538)
Control variables		
Age	-.002⁺	-.002⁺
	(.001)	(.001)
Size	-.028	-.027
	(.036)	(.033)
Independent variables		
PM Diversity		-.036
		(.117)
Search		.356***
		(.067)
Risk		.089***
		(.058)
Model statistics		
R^2	.001	.26
ΔR^2		.207***
F-value	2.298	12.61825***
N	145	145

Unstandardized regression coefficients are reported. Standard errors are in parentheses. Two-tailed tests for controls, one-tailed tests for hypothesized variables.
⁺ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

6.5.6.1 Direct effect of performance measurement diversity on product innovation (Hypothesis 1)

In hypothesis 1, I proposed that there is a positive relationship between performance measurement diversity and product innovation. Model 2 in Table 6-48 provides a test for this hypothesis. After controlling for the characteristics of the firms (Age and Size), the estimated coefficient for performance measurement diversity is negative which is not as predicted (-.036) and statistically insignificant (.117; p = .38). Therefore, the hypothesis is not supported.

6.5.6.2 Direct effects of innovation search on product innovation (Hypothesis 8)

In hypothesis 8, I proposed that there is a positive relationship between innovation search and product innovation. Model 2 in Table 6-48 provides a test for this hypothesis. After controlling for the characteristics of the firms (Age and Size), the estimated coefficient for Innovation Search is positive as predicted (+.356) and statistically significant (.067, $p \leq .001$). Therefore, the hypothesis is supported. The results show that product innovation increases by .356 for every one unit increase in innovation search

6.5.6.3 Direct effects of organizational risk taking on product innovation (Hypothesis 10)

In hypothesis 10, I proposed that there is a positive relationship between organizational risk taking and product innovation. Model 2 in Table 6-48 provides a test for this hypothesis. After controlling for the characteristics of the firms represented: Age and Size, the estimated coefficient for organizational risk taking is positive as predicted (+.089) and statistically significant (.058, $p \leq .10$). Therefore, the hypothesis is supported. The results show that product innovation increases by .089 for every one-unit increase in risk taking.

6.5.7 Summary of the results of Hypothesis Testing

Table 6-49 shows that all the hypotheses are supported except for hypothesis 1 which postulated that performance measurement diversity has positive relationship with product innovation and hypotheses 6a and 6b which postulated that diagnostic use of performance measurement moderates the relationship between performance measurement diversity and organizational risk taking.

Table 6-49: summary of the results of the hypothesis testing

No.	Hypothesis	Results
H1	There is a positive relationship between performance measurement diversity and product innovation.	Not supported
H2a	Diagnostic use moderates the form of the relationship between performance measurement diversity and product innovation.	Supported
H2b	When levels of diagnostic use are high, performance measurement diversity will be positively related to product innovation.	Supported
H2c	When levels of diagnostic use are low, performance measurement diversity will be negatively related to product innovation.	Supported
H3a	Interactive use moderates the form of the relationship between performance measurement diversity and product innovation.	Supported
H3b	When levels of interactive use are high, performance measurement diversity will be more negatively related to product innovation.	Supported
H4a	Diagnostic use moderates the form of the relationship between performance measurement diversity and innovation search.	Supported
H4b	When levels of diagnostic use are high, performance measurement diversity will be positively related to innovation search.	Supported
H4c	When levels of diagnostic use are low, performance measurement diversity will be negatively related to innovation search.	Supported
H5a	Interactive use moderates the form of the relationship between performance measurement diversity and innovation search.	Supported
H5b	When levels of interactive use are high, performance measurement diversity will be more negatively related to innovation search.	Supported
H6a	Diagnostic use moderates the form of the relationship between performance measurement diversity and organizational risk taking.	Not supported
H6b	When levels of diagnostic use are high, performance measurement diversity will be positively related to organizational risk taking.	Not supported
H7a	Interactive use moderates the form of the relationship between performance measurement diversity and organizational risk taking.	Supported
H7b	When levels of interactive use are high, performance measurement diversity will be more negatively related to organizational risk taking.	Supported
H8	There is a positive relationship between innovation search and product innovation.	Supported
H9	Innovation search mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.	Supported
H10	There is a positive relationship between organizational risk taking and product innovation.	Supported
H11	Organizational risk taking mediates the relationship between the moderating effect of performance measurement use on performance measurement diversity and product innovation.	Supported

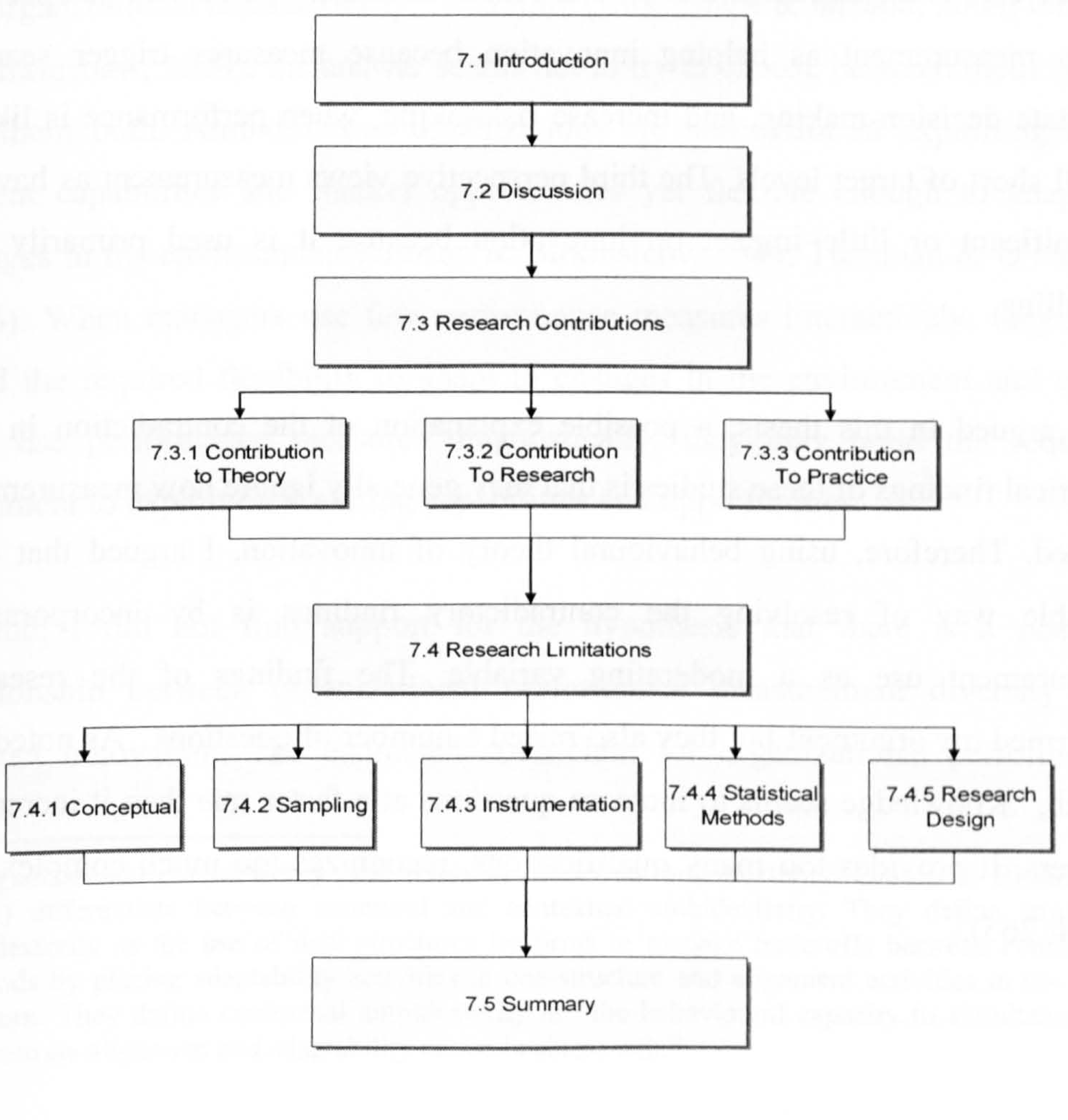
7 DISCUSSION, CONTRIBUTION, AND LIMITATIONS

7.1 Introduction

The previous chapter analysed the collected data and reported the results of the hypotheses testing. This chapter discusses the findings of the research reported in the previous chapter.

There are five sections in this chapter, as outlined in Figure 7-1. The first section introduces the chapter and subsequent sections. In the second section, I discuss the findings reported in the previous chapter. Next, I show how these findings contribute to theoretical, empirical, and practice literatures in several ways. The discussion then turns to the limitations of the research. In the final section, I summarize the chapter.

Figure 7-1: Outline of chapter 7



7.2 Discussion

This study set out to address the question of whether the effect of organizational performance measurement diversity on product innovation will differ depending on how performance measures are used. It was motivated by the strong empirical evidence showing that many companies who are successful today are less likely to be successful in the future because they fail to innovate. It was surprisingly then, that when everyone stresses the importance of innovation, there are many organizations adopting performance measurement systems, which may constrain their innovativeness.

As discussed in the second chapter, there are three differing perspectives on the effect of performance measurement on a firm's propensity to innovate. Moreover, each of these has empirical evidence to support its argument. The first perspective views measurement as constraining innovation because measures impede creativity, experimentation, and search in firms. The second perspective views measurement as helping innovation because measures trigger search, facilitate decision-making, and increase risk-taking, when performance is likely to fall short of target levels. The third perspective views measurement as having insignificant or little impact on innovation because it is used primarily for signalling.

As I argued in this thesis, a possible explanation of the contradiction in the empirical findings of these studies is that they generally ignore how measurement is used. Therefore, using behavioural theory of innovation, I argued that one possible way of resolving the contradictory findings is by incorporating measurement use as a moderating variable. The findings of the research confirmed my argument but they also raised a number of questions. As noted by March, "Knowledge seems to increase questions at a faster rate than it increases answers. It provides too many qualifications, recognizes too much complexity" (1994; 265).

First, an interesting finding of this study is that high levels of organizational performance measurement diversity coupled with high levels of diagnostic use resulted in innovation rates similar to those produced by lower levels of organizational performance measurement diversity coupled with higher levels of interactive use (See Figure 6-20 and Figure 6-24). On the one hand, given that the design, implementation and management of performance measures come at cost, these findings suggest that it may be more effective for firms to use fewer performance measures and to use them interactively as a learning lever to reduce strategic uncertainties. On the other hand, given that using performance measures interactively consume management time, these findings suggest that it may be more effective for firms to use more performance measures and to use them diagnostically as a control lever to reduce goal divergence. Based on the findings of this research, both configurations yield the same product innovation results at least in the short run but which of the two configurations a firm should choose. This research does not provide an answer to this question. However, drawing on the organizational ambidexterity[88] literature (Birkinshaw & Gibson, 2004; Gibson & Birkinshaw, 2004), the answer seems not to try to choose between them but to use them both. Ambidextrous organizations are successful in exploiting their current capabilities and market opportunities yet flexible enough to adapt to changes in the environment (Gibson & Birkinshaw, 2004; Tushman & O'Reilly, 1996). When managers use few performance measures interactively, they may build the required flexibility to adapt to changes in the environment and when they use performance measures diagnostically, they may build the required alignment to exploit their exiting capabilities and opportunities.

Second, I did not find support for the hypothesis that there is a positive relationship between organizational performance measurement diversity and product innovation. The estimated coefficient for organizational performance

[88] Organizational ambidexterity here refers to contextual ambidexterity. Gibson & Birkinshaw (2004) differentiate between structural and contextual ambidexterity. They define structural ambidexterity as the use of dual structures by firms to manage trade-offs between conflicting demands by placing adaptability activities in one-structure and alignment activities in the other structure. They define contextual ambidexterity as "the behavioural capacity to simultaneously demonstrate alignment and adaptability across business unit."

measurement diversity was negative which is not as predicted (B = -.036) and statistically insignificant (t = .117; p = .38). However, this finding may not be surprising if we consider the challenges in designing and managing non-financial measures. Neely (2006b) draws our attention to a number of these challenges: (1) increased organizational performance measurement diversity may lead to data overload; (2) non-financial measures are more difficult to design, measure, and collect compared to financial measures; and (3) if non-financial measures are not clearly defined, they may lead organizations to end up measuring the same concept in multiple different ways.

Third, I did not find support for the hypothesis that diagnostic use moderates the form of the relationship between organizational performance measurement diversity and organizational risk taking. The estimated coefficient for the interaction was positive as predicted (B = +.168) but it was statistically insignificant (t= .276, p = .30). This result may be due to measurement error because the risk construct was measured with one item, which is noted in the research limitations section. Another possible explanation for not finding a statistically significant interaction may be because the moderating effect was formed as a simple product term, which is only one of many possible functional forms of moderation. Therefore, a failure to obtain a statistically significant interaction may reflect the presence of an alternative form of interaction rather than the absence of a moderated relationship.

7.3 Contribution

By developing and testing a product innovation research model based on the behavioural theory of innovation that incorporates both performance measurement use and organizational performance measurement diversity in the British manufacturing context, this study contributes to existing theoretical, empirical, and practice literatures in several ways.

7.3.1 Contribution to Theory

This thesis has implications for behavioural theory of innovation (Greve, 2003b), strategic control theory (Simons, 1995), and theory of dynamic capabilities (Teece, Pisano, & Shuen, 1997).

7.3.1.1 Behavioural Theory of Innovation

The findings of this study confirm and extend behavioural theory of innovation in five important ways.

First, by making a distinction between the performance measures that are used diagnostically and those that are used interactively, I extend the behavioural theory of innovation by demonstrating that the effect of performance measures on the firm's propensity to innovate is contingent on how they are used. In its current formulation, behavioural theory of innovation would seem to apply to those performance measures that are used diagnostically. This distinction is important because it increases the explanatory and prediction power of behavioural theory of innovation in predicting when innovations are likely to occur.

Second, Greve (2003b) used the single financial measure of return on assets (ROA) as the performance variable[89]; this study extends the behavioural theory of innovation by examining the impact of multiple financial and non-financial performance measures on product innovation. The findings of the study suggest that the increase in multiple performance measures when used diagnostically will produce higher rates of product innovations, which provides an answer to Greve's question "[h]ow is innovation rate affected by the proliferation of profit centers and performance measures seen in large modern corporations?" and confirms his assertion that "[m]ore frequent and more specialized performance measures should produce greater variability in performance, which increases the probability that a given organizational unit will be taking risks" (2003b: 698).

[89] He also used return on sales (ROS) in a separate analysis and he found results consistent with the analysis that was reported using ROA.

Third, by including organizational risk taking explicitly as an explanatory variable in the product innovation model, I am able to test the postulated positive relationship between risk taking and product innovation in the theory directly. The findings of the study confirmed the claim that organizational risk taking is positively related to product innovation.

Fourth, by studying product innovation as an outcome variable, this study confirms Greve's (2003b) finding of organizational learning from performance feedback using a different research design – cross-sectional, probability sample, survey. This form of method triangulation adds to our confidence in the theory's ability to predict product innovation.

Fifth, given that, Greve (2003b) applied the behavioural theory of innovation to the Japanese shipbuilding industry, by applying it to the UK manufacturing sector; I increase the external validity of the theory by confirming the finding of organizational learning from performance feedback using a different research context.

7.3.1.2 Strategic control theory

The findings of this study confirm and extend strategic control theory in three important ways.

First, the findings of this study extend strategic control theory by making a distinction between the different levels of use (high/low) within each of diagnostic and interactive uses. The findings of the study demonstrated that high and low levels of diagnostic use or interactive use have different impact on product innovation.

Second, the findings of this study extend strategic control theory by hypothesizing interaction effects between the different levels of performance measurement diversity and the different levels of performance measurement use (diagnostic and interactive). Simons stresses that "[t]wo types of decisions must be made by the designer of a performance measurement system. The first decisions are about *design features*. What types of information should be collected and with what frequency of feedback? Second, decisions must be made about how to *use* performance measurement systems. Who should receive the data and what should they do and not do with it?" (2000: 7). Interestingly, research studies so far have either examined the impact of some design elements such as performance measure diversity (Hoque, 2004; Hoque & James) or they have only examined the performance measurement use (Bisbe & Otley; Henri, 2006b), but not the interaction between them. Furthermore, Simons (1995; 2000) seems to imply that this interaction is monotonic. This study shows that the interaction between diagnostic or interactive use with performance measurement is nonmontonic.

Third, the findings of the study confirm the existence of the two styles of performance measurement use as advanced by Simons (1995).

7.3.1.3 Theory of dynamic capabilities

This study also adds to the recent discussion of the theory of dynamic capabilities. Building on the concepts of standard operating procedures (Cyert & March, 1963) and evolutionary perspective of organizational routines (Nelson & Winter, 1982), Teece, Pisano, & Shuenand (1997: 516) define dynamic capabilities as "the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments". Eisenhardt & Martin (2000) offer a more specific and operational definition by defining dynamic capabilities as "[t]he firm's processes that use resources – specifically the processes to integrate, reconfigure, gain and release resources – to match and even create market change. Dynamic capabilities are thus the organizational and strategic routines by which active firms achieve new resource configurations as markets emerge, collide, split, evolve, and die" (2000: 1107). Furthermore, they argue that since the functionality of dynamic capabilities can be duplicated across firms, their value for competitive advantage lies in the resource configurations they create, not in the capabilities themselves. This study expands the theory of dynamic capabilities of firms by thoroughly examining one such capability: performance measurement and its impact on reconfiguring the product portfolio of the firm through the introduction of new products (Smith & Tushman, 2005; Tushman, Smith, Wood, Westerman, & O'Reilly, 2004). The results of this study indicate that firms differ in the way they use their performance measurement systems and these differences may provide an explanation of why the same organizational performance measurement system can produce differentiated product innovation outcomes.

7.3.2 Contribution to Research

The study contributes to empirical research in several ways.

First, although previous literatures on organizational performance measurement and innovation have highlighted the importance of performance measurement systems in enabling or hindering innovation in firms, very few studies have actually tested the impact of organizational performance measurement systems on product innovation at the firm level and recent studies (Bisbe & Otley, 2004; Davila, 2000) produced mixed results. I argued in this study that a possible explanation of the contradiction in the empirical findings of these studies is that they generally ignore how performance measurement systems are used "style of use". Therefore, I argued that one possible way of resolving the contradictory findings of these perspectives is by incorporating performance measurement use (diagnostic and interactive) as a moderating variable. The findings of this study lend support to my argument and seem to provide and explanation to the contradictory findings of the earlier empirical studies. I found that the impact of performance measurement on product innovation could be either positive or negative, depending upon the way performance measures are used.

Second, the empirical studies that have examined the role of performance measurement use on product innovation have examined only the interactive use.

This study extends earlier empirical studies by examining the role of not only interactive use but also the role of diagnostic use on product innovation. Different uses have different impacts on product innovation. My findings suggest that the extent to which a firm offers new products will be more positively (negatively) associated with performance measurement diversity when diagnostic use is high (low) holding interactive use constant and will be more negatively (positively) associated with performance measurement diversity when interactive use is high (low) holding diagnostic use constant.

Third, this study addresses the lack of empirical research examining the impact of performance measurement practices on product innovation in the UK manufacturing sector. The findings of this study suggest that performance measurement practices are one of the determinants of product innovation in the UK manufacturing sector.

7.3.3 Contribution to Practice

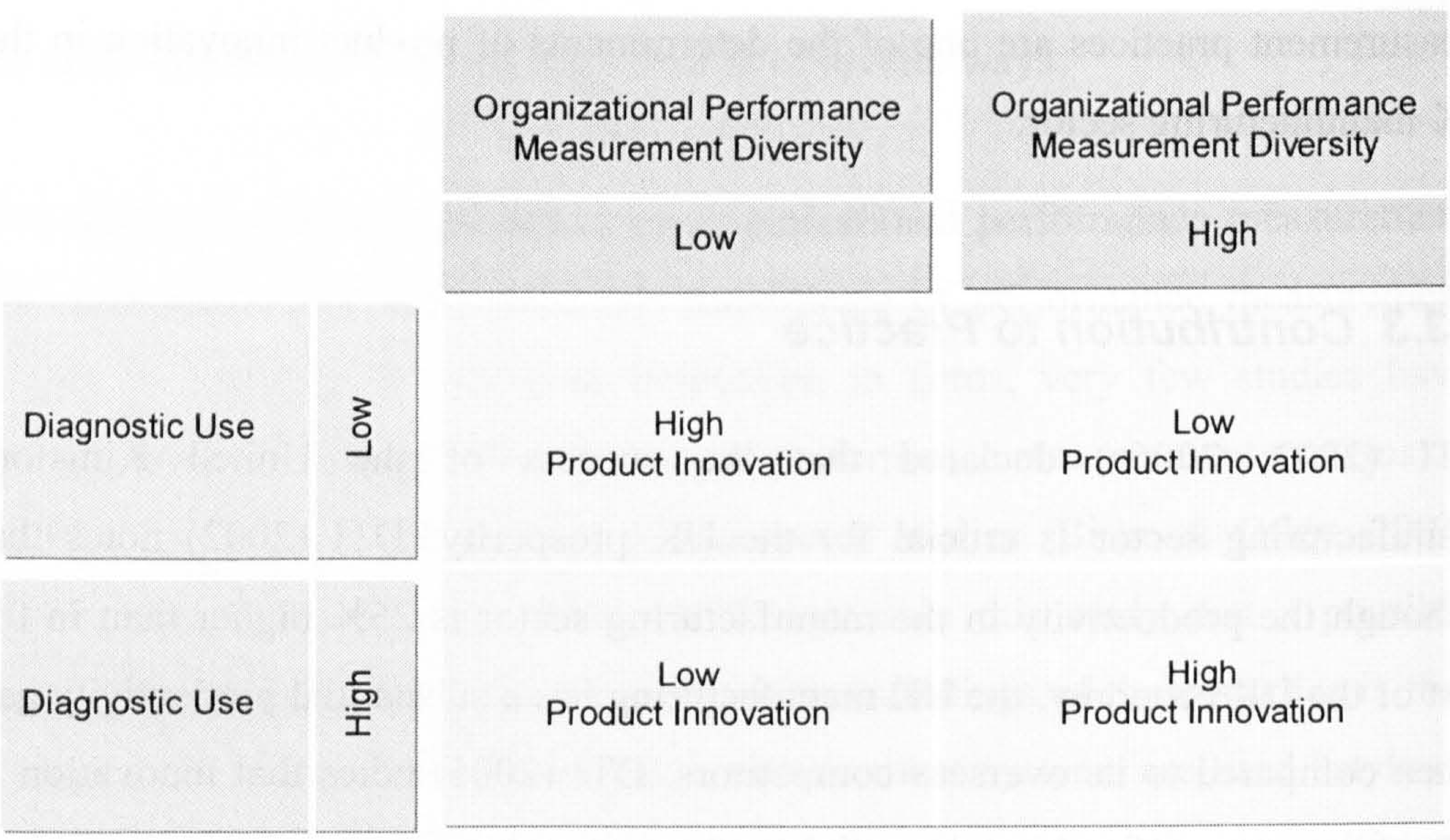
DTI (2002; 2006a) declared that the success of the United Kingdom manufacturing sector is crucial for the UK prosperity. DTI (2002) notes that although the productivity in the manufacturing sector is 25% higher than in the rest of the UK economy, the UK manufacturing has a substantial productivity gap when compared to its overseas competitors. DTI (2005) notes that innovation is one of the reasons for this substantial productivity gap.

By advancing a theoretically grounded and empirically tested product innovation model that enables managers to increase their understanding of how best they could use their performance measurement systems to increase the innovativeness of their firms, this study contributes to the innovation and performance measurement practice literatures. The study clearly demonstrates that performance measurement systems could be used to drive product innovations in manufacturing firms.

My findings provide insight as to how firms might become more innovative by properly designing and using their performance measurement systems. My results showed that the relationship between organizational performance measurement diversity and product innovations can be either positive or negative, depending upon the way the performance measurement system is used.

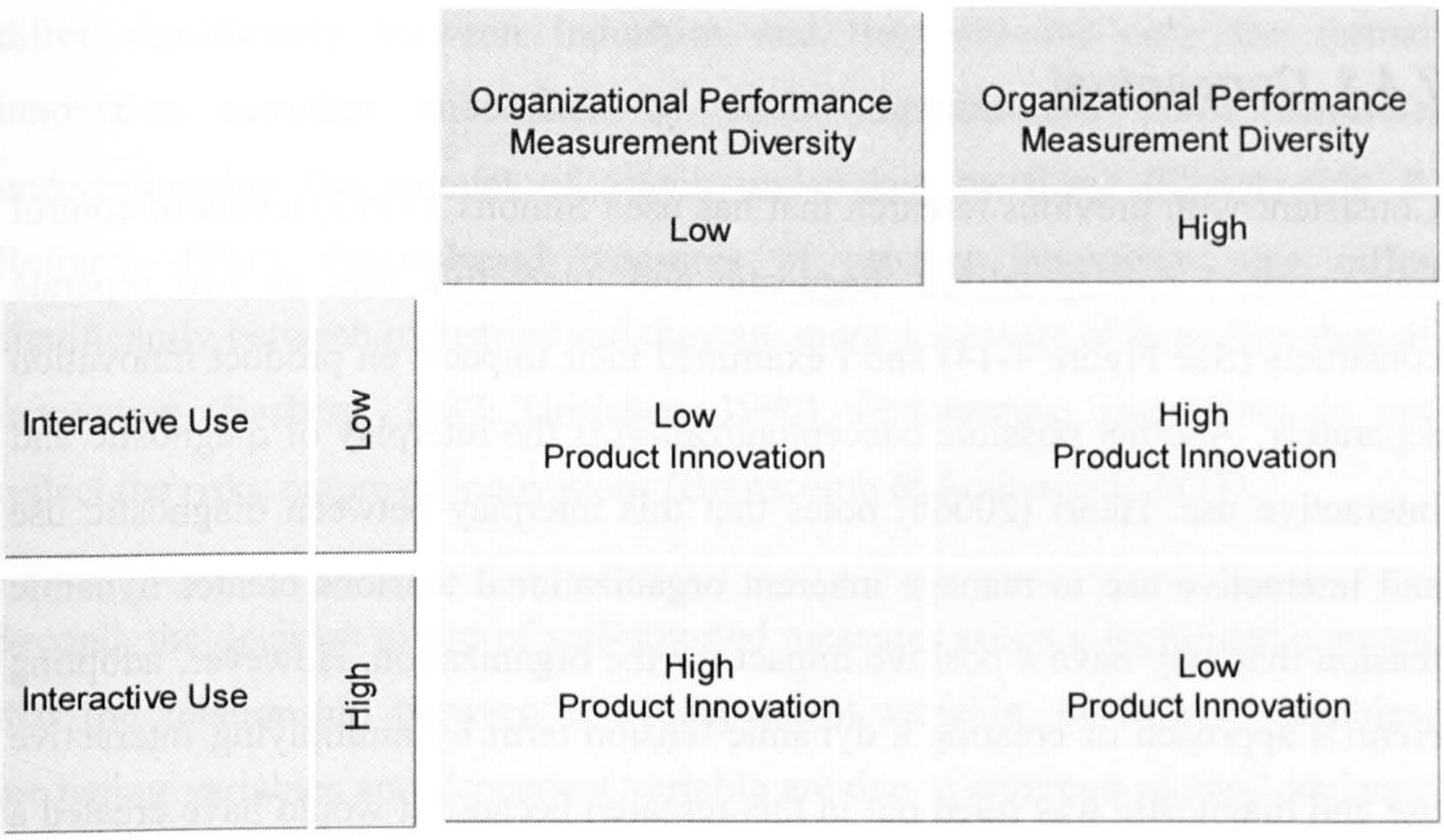
As shown in Figure 7-2, the findings of this study suggest firms using their performance measurement system diagnostically need to have diverse performance measures to achieve high product innovation rates.

Figure 7-2: Relationship between performance measurement diversity, diagnostic use, and product innovation



As shown in Figure 7-3, the findings of this study also suggest firms using their performance measurement system interactively need to have fewer performance measures to achieve high product innovation rates.

Figure 7-3: Relationship between organizational performance measurement, interactive use, and product innovation



My findings also encourage managers to think more broadly about their performance measurement systems. Given that, firms are making significant investments in their performance measurement systems and that they appear to pay less attention to the way they should use the data generated from these systems, the findings of this research should encourage mangers to invest in data analysis skills, processes, and infrastructure in their organizations (Neely, 2006a; Neely & Al Najjar, 2006).

Finally, the findings of this study suggest that using performance measures for control and learning purposes are conducive to innovation. Managers should possess the skills for managing performance measures diagnostically (control) and interactively (learning).

7.4 Research Limitations

Research designs entails making trade-off between research objectives of generalizability, accuracy, and simplicity (Weick, 1979). Therefore, all research designs are subject to limitations[90]. This research is no exception. It has several limitations.

7.4.1 Conceptual

Consistent with previous research that has used Simons' (1995) levers of control framework, I conceptualized diagnostic and interactive uses as two separate constructs (See Figure 4-14) and I examined their impacts on product innovation separately. Another possible conceptualization is the interplay of diagnostic and interactive use. Henri (2006b) notes that this interplay between diagnostic use and interactive use to manage inherent organizational tensions creates dynamic tension that may have a positive impact for the organization. However, adopting Henri's approach of creating a dynamic tension term by multiplying interactive use and diagnostic was ruled out in this research because it would have created a three-way interaction (diagnostic X interactive X performance measurement diversity) which is very hard to detect given the sample size of 145 (Aguinis, 1995; Aguinis & Stone-Romero 1997).

7.4.2 Sampling

Although the UK setting of our study is a strength given the lack of prior research in this context, I had no comparison country. This may limit the generalizability of the findings of the study. Future studies that replicate our model using multiple countries may enhance external validity.

7.4.3 Instrumentation

First, the decision to use subjective measures does not allow me to rule out the effects of short-comings associated with subjectivity. However, this measurement approach was deemed appropriate here because the constructs understudy here are inherently difficult to measure objectively. Of the constructs used in this study, two constructs could have been measured objectively- innovation search

[90] On noting the difficulties in making between trade-offs, Brinberg & McGrath (1985: 13) explains that "[v]alidity ... is a concept designating an ideal state – to be pursued, but not to be attained."

and product innovation. However, the objectives measures that are widely used to measure them were deemed inappropriate in this study. R&D-based measures differ significantly between industries and they account only for formal innovation activities undertaken by R&D organizational units, therefore underestimating the amount of search undertaken by firms (Kleinknecht & Reijnen, 1991). Patent-based measures of product innovation also differ significantly between industries and they are more a measure of invention than of innovation (Basberg, 1987; Griliches, 1990). Furthermore, inventions do not reflect the risky nature of innovations (Branscomb & Auerswald, 2001).

Second, the decision to use of self-reported measures raises a legitimate concern that the relationship between the independent variable, moderator variables, mediating variables and dependent variable are due to common method variance (CMV). CMV is present when the correlations between the measures are not due to substantive relationships between them but because they were provided by the same respondent (Podsakoff & Organ, 1986). As explained earlier, I assessed the existence of common method variance by conducting Harman's one factor test (Podsakoff & Organ, 1986). The results suggested common method variance does not exist in this study. In addition, the likelihood of common method variance is low in this study because the interactions in which my hypotheses were based on are relatively insensitive to the distortions due to mono-method variance (Aiken & West, 1991; Dooley & Fryxell, 1999; Evans, 1985). Moreover, I used several key informants strategies as outlined by Huber & Power (1985) and Mitchell (1994) to reduce the impact of CMV. For example, I selected members of top management teams as my key informants because they are usually the most knowledgeable, I promised key informants anonymity and confidentiality, I restricted the recall period to the previous two years, and I conveyed a sense of neutrality in the questions I asked by avoiding "agree/disagree" anchors and using anchors such as "To a very great extent/Not at all" (Dillman, 1978). These strategies seem to have worked.

Third, although I adapted my measure of risk taking from past research (Singh, 1986), the reliability of this construct was below the common threshold of .70 which forced me to measure risk taking using a single item. Therefore, related findings should be interpreted with some caution. Given that the scale was validated in the US context, one possible reason for the low reliability is that some of the wordings of the items such as "undo-the-competitors" may be more common in the US than in the UK.

7.4.4 Statistical Methods

First, I used maximum likelihood (ML) estimated confirmatory factor analysis (CFA) to evaluate the fit of the measurement model of the scales, which is not without weaknesses. Bagozzi et al. (1991) note two weaknesses associated with the use of CFA arising from two inherent assumptions behind the CFA model. First, the error term of the observed measures is assumed to contain two components: (1) measurement error analogous to random error in classical test-score theory and (2) specific true-score variance that is unique and different from the variance explained by traits and methods. However, CFA cannot separate these two components of the error term. Second, CFA models assume that the variation in observable measures is a linear combination of these sources of variance (i.e., traits, methods, and errors). The first weakness may not be a limitation in this research because the scales exhibited high reliabilities. The second weakness is a potential limitation of this research because I assessed the scales using a single method, which prevents me from assessing the linearity assumption. However, this limitation is consistent with other research that uses CFA to evaluate the fit of the measurement models.

Second, moderating (interaction) effects were created as a product of performance measurement use (moderator) and organizational performance measurement diversity (independent variable) which is widely used form in social research. However, as Jaccard & Turrisi (2003) note this simple product term is only one of many possible functional forms of the interaction effect and it is called bilinear interaction. This form of interaction indicates that the slope between dependent variable and independent variable changes as a linear function of scores on moderator variable. Therefore, a failure to obtain a statistically significant interaction may reflect the presence of an alternative form of interaction rather than the absence of a moderated relationship.

Third, structural equation modelling (SEM) offers several advantages over OLS regression and path analysis. SEM allows simultaneous estimation of multiple equation that represent the way constructs relate to observed indicators (Measurement model) as well as the way constructs relate to each other (structural model). One of the main advantages of SEM is its ability to model measurement error. However, SEM is a sensitive to small sample sizes. Although SEM methodologists do not agree on a formula for calculating the proper sample size, there is some consensus on having five cases per free parameter estimated (Kline, 2003). Given that, I had a sample size of 145; it precluded me from using SEM. However, given the high reliabilities of the scales, OLS regression and path analysis are deemed appropriate to test the hypotheses of this research (Hair et al., 2006).

7.4.5 Research Design

First, a limitation of this study that is common to all cross-sectional survey designs is its inability to predict causal relationships. Because the data were cross-sectional, I know that there were associations between the variables in the study, but I cannot conclude that the relationships are causal.

Second, I used firm size as a proxy measure of slack. George (2005) argues that controlling for firm size acts as a control for slack. Given that organizational slack has been found to increase firms' propensity to innovate (Singh, 1986, Greve, 2003b, 2004) and serves as an alternative explanation, our confidence in the findings of this study would have been strengthened if organizational slack were explicitly modelled.

Third, Simons (1995: 97) stresses that the unit of analysis in his framework is the system by stating that "[i]t is important to understand that the unit of analysis for these ideas is the "system," not the degree of interaction between organizational participants", which is the approach I followed here. However, empirical evidence suggests that performance measures and performance categories are other possible units of analysis (See Figure 2-2). For example, Marginson (2002), in his in depth study of a British telecom company at different managerial levels, "found an imbalance in how top management used the resulting information. Certain measures would be prioritized at different periods in time, while others were merely 'noted.'... Telco's top managers would involve themselves personally in some measures, while treating others as 'error based' or management-by-exception controls." It is clear from Marginson's discussion that the unit of analysis was the performance measures and not the systems. This research would have been richer if it could have adopted a multi-level research design.

Fourth, I adopted a cross-sectional research design that did not capture the temporal dynamics of performance measurement use. Empirical evidence reported by Ramos (2003; 2004) on a case study in a ceramic company showed that the use of performance measurement system shifted from diagnostic to interactive over time.

7.5 Summary

The chapter sought to discuss the findings of the research in relation to the research hypotheses and questions. It was found that indeed performance measurement use moderates the relationship between organizational performance measurement diversity and product innovation as has been argued. This chapter has shown that the findings of the research contribute to theory, research and practice despite its limitations. The findings were found to contribute to behavioural theory of innovation, strategic control theory and theory of dynamic capabilities.

8 CONCLUSIONS

8.1 Introduction

The aim of the research reported in this thesis is to enhance our understanding of the impact of performance measurement on product innovation. More specifically, it set out to answer the following research question:

What is the moderating effect of performance measurement use on the relationship between organizational performance measurement diversity and product innovation?

This question was in turn translated into four sub-questions as follows.

1- What are the different styles of use of performance measurement systems?
2- To what extent does the diagnostic use of performance measurement systems moderate the relationship between organizational performance measurement diversity and product innovation?
3- To what extent does the interactive use of performance measurement systems moderate the relationship between organizational performance measurement diversity and product innovation?
4- What are the mechanisms the mediate the impact of the interaction between organizational performance measurement uses (diagnostic and interactive) and organizational performance measurement diversity on product innovation?

To answer the research questions, I started by defining the key constructs of the study in the second chapter. Organizational performance measurement diversity is defined by the extent to which top management teams measure information related to a broad set of financial and non-financial measures and product innovation refers to the market introduction of a new good or service or a significantly improved good or service.

To answer the first question, I undertook an extensive review of the performance measurement literature and I chose Simons' (1995) conceptualisation of diagnostic and interactive uses of performance measurement system. Simons' conceptualisations was chosen for two reasons: (1) it provides one of the most comprehensive treatments of the concept of performance measurement use and (2) there is a substantial empirical literature stream that has built on Simons' conceptualization. Diagnostic use is defined by the extent to which top management teams use performance measures to monitor organizational outcomes to correct deviations from preset standards of performance (targets) and interactive use is defined by the extent they involve themselves regularly and personally in decision activities of subordinates to focus their subordinates' search on the strategic uncertainties facing their organizations.

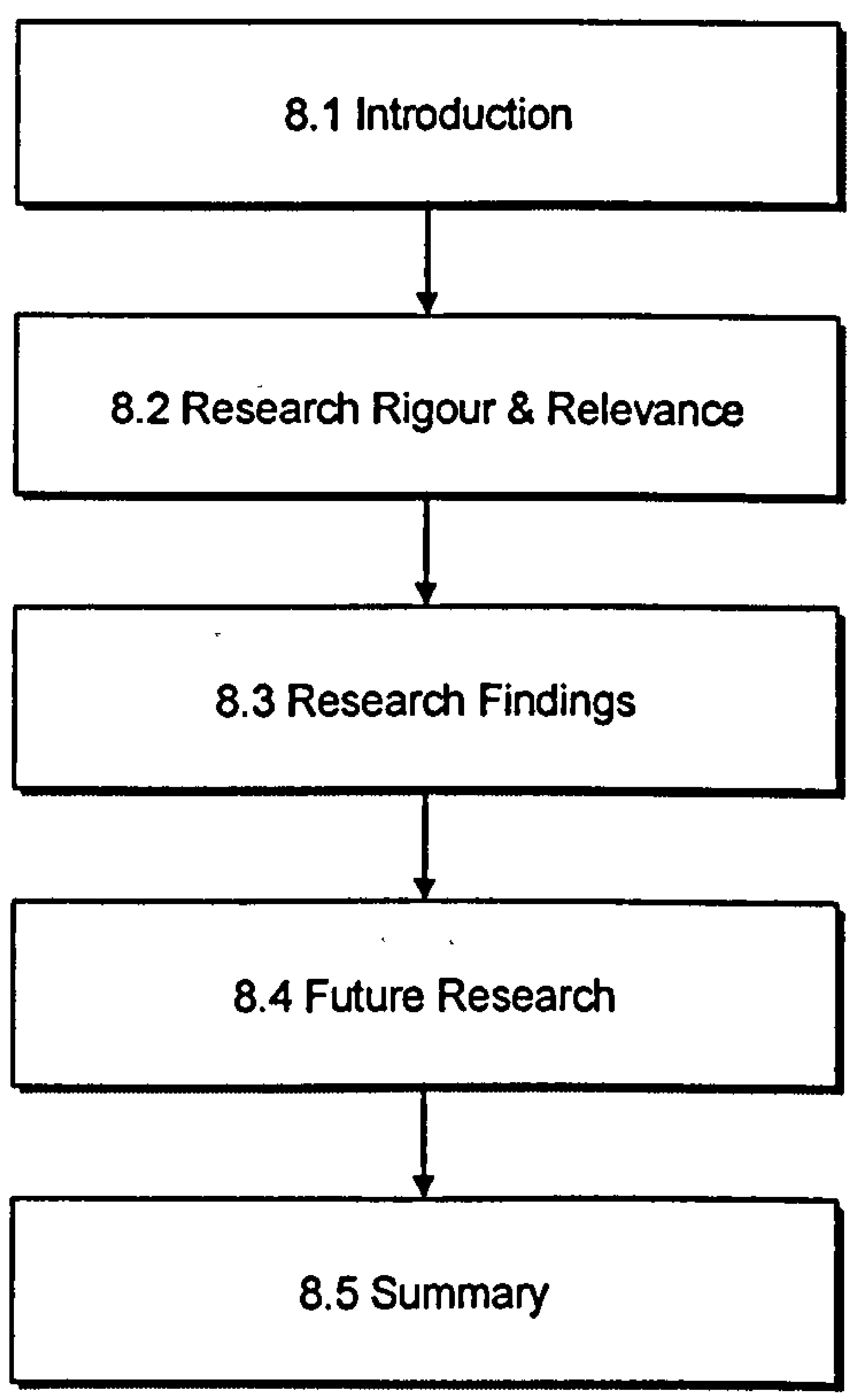
To answer the last three questions, I started by reviewing the literature that examined the impact of measurement on innovation in the second chapter. More specifically, I reviewed how measurement might constrain or help innovation by examining the theoretical and empirical literature from three perspectives: organizational contingency, creativity, and practice (rational) perspectives. In the third chapter, I developed a product innovation research model based on the behavioural theory of innovation (Greve, 2003) and eleven research hypotheses. These hypotheses postulated that performance measurement use moderates the relationship between organizational performance measurement diversity and product innovation through the mediating mechanisms of innovation search and organizational risk taking. In the fourth chapter, I explained the cross-sectional, probability sample survey research design used in this study and I operationalised the constructs using existing validated scales. I also specified the statistical models that would be used to test the research hypotheses. In the fifth chapter, I explained how the survey instruments were developed and administered. In the sixth chapter, I analysed the collected data and I reported the findings of the study, which confirmed that performance measurement use moderates the relationship between organizational performance measurement diversity and product innovation. Hence, I achieved the aim of the research.

