**Reflections on perceived online service quality:
Structure, antecedents, ontology, theory and
measurement**

By
Mary Tate

A thesis submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy in Information Systems

Victoria University of Wellington

2010

Acknowledgements

My sincerest thanks to the people who made this project possible and supported it along the way, especially to:

My supervisory team: Joerg Evermann, supervisor, colleague and friend, without whom this never would have happened; Beverley Hope for her continued support and Stuart Barnes for some important initial direction.

My colleagues, friends and students at the School of Information Management at Victoria University and in the Information Systems community: Sid Huff and Pak Yoong who provided invaluable mentoring at key points; and the people who make academic life stimulating, challenging and fun, especially Brian Harmer, Eusebio Scornavacca, Allan Sylvester, Mark Wilson, Hugo Gong, Haibo Yang and John Campbell.

My examiners: Guy Gable, Brian Corbitt, and Nick Ashill, for extraordinarily supportive, insightful and helpful feedback.

My extended family: this includes at least five PhD's and numerous other scholarly and educational achievements; thank you for a culture of scholarship and life-long learning.

Our friends: for my other life.

My special people: David, Sarah, Emily, Laura and Margaret, as always, thank you for everything.

Note

I would like to note that although I have used the pronoun “we” throughout the thesis, the work I have presented here is my own. The use of the term “we” in this report is mainly stylistic; and also indicates that the supervision team were consulted about important decisions and directions in the research.

Abstract

Online services are becoming increasingly ubiquitous, and this growth has been accompanied by increased business interest in measuring and managing online service quality. This interest is also reflected in a large number of academic studies. Despite this, there is very little consensus about the dimensions and antecedents of perceived online service quality (POLSQ).

We consider two possible reasons for this: first that the phenomenon of online service quality is changing as new technology affordances arise, so instability in the dimensionality would be a result of changes in the underlying phenomenon. Second, the theoretical approach and assumptions that studies of online service quality are usually founded on is flawed.

My research questions are: 1) what is the structure of perceived online service quality? 2) What are the antecedents of perceived online service quality? 3) What is the ontology of perceived online service quality? 4) What are the most appropriate modelling and measurement methods for measuring online service quality quantitatively, and what insights can be gained from psychometrics? 5) What insights does this offer IS researchers for the measurement of user attitudes and perceptions towards technologies?

We find that leading models and instruments tend to be based on exploratory factor analysis and have not been informed by advances in measurement theory, particularly co-variance-based structural equation models, and a sub-set of those models known as multi-indicator structural models.

We apply recent advances in measurement theory to a dataset. We apply and compare four different modelling methods, exploratory factor analysis, confirmatory factor analysis, formative models, and multi-indicator structural models, and discuss the theoretical foundations of each method.

We conclude that POLSQ may not have a separate ontology as a multi-dimensional construct, but overall affect towards the service quality of a website is likely to be the result of the perceived usefulness, perceived ease of use, and perceived trust towards the service. This finding supports the explanatory power of information systems theories over marketing and consumer behaviour theories when studying this phenomenon. We find that user's perceptions of detailed affordances of the service, such as the relevance and timeliness of the information are antecedent to overall affect towards the service, rather than being additional dimensions of POLSQ.

We find that widely used techniques such as exploratory factor analysis have serious drawbacks. We find that multi-indicator structural models provide an accurate and nuanced method for modelling the formation of attitudes and perceptions towards technology, which is also well-grounded in theory from social psychology.

Finally, we suggest that the approach we take to measuring POLSQ is has potential value for other research which aims to measure customer attitudes and perceptions towards technologies.