In chapter seven, I discussed the findings of the study and I identified the contribution to theory, research, and practice. I also discussed the limitations of the research

In this chapter, I tie together those that have gone before by summarising the findings of the research and identifying areas that require further work.

There are five sections in this chapter, as outlined in Figure 8-1. The first section introduces the chapter and its subsequent sections. In the second section, I evaluate how this study meets the criteria for research rigour and relevance. In the third section, I summarize the findings reported in this study. Next, I identify areas that require additional work. In the final section, I conclude the thesis with a summary.

Figure 8-1: Outline of chapter 8



8.2 Research Rigour & Relevance

Following Vermeulen's (2005) recommendations for synthesizing relevance and rigour, I combined relevance and rigour in this research by formulating research questions that are of importance to reality while not making concessions in terms of rigour in developing theory and empirical evidence.

Using the criteria advanced by Shrivastava (1987) and Varadarajan (2003), I construct Table 8-1 to assess the conceptual and methodological rigour of my research. As shown in Table 8-1, this research met all the criteria for conceptual and methodological rigour.

Table 8-1: Evaluation of research rigour (Source: adapted from Shrivastava, 1987 and Varadarajan, 2003)

Property	Description	Met? (Y/N)	How It was Met
Conceptual Rigour	Is the research well grounded in a basic discipline?	Y	It is grounded in strategic management and operations management disciplines.
	Does the research use a conceptual framework consistent with existing theories in the field?	Y	See section 3.3, section 3.4 and Figure 3-5
	Does the research review the current literature?	Y	See chapter 2.
	Does the research pay attention to definitional issues?	Y	See section 3.2.
	Does the research use evidence to support position – conceptual reasoning underlying conceptual model and hypotheses?	Y	See chapter 3.
	Does the research objectively treat complementing and competing perspectives?	Y	See chapter 2 and section 3.5.
Methodological Rigour	Does the research use the appropriate research design to answer the research questions?	Y	See section 4.6, Table 4-7, and Table 4-8.
	Does the research pay attention to the measurement related issues – construct operationalisation, validity, and reliability?	Y	See section 4.6, Figure 6-3, and Table 6-34.
	Does the research use the appropriate data to empirically	Y	See section 4.7, section

Property	Description	Met? (Y/N)	How It was Met
	examine the research questions?		6.2, and section 6.3.
	Does the research use the appropriate methods of analysis/statistical procedures to empirically examine research questions?	Y	See section 4.8 and section 6.2.2.
	Does the research accurately thoroughly report the results and procedures leading to the results?	Y	See chapter 6 and Figure 6-1.
	Does the research report the reliability and validity of the empirical findings?	Y	See section 4.6, section 6.5, Figure 6-3, and Table 6-34.

Using the criteria advanced by Shrivastava (1987) and Thomas & Tymon (1982), I construct Table 8-2 to assess the practical usefulness of my research. As shown in Table 8-2, this research met all the criteria for practical usefulness.

Table 8-2: Evaluation of practical usefulness (Source: adapted from Shrivastava, 1987 and Thomas & Tymon, 1982)

Property	Description	Met? (Y/N)	How It was Met
Meaningfulness	Is the research project meaningful, understandable, and adequately describe problems faced by managers?	Y	See section 1.1 and section 7.3.3.
Goal Relevance	Does the research project contain performance indicators that are relevant to manager's goals?	Y	Product innovation is major outcome that managers would like to influence.
Operational Validity (Actionability)	Does the research project have clear action implications that can be implemented using the causal variables used in the research?	Y	See Figure 7-2 and Figure 7-3.
Innovativeness (Non-obviousness)	Does the research project transcends "common sense" solutions and provides non-obvious insights into practical problems?	Y	See section 7.3.3, Figure 7-2 and Figure 7-3.
Cost of	Are the solutions suggested by the research feasible to	Y	See section 7.3.3.

Property	Description	Met? (Y/N)	How It was Met
Implementation	implement in terms of their costs and time?		
Timeliness	Are the solutions suggested by the research available in time for managers to use them?	Y	See section 7.3.3.

8.3 Research Findings

The main outcome of this research has been to show that the impact of organizational performance measurement diversity on product innovation differs depending on the way performance measurement systems are used.

Using data from a cross-sectional, large-scale, probability sample survey of 145 UK manufacturing firms, I show that indeed organizational performance measurement diversity interact with performance measurement use to determine product innovation through the mediating mechanisms of innovation search and organizational risk taking. More specifically, the findings of the survey research could be summarized as follows based on the research sub-questions.

To what extent does the diagnostic use of performance measurement systems moderate the relationship between organizational performance measurement diversity and product innovation?

1- Diagnostic use moderates the form of the relationship between organizational performance measurement diversity and product innovation.

2- When levels of diagnostic use are high, organizational performance measurement diversity will be positively related to product innovation.

3- When levels of diagnostic use are low, organizational performance measurement diversity will be negatively related to product innovation.

4- Diagnostic use moderates the form of the relationship between organizational performance measurement diversity and innovation search.

5- When levels of diagnostic use are high, organizational performance measurement diversity will be positively related to innovation search.

6- When levels of diagnostic use are low, organizational performance measurement diversity will be negatively related to innovation search.

To what extent does the interactive use of performance measurement systems moderate the relationship between organizational performance measurement diversity and product innovation?

7- Interactive use moderates the form of the relationship between organizational performance measurement diversity and product innovation.

8- When levels of interactive use are high, organizational performance measurement diversity will be more negatively related to product innovation.

9- Interactive use moderates the form of the relationship between organizational performance measurement diversity and innovation search.

10- When levels of interactive use are high, organizational performance measurement diversity will be more negatively related to innovation search.

11- Interactive use moderates the form of the relationship between organizational performance measurement diversity and organizational risk taking.

12- When levels of interactive use are high, organizational performance measurement diversity will be more negatively related to organizational risk taking.

What are the mechanisms the mediate the impact of the interaction between performance measurement uses (diagnostic and interactive) and organizational performance measurement diversity on product innovation?

13- There is a positive relationship between innovation search and product innovation.

14- Innovation search mediates the relationship between the moderating effect of performance measurement use on organizational performance measurement diversity and product innovation.

15- There is a positive relationship between organizational risk taking and product innovation.

16- Organizational risk taking mediates the relationship between the moderating effect of performance measurement use on organizational performance measurement diversity and product innovation.

8.4 Future Research

The work reported in this thesis provides the foundation for a variety of research projects. These are summarized below.

1- In this study, I examined the impact of the interaction effects of performance measurement use and organizational performance measurement diversity on product innovation; future studies may examine their impact on other organizational dependent variables. Predicting other organizational behaviours that constitute change, preferably strategic, requiring some form of organizational search and risk taking increase our confidence of the findings of this study. Some of the strategic outcomes that could be investigated are capital investment, resource reconfiguration, market entry, strategic reorientation.

2- In this study, I examined product launches; future studies may take into account the differences between developing and launching innovations and study the impact of performance measurement practices on these two different organizational behaviours. For example, the ambidextrous theory of innovation (Duncan, 1976; Zaltman, Duncan, & Holebeck, 1973) distinguishes between two stages of innovation: initiation and implementation. The initiation stage involves the activities pertaining to the knowledge awareness of innovation, formation of attitudes toward innovation and decision to adopt the innovation. The implementation sage involves the activities pertaining to initial implementation and continued-sustained implementation. The ambidextrous theory of innovation postulates that the initiation stage requires an organization structure that is characterized by high complexity, lower formalization, and lower centralization whereas the implementation stage requires an organization structure that is characterized by low complexity, higher formalization, and higher centralization. It would seem that interactive use of performance measures would be more appropriate in the initiation stage and diagnostic use of performance measures will be more appropriate in the implementation stage.

3- In this study, I did not differentiate between the different types of research strategies; future studies may investigate the impact of performance measurement systems and practices on the different types of search strategies. Innovation search has many dimensions: scope, depth, age, science, and geography. Search scope (local versus distant) refers to the degree to which it entails the exploration of new knowledge (Katila & Ahuja, 2002). Local search refers to the firm's search for solutions in the neighborhood of its current expertise or knowledge (Helfat, 1994; Martin & Mitchell, 1998; Rosenkopf & Nerkar, 2001; Stuart and Podolny, 1996). Search depth refers to the degree to which it entails revisiting a firm's prior knowledge (Katila & Ahuja, 2002). The age dimension of the search addresses the question of how firms search over time (Katila, 2000). Science search refers to firm's search of the science base to overcome the limitations of their current technology base. Geographic search refers to the firm's search across geographical boundaries to expand their technology base and to solve local technological problems (Ahuja & Katila, 2004).

4- Future studies may investigate the impact of performance measurement practices on the different types of product innovations: incremental, non-incremental, technology-based, and market-based innovations. Given that, non-incremental (discontinuous) innovations are riskier solutions to incremental innovations and may entail different innovation search strategies, how well the model advanced in this study predict these innovations. How well also the model would predict technology and market-based innovations given that they have different risk profiles (See Figure 2-3).

5- In this study, I examined the moderating effects of diagnostic and interactive use on the relationship between organizational performance measurement diversity and product innovation separately; future studies could examine the moderating effect of the interplay between diagnostic and interactive use on the relationship between organizational performance measurement diversity and product innovation.

6- Future studies may investigate the moderating effects of top management team characteristics on the impact of performance measurement systems and practices on strategic outcome variables. The upper echelons theory (Hambrick & Mason, 1984) posits that the top management teams (TMT) play a pivotal role in shaping the future of their organizations. Furthermore, empirical evidence has shown that TMT characteristics have impact on the strategic direction, growth and performance of firms (Carpenter, Geletkanycz, & Sanders, 2004; Finklestien & Hambrick, 1996).

7- I employed a cross-sectional design in my study; future studies may implement longitudinal designs. More specifically, panel data designs in single industries with lagged variables will increase the internal validity of the research design and strengthen our confidence in the findings of this study. They will also allow us to make stronger inferences about causal relationships between variables.

8- In this study, I examined the impact of organizational performance measures on product innovation, future studies could study the impact of specialized innovation performance measures on the organizations' propensity to innovate. Davila, Epstein, & Matusik (2004) investigated how managers choose what measures to pay attention to in managing innovation and they found that managers use performance measures about specific phases of the innovation process (focused pattern) rather than use a balanced set of measures across the innovation process (balanced pattern). Their study could be extended by investigating which measures are used diagnostically, interactively or both.

9- In this study, I adopted a variance research design employing cross-sectional survey method that did not capture the multilevel nature of performance measurement systems and the temporal dynamics of performance measurement use; future studies may adopt a process research design employing longitudinal case studies that examine the design and use (diagnostic, interactive, dynamic) of individual performance measures, performance measurement categories, and performance measurement systems at multiple organizational levels (top, middle, front-line) and over an extended period of time to capture the temporal dynamics between the design and use dimensions of performance measures and organizational levels (See Figure 8-2 and Figure 8-3).

Figure 8-2: Performance measurement: Spatial and temporal dimensions

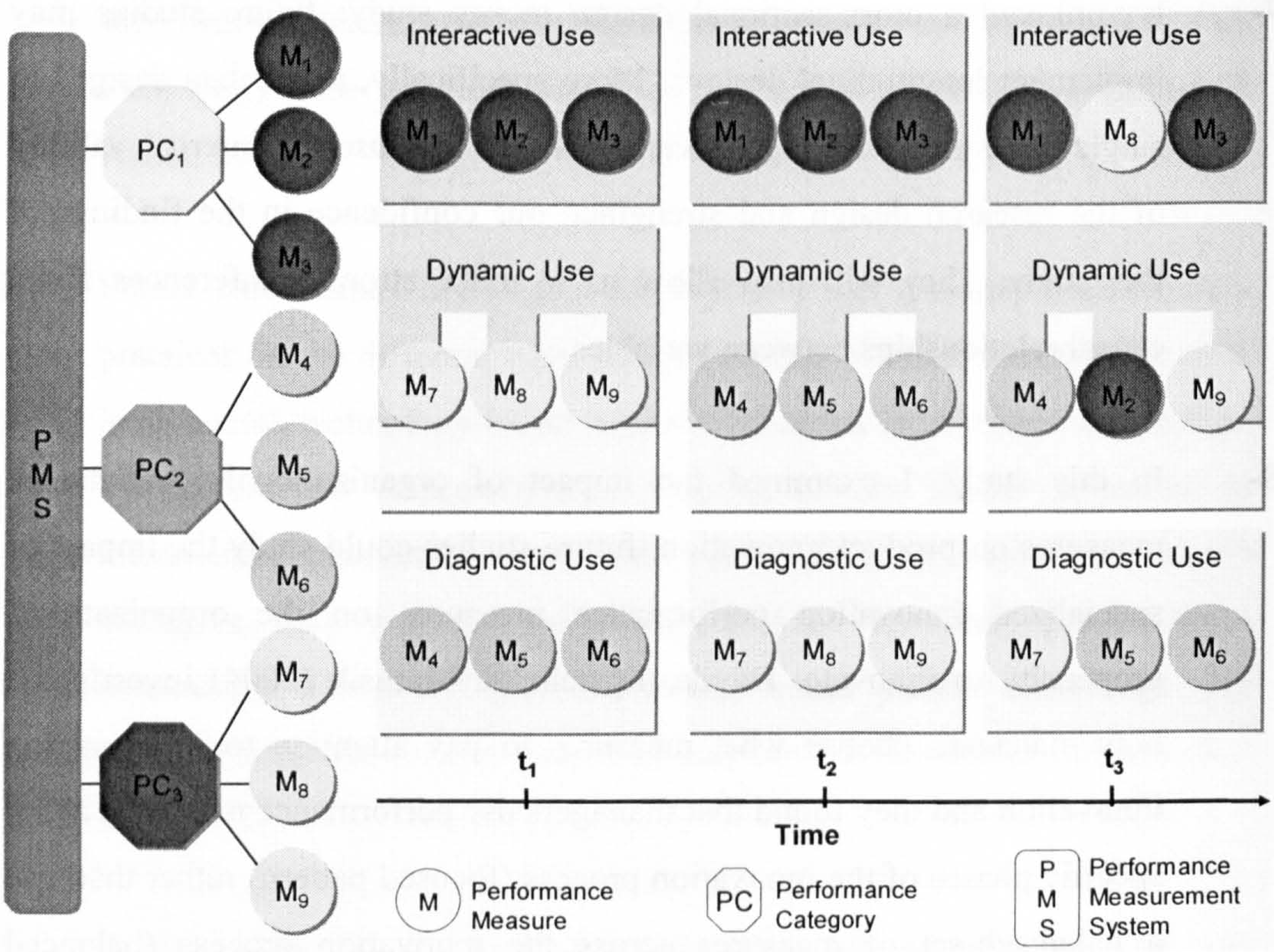
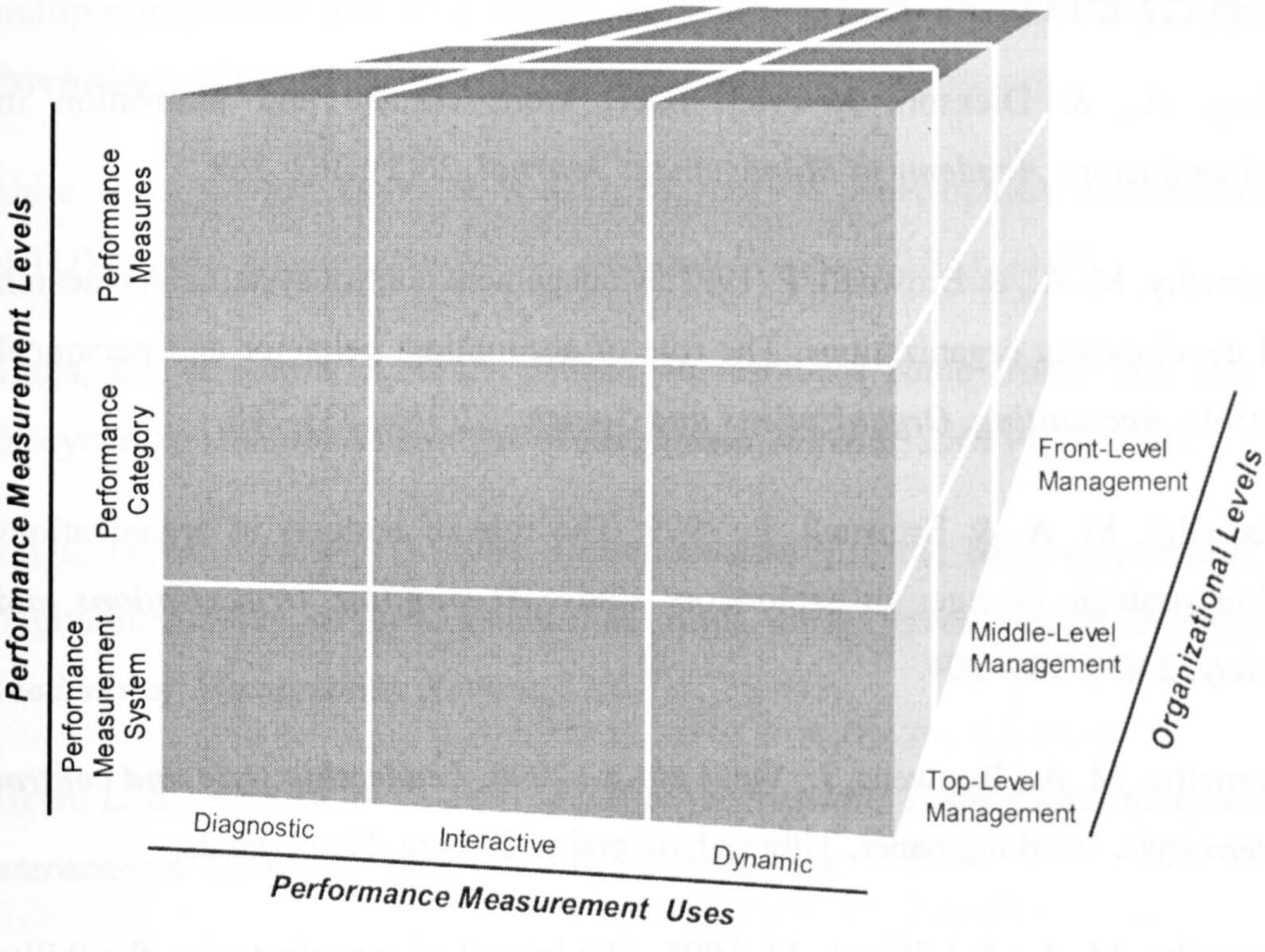


Figure 8-3: Performance measurement: Multilevel framework



8.5 Summary

The chapter sought to integrate and summarize the work presented in the rest of the thesis. In it both the research findings and areas requiring further work has been discussed.

The objective of this research was to show that the impact of organizational performance measurement diversity on the firm's propensity to innovate depends on how performance measures are used. This chapter has shown that not only this objective has been achieved, but also that the research reported in this thesis opens a potentially rich seam of future research topics.

9 REFERENCES

Abbey, A., & Dickson, J. 1983. R&D work climate and innovation in semiconductors. ***Academy of Management Journal,*** 26(2): 362–368.

Abernethy, M. A., & Brownell, P. 1997. Management control systems in research and development organizations: The role of accounting, behavior and personnel controls. ***Accounting, Organizations and Society***, 22(3/4), 233-248.

Abernethy, M. A. & Brownell, P. 1999. The role of budgets in organizations facing strategic change: an exploratory study, ***Accounting, Organizations and Society,*** 24(3): 189–204.

Abernethy, M. A., Bouwens, J., Van Lent, L. 2006. ***Leadership style and control system style.*** Working paper, Tilburg University, Tilburg, Netherlands.

Abernethy, M. A., & Lillis, A. M. 1995. The impact of manufacturing flexibility on management control system design. ***Accounting, Organizations and Society,*** 20(4): 241–258.

Abernethy, M. A., & Stoelwinder, J. U. 1991. Budget use, task uncertainty, system goal orientation and subunit performance: A test of the fit hypothesis in not-for-profit hospitals. ***Accounting Organizations and Society,*** 16(2): 105–120.

Academy of Management, 2006. ***Academy of management code of ethical conduct.*** Academy of Management Journal.

Adams, R., Bessant, J., & Phelps, R. 2006. Innovation management measurement. ***International Journal of Management Reviews,*** 8(1): 21-47.

Afuah, A. 2003. ***Innovation management.*** New York: Oxford University Press.

Aguinis H. 1995. Statistical power problems with moderated multiple regression in management research. ***Journal of Management,*** 12(6): 1141-1158.

Aguinis, H., & Stone-Romero, E. F. 1997. Methodological artifacts in moderated multiple regression and their effects on statistical power. ***Journal of Applied Psychology,*** 82(1): 192-206.

Ahire, S. L., Golhar, D. Y., & Walker, M. A. 1996. Development and validation of TQM implementation constructs. ***Decision Sciences,*** 27(1): 23-56.

Ahuja, G., & Katila, R. 2004. Where do resources come from? The role of idiosyncratic situations. ***Strategic Management Journal,*** 25(8/9): 887-907.

Ahuja, G., & Lampert, C. 2001. Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. ***Academy of Management Review,*** 22(6/7): 521-543.

Aiken, L. S., & West, S. G. 1991. ***Multiple regression: Testing and interpreting interactions.*** Thousand Oaks, CA: Sage Publications.

Aldrich, H. E., & Auster, E. 1986. Even dwarfs started small: Liabilities of age and size and their strategic implications. In L. L. Cummings & B. B. Staw (Eds.), ***Research in organizational behavior.*** Greenwich, CT: JAI Press.

Aldrich, H. E. 2001. Who wants to be an evolutionary theorist: Remarks on the occasion of the Year 2000 OMT Distinguished Scholarly Career Award Presentation. ***Journal of Management Inquiry,*** 10(2): 115-127.

Allen, M. J., & Yen, W. M. 2002. ***Introduction to measurement theory.*** Long Grove, IL: Waveland Press.

Amabile, T. M. 1983a. ***The social psychology of creativity.*** New York: Springer-Verlag.

Amabile, T. M. 1983b. Social psychology of creativity: A componential conceptualization. ***Journal of Personality and Social Psychology,*** 45(2): 357-377.

Amabile, T. 1988. A Model of Creativity and Innovation in Organizations. In B. M. Staw & L.L. Cummings (Eds.). In L. L. Cummings & B. B. Staw (Eds.), ***Research in organizational behavior:*** 123-167. Greenwich, CT: JAI Press.

Amabile, T. M. 1996. ***Creativity in Context: Update to the Social Psychology of Creativity.*** Boulder, CO: Westview Press.

Amabile, T. M. 1997. Motivating creativity in organizations: on doing what you love and loving what you do. ***California Management Review***, 40(1): 39-58.

Amabile, T. M., 1998. How to kill creativity. ***Harvard Business Review,*** Sept–Oct: 76–87.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron. 1996. Assessing the work environment for creativity. ***Academy of Management Journal***, 39(5): 1155-1184.

Amabile, T. M., & Gryskiewics, S. S. 1987. ***Creativity in the R&D laboratory.*** Technical Report no. 30, Center for Creative Leadership, Greensboro, NC.

Anderson, J. C., & Gerbing, D. W. 1988. Structural equation modeling in practice: a review and recommended two-step approach. ***Psychological Bulletin,*** 103: 411–423.

Antonelli, C. 1989. A failure-inducement model of research and development expenditure: Italian evidence from the early 1980s. ***Journal of Economic Behavior and Organization***, 12(2): 159-180.

Aram, J. D., & Salipante, P. F. 2003. Bridging scholarship in a management: Epistemological reflections. ***British Journal of Management***, 14(3): 189-205.

Arbuckle, J. L. 2005. ***AMOS 6.0 Users's Guide.*** Chicago, IL: SPSS, Inc.

Argote, L. 1999. ***Organizational learning: Creating, retaining and transferring knowledge.*** Norwell, MA: Kluwer Academic Publishers.

Argyris, C. 1992. ***On organizational learning.*** Oxford, UK: Blackwell Publishers.

Argyris, C., & Schon, D. 1978. ***Organizational learning: A theory of action perspective.*** Reading, MA: Addison-Wesley.

Armstrong, J. S., & Overton, T. S. 1977. Estimating nonresponse bias in mail surveys. ***Journal of Marketing Research,*** 14(3): 396–402

Atkinson, A. A. 1998. Strategic performance measurement and incentive compensation. ***European Management Journal,*** 16(5): 552-561.

Atkinson, A. A., Waterhouse, J. H., & Wells, R. B. 1997. A stakeholder approach to strategic performance measurement. ***Sloan Management Review,*** 38(3): 25-37.

Atuahene-Gima, K. 2004. ***Doing good research: Challenges and some strategies.*** Academy of Management Conference, New Orleans, US: August.

Atuahene-Gima, K., Slater, S. F., & Olson, E. M. 2005. The Contingent Value of Responsive and Proactive Market Orientations for New Product Program Performance. ***The Journal of Product Innovation Management,*** 22(6): 464-482.

Audia, P. G., & Greve, H. R. 2006. Less likely to fail: Low performance, firm size, and factory expansion in the shipbuilding industry. ***Management Science,*** 52(1): 83-94.

Audia, P. G., Locke, E. A., Smith, K. G. 2000. The paradox of success: An archival and a laboratory study of strategic persistence following a radical environmental change. ***Academy of Management Journal,*** 43(5): 837–853.

Babbie, E. 2001. ***The practice of social research (9th Edition).*** Stamford, CT: Wadsworth Thomson Learning.

Bacharach, S. 1989. Organizational theories: Some criteria for evaluation. ***Academy of Management Review,*** 14(4): 496-515.

Bagozzi, R. P., Yi, Y, & Phillips, L. W. 1982. Representing and testing organizational theories: A holistic construal. ***Administrative Science Quarterly,*** 27(3): 459-489.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. 1991. Assessing construct validity in organizational research. ***Administrative Science Quarterly,*** 36(3): 421–458.

Baird, I. S., & Thomas, H. 1990. What is risk anyway? Using and measuring risk in strategic management. In R. A. Bettis & H. Thomas (Eds.) ***Risk, strategy, and management, Vol. 5:*** 21-52, Greenwich, CT: JAI Press.

Baker, W. E., & Sinkula, J. M. 1999. Learning orientation, market orientation, and innovation: Integrating and extending models of organizational performance. ***Journal of Market – Focused Management,*** 4(4): 295

Baldwin, C. Y., & Clark, K. B. 2000. ***Design rules. Volume 1: The power of modularity.*** Cambridge, MA: MIT Press.

Balkin, D. B., Markman, G. D., & Gomez-Mejia, L. R. 2000. Is CEO pay in high-technology firms related to innovation? ***Academy of Management Journal,*** 43(6): 1118-1129.

Banker, R. D., Potter, G., & Srinivasan, D. 2000. An empirical investigation of an incentive plan that includes nonfinancial performance measures. ***Accounting Review,*** 75(1): 65–92.

Banker, R. D., Potter, G., & Srinivasan, D. 2005. Association of nonfinancial performance measures with the financial performance of a lodging chain. ***Cornell Hotel and Restaurant administration Quarterly,*** 46(4): 394-412.

Baron, R. M., & Kenny, D. A. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. ***Journal of Personality and Social Psychology,*** 51(6): 1173-1182.

Basberg, B. L. 1987. Patents and the measurement of technological change: a survey of the literature. ***Research Policy***, 16(2-4), 131–141.

Beer, S. 1981. **Brain of the firm**. Chichester, UK: John Wiley & Sons.

Beer, S. 1985. **Diagnosing the system for organizations.** Chichester, UK: John Wiley & Sons.

Beer, S. 1979. ***The heart of enterprise.*** Chichester, UK: John Wiley & Sons.

Benner, M., & Tushman, M. 2002. Process management and technological innovation: A longitudinal study of the photography and paint industries. ***Administrative Science Quarterly,*** 47(4): 676–706.

Bentler, P. M., & Bonnet, D. G. 1980. Significance tests and goodness of fit in the analysis of covariance structures. ***Psychological Bulletin,*** 88: 588-606.

Bentler, P. M. 1989. ***EQS Structural Equations Program Manual.*** Los Angeles, CA: BMDP Statistical Software.

Bentler, P. M. 1990. Comparative fit indexes in structural models. ***Psychological Bulletin,*** 107(2): 238–246.

Berggren, E., & Nacher, T. 2001. Introducing new products can be hazardous to your company: Use the right new solutions delivery tools. ***The Academy of Management Executive,*** 15(3): 92-101.

Bessant, J. 2003. Challenges in innovation management. In L. V. Shavinina (Ed.) ***The international handbook of innovation:*** 761-774, New York: Elsevier.

Bessant, J., Caffyn, S., & Gilbert, J. 1996. Learning to manage innovation. ***Technology Analysis and Strategic Management,*** 8(1): 59–70.

Bessant, J., Kaplinsky, R., & Lamming, R. 2003. Putting supply chain learning into practice. ***International Journal of Operations and Production Management,*** 23(2): 167–181.

Bhaskar, R. 1978. ***Realist theory of science.*** Brighton, UK: Harvester.

Birkinshaw, J., & Gibson, C. B. 2004. Building ambidexterity into an organization. ***Sloan Management Review***, 45(4), 47–55.

Birnberg, J. G. 1988. Discussion of an empirical analysis of the expenditure budget in research and development. ***Contemporary Accounting Research***, 4(2): 582-587.

Bisbe, J., & Otley, D. 2004. The effects of the interactive use of management control systems on product innovation, ***Accounting, Organizations and Society***, 29(8): 709–737.

Bititci, U. S., Carrie, A. S., & McDevitt, L. 1997. Integrated performance measurement systems: A development guide. ***International Journal of Operations and Production Management***, 17(5/6): 522–534.

Black, T. R. 1999. ***Doing quantitative research in the social sciences.*** London: Sage Publications.

Blaikie, N. 1993. ***Approaches to social enquiry.*** Cambridge, UK: Polity Press.

Blaikie, N. 2000. ***Designing social research.*** Cambridge, UK: Polity Press.

Bloom, N., & Van Reenen, J. 2006. ***Measuring and Explaining Management Practices Across Firms and Countries.*** CEP Discussion Paper No 716, London School of Economics, London, UK.

Boeker, W., & Goodstein, J. 1991. Organizational performance and adaptation: effects of environment and performance on changes in board composition. ***Academy of Management Journal***, 34(4): 805-826.

Boeker, W. 1997. Executive migration and strategic change: The effect of top manager movement on product-market entry. ***Administrative Science Quarterly***, 42(2): 213-236.

Boland, R. J., & Pondy, L. R. 1983. Accounting in organizations: a union of natural and rational perspectives. ***Accounting, Organizations and Society,*** 8(2/3): 223-234.

Bollen, K. A. 1998a. A new incremental fit index for general structural equation models. ***Sociological Methods and Research,*** 17(3): 303-316.

Bollen, K. A. 1998b. ***Structural equations with latent variables.*** New York: John Wiley & Sons.

Bolton, M. K. 1993. Organizational innovation and substandard performance: when is necessity the mother of innovation? ***Organization Science,*** 4(1): 57-75.

Bommer, M., & Jalajas, D. 2002. The innovation work environment of high-tech SMEs in the USA and Canada. ***R&D Management,*** 32(5): 379–377.

Bourgeois L. J., & Singh, J. V. 1983. Organizational slack and political behavior within top management teams. ***Academy of Management Proceedings,*** 43–47.

Bourne, M., Franco, M., Neely, A., Kennerly, M., & Marr, B. 2003. ***Corporate performance management.*** Marlow, UK: SAS.

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. 2000. Designing, implementing and updating performance measurement systems. ***International Journal of Operations & Production Management,*** 20(7): 754-771.

Bourne, M., Neely, A. D., Mils, J., & Platts, K. 2003. Implementing performance measurement systems: A literature review. ***International Journal of Operations and Production Management,*** 15(4): 80-116.

Bowen, H. P., & Wiersema, M. F. 1999. Matching method to paradigm in strategy research: Limitations of cross-sectional analysis and some methodological alternatives. ***Strategic Management Journal,*** 20(7): 625-636

Bowman, C., & Ambrocini, V. 1997. Using single respondents in strategy research. ***British Journal of Management,*** 8(2): 119-131.

Boyd, B. K., Gove, S., & Hitt, M. A. 2005. Construct measurement in strategic mangement research: Illusion or reality? ***Strategic Management Journal***, 26(3): 239-257.

Boyer, K. K., & Verma, R. 2000. Multiple raters in survey-based operations management research: A review and tutorial. ***Journal of Production and Operations Management,*** 9(2): 128-140.

Branscomb, L. M., & Auerswald, P. E. 2001. More ways to fail than to succeed: strategies for managing risks. In L. M. Branscomb & P. E. Auerswald (Eds.) ***Taking technical risks: How innovators, executives, investors manage high-tech risks:*** 96-124, Cambridge, MA: The MIT Press.

Brinberg, D., & McGrath, J. E. 1985. ***Validity and the research process.*** Beverly Hills, CA: Sage Publications.

Bromiley, P., & Johnson, S. 2005. Mechanisms and empirical research. In D. J. Ketchen & D. D. Bergh (Eds.), ***Research methodology in strategy and management: Vol. 2***: 15-29, Oxford, UK: Elsevier.

Bromiley, P. 1991. Testing a causal model of corporate risk taking and performance. ***Academy of Management Journal,*** 34(1): 37–59.

Bromiley, P. 2005. ***The behavioral foundations of strategic management.*** Oxford, UK: Blackwell Publishing.

Bromiley, P., Miller, K. D., & Rau, D. 2001. Risk in Strategic Management Research. In M. A. Hitt, R. E. Freeman, & J. S. Harrison (Eds.), ***The Blackwell Handbook of Strategic Management***: 259-288. Malden, Massachusetts: Blackwell.

Brouwer, E., & Kleinknecht, A. 1999. Innovative output and firm's propensity to patent: An exploration of CIS micro data. ***Research Policy,*** 28(6): 615-624.

Brown, J. S., & Duguid, P. 1991. Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. ***Organization Science,*** 2(1): 40-57.

Brown, S. L., & Eisenhardt, K. M. 1995. Product development: Past research, present findings, and future directions. ***Academy of management Review,*** 20(2): 343-378.

Brown, T. A. 2006. ***Confirmatory factor analysis for applied research.*** New York: The Guilford Press.

Browne, M. W., & Cudeck, R. 1993. Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), ***Testing structural equation models:*** 136 -162, Newbury Park, CA: Sage Publications.

Brownell, P. 1985. Budgetary systems and the control of functionally differentiated organizational activities. ***Journal of Accounting Research,*** 23(2): 502-512.

Bruderer, E., & Singh, J. 1996. Organization evolution, learning, and selection: A genetic-algorithm-based model. ***Academy of Management Journal,*** 39(5): 1322-1349.

Bruggeman, W. & Van der Stede, W. 1993. Fitting management control systems to competitive advantage. ***British Journal of Management,*** 4(3): 205-218.

Bruining, H., Bonnet, M., & Wright, M. 2004. Management control systems and strategy change in buyouts. ***Management Accounting Research,*** 15(2): 155–177.

Burchell, S., Clubb, C., Hopwood, A., Hughes, J., & Nahapiet, J. 1980. The roles of accounting in organizations and society. ***Accounting, Organization and Society,*** 5(1): 5–27.

Burgelman, R. A. 1983. A model of the interactions of strategic behavior, corporate context, and the concept of strategy. ***Academy of Management Review,*** 8(1): 61-71.

Burgelman, R. A. 2002. ***Strategy is destiny: How strategy-making shapes a company's future.*** New York: The Free Press.

Burgelman, R. A., Christensen, C. M., & Wheelwright, S. C. (Eds.). 2004. ***Strategic management of technology and innovation (4th Edition).*** New York: McGraw-Hill/Irwin.

Burgelman, R. A., & Sayles, L. R. 1988. ***Inside corporate innovation.*** New York: The Free Press.

Burns, T., & Stalker, G. M., 1961. ***The management of innovation.*** London: Tavistock Publications.

Burrell, C. & Morgan, G. 1979. ***Sociological paradigms and organizational analysis.*** London: Heinemann.

Bush, V. 1945. ***Science: The endless frontier.*** North Stratford, NH: Ayer Co.

Byrne, B. M. 1995. One application of structural equation modeling from two perspectives: Exploring the EQS and LSREL strategies. In R. H. Hoyle (Ed.) ***Structural equation modeling: Concepts, issues, and applications***: 138-157, Thousand Oaks, CA: Sage Publications.

Byrne, B. M. 2001. ***Structural equation modeling with AMOS.*** Mahwah, New Jersey: Lawrence Erlbaum Associates.

Camisón-Zornoza, C., Lapiedra-Alcami, R., Segarra-Cipres, M., & Boronat-Navarro, M. 2004. A Meta-analysis of innovation and organizational size. ***Organization Studies,*** 25(3): 331-361.

Campbell, D. T., & Fiske, D. W. 1959. Convergent and discriminant validation by the multitrait- multimethod matrix. ***Psychological Bulletin,*** 56(2): 81–105.

Campbell, D. T., & Stanely, J. C. 1966. ***Experimental and quasi-Experimental designs for research.*** Chicago: Rand McNally. [Reprinted from Campbell, D. T. & Stanely, J. C. 1963. Experimental and quasi-Experimental designs for research in teaching. In N. L. Gage (Ed.), ***Handbook of research on teaching***, Chicago: Rand McNally.]

Campbell, D. T. 1957. Factors relevant to the validity of experiments in social settings. ***Psychological Bulletin***, 54(4): 297-312.

Campbell, D. T. 1974. Evolutionary epistemology. In P. A. Schilpp (Ed.), ***The Philosophy of Karl Popper*** (*The Library of Living Philosophers*, (Vol. 14, I. & II): 413-463, La Salle, IL: Open Court. [Reprinted in G. Radnitzky & W. W. Bartley, III (Eds.). *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*: 47–89, La Salle, IL: Open Court.]

Campbell, D. T. 1988. Descriptive epistemology: Psychological, sociological, and evolutionary. In E. S. Overman (Ed). ***D. T. Campbell, Methodology and epistemology for social science: Selected papers***: 435-486, Chicago, IL: University of Chicago Press

Cangelosi, V. E., & Dill, W. R. 1965. Organizational learning: Observations towards a theory. ***Administrative Science Quarterly***, 10(2): 175-203.

Capon, N., Farley, J. U., Lehmann, D. R., & Hulbert, J. M., 1992. Profiles of product innovators among large US manufacturers. ***Management Science***, 38(2): 157–169.

Carmines, E. G., & Zeller, R. A. 1979. ***Reliability and validity assessment.*** Thousand Oaks, CA: Sage Publications.

Carpenter, M. A., Geletkanycz, M. A., & Sanders, W. G. 2004. Upper echelons research revisited: antecedents, elements, and consequences of top management team composition. ***Journal of Management***, 30(6): 749-778.

Cassiman, B., & Veugelers, R. 2002. Spillovers and R&D cooperation: some empirical evidence. ***American Economic Review***, 92(4): 1169–1184.

Cavalluzzo, K. S., & Ittner, C. D. 2004. Implementing performance measurement innovations: evidence from government. ***Accounting, Organizations & Society***, 29(3/4): 243-267.

Chandler, A. D. 1962. ***Strategy and Structure.*** New York: Doubleday.

Chandler, A. D. 1977. ***The visible hand: The managerial revolution in American business.*** Cambridge MA: Harvard University Press.

Chandler, G. N., & Lyon, D. W. 2001. Issues of research design and construct measurement in entrepreneurship research: The past decade. ***Entrepreneurship Theory and Practice***, 25(4): 101–113.

Chaney, P. K., & Devinney, T. M. 1992. New product innovations and stock price performance. ***Journal of Business Finance and Accounting***, 19(5): 677-695.

Chaney, P. K., Devinney, T. M., & Winter, R. S. 1991. The impact of new product introductions on the market value of firms. ***Journal of Business***, 64(4): 573-610.

Chenhall, 2003. Management control systems design within its organizational context: Findings from contingency-based research and directions for the future. ***Accounting, Organizations and Society***, 28(2/3): 127–168.

Christensen, C. M. 1997. ***The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.*** Boston, MA: Harvard Business School Press.

Christensen, C. M., & Raynor, M. E. 2003. ***The innovator's solution,*** Boston, MA: Harvard Business School Press.

Cohen, M. D., & Sproull, L. S. (Eds.) 1996. ***Organizational learning.*** Thousand Oaks, CA, Sage Publications.

Cohen, W. M., & Levinthal, D. A. 1990. Absorptive capacity: A new perspective on learning and innovation. ***Administrative Science Quarterly***, 35(1): 128-152.

Collier, P. M. 2005. Entrepreneurial control and the construction of relevant accounting. ***Management Accounting Research,*** 16: 321-339.

Collins, J. M., & Ruefli, T. W. 1996. ***Strategic risk: A state-defined approach.*** Norwell, MA: Kluwer Academic Publishers.

Cook, T. D., & Campbell, D. T. 1979. ***Quasi-Experimentation: Design and analysis issues for field settings.*** Boston: Houghton Mifflin Company.

Cook, T. D., Campbell, D. T., & Peracchio, L. 1990. Quasi Experimentation. In M. D. Dunnette. & L. M. Hough (Eds.), ***Handbook of Industrial and Organizational Psychology: Volume 1 (2nd Ed):*** 491-576, Palo Alto, CA: Consulting Psychologists Press.

Coombs, R., Narandren, & Richards, A., 1996. A literature-based innovation output indicator. ***Research Policy,*** 25(3): 403-413.

Covin, J. G., & Slevin, D. 1991. A conceptual model of entrepreneurship as firm behavior. ***Entrepreneurship: Theory and Practice,*** 16(1): 7-25.

Craggs, A. & Jones, P. 1998. UK results from the community innovation survey 2001, ***Economic trends. No. 539,*** (October), UK: Office for National Statistics.

Crocker, L., & Algina, J. 1986. ***Introduction to Classical and Modern Test Theory.*** Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. ***Psychometrika,*** 16(3): 297–334.

Cross, K. F., & Lynch, R. L. 1989. The SMART way to define and sustain success. ***National Productivity Review,*** 8(1): 23– 33.

Cross, K. F., & Lynch., R. L. 1992. For good measure. CMA, April: 20-23.

Curran, P. J., West, S. G., & Finch, J. F. 1996. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. ***Psychological Methods,*** 1(1): 16-29.

Cyert, R. M., & March, J. G. 1963. ***A Behavioral theory of the firm.*** Englewood Cliffs, New Jersey: Prentice-Hall.

Daft, R. L. 1978. A dual-core model of organizational innovation. ***Academy of Management Journal,*** 21(2): 193-210.

Daft, R. L., & Lewin, A. Y. 1990. Can organizational studies begin to breakout of the normal science straitjacket? An editorial essay. ***Organization Science,*** 1(1): 1-9.

Damanpour, F. 1991. Organizational innovation: A meta-analysis of effects of determinants and moderators. ***Academy of Management Journal,*** 34(3): 555–590.

Damanpour, F. 1992. Organizational size and innovation. ***Organization Studies,*** 13(3): 375–402.

Damanpour, F., & Gopalakrishnan, S. 1998. Theories of organizational structure and innovation adoption: the role of environmental change. ***Journal of Engineering and Technology Management,*** 15(1): 1-24.

Daniel, F., Lohrke, F. T., Fornaciari, C. J., & Turner, R. A. 2004. Slack resources and firm performance: a meta-analysis. ***Journal of Business Research,*** 57(6): 565-574.

Danneels, E. 2002. The dynamics of product innovation and firm competences. ***Strategic Management Journal,*** 23(12): 1095-1121

Davenport, T. H., & Markus, M. L. 1999. Rigor vs. relevance revisited: Response to Benbasat and Zmud, ***MIS Quarterly,*** 23(1): 19-23.

Davila, A. 2000. An empirical study on the drivers of management control systems design in new product development. ***Accounting, Organizations and Society,*** 25(4/5): 383–410.

Davila, T. 2005. The promise of management control systems for innovation and strategic change. In C. S. Chapman (Ed.), ***Controlling Strategy Management, Accounting, and Performance Measurement***: 37-61. Oxford, UK: Oxford University Press.

Davila, T., Epstein, M. J., & Matusik, S. 2004. Innovation strategy and the use of performance measures. J. Y. Lee & M. J. Epstein (Eds.), ***Advances in Management Accounting, Vol. 13:*** 27-58. Greenwich, Connecticut: JAI Press.

Davila, T., Epstein, M. J., & Shelton, R. 2006. ***Making innovation work: How to manage it, measure it, and profit from it.*** Upper Saddle River, New Jersey: Pearson Education.

Day, G. S. & Montgomery, D. B. 1983. Diagnosing the experience curve. ***Journal of Marketing***, 47(2): 44-58.

De Haas, M., & Kleingeld, A. 1999. Multilevel design of performance measurement systems: enhancing strategic dialogue throughout the organization. ***Management Accounting Research***, 10(3): 233-261.

De Toni, A., & Tonchia, S. 2001. Performance measurement system: models, characteristics and measures. ***International Journal of Operations and Production Management***, 21(1/2): 46-70.

Denyer, D., & Neely, A. 2004. Introduction to special issue: innovation and productivity performance in the UK. ***International Journal of Management Reviews***, 5/6(3/4): 131–135.

Dess, G., & Robinson, R. 1984. Measuring organizational performance in the absence of objective measures: the case of the privately-held firm and conglomerate business unit. ***Strategic Management Journal***, 5(3): 265-273.

DeVellis, R. F. 1991. ***Scale development: Theory and applications.*** Newbury Park, CA: Sage.

Dewar, R. D., & Dutton, J. E. 1986. The adoption of radical and incremental innovations: An empirical analysis. ***Management Science,*** 32(11): 1422-1433

Dillman, D. A. 1978. ***Mail and telephone surveys: The total design method.*** New York: John Wiley & Sons.

Dillman, D. A. 2000. ***Mail and internet surveys: The tailored design method.*** New York: John Wiley & Sons.

DiMaggio, P. J., & Powell, W. W. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. ***American Sociological Review,*** 48(2): 147-160.

Dina, G. J. 2006. ***A multi method investigation into the costs and into the benefits of measuring intellectual capital assets.*** Unpublished PhD Thesis, Cranfield School of Management, Cranfield University, UK.

Dixon, J. R., Nanni, A. J., & Vollmann, T. E. 1990. ***The new performance challenge: Measuring operations for world-class competition.*** Homewood, IL: Dow Jones-Irwin.

Dodgson, M. 2000. ***The management of technological innovation: The international and strategic approach.*** Oxford, UK: Oxford University Press.

Dodgson, M., Gann, D., & Salter, A. 2005. ***Think, play, do – technology, innovation, and organization.*** Oxford, UK: Oxford University Press.

Dooley, R. S., & Fryxell, G. E. 1999. Attaining decision quality and commitment from dissent: The moderating effects of loyalty and competence in strategic-decision-making teams. ***Academy of Management Journal,*** 42(4): 389-402.

Dosi, G. 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directives of technological change. ***Research Policy,*** 11: 147-62.

Dosi, G. 1988. The nature of the innovative process. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg & L. Soete (Eds.), ***Technical change and economic theory:*** 221-238. London: Pinter.

Dougherty, D. 1992. A practice-centered model of organizational renewal through product innovation. ***Strategic Management Journal,*** 13(5): 77-92.

Dougherty, D., & Hardy, C. 1996. Sustained product innovation in large, mature organizations: Overcoming innovation-to-organization problems. ***Academy of Management Journal,*** 39(5): 1120-1153.

Dougherty, D., & Heller T. 1994. The illegitimacy of successful product innovations in established Firms. ***Organization Science,*** 3(2): 200–218.

Downs, G. W., & Mohr, L. B. 1976. Conceptual Issues in the Study of Innovation. ***Administrative Science Quarterly,*** 21(4): 700-714.

Drazin, R., & Schoonhoven, C. B. 1996. Community, population, and organizational effects on innovation: A multilevel perspective. ***Academy of Management Journal,*** 39(5): 1065-1083.

DTI 2002. ***The government's manufacturing strategy.*** http://www.dti.gov.uk/manufacturing/strategy.htm. Accessed Mar. 15, 2006.

DTI 2004. ***PSA target metrics for the UK research base.*** London, UK: Department of Trade and Industry.

DTI 2005a. ***The UK innovation survey.*** London, UK: Department of Trade and Industry.

DTI 2005b. ***The 2005 Productivity & Competitiveness Indicators.*** London, UK: Department of Trade and Industry.

DTI 2006a. ***DTI's manufacturing gateway.*** http://www.dti.gov.uk/manufacturing/DTI - Manufacturing Gateway.htm. Accessed Mar. 15, 2006.

DTI 2006b. ***Driving force: Success and sustainability in the UK automotive industry.*** London, UK: Department of Trade and Industry.

Dubin, R. 1976. Theory building in applied areas. In M. Dunette (Ed.), ***Handbook of industrial and organizational psychology***: 17-39, Chicago, IL: Rand MacNally.

Duncan, R. 1976. The ambidextrous organization: Designing dual structures for innovation. In R. Kilman & L. Pondy (Eds.), ***The management of organizational design:*** 167-188. New York: North Holland.

Durbin, J., & Watson, G. S. 1951. Testing for serial correlation in least-squares regression, ***Biometrika,*** 38: 159-177.

Dutton, J. E., Fahey, L., & Narayanan, V. K. 1983. Toward understanding strategic issue diagnosis. ***Strategic management Journal,*** 4(4): 307-323.

Earl, M. J., & Hopwood, A. G. 1980. From management information to information management. In H. C. Lucas, F. F. Land, T. J. Lincoln, & K. Supper (Eds.), ***The information systems environment:*** 3-13. Amsterdam: North-Holland Publishing.

Easterby-Smith, M., Thorpe, R., & Lowe, A. 2002. ***Management research: An introduction.*** London: Sage.

Ebben, J. J., & Johnson, A. C. 2005. Efficiency, flexibility, or both? Evidence linking strategy to performance in small firms. ***Strategic Management Journal,*** 26(13): 1249-1259.

Eccles, R. G. 1991. The performance measurement manifesto. ***Harvard Business Review,*** January-February: 131-137.

Edvinsson, L., & Malone, M. 1997. ***Intellectual capital: realizing your companies true value by finding its hidden brainpower.*** New York: Harper Business.

Eisenhardt, K. M., & Martin J. A. 2000. Dynamic capabilities: what are they? ***Strategic Management Journal,*** 21(10/11): 1105-1121.

Ekvall, G. 1996. Organizational climate for creativity and innovation. ***European Journal of Work and Organizational Psychology,*** 5(1): 105–123.

Ekvall, G. 1997. Organizational conditions and levels of creativity. ***Creativity and Innovation Management,*** 6(4): 195–205.

Etzioni, A. 1964. ***Modern Organizations.*** Englewood Cliffs, New Jersey: Prentice-Hall.

Etzioni, A. 1975. ***A comparative analysis of complex organizations.*** New York: The Free Press.

Evans, M. G. 1985. A Monte Carlo study of the effects of correlated method variance in moderated multiple regression analysis. ***Organizational Behavior and Human Decision Processes,*** 36: 315-342.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. 1999. Evaluating the use of exploratory factor analysis in psychological research. ***Psychological Methods,*** *4*(3): 272-299.

Ferreira, A., & Otley, D. 2005. ***The design and use of management control systems: An extended Framework for analysis.*** Social Science Research Network.

Fiegenbaum, A., & Thomas, H. 1988. Attitudes toward risk and the risk return paradox: prospect theory explanations. ***Academy of Management Journal,*** 31(1): 85–106.

Finkelstein, S., & Hambrick, D. C. 1996. ***Strategic leadership: Top executives and their Effects on organizations.*** Minneapolis, MN: West publishing.

Fiol, C. M., & Lyles, M. A. 1985. Organizational learning. ***Academy of Management Review,*** 10(4): 803-813.

Fiol, C. M. 1996. Squeezing harder doesn't always work: Continuing search for consistency in innovation research. ***Academy of Management Review,*** 21(4): 1012-1021.

Fitzgerald, L., Johnston, R., Bringall, S., Silvestro, R., & Voss, C. 1991. ***Performance measurement in service business.*** London: CIMA.

Flynn, B. B., Sakakibara, S., Schroeder, R. G., Bates, K. A., & Flynn, E. J., 1990. Empirical research methods in operations management. ***Journal of Operations Management.*** 9(2): 250–284.

Forlani, D. 2002. Risk and rationality: The influence of decision domain and perceived outcome control on managers' high-risk decisions. ***Journal of Behavioral Decision Making,*** 15(2): 125-140.

Fornell, C., & Larker, D. F. 1981. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. ***Journal of Marketing Research,*** 18(1): 39-50.

Forza, C. 2002. Survey research in operations management: a process-based perspective. ***International Journal of Operations & Production Management.*** 22(2): 152–194.

Forza, C., & Salvador, F. 2000. Assessing some distinctive dimensions of performance feedback information in high performing plants. ***International Journal of Operations & Production Management,*** 20(3): 359-385.

Fowler, F. J. 1993. ***Survey Research Methods (Second Edition).*** Newbury Park, CA: Sage Publications.

Franco, M., Marr, B., Martinez, V., Gray, D., Adams, C., Micheli, P., Bourne, M., Kennerley, M., Mason, S., & Neely, A. 2004. ***Towards a definition of business performance measurement system,*** PMA 2004 Conference, Edinburgh: July.

Frank, J. D. 1935. Individual differences in certain aspects of the level of aspiration. ***American Journal of Psychology,*** 47(1): 119-128.

Freeman, C. 1974. ***The economics of industrial innovation.*** Harmondsworth, UK: Penguin Books.

Frenken, K. 2006. Technological innovation and complexity theory. ***Economics of Innovation and New Technology,*** 15(2): 137-155.

Galbraith, J. 1973. Designing complex organizations. Reading, MA: Addison-Wesley.

Galende, J. 2006. Analysis of technological innovation from business economics and management. ***Technovation,*** 26(3): 300-311.

GAO 2005. ***Performance Measurement and Evaluation: Definitions & Relationships,*** United States Government Accountability Office.

Garcia, R., & Calantone, R. 2002. Critical look at technological innovation typology and innovativeness terminology: A literature review. ***Journal of Product Innovation Management,*** 19(2): 110-132.

Garg, V. K., Walters, B. A., & Priem, R. L. 2003. Chief executive scanning emphases, environmental dynamism, and manufacturing firm performance. ***Strategic Management Journal,*** 24(4): 725-744

Garvin, D. A. 1993. Building a learning organization. ***Harvard Business Review,*** 71(4): 78-92.

Gates, S. 1999. ***Aligning strategic performance measures and results.*** New York: The Conference Board.

Gatignon, H., Tushman, M. L., Smith, W., & Anderson, P. 2002. A structural approach to assessing innovation: Construct development of innovation locus, type, and characteristics. ***Management Science,*** 48(9): 1103-1122.

George, G. 2005. Slack resources and the performance of privately held firms. *Academy of Management Journal*, 48(4): 661-676

Georgiou, P. 1973. The goal paradigm and notes towards a counter paradigm. *Administrative Science Quarterly*, 18(3): 291-310.

Gerbing, D. W., & Anderson, J. C. 1988. An updated paradigm for scale development incorporating unidimensionality and its assessment. ***Journal of Marketing Research,*** 25(5): 186-192.

Gerdin, J. 2005. ***Conceptualizations of contingency fit in management accounting research—correspondence between statistical models used and core contingency theory assumptions.*** Working paper, Örebro University, Örebro, Sweden

Ghalayini, A. M., & Noble, J. S. 1996. The changing basis of performance measurement. ***International Journal of Operations and Production Management,*** 16(8): 63-80.

Gibson, C. B., & Birkinshaw, J. 2004. The antecedents, consequences, and mediating role of organizational ambidexterity. ***Academy of Management Journal,*** 47(2): 209-226.

Glynn, M. A., Lant, T., & Mezias, S. J. 1991. Incrementalism, learning and ambiguity: An experimental study of aspiration level updating. J. L. Wall & L. R. Jauch, (Eds.). ***Best Paper Proceedings of the Academy of Management Meetings.*** Madison, WI: Omnipress.

Godener, A., & Soderquist, K. E. 2004. Use and impact of performance measurement results in R&D and NPD an exploratory study. ***R&D Management,*** 34(2): 191-219.

Godin, B. 2003. The emergence of S&T indicators: why did governments. ***Research Policy,*** 32(4): 679–691.

Godin, B. 2005a. ***Measurement and statistics on science and technology: 1920 to the present.*** London: Routledge.

Godin, B. 2005b. ***The linear model of innovation: The historical construction of an analytical framework.*** Working paper no. 30, Project on the history and Sociology of S&T Statistics, Quebec, Canada.

Goffin, K., & Mitchell, R. 2005. ***Innovation management: Strategy and implementation using the pentathlon framework.*** Basingstoke, UK: Palgrave Macmillan.

Gopalakrishnan S., & Damanpour, F. 1997. A Review of innovation research in economics, sociology and technology Management. ***Omega,*** 25(1): 15-28

Gopalakrishnan, S., & Bierly, P. 2001. Analyzing innovation adoption using a knowledge-based approach. ***Journal of Engineering and Technology Management,*** 18(2): 107-130.

Govindarajan, V. J. 1984. Appropriateness of accounting data in performance evaluation: An empirical examination of environmental uncertainty as an intervening variable. ***Accounting, Organization and Society,*** 9(2): 125–135.

Govindarajan, V. J. 1988. A contingency approach to strategy implementation at the business-unit level: Integrating administrative mechanisms with strategy. ***Academy of Management Journal,*** 31(4): 828-853.

Govindarajan, V., & Gupta, A. K. 1985. Linking control systems to business unit strategy: Impact on performance. ***Accounting, Organizations and Society,*** 10(1): 51–66.

Green, S. B. 1991. How many subjects does it take to do a regression analysis? ***Multivariate Behavioural Research,*** 26(3): 499-510.

Greve, H. R. 1998a. Performance, aspirations, and risky organizational change. ***Administrative Science Quarterly,*** 43(1): 58-86.

Greve, H. R. 1998b. Managerial cognition and the mimetic adoption of market positions: What you see is what you do. ***Strategic Management Journal,*** 19(10): 967-988.

Greve, H. R. 2003a. ***Organizational learning from performance feedback: A behavioral perspective on innovation and change.*** Cambridge, UK: Cambridge University Press.

Greve, H. R. 2003b. A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding. ***Academy of Management Journal,*** 46(6): 685-702.

Greve, H. R. 2003c. Investment and the behavioral theory of the firm: Evidence from shipbuilding. ***Industrial and Corporate Change,*** 12(5): 1051-1076.

Greve, H. R. 2004. ***More of a Journey: Antecedents of Explorative Product Innovation*** Working Paper, Norwegian School of Management BI, Oslo, Norway.

Greve, H. R., & Taylor, A. 2000. Innovations as catalysts for organizational change: Shifts in organizational cognition and search. ***Administrative Science Quarterly,*** 45(1): 54–80.

Griliches, Z. 1990. Patent statistics as economic indicators: A survey. ***Journal of Economic literature,*** 28(4): 1661–1707.

Gross, E. 1969. The definition of organizational goals. ***British Journal of Sociology,*** 20: 277-294.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. 2004. ***Survey methodology.*** Hoboken, New Jersey: John Wiley & Sons.

Grupp, H., & Mogee, M. E. 2004. Indicators for national science and technology policy: How robust are composite indicators. ***Research Policy,*** 33(9): 1373-1384.

Hage, J. T. 1999. Organizational innovation and organizational change. ***Annual Review of Sociology,*** 25(1): 597-622.

Hagedoorn, J., & Cloodt, M. 2003. Measuring innovative performance: is there an advantage in using multiple indicators? ***Research Policy,*** 32(8): 1365-1379.

Hair, J. F., Black, W. C., Anderson, R. E., & Tatham, R. L. 2006 ***Multivariate data analysis (Sixth Edition).*** New Jersey, USA: Pearson Prentice Hall.

Hambrick, D. C. 1994. What if the academy actually mattered? ***Academy of management review,*** 19(1): 11-16.

Hambrick, D. & Mason, P. 1984. Upper echelons: The organization as a reflection of its Top Managers. ***Academy of Management Review,*** 9(3): 193-206.

Hannan, M. T., & Freeman, J. 1977. The population ecology of organizations. ***American Journal of Sociology,*** 82(5): 929-964.

Hannan, M. T., & Freeman J. 1984. Structural inertia and organizational change. ***American Sociological Review,*** 49(2): 149-164.

Hannan, M. T., & Freeman, J. 1989. ***Organizational ecology.*** Cambridge, MA: Harvard University Press.

Harris, L. C. 2001. Market orientation and performance: objective and subjective empirical evidence from UK companies. ***Journal of Management Studies,*** 38(1): 17-34.

Hartmann, F. G. H., & Moers, F. 1999. Testing contingency hypotheses in budgetary research: An evaluation of the use of moderated regression analysis. ***Accounting, Organizations and Society,*** 24(4): 291–315.

Hartmann, F. G. H., & Moers, F. 2003. Testing contingency hypotheses in budgetary research using moderated regression analysis: A second look. ***Accounting, Organizations and Society,*** 28(7/8): 803-809.

Hauser, J., Tellis, G. J., & Griffin, A. 2005. ***Research on innovation: A review and agenda for marketing science.*** Special Report No. 05-200, Marketing Science Institute.

Hayes, D. D. 1977. The contingency theory of managerial accounting. ***The Accounting Review***, 52(1), 22-39.

Hayes, R. H., & Abernathy, W. J. 1980. Managing our way to economic decline. ***Harvard Business Review***, July-August: 67-77.

Helfat, C. 1994. Evolutionary trajectories in petroleum firm R&D. ***Management Science***, 40(12): 1720-1747.

Henderson, R., & Clark, K. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. ***Administrative Science Quarterly***, 35(1): 9-30.

Henri, J. 2006a. Organizational culture and performance measurement systems. ***Accounting, Organizations and Society***, 31(1): 77-103.

Henri, J. 2006b. Management control systems and strategy: A resource based perspective. ***Accounting, Organizations and Society***, 31(6): 529-558.

Hensley, R. L. 1999. A review of operations management studies using scale development techniques. ***Journal of Operations Management***, 17(3): 343-358.

Herriot, S. R., Levinthal, D. & March, J. G. 1985. Learning from experience in organizations. ***American Economic Review***, 75(2): 298-302.

Hill, C. W. L., & F. T. Rothaermel. 2003. The Performance of Incumbent Firms in the Face of Radical Technological Innovation. ***Academy of Management Review***, 28 (2): 257-274.

Hodgkinson, G. P., Herriot, P., & Anderson, N. 2001. Re-aligning the stakeholders in management research: lessons from industrial, work and organizational psychology. ***British Journal of management***, 12(SI): S41-S48.

Hollander, S. 1965. ***Sources of efficiency.*** Cambridge, MA: MIT Press.

Hoque, Z., & James, W. 2000. Linking balanced scorecard measures to size and market factors: Impact on organizational performance. ***Journal of Management Accounting Research***, 12: 1–17.

Hoque, Z. 2004. A contingency model of the association between strategy, environmental uncertainty and performance measurement: Impact on organizational performance. ***International Business Review***, 13(4): 485–502.

Hoque, Z., Mia, A., & Alam, M. 1997. ***Competition, new manufacturing practices, changes MAS and managerial choice of "balanced scorecard" approach to performance measures: an empirical investigation.*** Paper presented at the 1997 European Accounting Association Annual Congress in Graz, April.

Hoque, Z., Mia, L. & Alam, M. 2001. Market competition, computer-aided manufacturing and use of multiple performance measures: An empirical study. ***British Accounting Review***, 33(1): 23–45.

Hoskisson, R., & Hitt, M., 1988. Strategic control systems and relative R&D investment in large multiproduct firms. ***Strategic Management Journal***, 9(6), 605–621.

Hoyle, R. H., & Panter, A. T. 1995. Writing about structural equation models. In R. H. Hoyle (Ed.) ***Structural Equation Modeling: Concepts, Issues, and Applications***: 158-176, Thousand Oaks, CA: Sage Publications.

Hrebiniak, L. G. 1978. ***Complex organizations.*** St. Paul, MN: West Publishing Co.

Hu, L., & Bentler, P. M. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. ***Structural Equation Modeling***, 6(1): 1-55.

Huber, G. P. 1991. Organizational learning: The contributing processes and the literatures. ***Organization Science***, 2(1): 88-115.

Huber G. P., & Power, D. J., 1985. Retrospective reports of strategic-level managers: guidelines for increasing their accuracy. ***Strategic Management Journal***, 6(2): 171–180.

Hundley, G., Jacobson, C. K., & Park, S. H. 1996. Effects of profitability and liquidity on R&D intensity: Japanese and US companies compared. ***Academy of Management Journal***, 39(6): 1659-1674.

Ittner, C. D., Larcker, D. F., & Rajan, M. V. 1997. The choice of performance measures in annual bonus contract. ***The Accounting Review***, 72(2): 231–256.

Ittner, C. D., Larcker, D. F., & Randall, T. 2003. Performance implications of strategic performance measurement in financial service firms. ***Accounting, Organizations and Society***, 28(7/8): 715-741.

Iversen, E. 1998. Patents. In K. Smith (Ed.) ***Science, Technology and Innovation Indicators – A guide for policy makers***, IDEA Report 5, Oslo: STEP Group.

Jaccard, J., & Turrisi, R. 2003. ***Interaction effects in multiple regression (2nd Edition).*** Thousand Oaks, CA: Sage Publications.

Jackson, S. E., 1992. Consequences of group composition for the interpersonal dynamics of strategic issue processing. In P. Shrivastava, A. Huff, & J. Dutton (Eds.), ***Advances in strategic management, Volume 8***: 345-382, Greenwich, CT: JAI Press.

Jaikumar, R., & Bohn, R. 1992. A dynamic approach to operations management: An alternative to static optimization. ***International Journal of Production Economics***, 27(3): 265-282.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. 2003. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. ***Journal of Consumer Research***, 30(2): 199-218.

Johnson, H. T. 1990. Performance measurement for competitive excellence. In R. S. Kaplan (Ed.) ***Measures for manufacturing excellence.*** Boston, MA: Harvard Business School Press.

Johnson, P, & Duberley, J. 2000. ***Understanding management research.*** London: Sage.

Johnson, H. T., & Kaplan, R. S. 1987. ***Relevance lost: The rise and fall of management accounting.*** Boston, MA: Harvard Business School Press.

Jöreskog, K. G. 1993. Testing Structural Equation Models. In K. A. Bollen & J. S. Long (Eds.) ***In Testing Structural Equation Models.*** Newbury Park, CA: SAGE Publications.

Jöreskog, K. G., & Sorbom, D. 1979. ***Advances in factor analysis and structural equation models.*** Cambridge, MA: Abt Books.

Kahneman, D., & Tversky, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2): 263-291.

Kahneman, D., & Tversky, A. (Eds.) 2000. ***Choices values and frames.*** London: Cambridge University Press.

Kameda, T., & Davis, J. D. 1990. The function of the reference point in individual and group risk decision making. ***Organizational Behavior and Human Decision Processes,*** 46(1): 55-76

Kamien, M. I., & Schwartz, N. L. 1982. ***Market structure and innovation.*** Cambridge, MA: Cambridge University Press.

Kanter, R. 1988. When a thousand flowers bloom: Structural, collective, and social conditions for innovation in organizations. In B. M. Staw & L. L. Cummings (Eds.), ***Research in organizational behavior, Vol. 10:*** 160-211. Greenwich, CT: JAI Press.

Kaplan, R. S. 1984. The evolution of management accounting. ***The Accounting Review***, LIX(3): 390-418.

Kaplan, R. S. 1986. Accounting lag: The obsolescence of cost accounting systems. ***California Management Review***, 28(2): 174-199.

Kaplan, R., & Norton, D. 1992. The Balanced scorecard – Measures that drive performance, ***Harvard Business Review***, 70(1): 71–79.

Kaplan, R., & Norton, D. 1996. ***The balanced scorecard.*** Boston, MA: Harvard Business School Press.

Kaplan, R., & Norton, D. 2001. ***The strategy focused organization.*** Boston, MA: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. 2004. ***Strategy maps: Converting intangible assets into tangible outcomes.*** Boston, MA: Harvard Business School Press.

Katila, R. 2000. ***In search of innovation: Search determinants of new product introductions.*** Unpublished Ph. D. Dissertation, University of Texas at Austin, USA.

Katila, R. 2002. New product search overtime: Past ideas in their prime? ***Academy of Management Journal***, 45(5): 995-1010.

Katila, R., & Ahuja, G. 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. ***Academy of Management Journal***, 45(6): 1183-1194.

Kaynak, H. 2003. The relationship between total quality management practices and their effects on firm performance. ***Journal of Operations Management.*** 21(4): 405-435.

Kaynak, H., & Hartley, J. L. 2006. Using replication research for just-in-time purchasing construct development. ***Journal of Operations Management***, 24(6): 868-892.

Keats, B., & Hitt, M. 1988. A causal model of linkages among environmental dimensions, macro-organizational characteristics, and performance. ***Academy of Management Journal,*** 31(3): 570-598.

Keegan, D. P., Eiler, R. G., & Jones, C. R. 1989. Are your performance measures obsolete? ***Management Accounting,*** June: 45-50.

Kennerley, M., & Neely, A.D. 2002. A framework of the factors affecting the evolution of performance measurement systems. ***International Journal of Operations & Production Management,*** 22(2): 1222-1245.

Kennerley, M., & Neely, A.D. 2003. Measuring performance in a changing business environment. ***International Journal of Operations & Production Management,*** 23(2): 213-29.

Kerlinger, F. N., & Lee, H. B. 2000. ***Foundations of behavioral research.*** Stamford, CT: Wadsworth Thomson Learning.

Kerssens-van Drongelen, I. C., & Fisscher, O. A. M. 2003. Ethical dilemmas in performance measurement. ***Journal of Business Ethics,*** 45(1/2): 51-63.

Ketokivi, M. A., & Schroeder, R. G. 2004. Perceptual measures of performance: fact or fiction? ***Journal of Operations Management,*** 22(3): 247-264.

Khandwalla, P. N. 1972. The effect of different types of competition on the use of management controls. ***Journal of Accounting Research,*** 10(2): 275-285.

Khandwalla, P. N. 1973. Effect of Competition on the structure of top management control. ***Academy of Management Journal,*** 16(2): 285-295.

Khandwalla, P. N. 1977. ***The design of organizations.*** New York: Harcourt Brace Jovanovich.

Kleinknecht, A. 1987. Measuring R&D in small firms: How much are we missing? ***Journal of Industrial Economics,*** 36(2): 253-256.

Kleinknecht, A. 1993. Why do we need new innovation output indicators? An introduction. In A. Kleinknecht & D. Bain (Eds.), ***New concepts in innovation output measurement:*** 1-9. Basingstoke, UK: MacMillan Press.

Kleinknecht, A. 1996. New indicators and determinants of innovation: An introduction. In A. Kleinknecht (Ed.), ***Determinants of innovation: The new message from new indicators:*** 1-11. Basingstoke, UK: MacMillan Press.

Kleinknecht. A., & Reijnen, J. O. N. 1991. More evidence on the undercounting of small firm R&D. ***Research Policy***, 20(6): 579-587.

Kleinknecht, A., & Verspagen, B. 1989. R&D and market structure: The impact of measurement and aggregation Problems. ***Small Business Economics,*** 1(4): 297-302.

Kline, R. B. 2004. ***Principles and Practice of Structural Equation Modeling (2nd Ed.)***. New York: The Guilford Press.

Kline, S. J., & Rosenberg, N. 1986. An overview of innovation. In R. Landau & N. Rosenberg (Eds.), ***The positive sum strategy***: 275-305, Washington, DC: National Academy Press.

Knight, F. H. 1921. ***Risk, uncertainty, and profit.*** Houghton Mifflin (1971 Reprint). Chicago, IL: University of Chicago.

Knight, K. 1967. A descriptive model of the intra-firm innovation process. ***Journal of Business,*** 40(4): 478-496.

Knoke, D., Marsden, P. V., & Kalleberg, A. L. 2002. Survey research methods. In J. A. C. Baum (Ed.), ***Companion to Organization:*** 781-804, Oxford, UK: Blackwell Publishers.

Kuhn, T. S. 1962. ***The structure of scientific revolutions.*** Chicago, IL: University of Chicago Press.

Lam, A. 2005. Organizational innovation. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), ***The Oxford handbook of innovation:*** 115-147. Oxford, UK: Oxford University Press.

Lambert, D. M., & Harrington, T.C. 1990. Measuring nonresponse in customer service mail surveys. ***Journal of Business Logistics***, 11(2): 5-25.

Langer, E. J. 1989. ***Mindfulness.*** Reading, MA: Addison-Wesley.

Lant, T. K. 1992. Aspiration level adaptation: An empirical exploration. ***Management Science,*** 38(5): 623-644.

Lant, T. K., & Mezias, S. J. 1990. Managing discontinuous change: A simulation study of organizational learning and entrepreneurship. ***Strategic Management Journal,*** 11(4): 147-179.

Lant, T. K., & Mezias, S. J. 1992. An Organizational learning model of convergence and reorientation. ***Organization Science,*** 3(1): 47-71.

Lant, T. K., Milliken, F. J., & Batra, B. 1992. The role of managerial learning and interpretation in strategic persistence and reorientation: An empirical exploration. ***Strategic Management Journal,*** 13(8): 585-608.

Lant, T. K., & Montgomery, D. B. 1987. Learning from strategic success and failure. ***Journal of. Business Research.*** 15(6): 503–518.

Laursen, K., & Salter, A. 2006. Open for innovation: The role of openness in explaining innovation performance among U.K. manufacturing firms. ***Strategic Management Journal***, 27(2): 131-150.

Lave, C. A., & March, J. G. 1975. ***Introduction to Models in the Social Sciences.*** New York: Harper & Row.

Lawrence, P. R., & Lorsch, J. W. 1967. Differentiation and integration in complex organizations. ***Administrative Science Quarterly***, 12(1): 1-47

Lawrie, G., & Cobbold, I. 2004. Third-generation balanced scorecard: Evolution of an effective strategic control tool. ***International Journal of Productivity and Performance Management,*** 53(7): 611-623.

Lebas, M. 1995. Manufacturing performance measurement and performance management. ***International Journal of Production Economics,*** 41(1): 23-34.

Lee, C., & Sung, T. 2005. Schumpeter's legacy: A new perspective on the relationship between firm size and R&D. ***Research Policy,*** 34(6): 914-931

Leseure, M. J., Bauer, J., Birdi, B., Neely, A. & Denyer, D. 2005. Adoption of promising practices: A systematic review of the evidence. ***International Journal of Management Reviews,*** 5/6(3/4): 169–190

Lev, B. 1969. Industry averages as targets for financial ratios. ***Journal of Accounting Research,*** 7(2): 290-299.

Levinthal, D., & March, J. G. 1981. A model of adaptive organizational search. ***Journal of Economic Behavior and Organization,*** 2(4): 307-333.

Levinthal, D. A., & March, J. G. 1993. The Myopia of learning. ***Strategic Management Journal,*** 14(8): 95–112.

Levitt, B., & March, J. G. 1988. Organizational learning. ***Annual Review of Sociology,*** 14: 319-340.

Li, H., & Atuahene-Gima, K. 2001. Product innovation strategy and the performance of new technology ventures in china. ***Academy of Management Journal,*** 44(6): 1123-1134

Li, H., & Atuahene-Gima, K. 2002. The adoption of agency business activity, product innovation, and performance in Chinese technology ventures. ***Strategic Management Journal,*** 23(6): 469-490.

Lingle, J. H., & Schiemann, W. A. 1996. From balanced scorecard to strategic gauges: Is measurement worth it? ***Management Review,*** March: 56-61.

Lockamy, A. III, & Cox, J. F. III. 1994. ***Reengineering performance measurement: How to align systems to improve processes, products and profits.*** Burr Ridge, IL: Irwin.

Lomi, A., Larsen, E. R., & Ginsberg, A. 1997. Adaptive learning in organizations: A system dynamics-based exploration. ***Journal of Management,*** 23(4): 561-583.

Lopes, L. L. 1987. Between hope and fear: The psychology of risk. In L. Berkowitz (Ed.), ***Advances in experimental social psychology: Vol. 20***: 255-295, San Diego, CA: Academic Press.

Luce, R., & Raiffa, H. 1957. ***Games and decisions.*** New York: Wiley.

Lynch, R. L., & Cross, K. F. 1991. ***Measure up! How to measure corporate performance***, Cambridge, MA: Blackwell Publishers.

MacKenzie, S. B. 2003. The dangers of poor construct conceptualization. ***Academy of marketing Science Journal,*** 13(3): 323-326.

Mahdi, S. 2002. ***Search strategy in product innovation: theory and evidence from agrochemicals.*** Unpublished D.Phil. thesis, University of Sussex, UK.

Mahdi, S. 2003. Search strategy in product innovation process: Theory and evidence from the evolution of agrochemical lead discovery process. ***Industrial and Corporate Change,*** 12(2): 235-270

Mairesse J, Mohnen P. 2002. Accounting for innovation and measuring innovativeness: an illustrative framework and an application. ***American Economic Review,*** 92(2): 226–231.

Maisel, L. S. 2001. Performance measurement practices survey results. USA: AICPA.

Malhotra, M. K., & Grover, V. 1998. An assessment of survey research in POM: from constructs to theory. ***Journal of Operations Management,*** 16(4): 407-425.

Malina, M.A., & Selto, F. H. 2001. Communicating and controlling strategy: An empirical study of the effectiveness of the Balanced Scorecard. ***Journal of Management Accounting Research,*** 13: 47–90.

March, J. G. 1962. The business firm as a political coalition. ***Journal of Politics,*** 24: 622-678.

March, J. G. 1988. Variable risk preferences and adaptive aspirations. ***Journal of Economic Behavior and Organization,*** 9(1): 5-25.

March, J. G. 1989. ***Decisions and organizations.*** Oxford, UK: Basil Blackwell.

March, J. G. 1991. Exploration and exploitation in organizational learning. ***Organization Science,*** 2(1): 71-87.

March, J. G. 1994. ***A primer on decision making: How decisions happen.*** New York: Free Press.

March, J. G. 1997. Forward. In R. Garud, P. R. Nayyar, & Z. B. Shapira (Eds.) ***Technological innovation: Oversights and foresights***: ix-xii, Cambridge: Cambridge University Press.

March, J. G. 1999. ***The pursuit of organizational intelligence.*** Malden, MA: Blackwell Publishers.

March, J. G. 2006. Rationality, foolishness, and adaptive intelligence. ***Strategic Management Journal,*** 27(3): 201-214.

March, J. G., & Shapira, Z. 1987. Managerial perspectives on risk and risk taking. ***Management Science,*** 33(11): 1404–1418.

March, J. G., & Shapira, Z. 1992. Variable risk preferences and the focus of attention. ***Psychological Review,*** 99(1): 172–183.

March, J. G., & Simon, H. A. 1958. ***Organizations.*** New York: McGraw-Hill.