TABLE OF CONTENTS

1	INTRODUCTION.....	14
1.1	GROWTH IN ONLINE SERVICES AND DEMAND FOR QUALITY MEASURES.....	14
1.2	THE SCOPE OF ONLINE SERVICES INCLUDED IN THIS STUDY.....	15
1.3	THE STATE OF CURRENT RESEARCH.....	16
1.4	EXPLANATIONS FOR THE CURRENT LACK OF PROGRESS	17
1.5	RESEARCH QUESTIONS.....	18
1.6	BACKGROUND AND CHRONOLOGY OF THE STUDY	18
1.7	THE APPROACH TAKEN IN THIS STUDY	20
2	LITERATURE REVIEW.....	22
2.1	OVERVIEW	22
2.2	DEFINITIONS	23
2.2.1	<i>Service Quality</i>	23
2.2.2	<i>Service Quality and Customer Satisfaction</i>	25
2.2.3	<i>Service quality and service encounters</i>	28
2.3	SERVICE QUALITY AND ONLINE SERVICE QUALITY LITERATURE.....	30
2.3.1	<i>The service quality “whakapapa” or family tree</i>	30
2.3.2	<i>Face-to-face Service quality literature</i>	31
2.3.3	<i>Service quality in information systems</i>	41
2.3.4	<i>User satisfaction with end-user computing</i>	43
2.3.5	<i>Quality in web-sites</i>	44
2.3.6	<i>E-commerce and e-service quality literature</i>	49
2.3.7	<i>Domain-specific studies</i>	55
2.3.8	<i>Reflections on service quality and online service quality literature</i>	57
2.3.9	<i>Summary and Integration</i>	59
2.4	OTHER RELEVANT THEORY	63
2.4.1	<i>Expectation disconfirmation theory</i>	64
2.4.2	<i>Theories of user acceptance, adoption and continuance of technology use</i>	65
2.4.3	<i>Aligning expectations and perceptions of online service quality</i>	68
2.4.4	<i>Theories of Task Technology Fit (TTF)</i>	69
2.4.5	<i>E-commerce Success</i>	71
2.4.6	<i>Software Affordances</i>	72
2.4.7	<i>Summary</i>	76
2.5	REFLECTION.....	77
3	THEORY AND MEASUREMENT.....	80
3.1	OVERVIEW	80
3.2	THEORETICAL FOUNDATIONS	80
3.2.1	<i>Introduction</i>	80
3.2.2	<i>Philosophy of science</i>	81
3.2.3	<i>Theory in social sciences</i>	85
3.3	THEORY IN INFORMATION SYSTEMS AND ONLINE SERVICE QUALITY RESEARCH.....	86
3.3.1	<i>The Foundations of Information Systems Theory in Social Psychology</i>	86
3.3.2	<i>A Taxonomy of information systems theory</i>	91
3.3.3	<i>The quality of service quality theories</i>	93
3.3.4	<i>Issues with information systems and service quality theories</i>	95
3.3.5	<i>Summary</i>	101
3.4	ISSUES IN MEASUREMENT	102
3.4.1	<i>Introduction</i>	102
3.4.2	<i>Measurement Theory</i>	103
3.4.3	<i>Factor analytic approaches: Overview of formative and reflective variables</i>	104
3.4.4	<i>Statistical differences</i>	107
3.4.5	<i>Model fit</i>	109
3.4.6	<i>Measurement: Scale and index construction</i>	115
3.4.7	<i>Non-factor analytic approaches</i>	124
3.4.8	<i>Summary</i>	128
3.5	ONTOLOGICAL ISSUES AND ASSUMPTIONS OF MODELLING METHODS	129