March, J. G., & Olsen, J. P. 1976. ***Ambiguity and Choice in Organization.*** Bergen, Norway: Universitetsforlaget.

March, J. G., Schulz, M., & Zhou, X. 2000. ***The dynamics of rules: Change in written organizational codes.*** Stanford, CA: Stanford University Press.

Marinova, D., & Phillimore, J. 2003. Challenges in innovation management. In L. V. Shavinina (Ed.) ***Models of innovation:*** 44-53, New York: Elsevier.

Marr, B., & Schiuma, G. 2003. Business performance measurement – Past, present, future. ***Management Decision,*** 41(8): 680-687.

Martin, X., & Mitchell, W. 1998. The influence of local search and performance heuristics on new design introduction in a new product market. ***Research Policy,*** 26: 753-771.

Maruyama, G. M. 1997. ***Basics of Structural Equation Modeling.*** Thousand Oaks, CA: Sage Publications.

Maskell, B. H. 1991. ***Performance measurement for world class manufacturing.*** Cambridge, MA: Productivity Press.

McGee, J. V. 1992. ***What is strategic performance measurement?*** USA: Ernst & Young Center for Business Innovation.

McKee, D. O., Varadarajan, P. R., & Pride, W. 1989. Strategic adaptability and firm performance: A market-contingent perspective. ***Journal of Marketing,*** 53(3): 21-35.

McKelvey, B. 1999. Toward a Campbellian realist organization science. In B. McKelvey & J. A. C. Baum (Ed.), ***Variation in organization Science: In honor of Donald T. Campbell***: 383-411.Thousand Oaks, CA: Sage Publications.

McKelvey, B. 2002a. Model centered organization science epistemology. In J. A. C. Baum (Ed.), ***Companion to Organization***: 752-780. Oxford, UK: Blackwell Publishers.

McKelvey, B. 2002b. Glossary of epistemology terms. In J. A. C. Baum (Ed.), ***Companion to Organization***: 889-898. Oxford, UK: Blackwell Publishers.

McKelvey, B., & Baum, J. A. C. 1999. Donald T. Campbell's evolving influence on organization science. In B. McKelvey & J. A. C. Baum (Eds.), ***Variation in organization Science: In honor of Donald T. Campbell***: 1-15. Thousand Oaks, CA: Sage Publications.

Merchant, K. 1984. Influences on departmental budgeting: An empirical examination of a contingency model. ***Accounting, Organizations and Society***, 9 (3/4): 291-307.

Meyer, M., & Roberts, E. 1986. New product strategy in small technology-based firms: A pilot study. ***Management Science,*** 32(7): 806–821.

Mezias, S. 1988. Aspiration level effects: An empirical investigation. ***Journal of Economic Behavior and Organization,*** 10: 389-400.

Mezias, S. J., Chen, Y. & Murphy, P. 2002. Aspiration-level adaptation in an American financial services organization: A field study. ***Management Science,*** 48(10): 1285-1301.

Miles, R. E., & Snow, C. C. 1978. ***Organizational strategy, structure and process.*** New York: McGraw Hill.

Miller, K. D., & Bromiley, P. 1990. Strategic Risk and Corporate Performance: An Analysis of Alternative Risk Measures. ***Academy of Management Journal,*** 33 (4): 756-779.

Miller, K. D., & Friesen, P. H. 1983. Strategy-making and environment: The third link. ***Strategic Management Journal,*** 4(3): 221-235.

Mintzberg, H. 1973. Strategy making in three modes. ***California Management Review,*** 16(3): 294-316.

Mitchell, V. 1994. Using industrial key informants: some guidelines. ***Journal of the Market Research Society***, 36(2): 139-144.

Mohnen, P., & Dagenais, M. 2002. Towards an innovation intensity index: the case of CIS 1 in Denmark and Ireland. In A. Kleinknecht (Ed.), ***Innovation and firm performance: Econometric explorations of survey data:*** 3-29. Basingstoke, UK: Palgrave.

Mohr, L. B. 1982. ***Explaining Organizational Behavior***, San Francisco, CA: Jossey-Bass.

Moldoveanu, C. M., & Baum, J. A. C. 2002. Contemporary debates in organizational epistemology. In J. A. C. Baum (Ed.), ***Companion to Organization:*** 733-751, Oxford, UK: Blackwell Publishers.

Morecroft, J. D. 1985. Rationality in the analysis of behavioral simulation models. ***Management Science,*** 31(7): 900–916.

Myers, S., & Marquis, D. 1969. ***Successful industrial innovation.*** Washington, DC: NSF.

Naman, J. L., & Slevin, D. P. 1993. Entrepreneurship and the concept of fit: A model and empirical tests. ***Strategic Management Journal,*** 14(2): 137-153.

Neely, A. D. 1998. ***Measuring business performance: Why, what and how.*** London: Economist Books.

Neely, A. D. (Ed.) 2002. ***Business performance measurement: Theory and practice.*** Cambridge, UK: Cambridge University Press.

Neely, A. 2004. In search of a metric system for innovation. ***Financial Times,*** October 7.

Neely, A. 2005. The evolution of performance measurement research: Developments in the last decade and a research agenda for the next. ***International Journal of Operations and Production Management,*** 25(12):

1264-1277.

Neely, A. 2006a. The search for meaningful measures of performance. ***Financial Times***, April 6.

Neely, A. 2006b. ***The Challenges of Measuring Performance in Innovative Firms.*** Draft Book Chapter, London: Advanced Institute of Management Research (AIM).

Neely, A. D., Adams, C., & Kennerley, M. 2002. ***The performance prism: The scorecard for measuring and managing business success.*** London, UK: Pearson Education Ltd.

Neely, A., & Al Najjar, M. 2006. Management Learning Not Management Control: The true role of performance measurement. ***California Management Review***, 48(3): 99-114.

Neely, A., & Bourne, M. 2000. Why measurement initiatives fail. ***Measuring Business Excellence***, 4(4): 3-6.

Neely, A., Bourne, M., & Kennerly, M. 2000. Performance measurement system design: Developing and testing process-based approach. ***International Journal of Operations and Production Management***, 20(10): 1119-1145.

Neely, A., Filippini, R., Forza, C., Vinelli, A., & Hii, J. 2001. A framework for analyzing business performance, firm innovation and related contextual factors: Perceptions of managers and policy makers in two European regions. ***Integrated Manufacturing Systems***, 12(2): 114-124.

Neely, A. D., Gregory, M. J., & Platts, K. 1995. Performance measurement system design. A literature research and research agenda. ***International Journal of Operations and Production Management***, 15(4): 80-116.

Neely, A., Marr, B., Roos, G., Pike, P., & Gupta, O. 2003 - Towards the third generation of performance measurement. ***Controlling***, March/April: 129-135

Neely, A. D., Mills, J. F., Gregory, M. J., Richards, A. H., Platts, K. W., & Bourne, M. C. S. 1996. ***Getting the Measure of your Business.*** London: Findlay Publications.

Neely, A. D., Richards, H., Mills, J. F., Platts, K., & Bourne, M. C. S. 1997. Designing performance measures: A structured approach. ***International Journal of Operations & Production Management,*** 17(11/12): 1131-1152.

Nelson, R. R., & Winter, S. G. 1982. ***An evolutionary theory of economic change.*** Cambridge, MA: The Belknap Press.

Nilsson, F., & Kald, M. 2002. Recent advances in performance management: The Nordic case. ***European Management Journal,*** 20(3): 235-245.

Nixon, B. 1998. Research and development performance measurement: A case study. ***Management Accounting Research,*** 9(3): 329-355.

Nørreklit, H. 2000. The balance on the balanced scorecard: A critical analysis of some of its assumptions. ***Management Accounting Research***, 11(1): 65-88.

Nørreklit, H. 2003. The balanced scorecard: what is the score? A rhetorical analysis of the balanced scorecard. ***Accounting, Organizations and Society,*** 28(6): 591-619.

Nunnally, J. C., & Bernstein, I. H., 1994. ***Psychometric Theory (Third Edition).*** New York, NY: McGraw-Hill.

O'Leary-Kelly, S. W., & Vokurka, R. J. 1998. The empirical assessment of construct validity. ***Journal of Operations Management,*** 16(4): 387-405.

Ocasio, W. 1998. Towards an attention-based view of the firm. ***Strategic Management Journal,*** 18(SI): 187-206.

OCED. 1997. ***Proposed guidelines for collecting and interpreting technological innovation data: The 'Oslo Manual'.*** Paris: Organization for Economic Development and Co-operation.

Otley, D. 1999. Performance management: A framework for management control system design. ***Management Accounting Research,*** 10: 363–382.

Ouchi, W. G. 1977. The relationship between organizational structure and organizational control. ***Administrative Science Quarterly,*** 22(1): 95-113.

Ouchi, W. G. 1979. A conceptual framework for the design of organizational control mechanisms. ***Management Science,*** 25(9): 833-848.

Patel, P., & Pavitt, K. 1994. Uneven (and Divergent) Technological Accumulation among Advanced Countries: Evidence and a Framework of Explanation. ***Industrial and Corporate Change,*** 3(3): 759–787.

Pavitt, K., 1984. Sectoral patterns of technical change: towards taxonomy and a theory. ***Research Policy,*** 13: 343–373.

Pavitt, K. 1985. Patent statistics as indicators of innovative activities: possibilities and problems. ***Scientometrics,*** 7(1/2): 77–99.

Pavitt, K. 1988. Uses and abuses of patent statistics. In A. F. J. van Raan (Ed.), ***Handbook of quantitative studies of science and technology.*** London: Elsevier Science Publishers.

Pawson, R., & Tilly, N. 1997. ***Realistic Evaluation,*** London: Sage.

Pedhazur, E. J., & Schmelkin, L. P. 1991. ***Measurement, design, and analysis: An integrated approach.*** Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Penrose, E. 1959. ***The theory of the growth of the firm.*** New York: Oxford University Press.

Perera, S., Harrison, G., & Poole, M. 1997. Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: a research note. ***Accounting, Organizations and Society,*** 22(6): 557–572.

Perrow, C. 1961. The analysis of goals in complex organizations. ***American Sociological Review,*** 26(6): 854–866.

Peters, T. J. 1992. ***Liberation management: Necessary disorganization for the nanosecond nineties.*** New York: Alfred A. Knopf.

Pfeffer, J. 1992. ***Managing with power.*** Boston, MA: Harvard Business School Press.

Pfeffer, J. 1993. Barriers to the advancement of organizational science: Paradigm development as a dependent variable. ***Academy of Management Review,*** 18(4): 599-620.

Pfeffer, J. 1995. Mortality, reproducibility and the persistence of styles of theory. ***Organization Science,*** 6(6): 681-686.

Pfeffer, J., & Salancik, G. 1978. ***The external control of organizations: A resource dependence perspective.*** New York: Harper & Row Publishers.

Pierce, J. L. & Delbecq, A. L. 1977. Organization structure, individual attitudes and innovation. ***Academy of Management Review,*** 2(1): 27-37.

Pinsonneault, A., & Kraemer, K. 1993. Survey research methodology in management information systems: An assessment. ***Journal of Management Information Systems,*** 10(2): 75-105.

Podsakoff, P. M., & Organ, D. W. 1986. Self-Reports in organizational research: Problems and prospects. ***Journal of Management,*** 12(4): 531-544.

Podsakoff, P. M., MacKenzie, S. B., & Lee, J. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. ***Journal of Applied Psychology,*** 88(5): 879-903.

Poole, M. S., Van de Ven, A. H., Dooley, K., & Colmes, M. E. 2000. ***Organizational change and innovation processes.*** Oxford, UK: Oxford University Press.

Popper, K. R. 1972. ***The logic of scientific discovery*** (Revised Edition), London: Hutchinson.

Popper, K. R. 1979. ***Objective knowledge: An evolutionary approach*** (Revised Edition), Cambridge, UK: Oxford University Press.

Porter, M. E. 1979. The structure within industries and companies performance, ***Review of Economics & Statistics***, 61(2): 214-228.

Porter, M. E. 1985. ***Competitive advantage: Creating and sustaining superior performance.*** , New York: The Free Press.

Porter, M. E. 1985. ***Competitive strategy: Techniques for analyzing industries and competitors.*** New York: The Free Press.

Priem, R. L., Rasheed, A. M. A., & Kotulic, A. G. 1995. Rationality in strategic decision processes, environmental dynamism and firm performance. ***Journal of Management***, 21(5): 913-929.

Pun, K. F., & White, A. S. 2005. A performance measurement paradigm for integrating strategy formulation: A review of systems & frameworks. ***International Journal of Management Reviews***, 7(1): 49-71.

Ramos, M. M., & Hidalgo, F. G. 2003. From diagnostic to interactive style of management control. ***Management Research News***, 26(5): 21-31.

Ramos, M. M. 2004. Changes in the logistics management style through performance indicators: A case study. ***International Journal of Logistics: Research and Applications***, 7(4): 313-324.

Ranaweera, M. A. C. 2002. ***Drivers of customer retention in service industries with continuous purchasing: Evidence from telecommunication and banking sectors.*** Unpublished Ph. D. dissertation, Cambridge University, UK.

Rigby, D. 2001. Management tools and techniques: A survey. ***California Management Review***, 43(2): 139-160.

Roberts, P. W. 1999. Product innovation, product-market competition and persistent profitability in the U.S. pharmaceutical industry. ***Strategic Management Journal,*** 20(7): 655–670.

Robertson, T. S., & Gatignon, H. 1998. Technology development mode: A transaction cost conceptualization. ***Strategic Management Journal,*** 19(6): 515-531.

Robson, S., & Ortmans, L. 2006. First findings from the UK innovation survey, 2005, ***Economic Trends No. 628,*** (March), UK: Office for National Statistics.

Rockness, H. O., & Shields, M. D. 1984. Organizational control systems in research and development. ***Accounting, Organizations and Society,*** 9(2): 165-177.

Rockness, H. O., & Shields, M. D. 1988. An empirical analysis of the expenditure budget in research and development. ***Contemporary Accounting Research,*** 4(2): 568-581.

Rogers, S. 1990. ***Performance Management in Local Government,*** London, UK: Longman.

Rolstadas, A. 1998. Enterprise performance measurement. ***International Journal of Operations & Production Management,*** 18(9/10): 989-999.

Rosen, R. 1991. Research and development with asymmetric company sizes. ***Rand Journal of Economics,*** 22: 411-429.

Rosenkopf, L., & Nerkar, A. 2001. Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry. ***Strategic Management Journal,*** 22(4): 287-306.

Rossiter, J. R. 2002. The C-OAR-SE procedure for scale development in marketing. ***International Journal of Research in Marketing,*** 19: 305-335.

Roth, A. V., & Jackson, W. E. 1995. Strategic determinants of service quality and performance evidence from the banking industry. ***Management Science,*** 41(11): 1720-1733.

Rothwell, R., & Zegveld, W. 1985. ***Reindustrialization and technology.*** Harlow, UK: Longman.

Rothwell, R. 1992. Successful industrial innovation: Critical factors for the 1990s. ***R&D Management,*** 22(3): 221-239.

Rothwell, R. 1994. Towards the fifth-generation innovation process. ***International Marketing Review,*** 11(1): 7-31.

Ruefli, T. W., Collins, J. M., & Lacugna, J. R. 1999. Risk measures in strategic management research: Auld Lang Syne? ***Strategic Management Journal,*** 20(2): 167-194.

Rumelt, R. P. 1991. How much does industry matter? ***Strategic Management Journal,*** 12(3): 167-185.

Rungtusanatham, M. J., Choi, T., Hollingworth, D. G., Wu, Z., & Forza, C. 2003. Survey research in operations management: Historical analyses. ***Journal of Operations Management,*** 21(4): 475-488

Salazar, M., & Holbrook, A. 2004. A debate on innovation surveys. ***Science and Public Policy,*** 13(4): 254-266

Santarelli, E., & Piergiovanni, 1996. Analyzing literature-based innovation output indicators: the Italian experience. ***Research Policy,*** 25(5): 689-711.

Savotti, P. P. 1996. ***Technological evolution, variety and the economy.*** Cheltenham, UK: Edward Elgar.

Scandura, T. A., & Williams, E. A. 2000. Research methodology in management: current practices, trends, and implications for future research, ***Academy of Management Journal,*** 43(6): 1248-1264.

Schiemann, W. A., & Lingle, J. H. 1999. ***Bullseye!: hitting your strategic targets through high-impact measurement.*** New York: The Free Press.

Schmalensee, R. 1985. Do Markets Differ Much? ***American Economic Review,*** 75(3): 341-352.

Schmidt, J. B., & Calantone, R. J. 1998. Are really new product development projects harder to shut down? ***Journal of Product Innovation Management,*** 15(2): 111-123.

Schneiderman, A. 1999. Why balanced scorecards fail. ***Journal of Strategic Performance Measurement,*** SI: 6-11.

Schoonhoven, C. B. 1981. Problems with contingency theory: Testing assumptions hidden within the language of contingency 'Theory'. ***Administrative Science Quarterly,*** 26(3): 349-377.

Schoonhoven, C. B., Eisenhardt, K. M., & Lyman, K. 1990. Speeding products to market: Waiting time to first product introduction in new firms. ***Administrative Science Quarterly,*** 35(1): 177-207.

Schulz, M. 1998. Limits to bureaucratic growth: The density dependence of organizational rule births. ***Administrative Science Quarterly,*** 43(4): 845-876.

Schulz, M. 2002. Organizational learning. In J. A. C. Baum (Ed.), ***Companion to Organizations:*** 415-441, Oxford, UK: Blackwell Publishers.

Schumacker, R. E., & Lomax, R. G. 2004. ***A beginner's guide to structural equation modeling.*** Mahwah, New Jersey: Lawrence Erlbaum Associates.

Scott, W. R. 1981. ***Organizations: Rational, Natural, and Open Systems.*** Upper Saddle River, New Jersey: Pearson Education.

Scott, W. R. 2003. ***Organizations: Rational, Natural, and Open Systems (Fifth Edition).*** Upper Saddle River, New Jersey: Pearson Education.

Senge, P. M. 1990. ***The fifth discipline: The art and practice of the learning organization.*** New York: Doubleday.

Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. ***Experimental and quasi-experimental designs for generalized causal inference.*** New York: Houghton Mifflin.

Shafer. J. L. & Graham, J. W. 2003. Missing data: Our view of the state of the art. ***Psychological Methods,*** 7(2): 147-177.

Shapira, Z. 1994. ***Risk taking: A managerial perspective.*** New York: Russell Sage Foundation.

Sharma, S., Durand, R. M., & Gur-Arie, O. 1981. Identification and analysis of moderator variables. ***Journal of Marketing Research,*** 18(3): 291–300.

Shneider, S. L. 1992. Framing and conflict: aspiration level contingency, the status quo, and current theories of risky choice. ***Journal of Experimental Psychology: Learning, Memory, and Cognition.*** 18(5): 1040-1057.

Shortell, S. M., & Zajac, E. J. 1990. Perceptual and archival measures of Miles and Snow's strategic types: A comprehensive assessment of reliability and validity. ***Academy of Management Journal,*** 33(4): 817-832.

Shrivastava, P. 1987. Rigor and practical usefulness of research in strategic management. ***Strategic Management Journal,*** 8(1): 77–92.

Siggelkow, N., & Rivkin, J. W. 2006. When exploration backfires: Unintended consequences of multi-level organizational search. ***Academy of management journal,*** 49(4): 779-795.

Silk, S. 1998. Automating the balanced scorecard. ***Management Accounting,*** 79(11): 38-44.

Simon, H. A. 1947. ***Administrative Behavior: A Study of Decision Making Processes in Administrative Organizations.*** New York: Macmillan.

Simon, H. A. 1955. A behavioral model of rational choice. ***Quarterly Journal of Economics***, 69(1): 99-118.

Simon, H. A. 1956. Rational Choice and the Structure of the Environment. ***Psychological Review***, 63(2): 129–38.

Simon, H. A. 1965. On the concept of organizational goal. ***Administrative Science Quarterly***, 9(1): 1-22.

Simon, H. A. 1967. Understanding creativity. In C. Gowan, G. D. Demos, & E. P. Torrance (Eds.), ***Creativity: Its educational implications:*** 3-29. New York: Wiley.

Simon, H. A. 1997. ***Models of Bounded Rationality, Volume 3. Empirically Grounded Economic Reason.*** Cambridge, Massachusetts: The MIT Press.

Simon, H. A., Guetzkow H., Kozmetsky G, & Tyndall G. 1954. ***Centralization Vs. Decentralization in Organizing the Controller's Department.*** New York: Controllership Foundation Inc.

Simons, R. 1987a. Accounting control systems and business strategy: an empirical analysis. ***Accounting, Organizations and Society***, 12(4): 357–374.

Simons, R. 1987b. Planning, control, and uncertainty: A process view. In W. J. Bruns & R. S. Kaplan (Eds.), ***Accounting & Management Field Study Perspectives:*** 339-362. Boston, MA: Harvard Business School Press.

Simons, R. 1990. The role of management control systems in creating competitive advantage: New perspectives. ***Accounting, Organizations and Society***, 15(1/2): 127–143.

Simons, R. 1991. Strategic orientation and top management attention to control systems. ***Strategic Management Journal***, 12(1): 49–62.

Simons, R. 1992. The Strategy of Control. ***CA Magazine***, March: 44-50

Simons, R. 1994. How new top managers use control systems as levers of strategic renewal. ***Strategic Management Journal***, 15(3): 169–189.

Simons, R. 1995. ***Levers of control: How managers use innovative control systems to drive strategic renewal.*** Cambridge, MA: Harvard Business School.

Simons, R. 2000. ***Performance measurement & control systems for implementing strategy***, New Jersey: Prentice Hall

Singh, J. V. 1986. Performance, slack, and risk taking in organizational decision making. ***Academy of Management Journal***, 29(3): 562-585.

Singh, J., Goolsby J. R., & Rhoads, G. K. 1994. Behavioral and psychological consequences of boundary spanning burnout for customer service representatives. ***Journal of Marketing Research***, 31(4): 558–569.

Singleton, R. A., & Straits, B. C. 2005. ***Approaches to social research (4th Edition).*** New York: Oxford University Press.

Sink, D. S. & Tuttle, T.C. 1989. ***Planning and measurement of in your organization of the future.*** Norcross, GA: Industrial Engineering and Management Press.

Slappendel, C. 1996. Perspectives on innovation in organizations. ***Organizational Studies***, 17(1): 107-129.

Slater, S. F., & Atuahene-Gima, K. 2004. Conducting survey research in strategic management. In D. J. Ketchen & D. D. Bergh (Eds.), ***Research methodology in strategy and management: Vol. 1***: 227-249, Oxford, UK: Elsevier.

Smith, K. 2005. Measuring innovation. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), ***The Oxford handbook of innovation:*** 148-177. Oxford, UK: Oxford University Press.

Smith, W. K., & Tushman, M. L. 2005. Managing strategic contradictions: A top management model for managing innovation streams. ***Organization Science*,** 16: 522-536.

Snow, C. C., & Hrebiniak, L. G. 1980. Strategy, distinctive competence, and organizational performance. ***Administrative Science Quarterly*,** 25(2): 317-336.

Sørensen, J. B., & Stuart, T. E. 2000. *Aging, Obsolescence, and Organizational Innovation.* ***Administrative Science Quarterly***, 45(1): 81-112.

Southwood, K. E. 1978. Substantive theory and statistical interaction: Five models. ***American Journal of Sociology*,** 83(5): 1154-1203.

Spaienza, H. J., Smith, K. G., & Gannon, M. J. 1988. Using subjective evaluations of organizational performance in small business research, American Journal of Small Business, ***Journal of Small Business*,** 12(3): 45-53.

Spanos, Y. E., & Lioukas, S. 2001. An examination into the causal logic of rent generation: Contrasting Porter's competitive strategy framework and the resource-based perspective. ***Strategic Management Journal*,** 22(10): 907-934.

Starbuck, W. H. 1992. Learning by knowledge intensive firms. ***Journal of Management Studies*,** 29(6): 713-740.

Starkey, K., & Madan, P. 2001. Bridging the relevance gap: Aligning stakeholders in the future of management research. ***British Journal of Management*,** 12(SI): S3-S26.

Staw, B. M., Sandelands, L. E., & Dutton, J. E. 1981. Threat-Rigidity Effects in Organizational Behavior: A Multi-Level Analysis. ***Administrative Science Quarterly***, 26(4): 501-524.

Sterman, J. D. 1989. Modeling managerial behaviour: Misperceptions of feedback in a dynamic decision making experiment. ***Management Science*,** 35(3): 321–339.

Sternberg, R. J., & Lubart, T. I. 1995. ***Defying the crowd: Cultivating creativity in a culture of conformity.*** New York: Free Press.

Sternberg, R. J., & Lubart, T. I. 1991. An investment theory of creativity and its development. ***Human Development, 34,*** 1-31.

Stevens, E., & Dimitriadisb, S. 2004. New service development through the lens of organisational learning: Evidence from longitudinal case studies. ***Journal of Business Research,*** 57(10): 1074-1185.

Stockdale, B. 2002. UK innovation survey 2001. ***Economic Trends No. 580,*** (March), UK: Office for National Statistics.

Stone-Romero, E. 2002. The relative validity and usefulness of various empirical research designs. In S. G. Rogelberg (Ed.), ***Handbook of research methods in industrial and organizational psychology:*** 77-98, Oxford, UK: Blackwell Publishing.

Stuart, T. E., & Podolny, J. M. 1996. Local search and the evolution of technological capabilities. ***Strategic Management Journal,*** 17(SI): 21-38.

Studenmund, A. H. 2001. Using econometrics: A practical guide. Boston, MA: Addison Wesley Longman.

Sundgren, M., Dimenäs, E., Gustafsson, J., & Selart, M. 2005. Drivers of organizational creativity: A path model of creative climate in pharmaceutical R&D. ***R & D Management,*** 35(4): 359-374.

Suppe, F 1977. ***The structure of scientific theories (2nd Edition).*** Urbana, IL: University of Illinois Press.

Tabachnick, B. G., & Fidell, L. S. 1996. ***Using multivariate statistics (3rd Ed.).*** New York: Harper-Collins.

Teece, D. J., Pisano, G., & Shuen, A. 1997. Dynamic capabilities and strategic management. ***Strategic Management Journal,*** 18(7): 509-533.

Thomas, K. W., & Tymon, W. G. 1982. Necessary properties of relevant research: Lessons from recent criticisms of the organizational sciences. ***Academy of Management Journal,*** 7(3): 345-352.

Thompson, J. D. 1967. ***Organizations in action.*** New York: McGraw-Hill.

Thomson, G., & Abernethy, M. A. 1998. ***Product innovation and management control system design: the concept of fit.*** Paper presented at the 4th international management control systems research conference. Reading, UK: July.

Thurstone, L. L. 1947. ***Multi-factor analysis.*** Chicago: University of Chicago Press.

Tidd, J., Bessant, J., & Pavitt, K. 2005 ***Managing innovation: Integrating technological, market and organizational change (3rd Ed.).*** Chichester, UK: John Wiley & Sons.

Tiger, L., Nicholls, J. A. F., & Sydney, R. 1999. The relationships between market-driven learning and new product success in export markets. ***International Marketing Review,*** 16(6): 476-503

Tippins, M. J., & Sohi, R. S. 2003. IT competency and firm performance: Is organizational learning a missing link? ***Strategic Management Journal,*** 24(8): 745-761.

Tranfield, D., & Starkey, K. 1998. The natural, social organization and promotion of management research: Towards policy. ***British Journal of management,*** 9(4): 341-353.

Tranfield, D. 2004. ***Nature of management research.*** Cranfield School of Management Presentation, Cranfield, UK: October.

Tranfield, D., Denyer, D., & Smart, P. 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. ***British Journal of Management,*** 14(3): 207-222.

Tripsas, M., & Gavetti, G. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. ***Strategic Management Journal,*** 21(10/11): 1147–1161.

Trott, P. 2002. ***Innovation management and new product development (2nd Ed.).*** Essex, UK: Financial times – Prentice Hall.

Tsai, W., & Ghoshal, S. 1998. Social capital and value creation: The role of intrafirm networks. ***Academy of Management Journal,*** 41(4): 464-476.

Tsikriktsis, N. 2005. A review of techniques for treating missing data in OM survey research. ***Journal of Operations Management,*** 24(1): 53-62.

Tucker, L. R., & Lewis, C. 1973. A reliability coefficient for maximum likelihood factor analysis. ***Psychometrika,*** 38(1): 1–10.

Tuomela, T. 2005. The interplay of different levers of control: A case study of introducing a new performance measurement system. ***Management Accounting Research,*** 16(3): 293-320.

Tushman, M. L., & O'Reilly, C. 1997. ***Winning through innovation.*** Boston, MA: Harvard Business Press.

Tushman, M. L., & Anderson, P. C. 1986. Technological discontinuities and organizational environments. ***Administrative Science Quarterly,*** 31(3): 439-465.

Tushman, M. L., & Smith, W. 2002. Organizational technology. In J. A. C. Baum (Ed.), ***Companion to Organizations.*** Oxford, UK: Blackwell Publishers.

Tushman, M. L., & Nadler, D. 1978. Information processing as an integrating concept in organizational design. Academy of Management Review, 3, 613-624.

Tushman, M. L., Smith, W., Wood, R., Westerman, G., & O'Reilly, C. 2004. ***Innovation streams and ambidextrous organizational designs: On building dynamic capabilities.*** Working paper no. 03-106, Harvard Business School, Boston, MA.

Utterback, J. M. 1971. The process of technological innovations within the firm.

Academy of management journal, 14(3): 75-88.

Van de Ven, A. H. 1986. Central problems in the management of innovation. ***Management Science,*** 32(5): 590-607.

Van de Ven, A. H. (2007, forthcoming). ***Engaged scholarship: A guide for organizational and social research.*** Oxford University Press.

Van de Ven, A. H., Angle, H. L., & Poole, M. S. (Eds.) 2000. ***Research on the management of innovation: The Minnesota studies.*** New York, Oxford University Press.

Van der Stede, W. A., Young, S. M., & Chen, C. X. 2005. Assessing the quality of evidence in empirical management accounting research: The case of survey studies. ***Accounting, Organizations and Society,*** 30(7/8): 655-684.

Van Maanen, 1995a. Fear and loathing in organization studies. ***Organization Science,*** 6(6): 687-692.

Van Maanen, 1995b. Style as a theory. ***Organization Science,*** 6(1): 132-143.

Vandenbosch, B. 1999. An empirical analysis of the association between the use of executive support systems and perceived organizational competitiveness. ***Accounting Organizations and Society,*** 24(1): 77–92.

Varadarajan, P. R. 2003. Musings on relevance and rigor in scholarly research in marketing. Journal of the Academy of marketing Science, 31(4): 368-376.

Venkatraman, N., & Ramanujam, V. 1986. Measurement of business performance in strategy research: A comparison of approaches. ***Academy of Management Review,*** 11(4): 801–814.

Venkatraman, N. 1989. The concept of fit in strategy research: Toward verbal and statistical correspondence. ***Academy of Management Review,*** 14(1): 423-444.

Vermeulen, F. 2005. On rigor and relevance: Fostering dialectic progress in management research. ***Academy of Management Journal***, 48(6): 978-982.

Verona, G. 1999. A resource-based view of product development. ***Academy of management review***, 24(1): 132-142.

Vincent, L. H., Bharadwaj, S. G., & Challagalla, G. N. 2005. ***Antecedents, consequences, and the mediating role of organizational innovation: Empirical generalizations,*** Working Paper Technological Innovation: Generating Economic Results (TI:GER®).

Von Hippel, E., & Tyre, M. 1995. How learning by doing is done: Problem identification in novel process equipment. ***Research Policy,*** 24(1): 1-12.

Wade, J. 1996. A communication-level analysis of sources and rates of technological variation in the microprocessors market. ***Academy of Management Journal,*** 39(5): 1218–1244.

Wall, T. D., Michie, J., Patterson, M., Wood, S. J., Sheehan, M., Clegg, C. W., & West, M. 2004. On the validity of subjective measures of company performance. ***Personnel Psychology,*** 57(1): 95-118.

Wallace, W. L. 1975. Structure and action in the theories of Coleman and Parsons. In P. M Blau. (Ed.), ***Approaches to the study of social structure:*** 121-134. New York: Free Press.

Wan, T. W. D., Ong, C. H., & Lee, W. S. F. 2000. The impact of firm characteristics of firm innovation. ***Management of Innovation and Technology,*** 16(1): 59-91.

Weber, M. 1947. ***The theory of social and economic organization.*** Translated by A. M. Henderson & T. Parsons, Glencoe, IL: The Free Press.

Webster, C. 2006. ***An empirical analysis of the relationships between the interactive use of performance measurement systems, creativity and***

performance: the intervening role of psychological empowerment. Working Paper, Monash University, Victoria, Australia.

Weerawardena, J., O'Cass, A., & Julian, C. 2006. Does industry matter? Examining the role of industry structure and organizational learning in innovation and brand performance, ***Journal of Business Research***, 59(1): 37-45.

Weick, K. E. 1979. ***The Social Psychology of Organizing (Second Edition).*** New York: McGraw-Hill Inc.

Werts, C. E., Linn, R. L., & Jöreskog, K. G. 1974. Intraclass reliability estimates: Testing structural assumptions. ***Educational and Psychological Measurement***, 34(1): 25-33.

West, S. G., Finch, J. F., & Curran, P. J. 1995. Structural equation modeling with non-normal variables. In R. Hoyle (Ed.) ***Structural equation modeling: Concepts, issues and applications:*** 56-75, Thousand Oaks, CA: Sage Publications.

Widener, S. K. 2005. ***An empirical analysis of levers of control framework.*** Working paper, Rice University, Houston, TX.

Wiersema, M. F., & Bowen, H. P. 1997. Matching method to paradigm in strategy reserach: Limitations of cross-sectional analysis and some methodological alternatives. In M. Ghertman, J. Obadia, & J. Arregle (Eds.), ***Staistical models for strategic management***: 201-21, Dordrecht, Netherlands: Kluwer Academic Publishers.

Winter, S. 1984. Schumpeterian competition in alternative technological regimes. ***Journal of Economic Behavior and Organizations***, 5(3/4): 287-320.

Wolfe, R. A. 1994. Organizational innovation: Review, critique and suggested research directions. ***Journal of Management Studies***, 31(3): 405-431.

Woodman, R. W., & Schoenfeldt, L. F. 1989. Individual differences in creativity: An interactionist perspective. In: J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), ***Handbook of Creativity***: 77-92, New York: Plenum.

Woodman, R. W., & Schoenfeldt, L. F. 1990. An interactionist model of creative behavior. ***Journal of Creative Behavior,*** 24(1): 10-20.

Woodman, R. W., Sawyer, J. E., & Griffin, R. W. 1993. Toward a theory of organizational creativity. ***Academy of Management Review***, 18(2): 293-321.

Woodward, J. 1958. ***Management and technology.*** London: HSMO.

Yaghi, B. 2005. ***Organizational learning from performance feedback and product innovations: A systematic review of the empirical evidence.*** Unpublished MRes. Dissertation, Cranfield School of Management, Cranfield University, UK.

Yelle, L. E. 1979. The learning curve: Historical review and comprehensive survey. ***Decision Sciences***, 10(2): 302–328.

Young, S. M. 1996. Survey research in management accounting: A critical assessment. In A. J. Richardson (Ed.), ***Research methods in accounting: Issues and debates***: 55-68. CGA: Canada Research Foundation.

Zahra, S. 1993a. New product innovation in established companies: associations with industry and strategy variables. ***Entrepreneurship: Theory & Practice,*** 18(2): 47–70.

Zahra, S. 1993b. Environment, corporate entrepreneurship and financial performance: A taxonomic approach. ***Journal of Business Venturing***, 8(4): 319-340.

Zahra, S. 1996. Governance, ownership, and corporate entrepreneurship: The moderating impact of industry technological opportunities. ***Academy of Management Journal,*** 39(6): 1713-1735.

Zaltman, G., Duncan, R., & Holbeck, J. 1973. ***Innovations and organizations.*** New York: John Wiley & Sons.

Zhou, X. 1993. The dynamics of organizational rules. ***American Journal of Sociology,*** 98(5): 1134-1166.

Zhou, Z. K., Yim, C. K., & Tse, D. K. 2005. The effects of strategic orientations on technology and market-based breakthrough innovations. ***Journal of Marketing***, 69(2): 42-60

Zollo, M., & Winter, S. G. 2002. Deliberate learning and the evolution of dynamic capabilities. ***Organization Science,*** 13(3): 339–351.

10 APPENDICES

10.1 Appendix A: Historical Review of Performance Measurement Research

Ghalayini & Noble (1986) observe that performance measurement has had two main phases. The first phase began in the late 1880s and went through the 1980s and the second phase started in late 1980s. The first phase started when the demand for information for internal planning and control arose in the first half of the 19th century as a result of the rise of firms (e.g., textile mills and railroads) that needed to manage multiple processes involved in the performance of its basic activity (Kaplan, 1984). Thus, the demand for management control systems that provided information about transactions occurring within the firms was to support the firms that have adopted multiple processes and hierarchies to manage activities that otherwise would have been transacted in the marketplace (Johnson & Kaplan, 1987). In the mid 1980s, Kaplan (1984) was claiming that virtually all of the management accounting practices employed today had been developed by 1925.

The first phase of performance measurement system was characterized by financially oriented measurement systems (Kaplan, 1986). By the 1980s, it was clear that these financially oriented performance measurement systems were increasingly failing to provide the accurate information on the efficiency and profitability of internally managed transactions (Johnson & Kaplan, 1987).

Neely (1998) reports that by the mid-1980s there were many vocal voices criticizing the traditional financial measures that started phase 2 in the evolution of performance measurement systems. Neely (2005) referred to this first stage of the second phase as the problem identification stage. Although accrual-based financial measures have served some firms fairly well before the 1980s, they were becoming now at best obsolete even more harmful (Eccles, 1991). Neely (1998) posited the question why now and gave seven reasons that contributed to the demise of their use:

1. The change in the nature of work
2. The increase in global competition
3. The implementation of quality improvement management techniques
4. The proliferation of national and international quality awards (e.g. Malcolm Baldridge National Quality Award in USA, European Foundation of Quality Management (RFQM) in Europe, Deming prize in Japan)
5. The change in organizational roles
6. The change in external demands
7. The increase in the power of information technology (IT).

These new demands on firms exposed the limitations of their financially based measurement systems. These limitations could be summarized as follows:

1. *Lagging metrics.* The income-based measures are better at better at measuring the consequences of past decisions than predicting future performance (Eccles, 1991; Kaplan & Norton 1992).
2. *Short-termism.* Financial measures encourage short-term thinking because short term profit measures will not signal a decrease in firms' value when the needed capital investments or investments in innovation are reduced (Hayes & Abernathy, 1980; Kaplan 1986). On the contrary, short-term financial profitability measures will increase when strategic investments are reduced (Kaplan, 1986). Consequently, the heavy focus on the quarterly financial reports by the firms and the financial investment community may pressure managers into manipulating the figures (Eccles, 1991). Neely (1998) concludes that short-termism encourages gaming.
3. *Cost distortion.* Maskell (1991) notes that the traditional financially oriented systems were based on cost elements that have changed drastically. The difference between fixed and variable costs and direct and indirect costs are blurring. Furthermore, traditional procedures for allocating overheads are distorting product costs.
4. *Inflexibility.* Traditional financially oriented measurement systems are inflexible because they produce standard reports across manufacturing

plants and divisions that do not reflect the different value drivers of the different operations (Maskell, 1991).

5. *Untimely.* Since financial reports are based on cost accounting, they are usually received too late to be of value (Maskell, 1991).
6. *Lack strategic focus.* Traditional financial measures lack strategic focus as they fail to provide information on quality, responsiveness, and flexibility (Neely, 1998).
7. *Lack relevance to practice.* Since financial measures try to quantify performance in financial terms, they cannot quantify many measures that are beneficial such as decreased new product launch times, product quality, and customer satisfaction (Ghalayini & Noble, 1996; Kaplan, 1986).
8. *Lack stakeholder information.* Traditional financial measures lack information on stakeholders other than the shareholders (Atkinson et al., 1997). They usually do not provide information on employees, customers, suppliers (Kaplan & Norton, 1992).

By the late 1980s, firms were becoming increasingly dissatisfied with their financially oriented performance measurement systems and they started incorporating non-financial measures in their performance measurement systems (Neely & Bourne, 2000). In 1991, Eccles declared that the performance measurement revolution has begun. Eccles explained that it is not the mere adoption of non-financial measures that constituted the revolution as many managers can claim that they have been doing this for sometime but the equal (or even greater) importance given to the non-financial measures in determining strategic decisions and rewards that constituted the revolution.

In the early to mid 1990s, the emphasis shifted from piecemeal approaches to integrated approaches to measuring performance given rise to the performance measurement frameworks (Neely & Bourne, 2000). Some of the performance measurement frameworks are as follows: Kaplan & Norton's Balanced Scorecard (1992; 1996; 2001; 2004); Neely, Adams, & Kennerly's Performance

Prism (2002).; Lynch & Cross' SMART Performance Pyramid (1991; 1992); Keegan, Eiler, Jones' Performance Measurement Matrix (1989); Fitzgerald, Johnston, Bringall, Silvestro, & Voss' Results and Determinants Framework (1991); Dixon, Nanni, & Vollmann's Performance measurement Questionnaire (PMQ) (1990); Bittici, Carrie, & McDevitt's Integrated Performance Measurement System (1997); Sink & Tuttle's Performance Model (1989); Rolstadas' TOPP Performance Model (1989).

By the late 1990s, the agenda shifted from measurement design to measurement implementation (Neely, 2005). The work now focused in finding the best methods and approaches for implementing performance measurement systems.

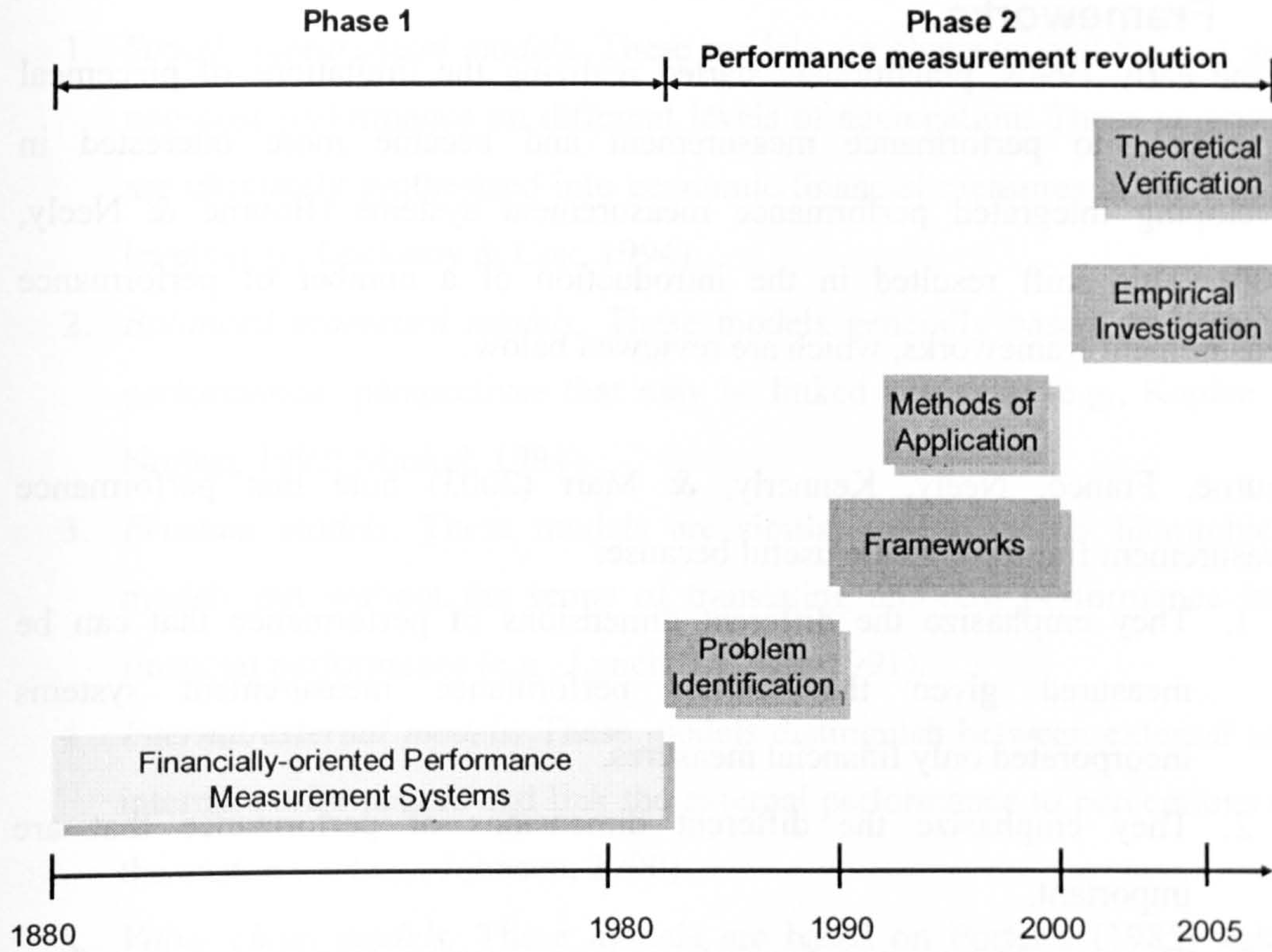
By the early 2000s, there was a huge uptake of performance measurement systems (Marr & Sciuma, 2003; Neely, 2005; Rigby, 2001; Silk, 1998) and the agenda now shifted to the quantification of the benefits of adopting such systems. This empirical investigation phase (Neely, 2005) resulted in inconsistent results. Many voices started now to question the theoretical foundations of the performance measurement systems especially the Balanced Scorecard (for example, Neely, Kennerly & Martinez, 2004; Nørreklit, 2000; 2003).

Based on Neely's (2005) recent review of performance measurement research, I constructed Table 10-1 to summarize the evolution of the major performance measurement research streams. It defines the timeline, salient characteristics of each research stream, and representative studies in it. Additionally, Figure 10-1 illustrates the evolution of the field of performance measurement based on Neely's (2005) evolution cycle.

Table 10-1: Performance measurement evolution and the development of performance measurement research streams (Source: Neely, 2005)

Research Stream	Timeline	Description	Representative Studies
Problem identification	1980s	This research stream highlighted the problems associated with financial-based performance measurement systems. It highlighted the need for new systems that incorporated non-financial measures because economies have evolved from being asset-based to knowledge-based.	Eccles, 1991; Hayes & Abernathy, 1980; Johnson & Kaplan, 1987
Frameworks	Early 1990s	In this research stream, solutions were offered to meet the challenges of new measurement requirements. These solutions resulted in numerous frameworks	Neely et al., 1995; Neely, Mills, Gregory, Richards, Platts, & Bourne, 1996; Kaplan & Norton, 1992, 1996, 2001
Methods of applications	Late 1990s	In this research stream, many methods of implementing, using and updating performance measurement systems were proposed.	Bourne, Mills, Wilcox, Neely, & Platts, 2000; Bourne, Neely, Mills, & Platts, 2003; Kennerley & Neely, 2003; Neely, Bourne, & Kennerly, 2000; Kennerly & Neely, 2002
Empirical investigation	2000s	Given that there were huge uptake of performance measurement systems, this stream sought to quantify the benefit of these systems and analyse the contingency factors leading to its success.	Banker, Potter, Srinivasan, 2000, 2005; Hoque & james, 2000; Ittner, Larcker, & Randall, 2003
Theoretical verification	2000s	Given the mixed results from the empirical investigations, some researchers are examining the theoretical foundations of some of the well-known frameworks (such as, Balanced Scorecard).	Nørreklit, 2000, 2003; Neely, Kennerly & Martinez, 2004

Figure 10-1: Performance measurement research streams: Mapping the field (Source: Adapted from Neely, 2005)



10.2 Appendix B: Review of Performance Measurement Frameworks

In the early 1990s, practitioners started realizing the limitations of piecemeal approaches to performance measurement and became more interested in developing integrated performance measurement systems (Bourne & Neely, 2000). This shift resulted in the introduction of a number of performance measurement frameworks, which are reviewed below.

Bourne, Franco, Neely, Kennerly, & Marr (2003) note that performance measurement frameworks are useful because:

1. They emphasize the different dimensions of performance that can be measured given that earlier performance measurement systems incorporated only financial measures.
2. They emphasize the different dimensions of performance that are important.
3. They provide structure for performance measures. This is important given the limited number of measures managers could remember. Thus by creating categories more measures could be incorporated. Performance categories also allow linkages to be made between them.

They emphasize that performance frameworks do not serve as an end in themselves but as tools to help emphasize, categorize, and communicate performance and they generally do not specify what measures should be chosen as these need to be chosen to meet the requirements of the adopting organization.

Performance measurement frameworks use different approaches to achieve these objectives. Some performance measurement researchers categorized these frameworks based on their historical development (e.g., Lawrie & Cobbold, 2004; Neely, Marr, Roos, Pike & Gupta, 2003), architecture (e.g., De Toni & Tonchia, 2001), or strategic linkages (Ittner, et al., 2003). One useful classification based on the architecture of the framework is De Toni & Tonchia's

(2001) classification of performance frameworks into five typologies as illustrated in Figure 10-2 and explained as follows:

1. *Strictly hierarchical models.* These models are characterized by cost and non-cost performance on different levels of aggregation. These measures are ultimately synthesized into economic/financial measures at the higher levels (e.g., Lockamy & Cox, 1994).
2. *Balanced scorecard models.* These models generally based on separate performance perspectives that may be linked vertically (e.g., Kaplan & Norton, 1992; Maskel, 1991)
3. *Frustum models.* These models are similar to the strictly hierarchical models but without the scope of translating non-cost performance into financial performance (e.g., Lynch & Cross, 1991).
4. *Internal/external models.* These models distinguish between external and internal performances and link the external performance to perceptions of the customers (e.g., Johnson, 1990).
5. *Value chain models.* These models are based on Porter's (1985) value concept (e.g., Sink & Tuttle, 1989).

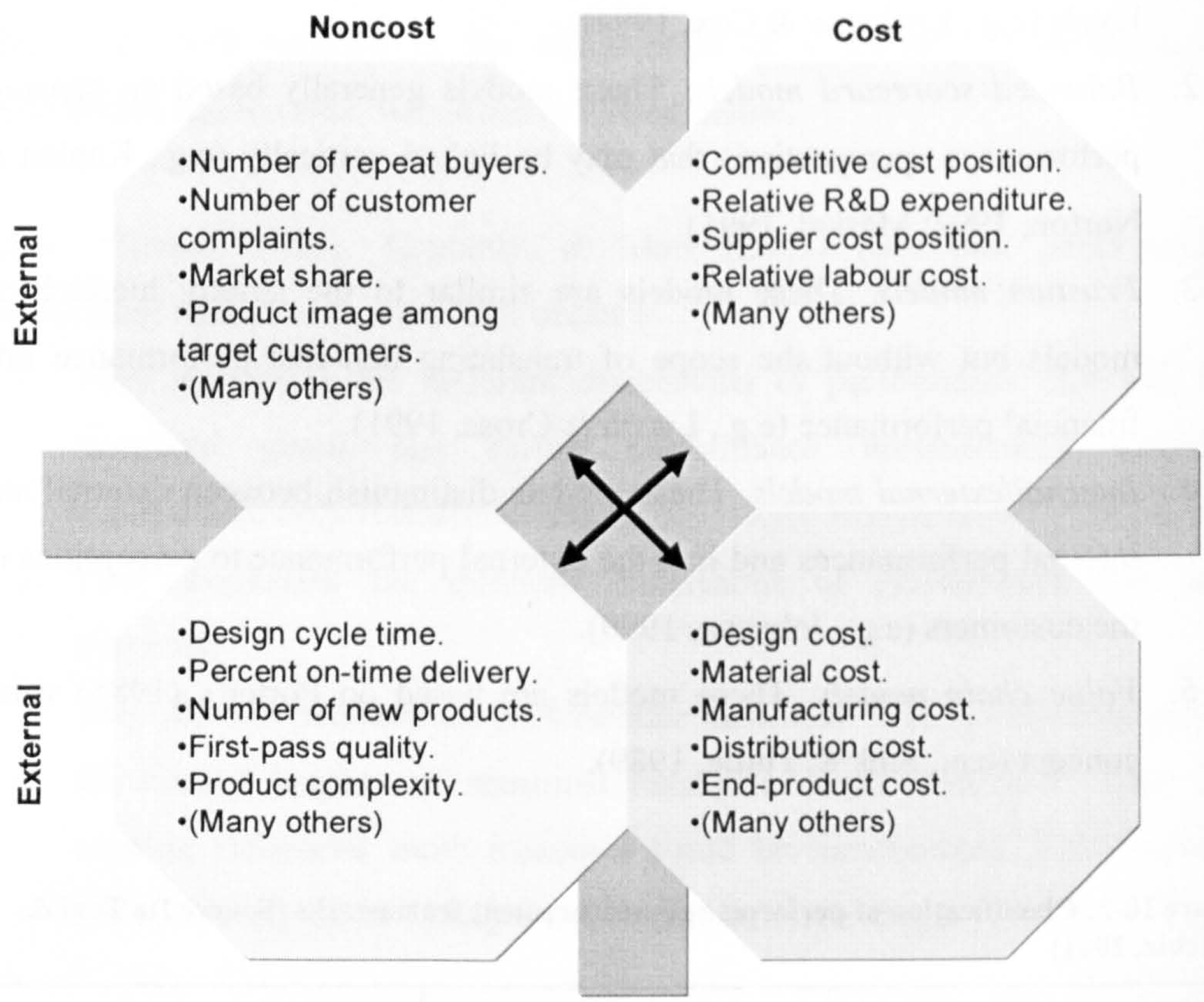
Figure 10-2: Classification of performance measurement frameworks (Source De Toni & Tonchia, 2001)

<table>
<tr><td>Architecture Vertical</td><td>Strictly hierarchical models</td><td></td><td rowspan="2">Frustum models</td><td colspan="2"></td></tr>
<tr><td>Architecture Balanced</td><td rowspan="2"></td><td>Balanced scorecard models</td><td rowspan="2">Internal/external models</td><td></td></tr>
<tr><td>Architecture Horizontal (By process)</td><td colspan="2"></td><td>Value chain models</td></tr>
</table>

10.2.1 Performance Measurement Matrix

Keegan, Eiler, Jones (1989) developed the performance measurement matrix to support the companies' multidimensional environments. This framework is also called the supportive performance measures framework. It incorporates four different types of performance measures: external, non-external, cost and non-cost. These performance measures were placed in 2X2 matrix consisting of two dimensions: cost/non-cost and external/internal as illustrated in Figure 10-3.

Figure 10-3: Performance measurement matrix (Source: Keegan, Eiler, Jones, 1989)



10.2.2 Performance Measurement Questionnaire (PMQ)

Dixon, Nanni, & Vollmann, (1990) developed the performance measurement questionnaire (PMQ). They state three objectives for developing PMQ: (1) to provide a means by which an organization can articulate its improvement needs; (2) to determine the extent to which its existing set of measurements is supportive of the necessary improvements; and (3) to establish an agenda for improving the measures so they better support the achievement of the improvements.

The PMQ consists of four parts. The first part profiles the organization. The second part focuses on competitive priorities and performance measurement systems (See Figure 10-4 for an illustration). The third part focuses on generic performance factors (performance measures). Since these performance measures are generic, some of them may not be in use by some organizations (See Figure 10-5 for an illustration). The fourth part is titled personal performance metrics and asks respondents to record their perceptions of the most important measures against which their individual performance is judged in each of five periods: daily, weekly, monthly, quarterly, and annually.

After the PMQ has been administered, the results are analysed in four different ways:

Alignment analysis. It is conducted to investigate the extent to which a company's strategy, actions, and measures line up with each other. It provides a general overview of the consistency between strategy, actions, and measures.

Congruence analysis. It provides a detailed look at how well the performance measurement system supports and organization's actions and strategies. It is done by examining the difference between the left-side response and the right-side response for each item in part II and III. If the left side of an item is higher than the right side, it signals the existence of gap that calls for the increased support from the performance measurement in that improvement area.

Consensus analysis. It is conducted to contrast the perceptions between hierarchical levels and across functional organizations.

Confusion analysis. It is conducted to determine the relative extent of consensus in opinions on each improvement area and their corresponding performance measures.

Figure 10-4: Performance measurement quessionaire : Part II (Source: Dixon, Nanni, Vollmann, 1990)

Long-Run Importance of Improvement	Improvement Areas	Effect of Current Performance Measures on Improvement
None >>> Great		Inhibit >>> Support
1 2 3 4 5 6 7	Quality	1 2 3 4 5 6 7
1 2 3 4 5 6 7	Labor efficiency	1 2 3 4 5 6 7
1 2 3 4 5 6 7	Machine efficiency	1 2 3 4 5 6 7

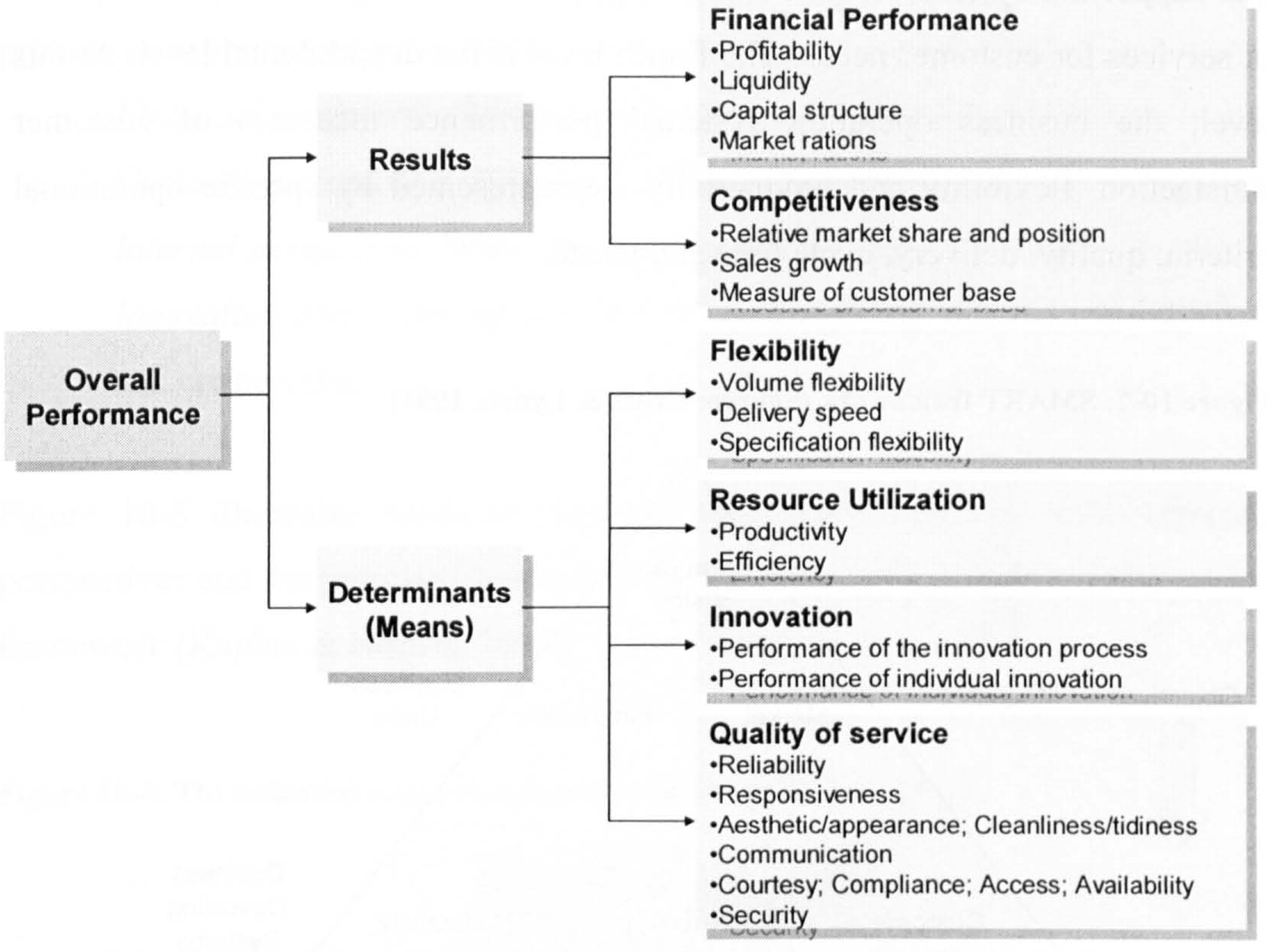
Figure 10-5: Performance measurement quessionaire : Part III (Source: Dixon, Nanni, Vollmann, 1990)

Relative Importance to the Company	Performance Factors	Emphasis of Measurement
Very Unimportant >>> Very Important		No Emphasis >>> Major Emphasis
1 2 3 4 5 6 7	Inventory turnover	1 2 3 4 5 6 7
1 2 3 4 5 6 7	Conformance to specifications	1 2 3 4 5 6 7
1 2 3 4 5 6 7	Cost of quality	1 2 3 4 5 6 7

10.2.3 Results and Determinants Framework

Fitzgerald, Johnston, Bringall, Silvestro, & Voss (1991) developed the results and determinants framework based on their research in the service industry. This framework group performance measures into two categories: results and determinants (means) as illustrated in Figure 10-6. The results category consists of two subcategories: financial performance and competitiveness. The determinants category consists of four subcategories: flexibility, resource utilization, innovation, and quality of service.

Figure 10-6: Results and determinants framework (Fitzgerald, Johnston, Bringall, Silvestro, & Voss, 1991)



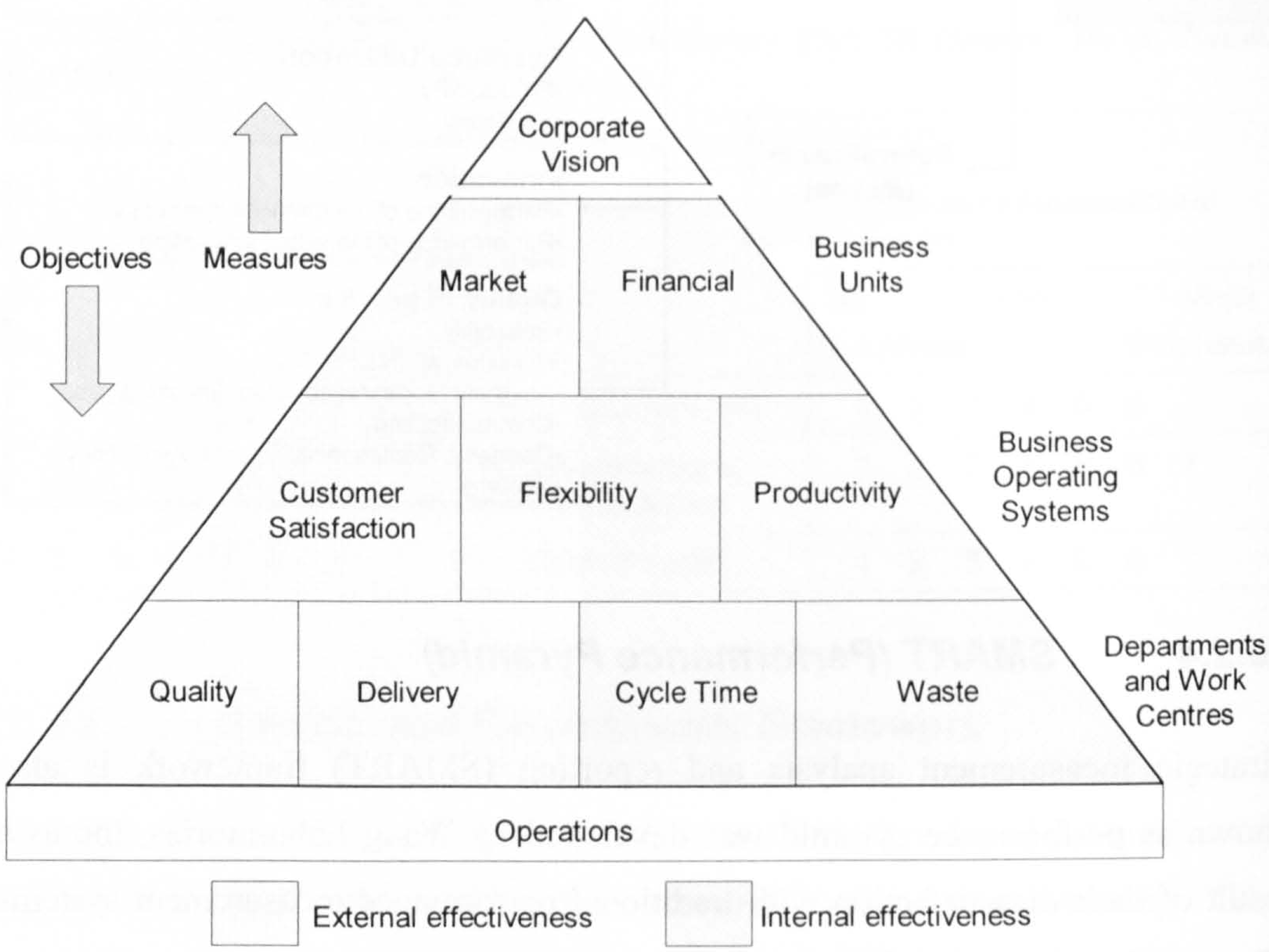
10.2.4 SMART (Performance Pyramid)

Strategic measurement analysis and reporting (SMART) framework is also known as performance pyramid was developed by Wang Laboratories, Inc as a result of their dissatisfaction with traditional performance measurement systems (Cross & Lynch, 1989; 1992; Lynch & Cross, 1991).

As illustrated in Figure 10-7, the SMART framework consists of four hierarchical levels with their associated objectives and performance measures making up a pyramid of measures. The organizational unit at the top is corporate. Corporate determines the corporate vision or strategy and manages the organization as a portfolio of business units. The second level is business unit where objectives are defined in terms of short-term financial results and long-term market results. The next level is business operating systems that bridge the gap between top-level and day-to-day operational measures (customer satisfaction, flexibility, productivity).

These systems include all internal functions, activities, policies, and procedures, and supporting systems needed to develop, produce, and provide specific goods or services for customer needs. The fourth level is the departmental level. At this level, the business operation systems' performance measures of customer satisfaction, flexibility and productivity are represented by specific operational criteria: quality, delivery, cycle time and waste.

Figure 10-7: SMART framework (Source: Cross & Lynch, 1991)



10.2.5 Balanced Scorecard

Kaplan & Norton (1992) introduced the concept of balanced scorecard, which was based on the work carried out by Analog Devices in the late 1980s (Schneiderman, 1999). Balanced scorecard is recognized as the most widely used performance measurement framework (Neely, 2005; Rigby, 2001; Silk, 1998) and it has gone through several iterations (Kaplan & Norton, 1996; 2001). The balanced scorecard emphasizes the diversity of performance measures (for example; lag versus predictive; outcome versus activity; financial versus non-financial) and it groups performance measures into four perspectives: financial, customer, internal, and innovation and learning. The four perspectives aim to provide answers to four basic questions:

Financial perspective. How do we look to our shareholders?

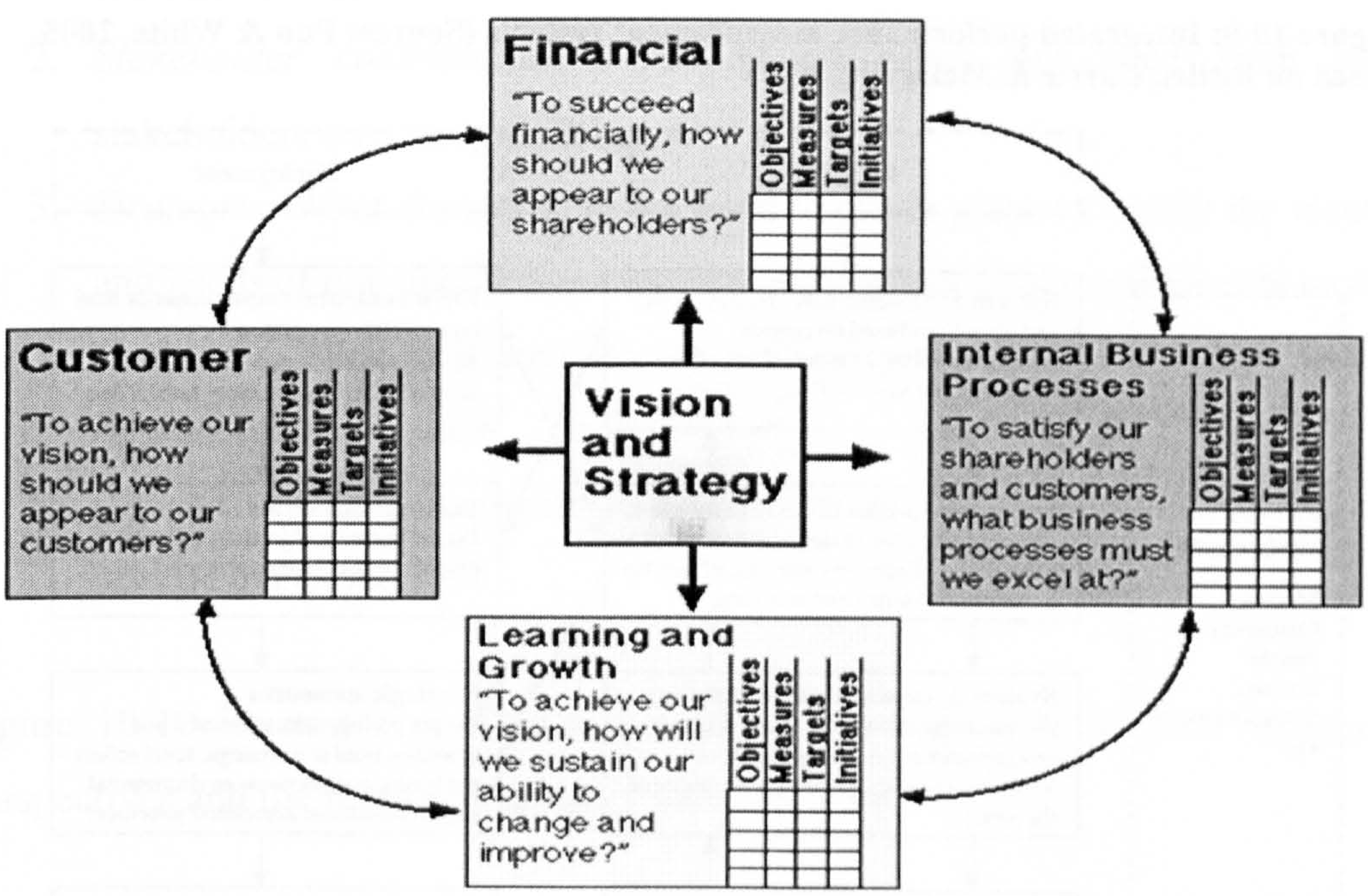
Customer perspective. How do we look to our customers?

Internal perspective. What we must excel at?

Innovation and Learning perspective. How can we continue to improve and create value?

Figure 10-8 illustrates balanced scorecard framework with its four different perspectives and the proposed linkages between them in the second version of the framework (Kaplan & Norton, 1996).

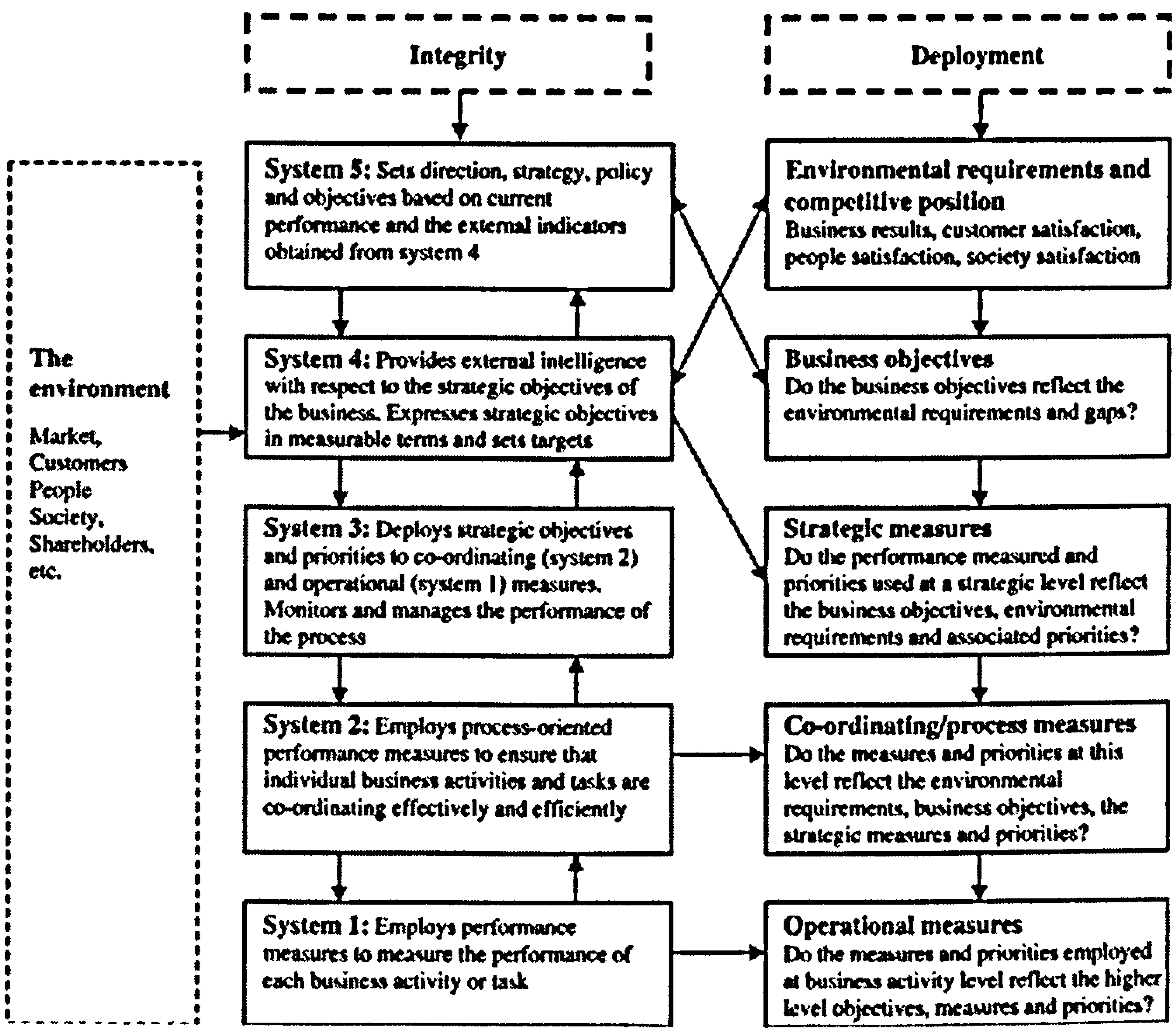
Figure 10-8: The balanced scorecrad (Source: Kaplan & Norton, 1996)



10.2.6 Integrated Performance Measurement System

The integrated performance measurement system (IPMS) was developed by researchers at the University of Strathclyde's manufacturing systems group (Bittici, Carrie, & McDevitt's, 1997). IPMS is based on two facets of the performance measurement systems: integrity of the system and deployment. Integrity refers to the ability of the performance measurement system to promote integration between various areas of the business. The concept of integrity is based on the viable systems model (Beer, 1979; 1981; 1985). The viable system model (VSM) states that for any system to be viable it must have five systems as illustrated in Figure 10-9. In addition, it must have a meta system which is a combination of systems 3, 4, and 5. The meta system is responsible for identifying and implementing change. Integrity is complemented by deployment which refers to the deployment of business objectives and policies throughout the hierarchical structure of the organization as illustrated in Figure 10-9. The purpose of the deployment component of the model is to ensure that performance measures used at various levels of the organization reflect the business objectives and policies through the hierarchy of the organization.

Figure 10-9: Integrated performance measurement systems (Source: Pun & White, 2005 based on Bittici, Carrie & McDevitt, 1997)



10.2.7 Performance Prism

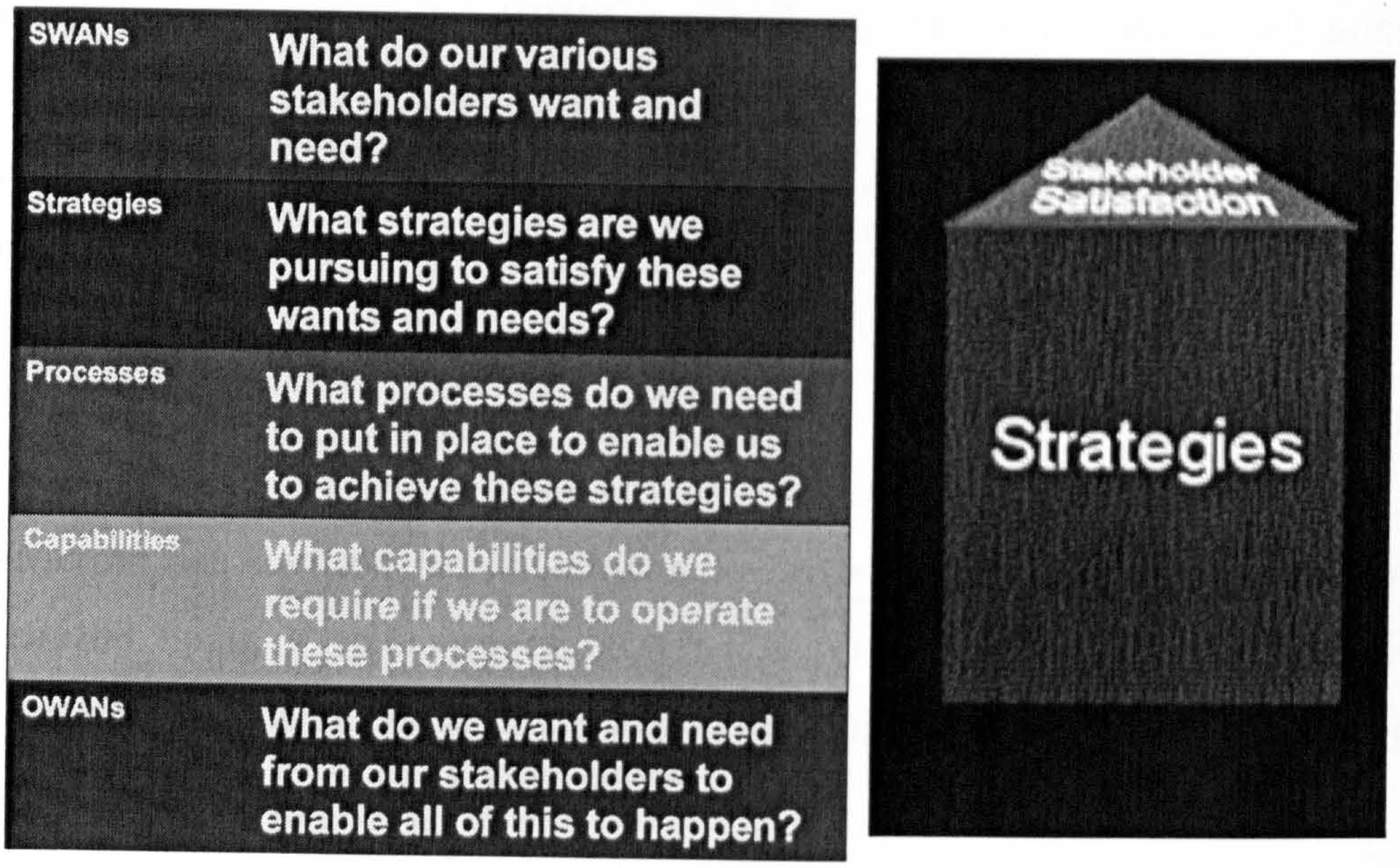
The performance prism a framework developed jointly by Accenture and the Centre for business Performance at Cranfield School of management (Neely, Adams, & Kennerly, 2002). The performance prism incorporates multi-stakeholder approach in its design and the critical elements of the strategy (success map). Therefore, it extends and improves on the concept of balanced scorecard that was questioned for its questionable assumptions and logical flow (Nørreklit, 2000; 2003; Neely, Kennerly & Martinez, 2004).

The performance prism consists of five interrelated perspectives and provides a comprehensive framework that can be used to articulate a given business's operating model. The five perspectives pose specific vital questions as follows:

1. *Stakeholder satisfaction.* Who are our key stakeholders and what do they want and need?
2. *Stakeholder contribution.* What do we want and need from our stakeholders on a reciprocal basis?
3. *Strategies.* What strategies do we need to put in place to satisfy the wants and needs of our stakeholders while satisfying our own requirements too?
4. *Processes.* What processes do we need to put in place to enable us to execute our strategies?
5. *Capabilities.* What capabilities do we need to put in place to allow us to operate our processes?