3.5.1	<i>The explanations compatible with reflective models</i>	130
3.5.2	<i>The explanations compatible with formative models</i>	132
3.5.3	<i>The explanations compatible with multiple indicator structural models</i>	134
3.5.4	<i>Debates in psychometrics and their relevance to the measurement of online service quality</i>	137
3.6	MODELLING AND MEASUREMENT IN ONLINE SERVICE QUALITY RESEARCH	140
3.6.1	<i>ServQual</i>	141
3.6.2	<i>E-commerce service quality</i>	148
3.6.3	<i>eQual</i>	150
3.7	SYNTHESIS	154
4	METHODOLOGY	158
4.1	OVERVIEW	158
4.2	RESEARCH QUESTIONS AND MODELLING METHODS	158
4.3	INSTRUMENT DEVELOPMENT AND EXPLORATORY FACTOR ANALYSIS (PRESENTED IN CHAPTER 5)	161
4.3.1	<i>Instrument development, pre-testing, and survey administration</i>	161
4.3.2	<i>Exploratory factor analysis</i>	161
4.4	CONFIRMATORY FACTOR ANALYSIS (PRESENTED IN CHAPTER 6)	162
4.5	FORMATIVE MODELS (PRESENTED IN CHAPTER 7)	163
4.6	MULTI-INDICATOR STRUCTURAL MODELS (PRESENTED IN CHAPTER 8)	164
4.7	THE CONTEXT OF THE STUDY	164
4.8	THE DATASET USED FOR THE STUDY	167
4.9	REFERENCES TO THE DATASET IN THE TEXT	167
4.10	THE ADEQUACY OF THE SAMPLE	168
4.10.1	<i>Overview</i>	168
4.10.2	<i>Sample size and exploratory quantitative research</i>	169
4.10.3	<i>Sample size and SEMs</i>	170
4.11	ISSUES AND LIMITATIONS OF THE APPROACH	171
5	EXPLORATORY FACTOR ANALYSIS	173
5.1	OVERVIEW	173
5.2	INSTRUMENT DEVELOPMENT	173
5.2.1	<i>Overview of Our Approach to Instrument Development</i>	173
5.2.2	<i>Specification</i>	175
5.2.3	<i>Scale Development</i>	176
5.2.4	<i>Pre-testing and initial survey administration</i>	182
5.3	OUR APPROACH TO EXPLORATORY FACTOR ANALYSIS	184
5.3.1	<i>Conducting the Survey</i>	184
5.3.2	<i>Descriptive statistics</i>	185
5.3.3	<i>Descriptive Statistics</i>	186
5.3.4	<i>Exploratory Factor analysis (EFA)</i>	188
5.3.5	<i>Overview of our approach to EFA</i>	188
5.3.6	<i>Determine the overall suitability of the data for factor analysis</i>	190
5.3.7	<i>Ensure each variable is suitable for inclusion</i>	192
5.3.8	<i>Determine the number of factors to extract</i>	192
5.3.9	<i>Determine the most appropriate method of factor analysis and factor rotation</i>	197
5.3.10	<i>Consider Item and Factor loadings</i>	199
5.3.11	<i>Refine the analysis</i>	204
5.4	QUESTIONS OF VALIDITY AND RELIABILITY	209
5.4.1	<i>Overall evaluation of the instrument development process</i>	209
5.4.2	<i>Reliability</i>	210
5.4.3	<i>Validity</i>	213
5.5	DISCUSSION	216
5.5.1	<i>The four-factor structure</i>	216
5.5.2	<i>Factor structure compared to eQual</i>	217
5.5.3	<i>Implications of the EFA</i>	221
5.6	CONCLUSION	224
6	FACTOR-ANALYSIS-BASED SEMS	226

6.1	OVERVIEW	226
6.2	SEM INTRODUCTION.....	227
6.2.1	<i>The SEM Measurement Model.....</i>	227
6.2.2	<i>The SEM Structural Model.....</i>	228
6.2.3	<i>Two step modelling.....</i>	229
6.2.4	<i>SEM modelling approaches.....</i>	229
6.3	SEM CONCEPTS	230
6.3.1	<i>Model specification.....</i>	230
6.3.2	<i>Model parsimony.....</i>	230
6.3.3	<i>Metrics.....</i>	231
6.3.4	<i>Error terms.....</i>	231
6.3.5	<i>Parameter Estimates.....</i>	231
6.3.6	<i>Structural or path coefficients.....</i>	231
6.4	OUR APPROACH TO SEM MODELLING	232
6.4.1	<i>Step 1: Select approach to SEM.....</i>	232
6.4.2	<i>Step 2: Model specification.....</i>	232
6.4.3	<i>Step 3: Evaluate data.....</i>	238
6.4.4	<i>Step 4: Select metrics.....</i>	240
6.4.5	<i>Step 5: Evaluate individual model parameters.....</i>	246
6.4.6	<i>Step 6: Evaluate measurement models using CFA.....</i>	253
6.4.7	<i>Step 7: Specify and Evaluate structural model.....</i>	255
6.4.8	<i>Step 8: Plan next steps.....</i>	255
6.5	DISCUSSION.....	258
6.6	CONCLUSION.....	258
7	FORMATIVE MODELS.....	259
7.1	OVERVIEW	259
7.2	FORMATIVE MODEL CONCEPTS.....	259
7.2.1	<i>Co-variance and partial-least squares regression “SEMs”.....</i>	259
7.2.2	<i>What are Formative Indicators?.....</i>	260
7.2.3	<i>Achieving Model Identification for a Formative Model (co-variance-based).....</i>	262
7.2.4	<i>The Disturbance term of a Formative Construct (co-variance-based).....</i>	263
7.3	ISSUES WITH FORMATIVE MODELS.....	264
7.3.1	<i>The Ontology of Formative Models.....</i>	265
7.3.2	<i>Issues with error free measures.....</i>	266
7.3.3	<i>Issues with constraining the model to achieve identification.....</i>	267
7.3.4	<i>Interpretational Confounding.....</i>	267
7.3.5	<i>Co linearity and correlation.....</i>	269
7.3.6	<i>Summary of issues and our approach.....</i>	271
7.4	OUR APPROACH TO FORMATIVE MODEL DEVELOPMENT AND EVALUATION	272
7.4.1	<i>Evaluating Formative Models.....</i>	272
7.4.2	<i>Summary: Evaluation criteria for formative models.....</i>	276
7.5	SPECIFYING THE MODELS	279
7.5.1	<i>Census of indicators.....</i>	280
7.5.2	<i>Identify causality.....</i>	281
7.5.3	<i>Identify reflective Indicators to enable model identification.....</i>	282
7.5.4	<i>Overview of the Models.....</i>	283
7.5.5	<i>Evaluate Models.....</i>	284
7.6	FORMATIVE MODEL 1 (PLS AND SPSS): EVALUATE MEASURES.....	284
7.6.1	<i>Results: Construct Validity.....</i>	285
7.7	FORMATIVE MODEL 2 (COVARIANCE-BASED): PURIFIED MIMIC MODEL.....	289
7.7.1	<i>Results: Model Quality.....</i>	291
7.7.2	<i>Model Fit Summary.....</i>	292
7.8	DISCUSSION.....	294
7.8.1	<i>Model 1.....</i>	294
7.8.2	<i>Model 2.....</i>	296
7.8.3	<i>Revisiting the issues with formative models.....</i>	296
7.9	CONCLUSION.....	300
8	MULTI-INDICATOR STRUCTURAL MODELS.....	301