Figure 10-10 illustrates performance prism framework with its five different perspectives and the proposed linkages between them.

Figure 10-10: The performance prism (Source: Neely, Adams, & Kennerly, 2002)



10.3 Appendix C: Cover Letter and Mail-based Survey Instrument

In this appendix, I present the cover letter that was mailed with survey (It been reduced to fit in the page) followed by the scales used in this study.

Cranfield Survey: Impact of performance measurement and management systems

Survey objectives

The main objective of this survey is to study the *positive and negative effects* of performance measurement and management systems on the way organisations perform. It also aims to analyse the *factors* that influence the effects of such systems.

In recognition of the importance of this research issue to UK industry, the government through the Engineering and Physical Science Research Council **(EPSRC)** and the Department of Trade and Industry **(DTI)** supports this research [under the research grant:GR/S28846]. Hence, UK private industry is the focal point of this research, particularly the manufacturing, retail and services sectors.

Benefits for your organisation

By participating in this survey, your organisation will obtain a free copy of the study report. In addition, the first 100 organisations returning the completed survey will get a free ticket for the conference where the results of this research will be presented. It is important to mention that this is the first survey in the UK focused on the study of the positive and negative effects of performance measurement and management systems. *You responses will be treated in strict confidentiality.*

By participating in this study you will contribute to the science of understanding performance management and hence improve the ability of your company to promote desired business outcomes.

Survey instructions

1) Please fill out the questionnaire - the questions are simple and self-explanatory. This will take you approximately *15mins*.
2) If your company is part of a multinational, please answer this questionnaire according to your UK company's experiences/performance.
3) Return the completed questionnaire to the Centre for Business Performance in the pre-coded, pre-paid envelope provided. Address: to the Centre for Business Performance; Cranfield School of Management; Cranfield; Bedfordshire; MK43 0AL
4) The deadline to return this survey is the: **Feb 24, 2006**
5) If you have any queries please call us at 01234 75 1122 ext 2923, and contact Dr. Veronica Martinez or Dr. Mike Kennerley.

About our research centre

The Centre for Business Performance (CBP) at Cranfield School of Management has extensive experience in performance measurement and management systems. The centre is a leading research institution in this field and supports the dissemination of performance measurement and management practices and innovations through two business round tables and a biennial conference. Its close relationship with industry allows CBP research to focus on current problems that UK organisations are experiencing.

EPSRC Engineering and Physical Sciences Research Council

*Please return this questionnaire before **24th Feb 2006** to the Centre for Business Performance; Cranfield School of Management; Cranfield; Bedfordshire; MK43 0AL in the pre-paid enveloped supplied. Or by fax: 01234 75 7409/e-mail: v.martinez@cranfield.ac.uk*

Company name:	
Respondent's name:	Respondent's job title:

My organisation is part of a multinational: ☐ Yes ☐ No

If yes, please answer this questionnaire according to your UK business unit's experiences/performance.

Employee size

How many employees does your organisation employ in the UK?

☐ Fewer than 500 ☐ 500 – 999 ☐ 1,000 – 1,499 ☐ 1,500 – 1,999

☐ 2,000 – 2,999 ☐ 3,000 – 9,999 ☐ More than 10,000

Company's age

My organisation's age is: ______________________ years

Focus attention

Please rate the extent to which your top management team currently uses performance measures to…

Please mark one response for each line	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Integrate the organisation- i.e. tie the organisation together	☐	☐	☐	☐	☐	☐
Enable the organisation to focus on common issues	☐	☐	☐	☐	☐	☐
Enable the organisation to focus on your critical success factors	☐	☐	☐	☐	☐	☐
Develop a common vocabulary in the organisation	☐	☐	☐	☐	☐	☐
Provide a common view of the organisation	☐	☐	☐	☐	☐	☐
Enable discussion in meetings of superiors, subordinates and peers	☐	☐	☐	☐	☐	☐
Enable continual challenge and debate underlying results, assumptions and action plans	☐	☐	☐	☐	☐	☐

Monitoring

Please rate the extent to which your top management team currently uses performance measures to...

Please mark one response for each line	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Track progress towards goals	☐	☐	☐	☐	☐	☐
Review key measures	☐	☐	☐	☐	☐	☐
Monitor results	☐	☐	☐	☐	☐	☐
Compare outcomes to expectations	☐	☐	☐	☐	☐	☐

Diversity of measures

Please rate the following items...

	*To what extent does your top management team **use** the following categories to measure:*					
Please mark one response for each line	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Short term financial results- e.g. operating income, sales growth,etc	☐	☐	☐	☐	☐	☐
Customer relations- e.g. market share, customer satisfaction, etc.	☐	☐	☐	☐	☐	☐
Employee relations- e.g. employee satisfaction, safety, etc	☐	☐	☐	☐	☐	☐
Operational performance- e.g. productivity, lead times, etc.	☐	☐	☐	☐	☐	☐
Quality – e.g. quality performance, defect rates, etc.	☐	☐	☐	☐	☐	☐
Innovation and learning- e.g. number of new products and /or services launched, training, etc.	☐	☐	☐	☐	☐	☐
Supplier relations- e.g. on-time delivery, suppliers' integration etc	☐	☐	☐	☐	☐	☐
Alliances- e.g. joint marketing, joint product designs, etc.	☐	☐	☐	☐	☐	☐
Environmental performance- environmental compliances, etc.	☐	☐	☐	☐	☐	☐
Community- public image, community involvement, etc.	☐	☐	☐	☐	☐	☐

Product/ Service Innovation

A product innovation is the market introduction of a *new* good or service or a *significantly* improved good or service. The innovation must be new to your enterprise, but it does not need to be new to your market. It does not matter if the innovation was originally developed by your enterprise or by other enterprises.

During the two-year period 2004-2005, to what extent did your organisation introduce...

Please mark one response for each line	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Significantly improved goods or services	☐	☐	☐	☐	☐	☐
New goods or services onto your market before your competitors *(New to your market)*	☐	☐	☐	☐	☐	☐
New goods or services that was essentially the same as a product already available from your competitors in your market *(Only new to your company)*	☐	☐	☐	☐	☐	☐
The percentage of new products or services significantly in your portfolio	☐	☐	☐	☐	☐	☐

Product Innovation Activities

During the two-year period 2004-2005, to what extent did your organisation engage in the following innovation activities?

Please mark one response for each line	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Intramural (in-house) R&D - Creative work undertaken within your enterprise on an occasional or regular basis to increase the stock of knowledge and its use to devise new and improved goods, services and processes	☐	☐	☐	☐	☐	☐
Acquisition of R&D (extramural R&D) - Same activities as above, but purchased by your enterprise and performed by others.	☐	☐	☐	☐	☐	☐
Acquisition of machinery, equipment and software -Acquisition of advanced machinery, equipment and computer hardware or software to produce new or significantly improved goods and services.	☐	☐	☐	☐	☐	☐
Acquisition of external knowledge - Purchase or licensing of patents and non-patented inventions, know-how, and other types of knowledge from other enterprises or organisations	☐	☐	☐	☐	☐	☐
Training - Internal or external training for your personnel specifically for the	☐	☐	☐	☐	☐	☐

development and/or introduction of innovations						
All forms of Design - Expenditure on design functions for the development or implementation of new or improved goods, services and processes. Expenditure on design in the R&D phase of product development should be excluded.	☐	☐	☐	☐	☐	☐
Market introduction of innovations - Activities for the market preparation and introduction of new or significantly improved goods and services, including market research and launch advertising	☐	☐	☐	☐	☐	☐

Risk Taking

18. My company has an operating top management philosophy of:

Please mark one response for each line	**To a very great extent**	**To a great extent**	**To some extent**	**To a little extent**	**To a very little extent**	**Not at all**
Strong emphasis on research and development, technological leadership, and innovations	☐	☐	☐	☐	☐	☐
Strong proclivity to high risk, high return investments	☐	☐	☐	☐	☐	☐
Growth strategy primarily through external financing (borrowings, capital issues, etc.)	☐	☐	☐	☐	☐	☐
Very competitive, "undo-the-competitors" philosophy	☐	☐	☐	☐	☐	☐

*Please return this questionnaire before **24th Feb 2006** to the Centre for Business Performance; Cranfield School of Management; Cranfield; Bedfordshire; MK43 0AL in the pre-paid enveloped supplied. Or by fax: 01234 75 7409/e-mail: v.martinez@cranfield.ac.uk*

10.4 Appendix D: Cover Letter and Web-based Survey Instrument

In this appendix, I present the cover letter that was emailed with survey followed by the scales used in this study.

Innovation & Measurement: What works and what does not

Dear <Title Name>,

Although we know that all organisations use performance measures to manage some aspects of their operations and they are increasingly adopting performance measurement and management frameworks, such as Balanced Scorecard, EVA, Performance Prism, and Business Excellence Model, surprisingly, we know very little about their impact on organisational innovativeness. Do performance measurement systems enable or hinder organisational innovativeness? What are the performance measurement practices that enable innovations and what are the ones that don't and how they impact organisational innovativeness? This study by the leading research authority on business performance measurement seeks to answer these questions. It is hoped that the insights from this study will enable organisations become more innovative.

Your organization was one of few firms that were randomly selected to participate in this survey. Your participation is crucial to the success of this study. It will take approximately 10-15 minutes to complete the web-based survey. Please access the survey at
http://www.performancemeasurementsurvey.co.uk

However, if you prefer to receive an Excel or PDF copy of the survey, so you could complete it offline, please let me know.

By participating in this survey, you will receive a free copy of the study report. Your response will be treated in strict confidentiality and it will not be analysed individually. Only aggregated data will be used and reported.

I look forward to your participation. If you have any questions, please do not hesitate to contact me at your convenience.

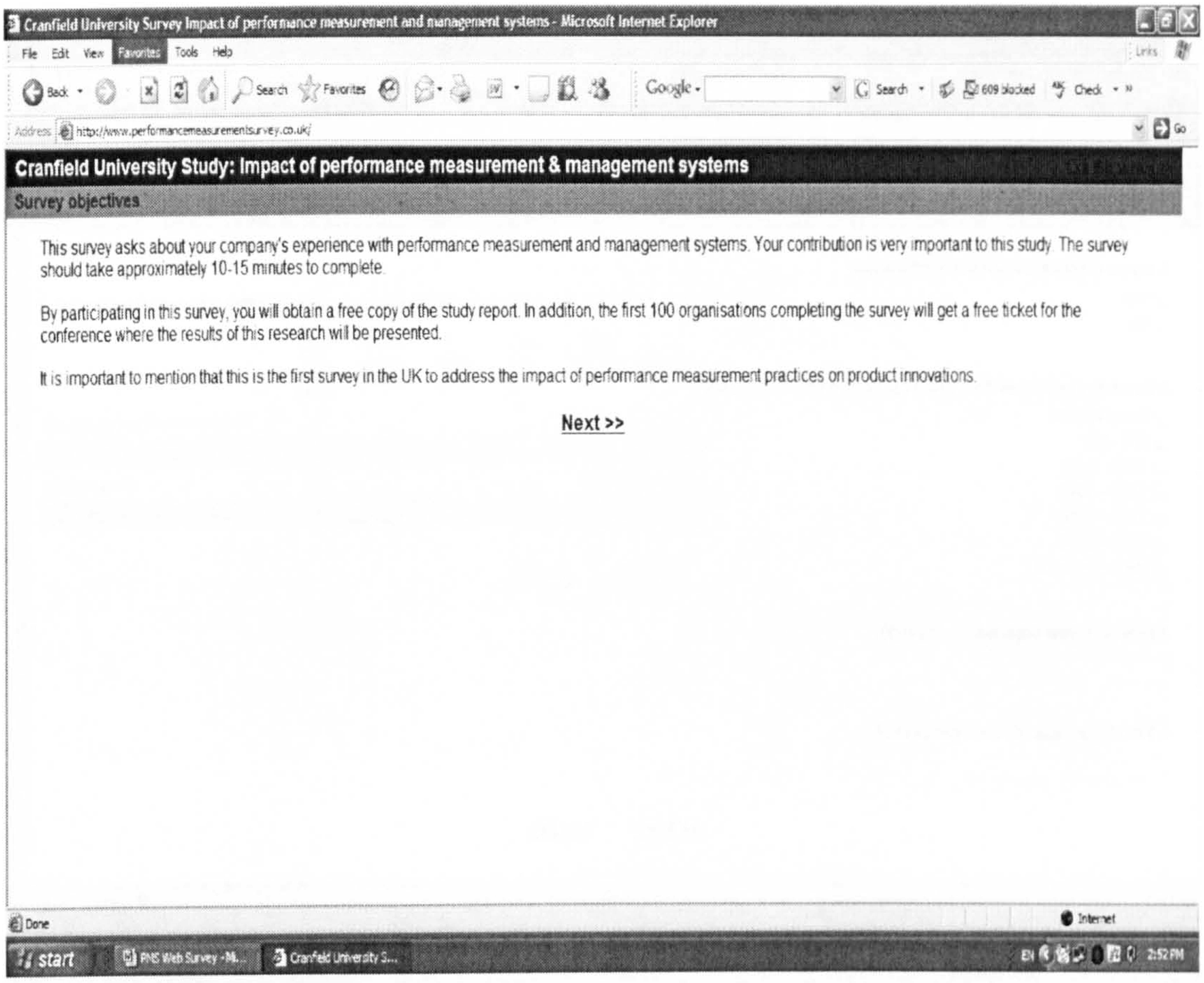
Sincerely Yours,

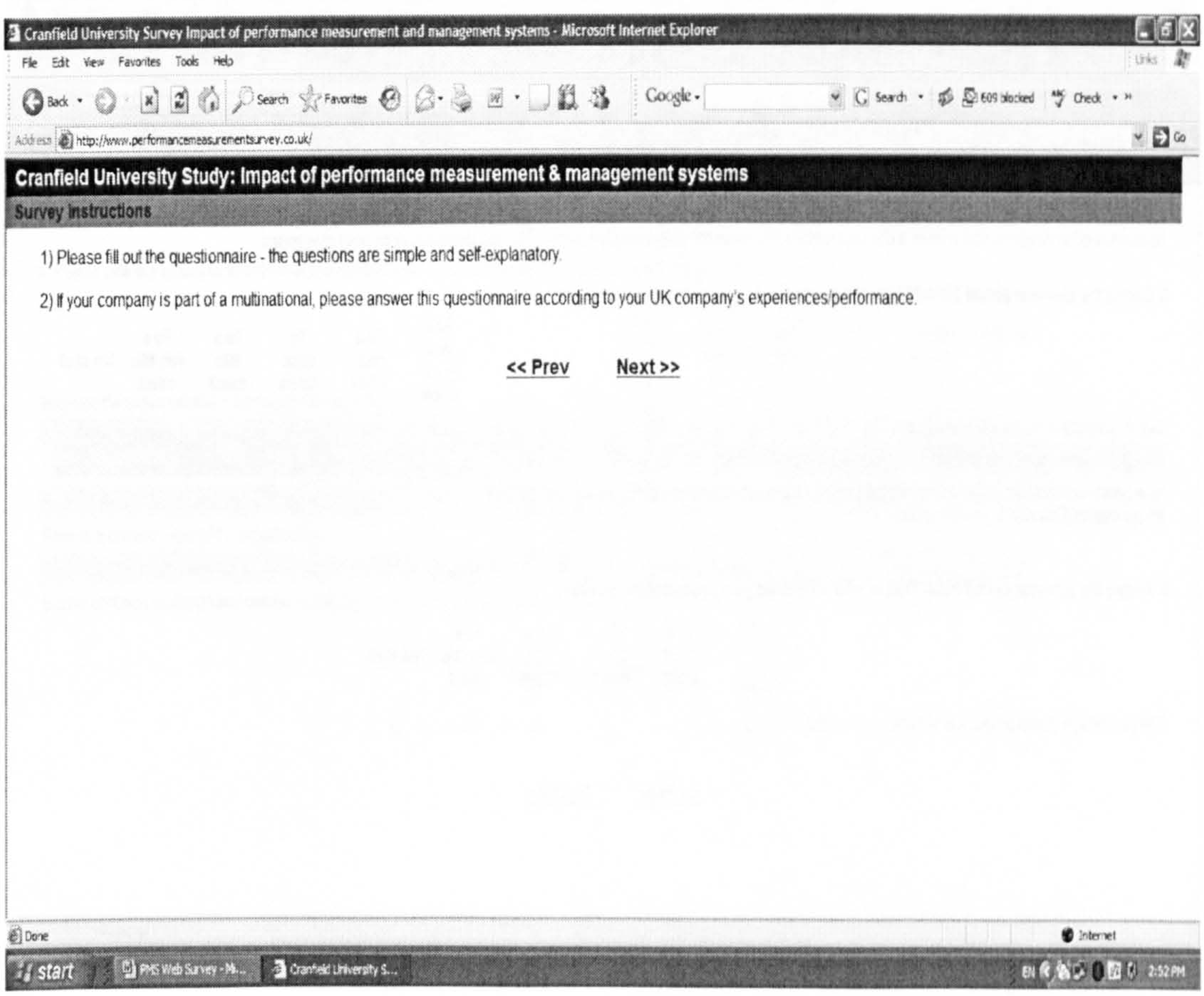
Bassil Yaghi
Doctoral Researcher

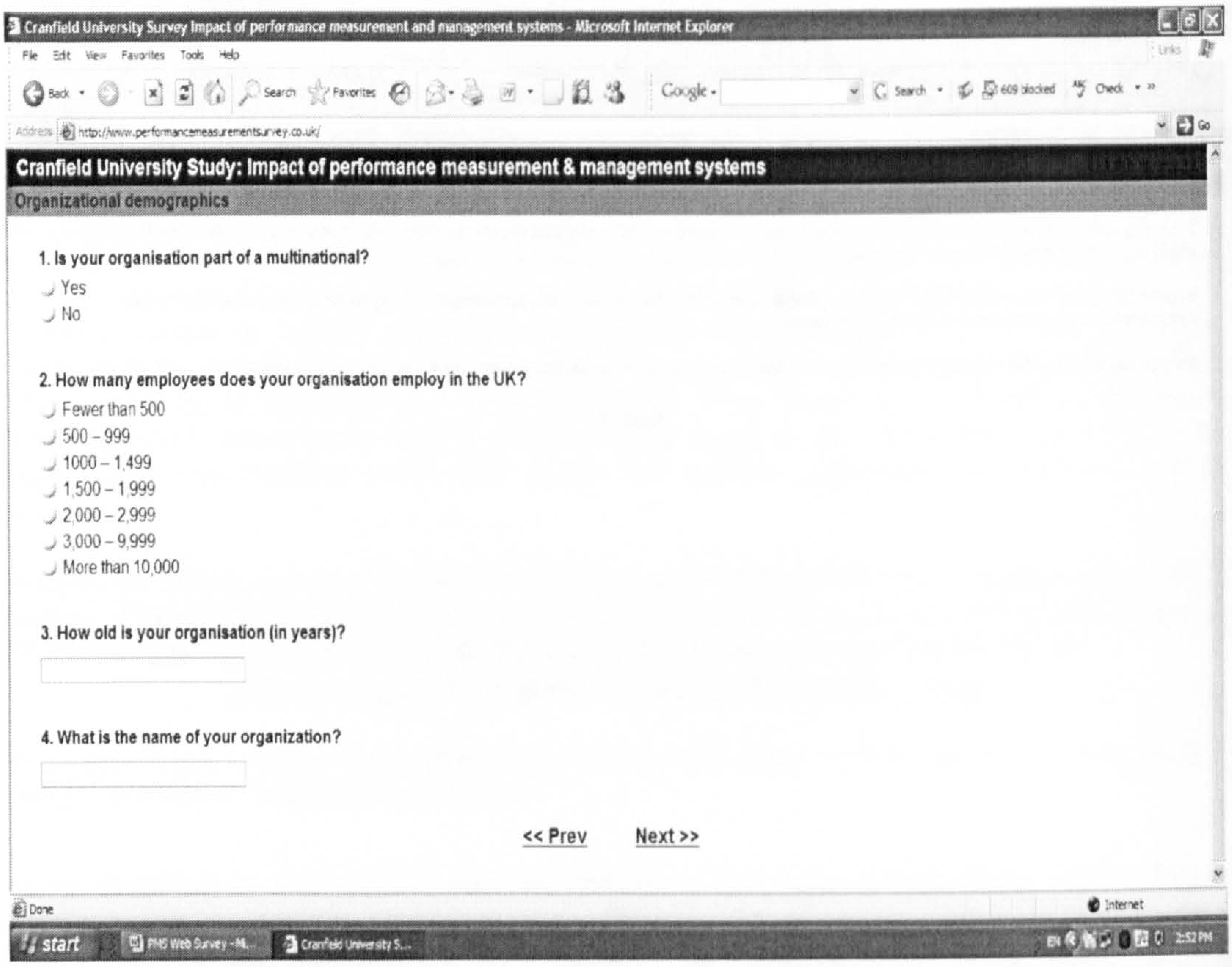
Centre for Business Performance
Cranfield School of Management
Building 31, Cranfield University
Cranfield, Bedfordshire MK43 0AL
United Kingdom

About our research centre

The Centre for Business Performance (CBP) at Cranfield School of Management has extensive experience in performance measurement and management systems. The centre is a leading research institution in this field and supports the dissemination of performance measurement and management practices and innovations through two business round tables and a biennial conference. Its close relationship with industry allows CBP research to focus on current problems that UK organisations are experiencing.


Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer
http://www.performancemeasurementsurvey.co.uk/
Cranfield University Study: Impact of performance measurement & management systems
Survey objectives
This survey asks about your company's experience with performance measurement and management systems. Your contribution is very important to this study. The survey should take approximately 10-15 minutes to complete.
By participating in this survey, you will obtain a free copy of the study report. In addition, the first 100 organisations completing the survey will get a free ticket for the conference where the results of this research will be presented.
It is important to mention that this is the first survey in the UK to address the impact of performance measurement practices on product innovations.
Next >>


Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer
http://www.performancemeasurementsurvey.co.uk/
Cranfield University Study: Impact of performance measurement & management systems
Survey instructions
1) Please fill out the questionnaire - the questions are simple and self-explanatory.
2) If your company is part of a multinational, please answer this questionnaire according to your UK company's experiences/performance.
<< Prev
Next >>



Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

Cranfield University Study: Impact of performance measurement & management systems

Organizational demographics

1. Is your organisation part of a multinational?

- Yes
- No

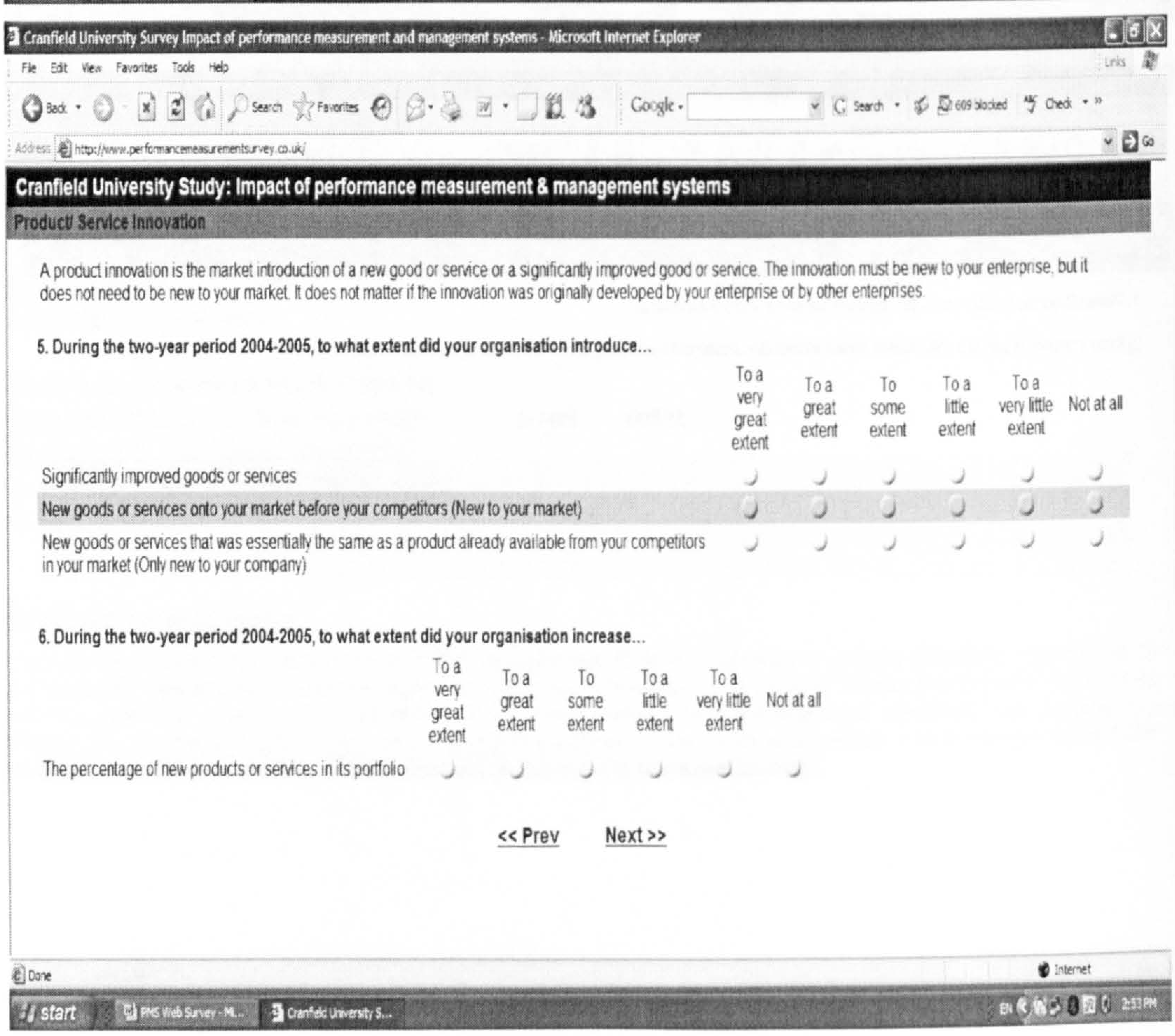
2. How many employees does your organisation employ in the UK?

- Fewer than 500
- 500 – 999
- 1000 – 1,499
- 1,500 – 1,999
- 2,000 – 2,999
- 3,000 – 9,999
- More than 10,000

3. How old is your organisation (in years)?

4. What is the name of your organization?

<< Prev Next >>



Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

Cranfield University Study: Impact of performance measurement & management systems

Product/ Service Innovation

A product innovation is the market introduction of a new good or service or a significantly improved good or service. The innovation must be new to your enterprise, but it does not need to be new to your market. It does not matter if the innovation was originally developed by your enterprise or by other enterprises.

5. During the two-year period 2004-2005, to what extent did your organisation introduce...

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Significantly improved goods or services						
New goods or services onto your market before your competitors (New to your market)						
New goods or services that was essentially the same as a product already available from your competitors in your market (Only new to your company)						

6. During the two-year period 2004-2005, to what extent did your organisation increase...

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
The percentage of new products or services in its portfolio						

<< Prev Next >>

Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.performancemeasurementsurvey.co.uk/

Cranfield University Study: Impact of performance measurement & management systems

Nature of performance measurement's use

Monitoring

7. Please rate the extent to which your top management team currently uses performance measures to...

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Track progress towards goals						
Review key measures						
Monitor results						
Compare outcomes to expectations						

<< Prev Next >>

Done Internet

start PMS Web Survey - M... Cranfield University S... 2:53 PM

Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.performancemeasurementsurvey.co.uk/

Cranfield University Study: Impact of performance measurement & management systems

Nature of performance measurement's use

Focus attention

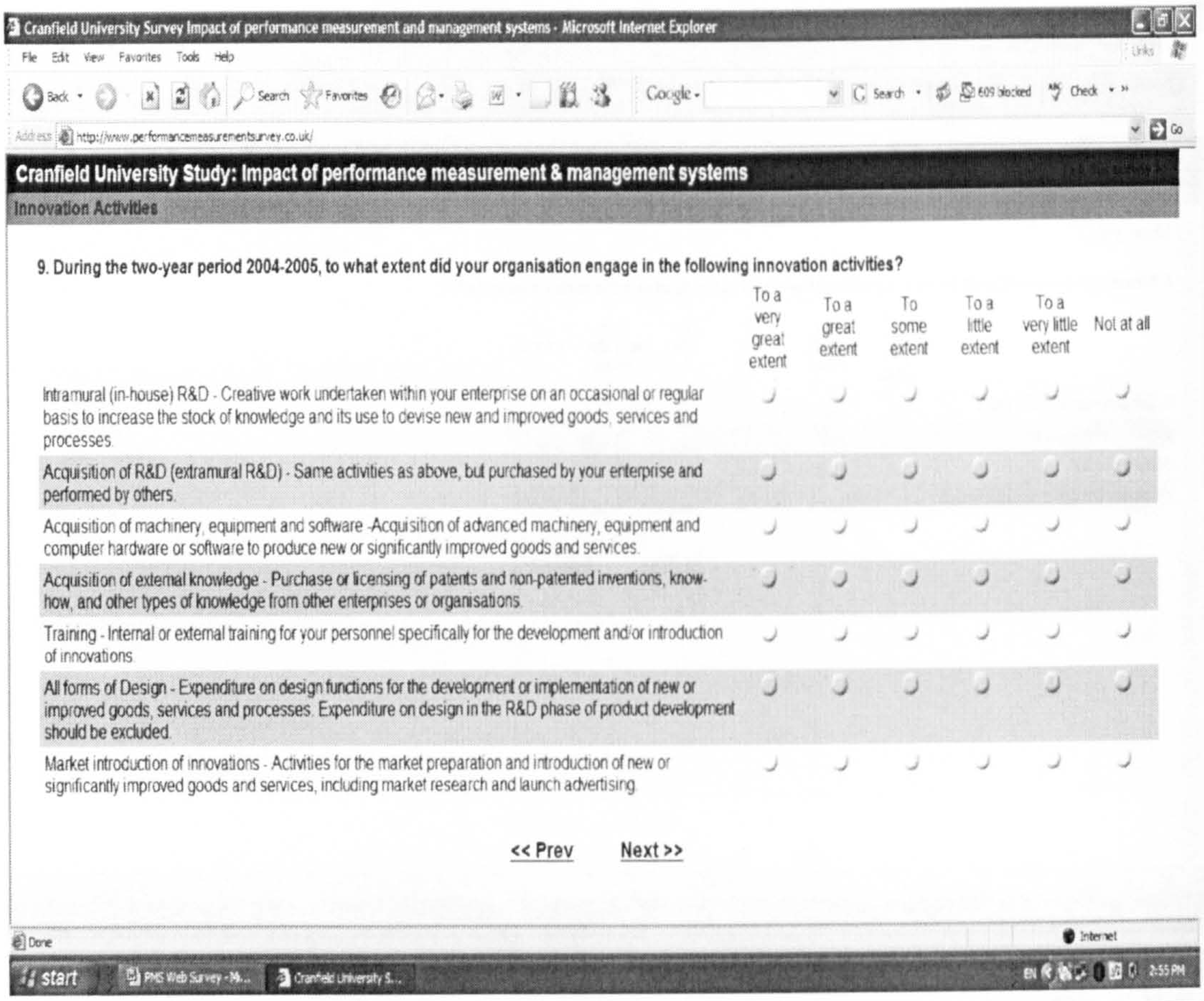
8. Please rate the extent to which your top management team currently uses performance measures to...

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Integrate the organisation- i.e. tie the organisation together						
Enable the organisation to focus on common issues						
Enable the organisation to focus on your critical success factors						
Develop a common vocabulary in the organisation						
Provide a common view of the organisation						
Enable discussion in meetings of superiors, subordinates and peers						
Enable continual challenge and debate underlying results, assumptions and action plans						

<< Prev Next >>

Done Internet

start PMS Web Survey - M... Cranfield University S... 2:54 PM



Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

http://www.performancemeasurementsurvey.co.uk/

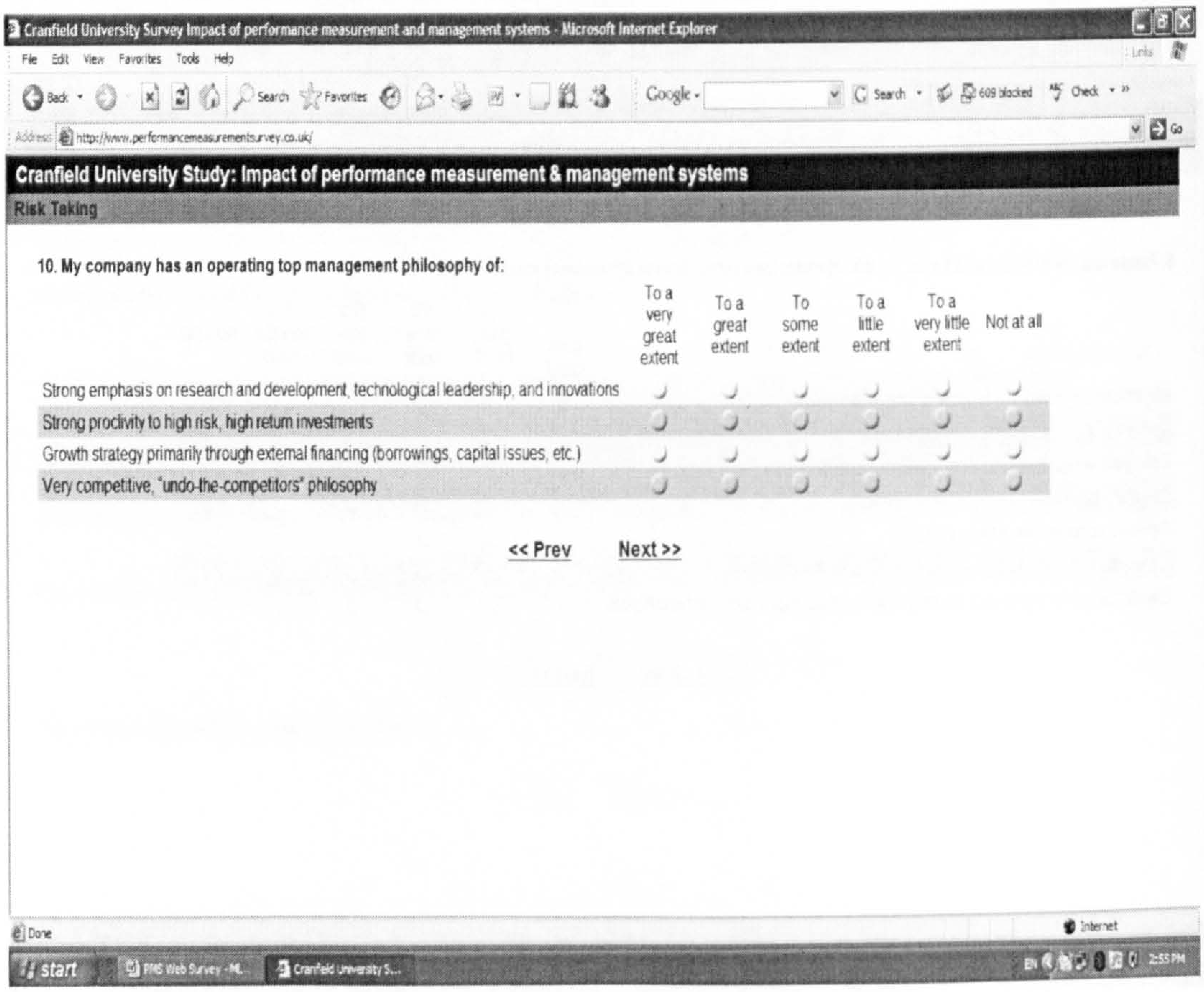
Cranfield University Study: Impact of performance measurement & management systems

Innovation Activities

9. During the two-year period 2004-2005, to what extent did your organisation engage in the following innovation activities?

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Intramural (in-house) R&D - Creative work undertaken within your enterprise on an occasional or regular basis to increase the stock of knowledge and its use to devise new and improved goods, services and processes.						
Acquisition of R&D (extramural R&D) - Same activities as above, but purchased by your enterprise and performed by others.						
Acquisition of machinery, equipment and software -Acquisition of advanced machinery, equipment and computer hardware or software to produce new or significantly improved goods and services.						
Acquisition of external knowledge - Purchase or licensing of patents and non-patented inventions, know-how, and other types of knowledge from other enterprises or organisations.						
Training - Internal or external training for your personnel specifically for the development and/or introduction of innovations.						
All forms of Design - Expenditure on design functions for the development or implementation of new or improved goods, services and processes. Expenditure on design in the R&D phase of product development should be excluded.						
Market introduction of innovations - Activities for the market preparation and introduction of new or significantly improved goods and services, including market research and launch advertising.						

<< Prev   Next >>



Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

http://www.performancemeasurementsurvey.co.uk/

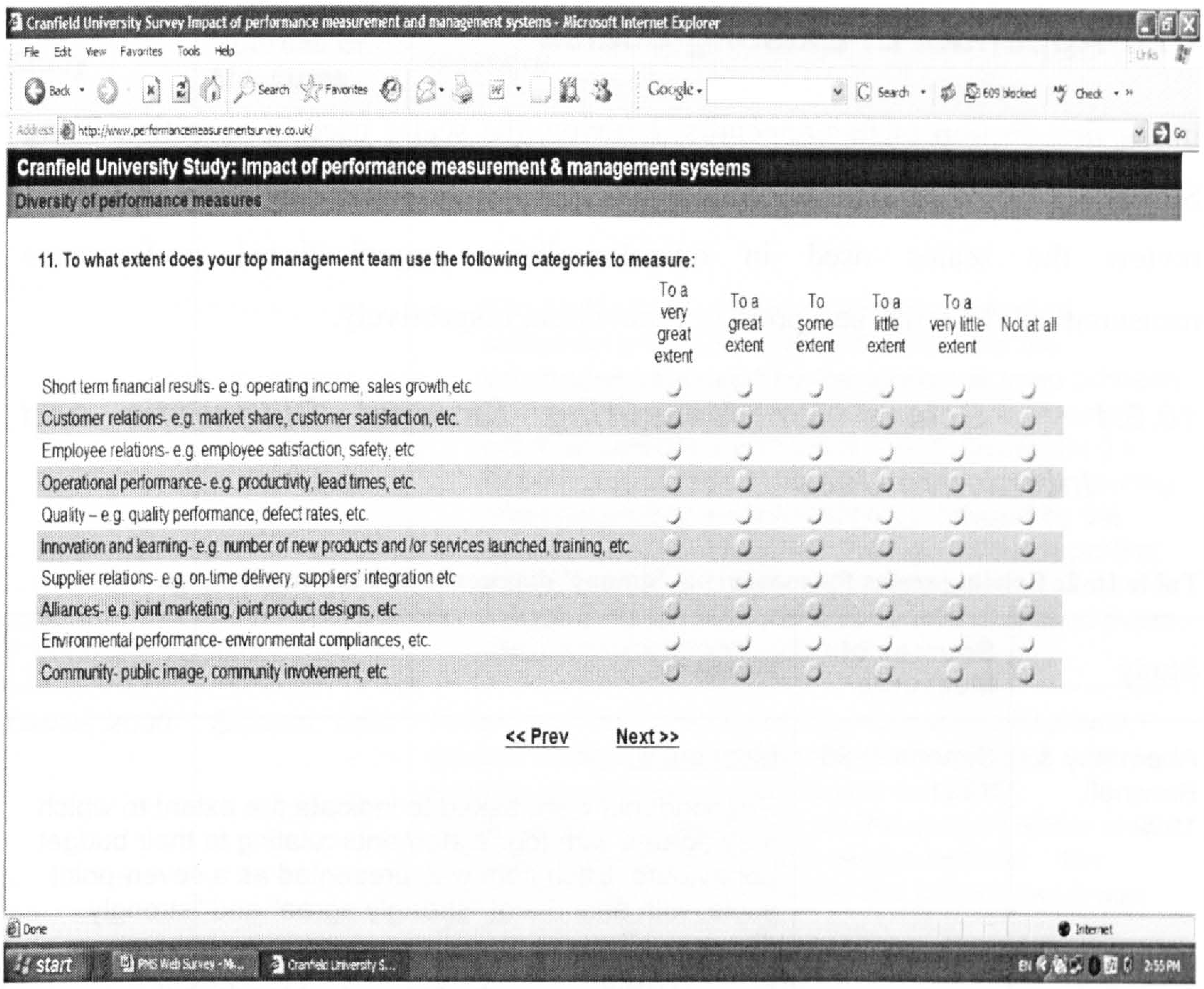
Cranfield University Study: Impact of performance measurement & management systems

Risk Taking

10. My company has an operating top management philosophy of:

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Strong emphasis on research and development, technological leadership, and innovations						
Strong proclivity to high risk, high return investments						
Growth strategy primarily through external financing (borrowings, capital issues, etc.)						
Very competitive, "undo-the-competitors" philosophy						

<< Prev   Next >>



Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.performancemeasurementsurvey.co.uk/

Cranfield University Study: Impact of performance measurement & management systems

Diversity of performance measures

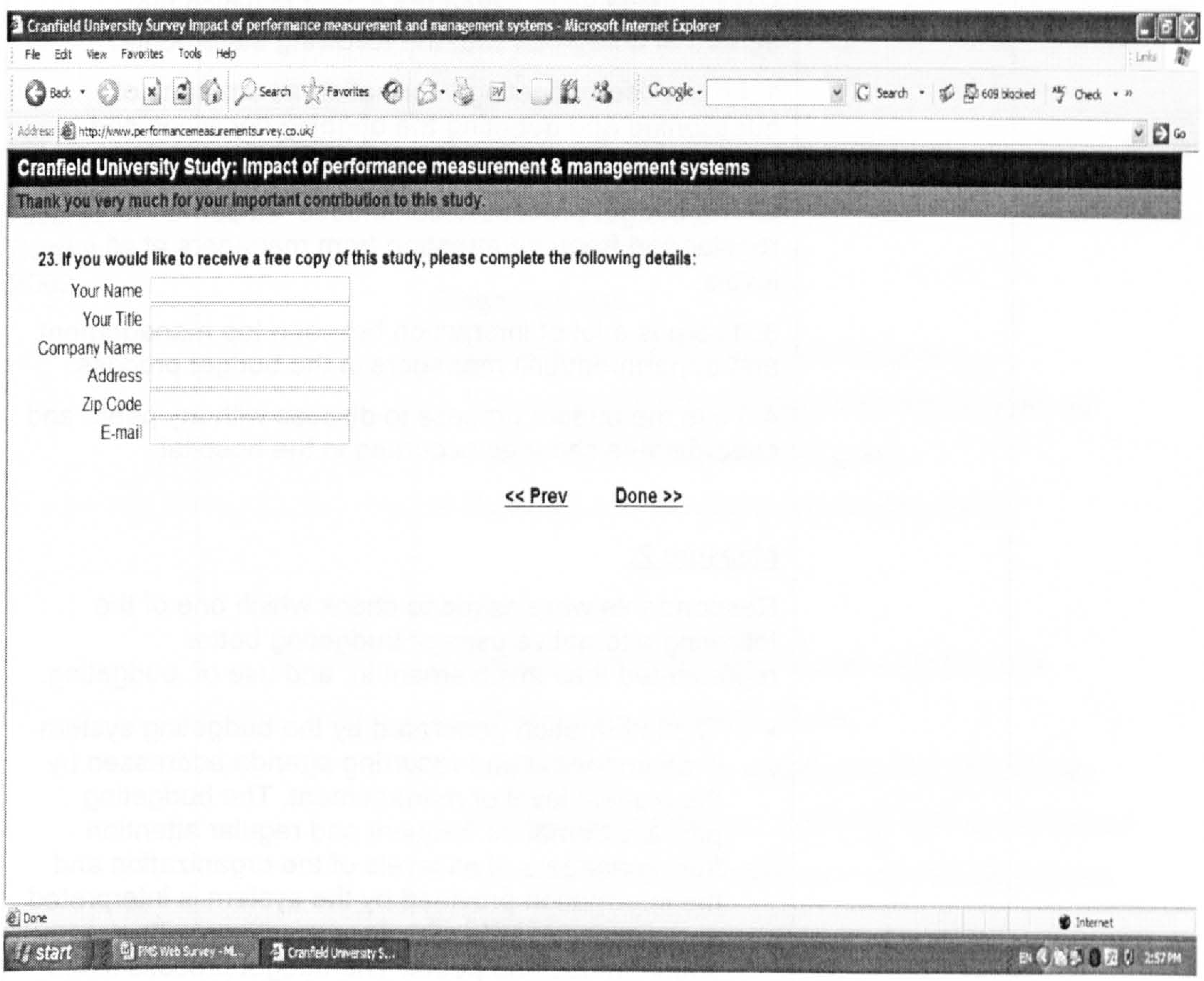
11. To what extent does your top management team use the following categories to measure:

	To a very great extent	To a great extent	To some extent	To a little extent	To a very little extent	Not at all
Short term financial results- e.g. operating income, sales growth,etc						
Customer relations- e.g. market share, customer satisfaction, etc.						
Employee relations- e.g. employee satisfaction, safety, etc						
Operational performance- e.g. productivity, lead times, etc.						
Quality – e.g. quality performance, defect rates, etc.						
Innovation and learning- e.g. number of new products and /or services launched, training, etc.						
Supplier relations- e.g. on-time delivery, suppliers' integration etc						
Alliances- e.g. joint marketing, joint product designs, etc.						
Environmental performance- environmental compliances, etc.						
Community- public image, community involvement, etc.						

<< Prev Next >>

Done Internet

start PMS Web Survey - M... Cranfield University S... 2:55 PM



Cranfield University Survey Impact of performance measurement and management systems - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.performancemeasurementsurvey.co.uk/

Cranfield University Study: Impact of performance measurement & management systems

Thank you very much for your important contribution to this study.

23. If you would like to receive a free copy of this study, please complete the following details:

Your Name

Your Title

Company Name

Address

Zip Code

E-mail

<< Prev Done >>

Done Internet

start PMS Web Survey - M... Cranfield University S... 2:57 PM

10.5 Appendix E: Existing Scales

In the first section of this appendix, I review the scales used in operationalising Simons' (1995) diagnostic and interactive uses. In the second and third sections, I review the scales used in operationalising organizational performance measurement diversity and product innovation, respectively.

10.5.1 Scales for Measuring Simons' Diagnostic and Interactive Uses

Table 10-2: Existing scales for measuring Simons' diagnostic and interactive uses.

Study	Sources of Measures	Measure
Abernethy & Brownell, 1999	Simons 1990; 1991	Measure 1: Respondents were asked to indicate the extent to which they agreed with four statements relating to their budget behaviours. Each item was presented as a seven-point scale, with anchors of "strongly agree" and "strongly disagree". Respondents were asked the extent to which they agreed or disagreed with the following statements. 1. I often use budgeting information as a means of questioning and debating the ongoing decisions and actions of department/clinical managers. 2. The budget process is continuous – it demands regular and frequent attention from managers at all levels. 3. There is a lot of interaction between top management and department/unit managers in the budget process. 4. I use the budget process to discuss with my peers and subordinates changes occurring in the hospital. Measure 2: Respondents were asked to check which one of the following alternative uses of budgeting better represented their involvement in, and use of, budgeting. • The information generated by the budgeting system is an important and recurring agenda addressed by the highest level of management. The budgeting process demands frequent and regular attention from managers at all levels of the organization and the information provided by the system is interpreted and discussed in face-to-face meetings with subordinates and peers. The budget process relies

Study	Sources of Measures	Measure
		on the continual challenge and debate of underlying data, assumptions and action plans (description of interactive use) or • The budgeting system is a process aimed at achieving predetermined outcomes and the information produced by the system is used primarily to inform top managers if actions or outcomes are not in accordance with plans. Staff specialists (i.e. finance departments) play a pivotal role in preparing and interpreting the information produced by the system. Data are reported through formal reporting procedures and top managers tend to be involved in the process infrequently and on an exceptions basis (description of diagnostic use)
Davila, 2000	Simons, 1995	For each of the six types of information (product cost, product design, time-related, customer-related, resource input [budgets], and profitability) in the MCS, respondents were asked to rate the usage pattern on a five-point scale anchored with two sentences: "the information was used to monitor the project, but it was not discussed with my team except when it reported events that fell below plans or expectations" (diagnostic system) and "the information was used constantly in the interactions with my team. Frequently it was the main topic of our conversation" (interactive system).
Gil, Hartmann, & Alvarez-Dardet, 2003	Abernethy & Brownell, 1999	According to the following sentences, please indicate your general use of your Management Control System, considering it as the whole of management accounting and control techniques: Diagnostic use: 1- Set and negotiate goals and targets 2- Follow up significant exceptions and deviations 3- Follow up preset plans and goals 4- Evaluate and Control subordinates tightly Interactive use: 1- Debate data assumptions and actions plans 2- Signaling key strategic areas 3- Challenge news ideas and ways for doing tasks 4- Involve in a permanent attention with subordinates 5- Align performance measures with strategic goals 6- Learning tool

Study	Sources of Measures	Measure
Bisbe & Otley, 2004	Abernethy & Brownell, 1999; Davila, 2000	Respondents were asked to rate the items related to the interactive use of MCS on 1–7 Likert scales. 1- The main aim of MCS is (1) to ensure that previously established objectives are met vs. (7) to force us to continually question and revise the assumptions upon which we base our plans. 2- (1) Only when there are deviations from planned performance are budget tracking reports the main subject for face-to-face discussion with my executive team vs. (7) Whether there are deviations from planned performance or not, budget tracking reports are the main subject for face-to-face discussion with my executive team. 3- (1) I pay periodic or occasional attention to budgets (e.g. setting objectives, analysing periodic tracking reports,...) vs. (7) I pay regular and frequent attention to budgets. I use them permanently. 4- (1) For many managers in my company, budgets require periodic or occasional attention, but not permanent attention vs. (7) In my company, budgets require permanent attention from all managers.
Abernethy, Bouwens & Van Lent, 2006	Abernethy & Brownell, 1999; Simons, 1990; 1991	Respondents were asked to indicate the extent to which they agree with the following statements: (1=strongly agree, 7= totally disagree): 1- My superior and I often use the planning and control system as a means of questioning and debating the factors affecting the strategy. 2- The planning and control system is continuous – it demands regular and frequent attention from managers at all levels. 3- I use the planning and control system to discuss with my peers and subordinates changes occurring in my organization. 4- The planning and control system is used throughout the year to develop and present new programs, services, and strategies.

10.5.2 Existing Scales for Measuring Organizational Performance Measurement Diversity

Table 10-3: Existing scales for measuring organizational performance measurement diversity.

Study	Sources of Measures	Measure
Abernethy & Lillis, 1995		*Performance Measurement System:* Respondents were asked to rate the extent to which each of the following measures was used for performance evaluation in the manufacturing department of their organization. 1- On time delivery performance records. 2- Number of customer complaints. 3- Labour utilization/efficiency statistics. 4- Incidences of product defects. 5- Number of warranty claims. 6- Inventory turnover ratio. 7- Reports on whether standard product costs are met. 8- Survey of customer satisfaction. 9- Measurement of cost reduction due to quality improvements. 10- Evaluation of whether people in manufacturing are co-operative and responsive to the demands of other departments within the organization. 11- Customer satisfaction with problem solving and sales assistance. 12- Rate of introduction of new products. 13- Evaluation of the ability to vary product characteristics. 14- Length of cycle time from order to delivery (for standard products). 15- Length of cycle time for customer-requested product variations. 16- Material purchase price variance. 17- Rate of material scrap loss. 18- Measurement of machine utilization and down time.
Hoque & James, 2000	Hoque, Mia, & Alam, 1997	*BSC usage:* Respondents were asked to indicate the extent to which each item was used to assess their organization's performance on a fully anchored, five-point Likert scale ranging from 1 (not at all) to 5 (to a great extent). 1- Operating income 2- Sales growth

Study	Sources of Measures	Measure
		3- Return on investment 4- Labor efficiency variance 5- Rate of material scrap loss 6- Material efficiency variance 7- Manufacturing lead time 8- Ratio of good output to total output 9- Percent defective products shipped 10- Numbers of new product launches 11- Number of new patents 12- Time to market new products 13- Survey of customer satisfaction 14- Number of customer complaints 15- Market share 16- Percent shipments returned due to poor quality 17- On-time delivery 18- Warranty repair cost 19- Customer response time 20- Cycle time from order to delivery

10.5.3 Existing Scales for Measuring Product Innovation

Table 10-4: Existing scales for measuring product innovation.

Study	Sources of Measures	Measure
Miller and Friesen 1982		To what extent are the following activities carried out? 1- (1) There is a strong emphasis on the marketing of true and tried products or services. (7) There exists a very strong emphasis on R&D, technological leadership, and innovation. How many new lines of products or services has your firm marketed in the past 5 years? Please exclude mere minor variations. 2- (1) No new lines of product or services in past 5 years. (7) Hundreds of new lines of products or services in past 5 years. 3- (1) Changes in product lines have been mostly of a minor nature (e.g. putting in towel with the soap). (7) Changes in product lines have usually been dramatic

Study	Sources of Measures	Measure
		(e.g. changing from mechanical to electric calculators).
Miller & Friesen, 1983		Respondents were asked to rate the extent to which product-market innovation changed over the past 5 years. 1- The rate, relative to competitors, of new product/service introduction by firm. (1) Has decreased very much; (4) No change; (7) Has increased very much. 2- The rate of change in your methods of production or rendering of service. (1) Rate of change has declined much; (4) No change; (7) Change has accelerated rapidly.
Capon, Farley, Lehmann, & Hulbert, 1992		1- Please estimate the fraction of total corporate revenues in 1979 from the introductory and growth stages of the product life cycle? 2- What percent of this year's sales will be generated by products dependent on technology which did not exist or was not commercially feasible in 1970? 3- In new product and service introductions, how often is your company first-to-market with new products and services (1 = never, 5 = always)? 4- In new product and service introductions, how often is your company at the cutting edge of technology (1 = never, 5 = always)
Bisbe & Otley, 2004	Capon, Farley, Lehmann, & Hulbert, 1992; Thomson & Abernethy, 1998	In comparison with the industry average, 1- During the last three years we have launched, (1) many new products vs. (7) few new products. 2- During the last three years we have launched (1) many modifications to already existing products vs. (7) few modifications to already existing products. 3- In new products, we are (1) very often first-to-market vs. (7) very rarely first-to-market. 4- The percentage of new products in our product portfolio is (1) much higher than the industry average vs. (7) is much lower than the industry average.

10.6 Appendix F: Simons' Levers of Control – Empirical Research

Table 10-5: Empirical research using Simons' levers of control framework

Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
Simons, 1990	Case study, 2-year longitudinal	16 firms in a single American industry. Study reports on 2 firms (A & B)	Business strategy	Multiple control systems - Interactive - Diagnostic - Separate	MCS	A dynamic process model of interactive control was developed comprising 4 concepts: business strategy (BS), strategic uncertainties (SU), interactive controls & organizational learning (OL). While all large firms have similar MCSs, top managers make selected control systems interactive to personally monitor SUs that they believe to be critical to achieving BS. The choice by top managers to make certain MCS interactive and others programmed provides signals to employees about what should be monitored & where new ideas should be proposed and tested. This signal activates OL & through debate & dialogue, that surrounds interactive management control process, new strategies & tactics emerge over time.	
Simons, 1991	Case study	30 business units in the US heath care products industry from 16 companies	Business unit	Multiple control systems - Interactive - Diagnostic	Business strategy	MCS are not only used to implement strategies but also to give the emergence of new ones. While control systems may appear to be similar across settings, the study suggests that there are fundamental differences in the way that policy-making managers use control systems. Top managers decide which formal processes to use interactively and which to use diagnostically based on their sense of purpose for the organization and their personal assessment of associated uncertainties. Top managers with a clear sense of strategic vision choose	

Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
				- Separate		one MCS to use interactively. Top managers use multiple MCSs interactively only during periods of crisis. Top managers without a strategic vision (or an urgency to create a strategic vision) do not use control systems interactively. Three factors influence the design and use of interactive control systems: technological dependence within product markets, complexity of the value chain, and ability of competitors to respond to product market initiatives.	
Simons, 1992	Conceptual [91]			Accounting-based systems - Interactive - Diagnostic - Boundary systems - Seperate		The boundary systems lever is introduced. First type of boundary systems communicates the domain of permissible activity to all employees by dictating acceptable rules of competition. The second type formally states what strategic opportunities to be avoided.	
Simons, 1994	Case study, 18-months longitudinal	10 newly appointed managers in multiple industries	Business unit	Multiple control systems - Belief systems - Boundary systems - Interactive - Diagnostic	Strategic change	Based on the mandate to change, top managers were divided into two clusters: strategic turnaround & strategic evolution. MCSs were instrumental in advancing the agendas of all managers. Management control process can, depending on the context, act as agents for both intended change and autonomous, emergent change.	

[91] Following Chandler & Lyon (2001), empirical studies were classified as those that included some kind of data or data analysis in the study. These included both statistical and qualitative analyses. Literature review, untested theoretical models, and proposed mathematical models were defined as conceptual studies. Studies that both presented and tested theory with empirical data were counted as empirical studies.

Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
				- Interplay		Managers belonging to the two clusters used MCSs to: - Overcome organizational inertia. - Communicate the substance of their new agenda - Structure implementation timetables & targets - Ensure continuing attention through incentives - Focus organizational learning on strategic uncertainties associated with their vision for the future. Organizational context & the urgency of the mandate often resulted in levers being used differently between two clusters. The use of MCS progressed through stages to foster learning & unlearning. MCS appear to be important in building credibility & selling a new strategy to various constituents.	
Abernethy & Bouwens, 1999	Cross-sectional survey	63 Australian public hospitals	Hospital	Budget - Interactive - Diagnostic - Opposite poles	Strategic change	Style of budget use moderates the relationship between strategic change and performance. The relation between strategic change and performance is more positive for interactive use compared to diagnostic use. Performance is highest when there is a match between style of use and the type of strategic change. Performance is highest when strategic change is low and budget use is diagnostic and when strategic change is high and budget use is interactive. Performance is lowest when there is a mismatch. Performance is lowest when strategic change is high and budget use is diagnostic and when strategic change is low and budget use is interactive.	Interactive use: Simons, 1990, 1991
Davila, 2000	Case study & Cross-sectional survey	Case studies: 12 business units in 7 companies Survey: 56 project managers from 11 companies.	Project manager	- Project management system - Interactive - Diagnostic	MCS; project performance	Management control systems' design is related to performance, in contrast to other variables including product strategy that have no such relationship. Cost and design information has a positive effect upon performance. In contrast, time information hinders performance.	Interactive use: Simons, 1995

Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
		The companies in both studies are from medical devices industry in Europe & US		- Opposite poles			
Marginson, 2002	Case study	British-based company in the British telecom industry	Firm	Administrative controls - Performance measurement systems - Belief systems - Boundary systems - Interactive - Diagnostic Interplay		Informed by Simons' theoretical model of the strategy process–MCS relationship, the study examined the nature and extent of this relationship at middle- and lower-management levels. Of particular interest were the effects that the design and use of three groups of MCS (value systems, administrative control, and performance measurement systems) have on the development of new ideas and initiatives. Findings suggest that beliefs systems influence managers' initiation or 'triggering' decisions, the use of administrative controls affects the location of strategic initiatives and may lead to the polarization of roles, and simultaneous emphasis on a range of key performance indicators can create a bias towards one set of measures and against another.	
Nilsson & Kald, 2002	Survey	209 firms in Nordic countries (Denmark, Finland, Norway, Sweden)	Firm	Performance measurement systems - Interactive - Diagnostic - opposite poles		The findings show that the performance management systems of the Nordic firms under study are used both diagnostically and interactively. The two principal applications are in decision support at the top management and the operating levels. Furthermore, in the design of the performance management systems, both financial and nonfinancial measures are quite significant. These tendencies are particularly noticeable at the Danish, Finnish, and Swedish companies. At the Norwegian companies, on the other hand, the design and uses of performance management are somewhat more traditional.	Interactive & Diagnostic: Simons, 1995
Gil, Hartmann, & Alvarez-Dardet, 2003	Cross-sectional survey	92 top management teams (TMT) from 218 Spanish public hospitals	Firm	- Management accounting systems (MAS) - Interactive - Diagnostic	Strategy implementation: - Flexibility strategy - Cost strategy	The more the professional is the TMT orientation, the more they use MAS in an interactive way and the more they use non-financial information. The more administrative is the TMT's orientation, the more they use MAS in a diagnostic way and the more they use of financial information. The findings show positive and significant relationships between the different uses of MAS and the two types of strategy implementation,	Interactive & diagnostic use: Abernethy & Brownell, 1999; Davila,

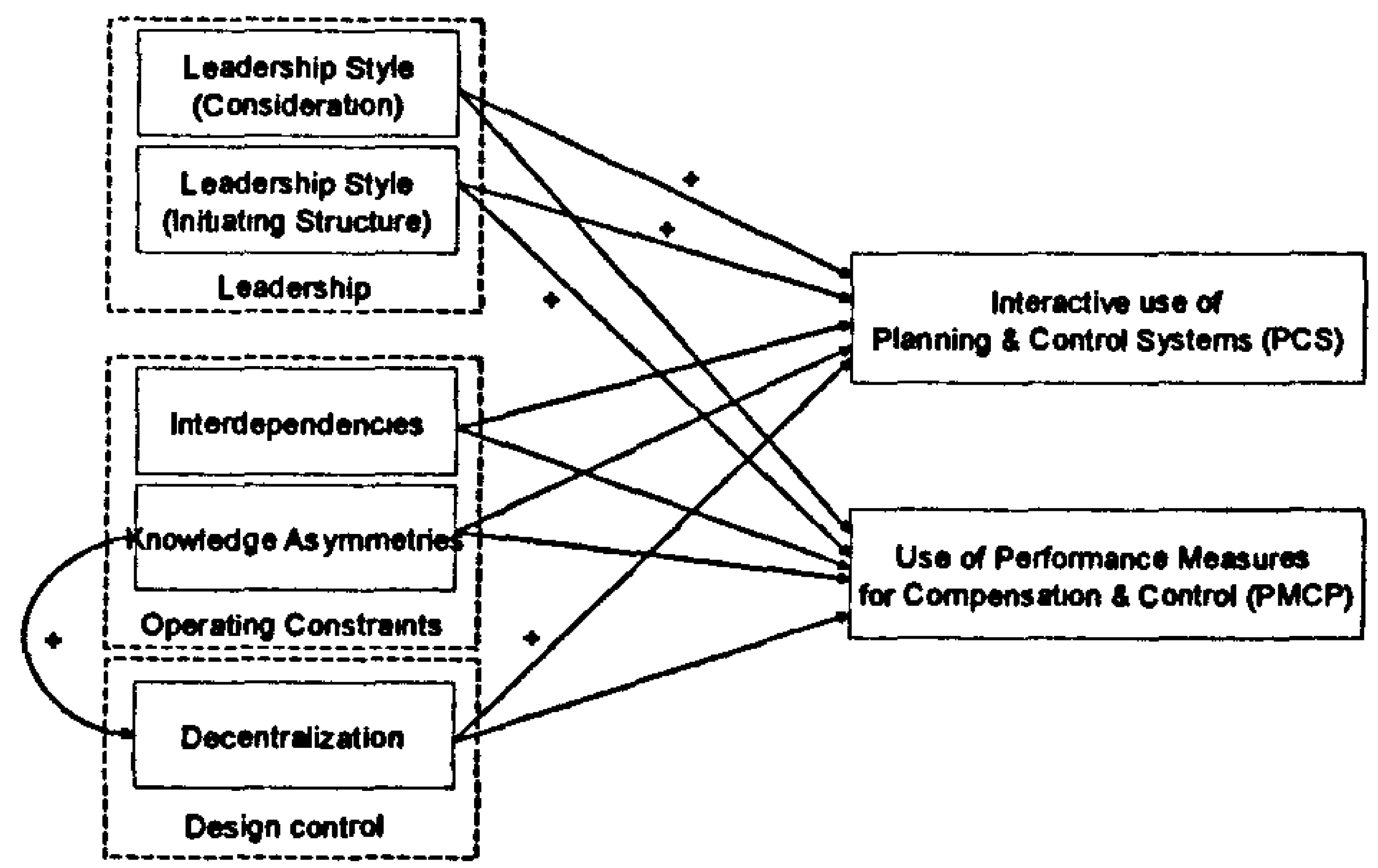
Study	Method	Sample	Unit	MCS Levers Operationalisation	Dependent	Findings	Source of measures
				- Separate		focus on flexibility and focus on cost.	2000
Bisbe & Otley, 2004	Cross-sectional survey	58 medium sized, mature Spanish manufacturing firms	Firm	- Balanced scorecard - Budgets - Project management systems	Product innovation; organizational performance	The evidence does not support the proposition that the more interactive the use of MCS by top managers, the higher the product innovation (and acting through innovation, the better the performance). The evidence suggests that interactive use may favour innovation in low-innovating firms, while the effect appears to be in the opposite direction in high-innovating firms. The evidence supports the proposition that the impact of product innovation on performance is moderated by the style of use of MCS.	Interactive use: Abernethy & Brownell, 1999; Davila, 2000
				- Interactive - Diagnostic			
				- Opposite poles			
Bruining, Bonnet, & Wright, 2004	Case study	2 firms - Rehab and Packaging	Firm	MCS		The study concludes that there is a need in post-MBO to develop coherence between a change in strategy & the application of levers of control. The coherence of management control & management accounting systems, which was lacking pre-buy-out, developed dynamically post-buy-out. They suggest that the development of belief and interaction control systems are of particular importance in more entrepreneurial buy-outs and serve as a valuable complement & extension to but not replacement of the more traditional diagnostic control systems post-MBO.	
				- Belief systems - Boundary systems - Interactive - Diagnostic			
				Interplay			
Collier, 2005	Case study, 10-year longitudinal	Owner owned Australian-based packaging equipment supplier	Firm	Multiple control systems		The role of an entrepreneur in constructing a control system contained a formal, spreadsheet-based control model and a form of social control, both modes of control being inseparable. The research found that Simons (1995) framework is more helpful than Ferreira and Otley (2005) as it reflects the importance, demonstrated in this study, of belief systems reflected in the social controls exercised by entrepreneur owner.	
				- Belief systems - Boundary systems - Interactive - Diagnostic			
				Interplay			

Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
Ferreira & Otley, 2005	Case study	4 Portuguese firms – (1) public and (3) non-public	Firm	MCS - Belief systems - Boundary systems - Interactive - Diagnostic Interplay		The study proposes Performance Management & Control (PMC) framework as a research tool for describing the structure & operation of MCS. The framework extends frameworks proposed by Simons (1995) & Otley (1999). The case studies have also unveiled three issues in using Simons' (1995) framework. The first is that it was not always possible to establish whether a particular control tool is part of diagnostic control system or of the interactive control system. Similarly, it proves difficult to ascertain definitely whether certain control structures belong to the beliefs system or to the boundary system, thus illustrating how different are theoretical formulations from practical applications. The second weakness is that meanings and associations of the concepts embedded in levers of control (e.g. core values) are diffuse. This creates scope for subjective interpretation and thus reduces the power of the conceptual model. In particular, they note the ambiguity in the definition of 'interactive controls' and suggest that this concept is split into two distinct components: interactive use of controls, and strategic validity controls. The third weakness is that strictly speaking, the model is not of general applicability and this is so for two reasons. Firstly, in some types of organizations such as subsidiaries of larger groups, the design of belief & boundary systems are largely beyond the organization's domain. Secondly, the greater the relative importance of informal controls in organizations the lower the contribution of Simons' framework to the understanding of the operation of the control system.	
Henri, 2006b	Survey	383 Canadian manufacturing firms	Firm	- Performance measurement systems (PMS)	Capabilities; Organizational performance	Diagnostic use of PMS is significantly and negatively related to capabilities of market orientation, entrepreneurship, innovativeness and organizational learning.	Vandenbosch, 1999

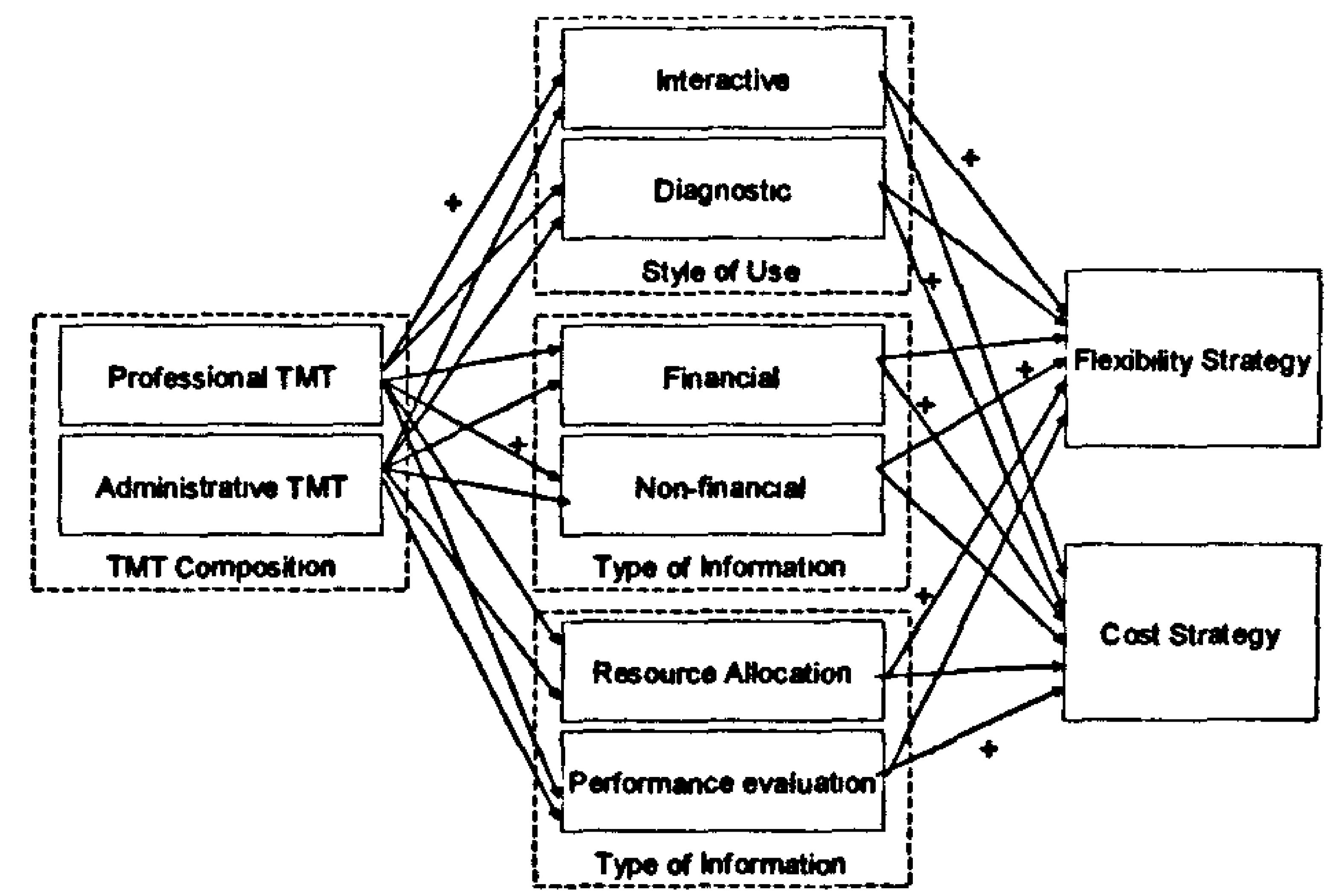
Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
				- Interactive - Diagnostic		Interactive use of PMS is significantly and negatively related to capabilities of market orientation, entrepreneurship, innovativeness and organizational learning. The direct effect of diagnostic use of PMS on financial performance is positive but not significant. The direct effect of interactive use of PMS on financial performance is negative but not significant.	
				- Separate - Combined		Dynamic tension has a direct positive and significant impact on performance.	
Tuomela, 2005	Case study, 4-year longitudinal	A subsidiary of ABB Finland	Firm	Performance measurement system - Belief systems - Boundary systems - Interactive - Diagnostic Interplay		Using Simons' levers of control as a theoretical frame of reference, the study finds that strategic performance measurement systems can be used both diagnostically and interactively, but such systems have implications for beliefs control and boundary control as well. Interactive use of performance measures is apt to improve the quality of strategic management and to increase commitment to strategic targets. On the other hand, interactive discussion of specific performance metrics increases the visibility of actions, which may initiate resistance. In addition, interactive use of performance measures may be costly in terms of time consumption both when collecting the data and when discussing the results. Two major differences were found in the actual use of strategic performance measures when compared to the normative literature. First, in contrast to ascertaining certain cause-and-effect relationships before implementing new measures, it was perceived that the measures themselves would be used over time to confirm or reject alleged relationships. Second, no tight connections between the new measurement system and managerial bonuses were made. This was mostly due to the development process, during which the top managers themselves developed the measures to reflect their belief about the best way of achieving the ultimate financial targets.	
Widener, 2005	Survey	122 US firms	Firm	Performance measurement	Performance; Organizational	The results also show that the performance measurement system can be used in dual roles simultaneously indicating	Interactive & Diagnostic

Study	Method	Sample	Unit	MCS Levers Operational-isation	Dependent	Findings	Source of measures
				system - Belief systems - Boundary systems - Interactive - Diagnostic Separate	learning; Attention	that pieces of the PM system are used interactively while other components may be used in a more diagnostic manner. Surprisingly, interactive use of performance measurement system is not associated with organizational learning. This study shows that reliance on control systems influence performance through their affect on learning and management attention. Sensitivity tests demonstrated that the relation between control systems and performance is weak; however, the effects become apparent when organizational learning & attention are included in the model. The study finds that the two types of strategic uncertainties (competitive uncertainty & operational uncertainty) are associated with the importance of the four control systems. Moreover, operational uncertainties have the largest effect on diagnostic controls, belief systems, & the boundary system, while competitive uncertainties drive interactive controls. This implies that the interactive control system is used to scan the external environment while the other three types of controls.	Henri, 2006a, 2006b
Abernethy, Bouwens & Van Lent, 2006	Cross-sectional survey	128 profit cent of firms in the service & manufacturing sectors in the Netherlands	Lower level managers	- Planning & control systems (PCS) - Interactive - Diagnostic - Opposite poles	- Interactive use of (PCS) - interactive use of Performance Measures for Compensation & Control (PMCP)	Neither leadership style is associated with decentralization. Both considerate and initiating structure styles of leadership are positively associated with interactive use of PCS. Initiating structure is positively associated with PMCP. However, considerate leadership does not affect PMCP. Decentralization is positively with the interactive use of PCS and not associated with PMCP. The evidence with regard to the two constructs that capture operating constraints is mixed. Knowledge asymmetries are positively associated with decentralization. The evidence is only marginally consistent with a negative association between interdependencies and decentralization. There is no evidence that operating constraints impact directly on the interactive use of PCS or PMCP.	Interactive use: Abernethy & Brownell, 1999

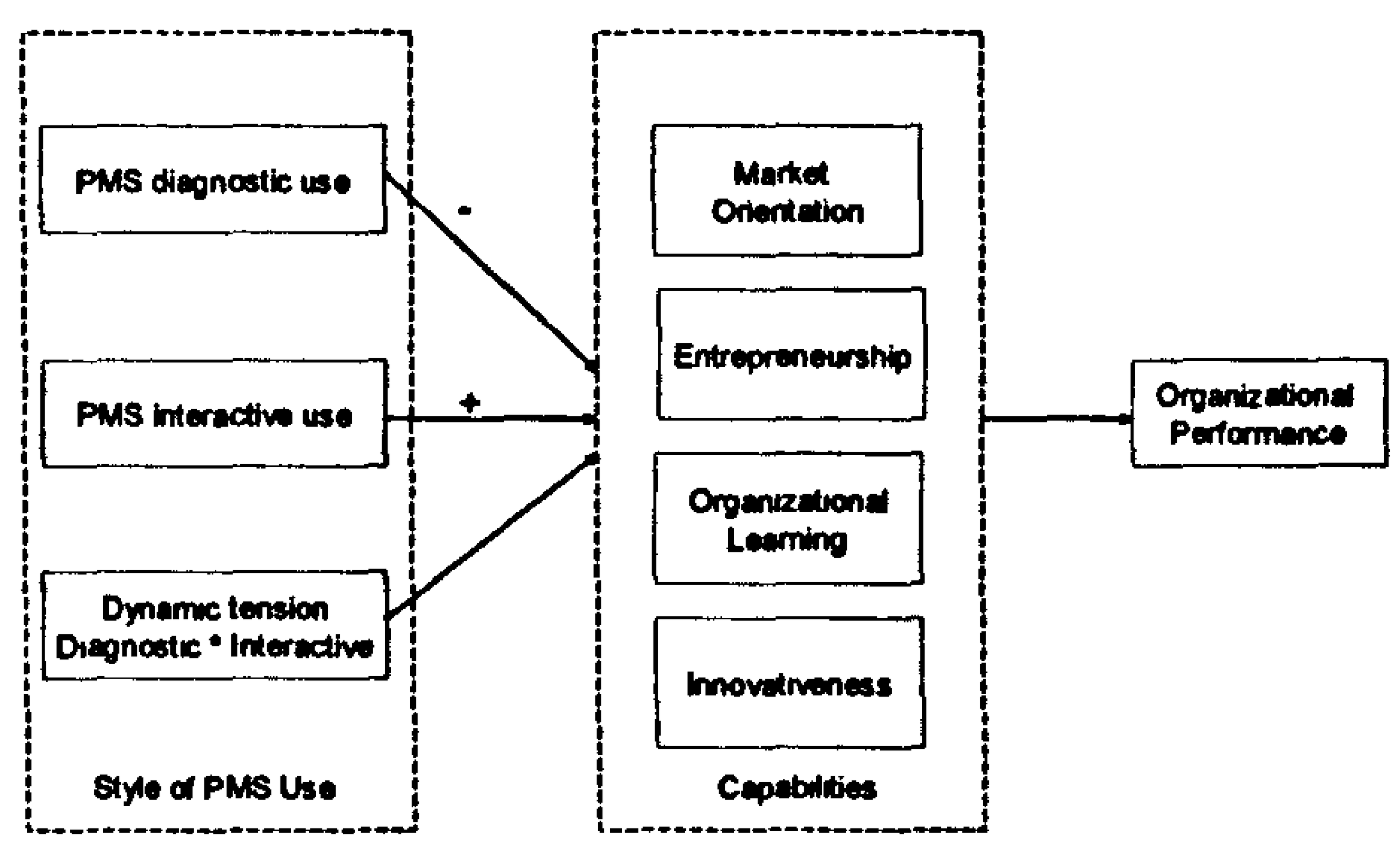
Table 10-6: Research models for empirical research using Simons' levers of control framework



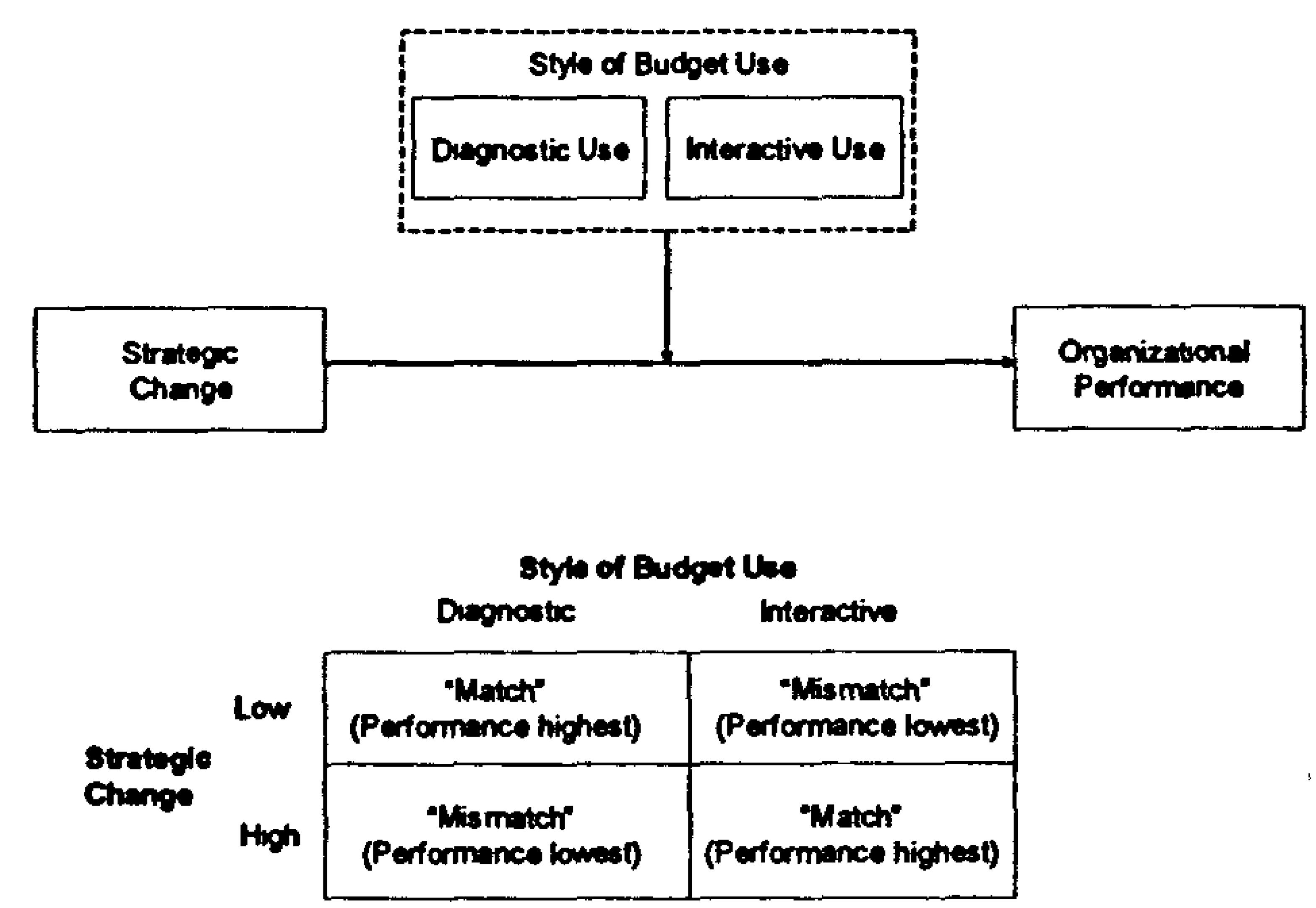
Source: Adapted from (Abernethy, Bouwens & Van Lent)



Source: Adapted from (Gil, Hartmann & Avarez-Dardet, 2003)

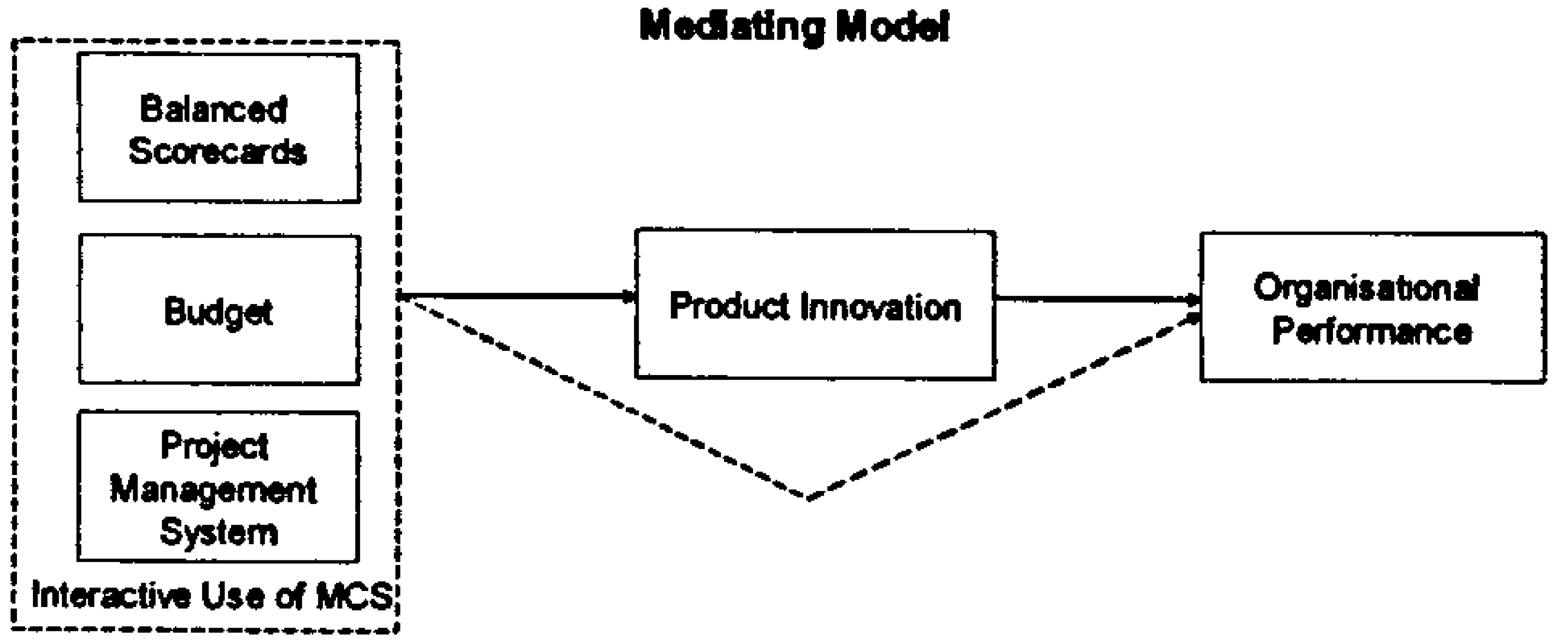


Source. Adapted from (Henri, 2005)

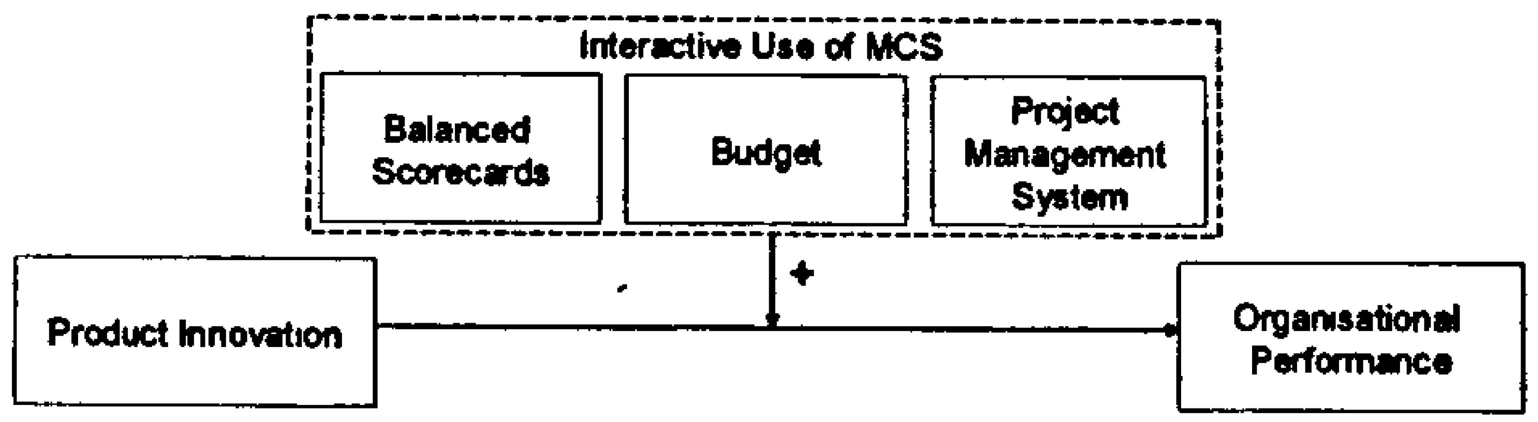


Source: Adapted from (Abernethy & Brownell, 1999)

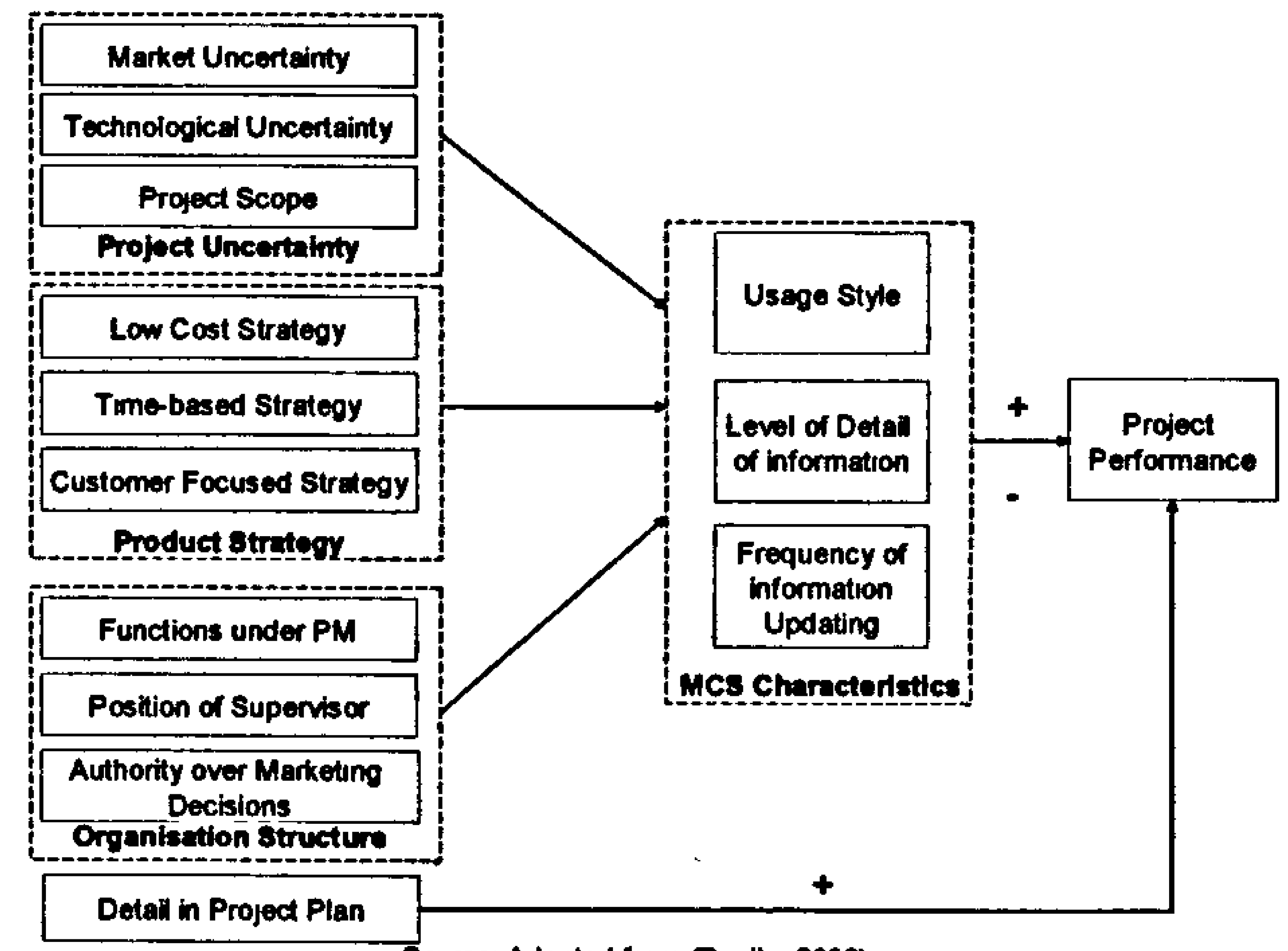
Mediating Model



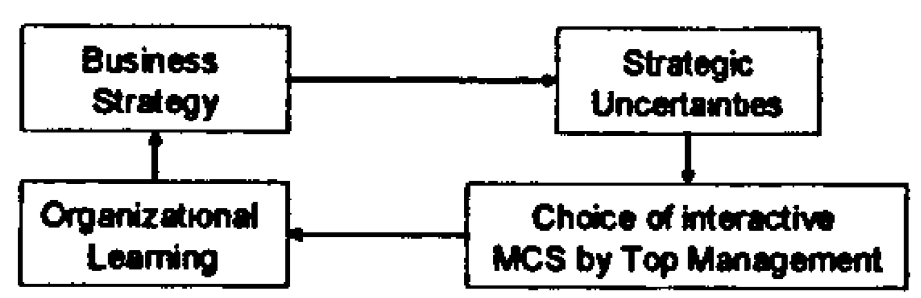
Moderation Model



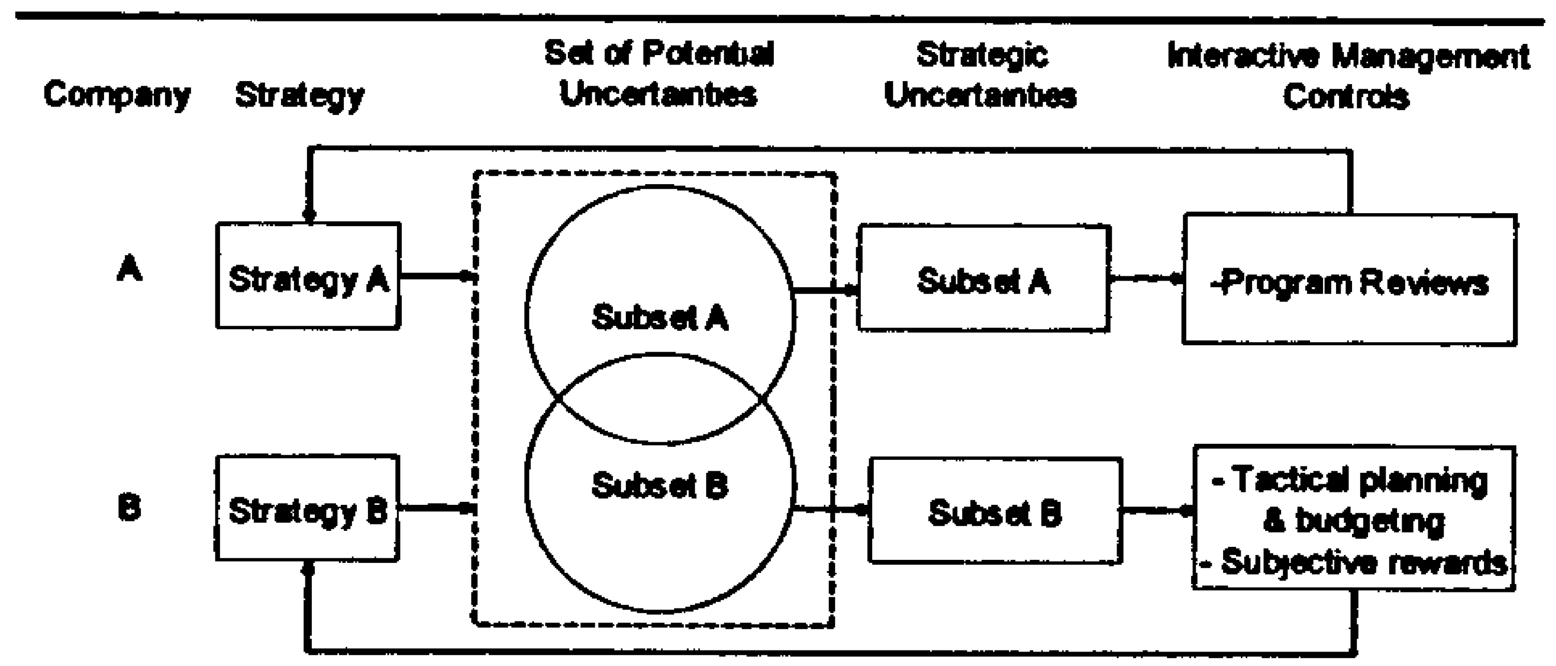
Source: Adapted from (Bisbe & Otley, 2004)



Source: Adapted from (Davila, 2000)

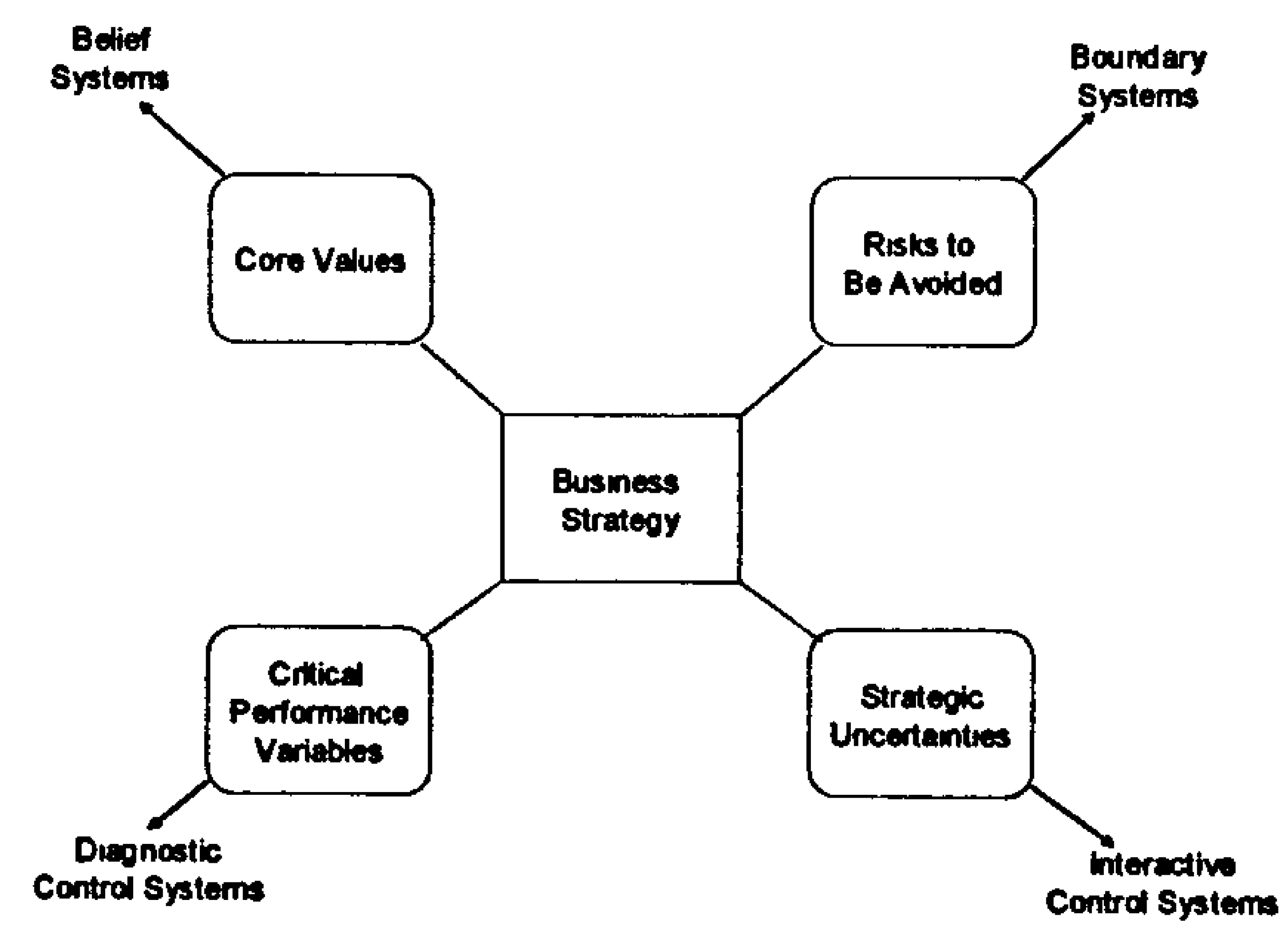


Process Model of Relationship Between Business Strategy & MCS



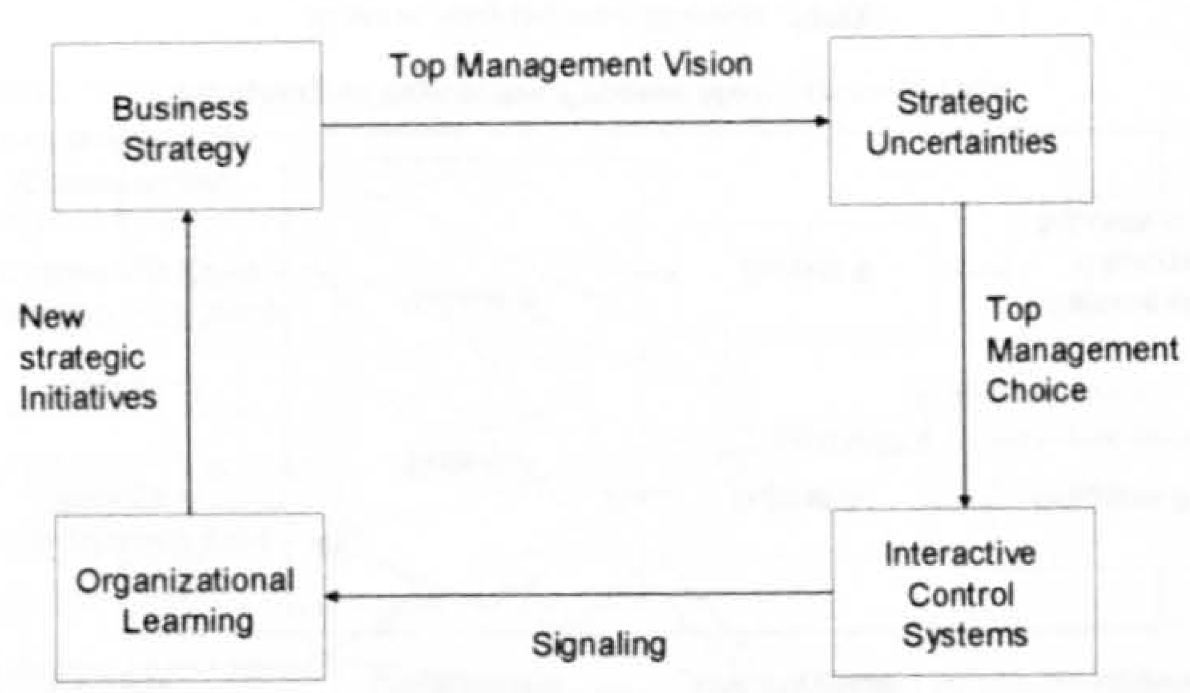
Summary of Interactive Process Model in two Firms

Source: Adapted from (Simons, 1990)



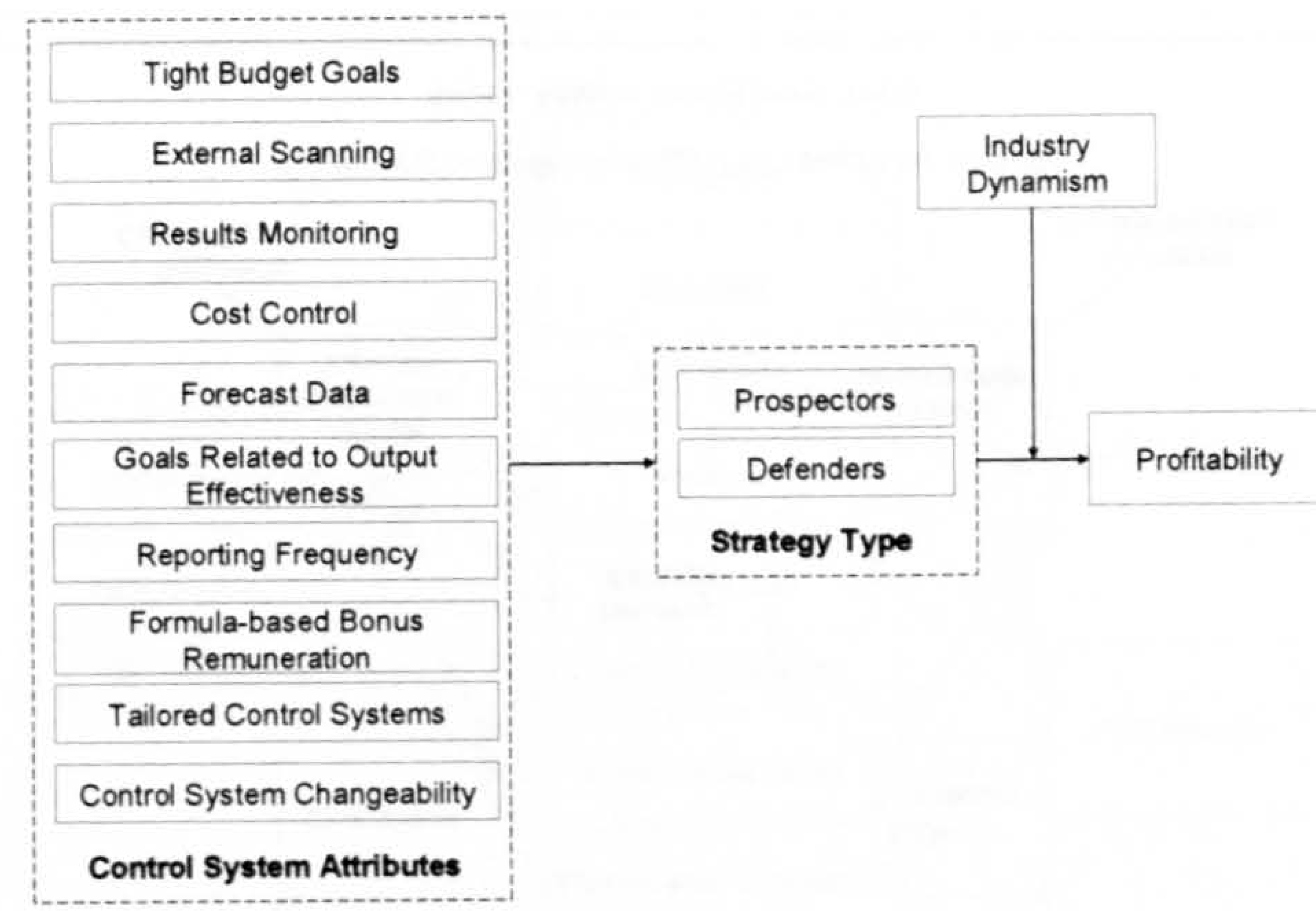
Controlling business strategy: Framework for analysis

Source: Adapted from (Simons, 1994)

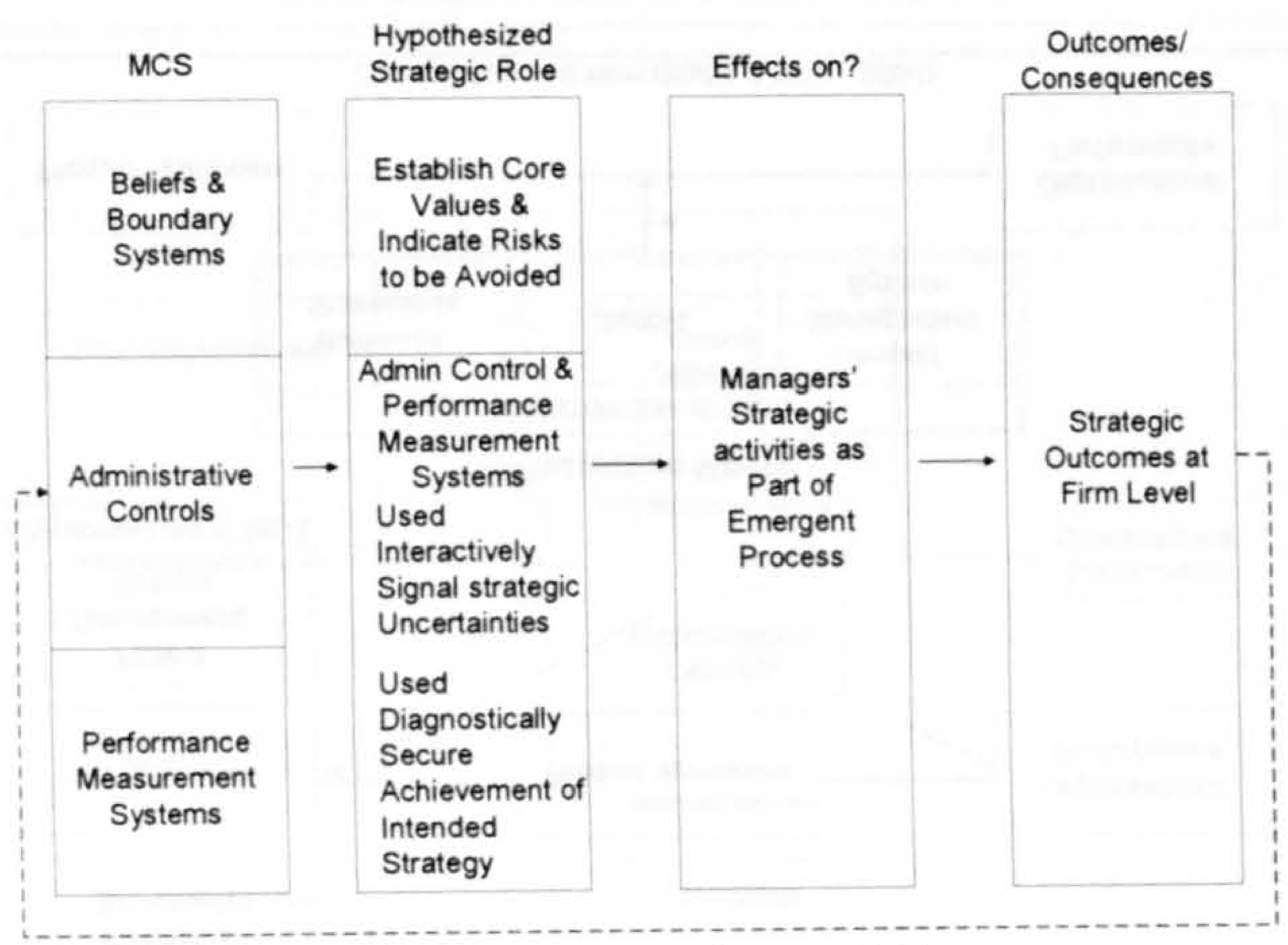


Process Model of Relationship Between Business Strategy & MCS

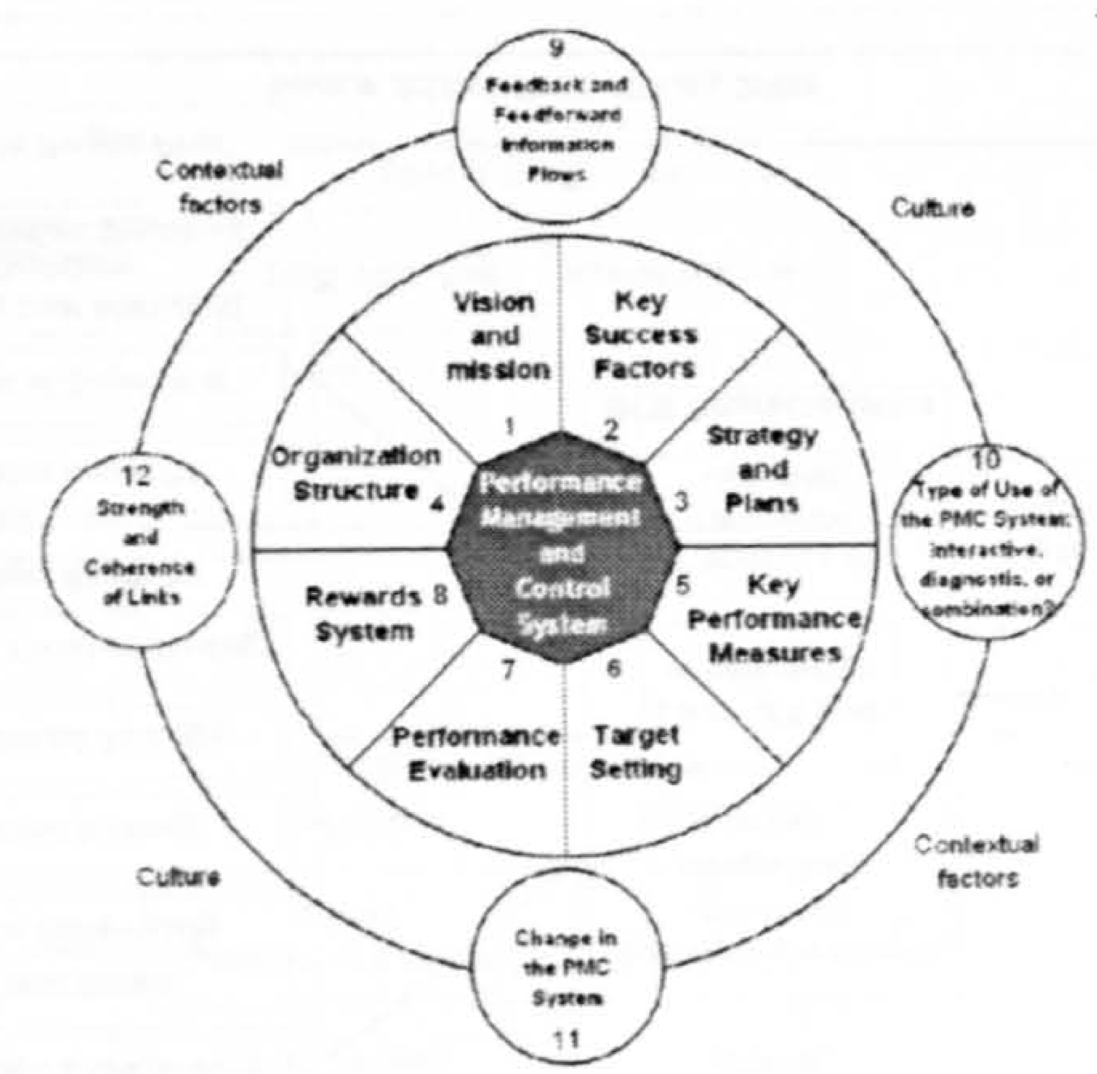
Source: Adapted from (Simons, 1991)



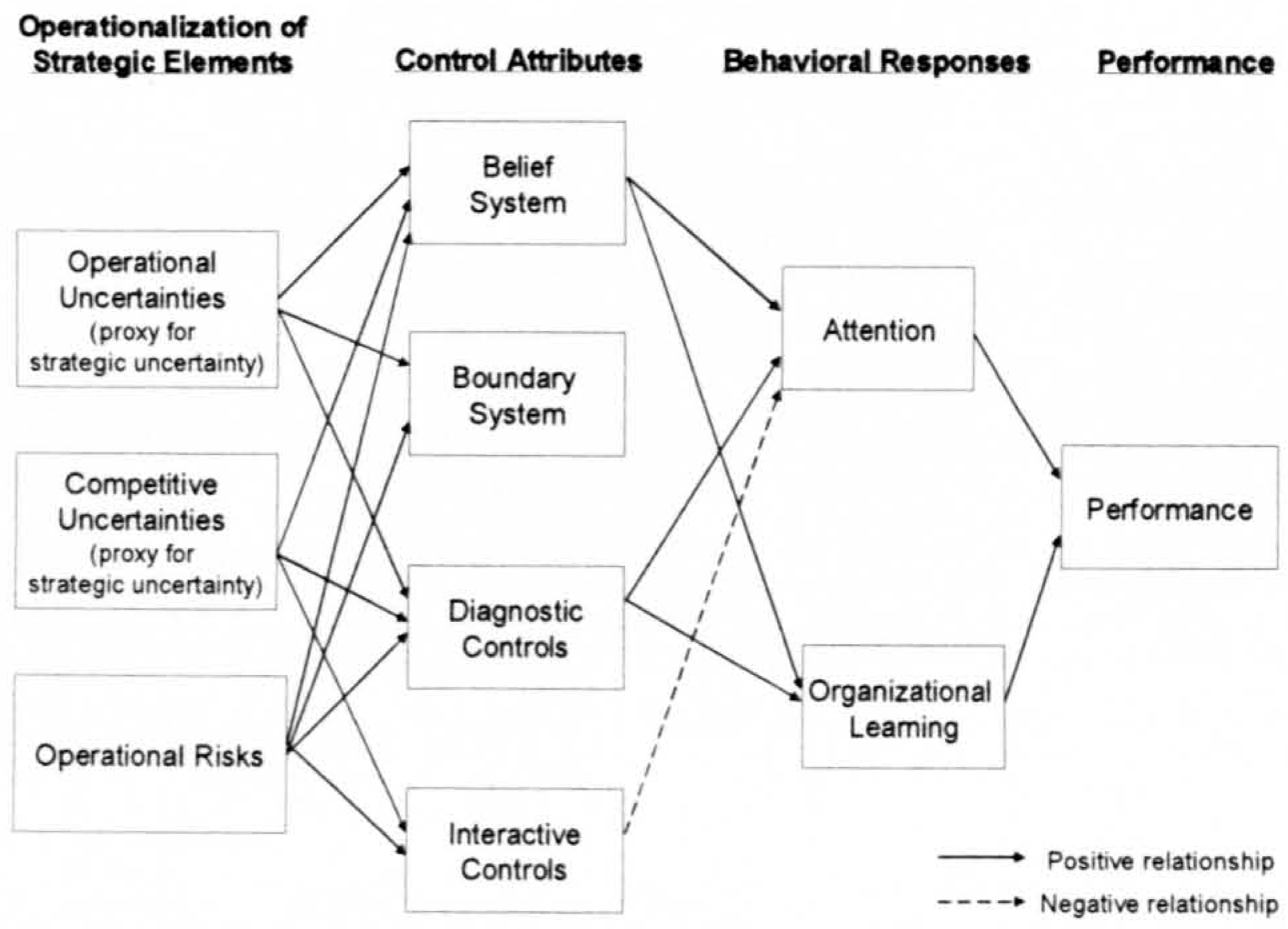
Source: Adapted from (Simons, 1987)



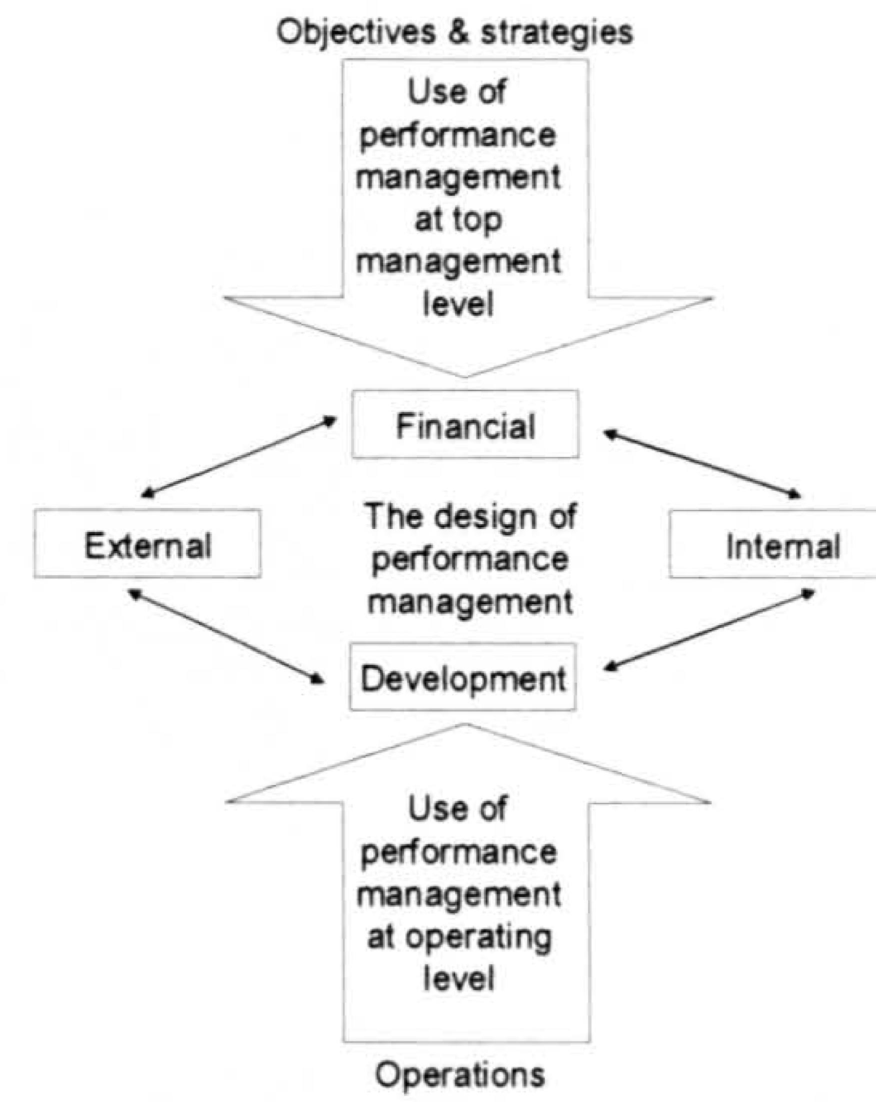
Source: Adapted from (Marginson, 2002)



Source: Adapted from (Ferreira & Otley, 2005)



Source: Adapted from (Widener, 2005)



Source: Adapted from (nilsson & Kald, 2002)