8.1	OVERVIEW	301
8.1.1	<i>Non-factor analytic models</i>	301
8.2	OUR APPROACH TO MODEL DEVELOPMENT	303
8.2.1	<i>Revised eQual instrument and items</i>	303
8.3	MULTI-INDICATOR STRUCTURAL MODELS	308
8.3.1	<i>Overview</i>	308
8.3.2	<i>MISM-1</i>	311
8.3.3	<i>MISM-2</i>	312
8.3.4	<i>MISM-3</i>	313
8.3.5	<i>MISM-4</i>	314
8.3.6	<i>MISM-5</i>	315
8.3.7	<i>MISM-6</i>	316
8.3.8	<i>MISM-7</i>	317
8.4	RESULTS.....	319
8.4.1	<i>Model Comparison</i>	323
8.5	DISCUSSION.....	324
8.5.1	<i>Overview</i>	324
8.5.2	<i>Model fit</i>	324
8.5.3	<i>Comparison with our other models</i>	327
8.5.4	<i>Comparison with models from literature</i>	329
8.6	CONCLUSION	329
9	DISCUSSION.....	332
9.1	OVERVIEW	332
9.2	THE STRUCTURE OF POLSQ	332
9.3	THE ANTECEDENTS OF POLSQ.....	334
9.3.1	<i>An unstructured set of antecedents</i>	335
9.3.2	<i>Formative model of trust, usefulness and ease of use</i>	336
9.3.3	<i>More nuanced models of the antecedents</i>	337
9.3.4	<i>Summary</i>	338
9.4	THE ONTOLOGY OF POLSQ.....	339
9.4.1	<i>Ontological assumptions</i>	339
9.4.2	<i>These assumptions applied to POLSQ</i>	342
9.5	THEORETICAL, PARADIGMATIC AND MEASUREMENT APPROACHES	343
9.5.1	<i>What is the most appropriate theoretical approach?</i>	343
9.5.2	<i>What is the most appropriate paradigmatic approach?</i>	347
9.5.3	<i>Measurement approaches for POLSQ</i>	348
9.5.4	<i>What insights can be gained from psychometrics?</i>	353
9.6	IMPLICATIONS FOR IS RESEARCH	355
9.6.1	<i>Use of Formative Latent Variables in Theories of Attitudes and Behaviour towards Technology</i>	355
9.6.2	<i>The Scope and Nature of Information Systems Theories of Attitudes and Perceptions Towards technology</i>	356
9.6.3	<i>Approaches for information systems researchers</i>	357
9.6.4	<i>The provenance and further development of theories of attitudes and perceptions towards technology</i>	358
9.7	IMPLICATIONS FOR PRACTICE.....	359
9.8	LIMITATIONS	360
9.8.1	<i>Scope</i>	360
9.8.2	<i>Research Design</i>	361
9.8.3	<i>Data analysis</i>	361
9.8.4	<i>Alternatives</i>	363
10	CONCLUSION.....	364
	REFERENCES	372

APPENDICES

APPENDIX I:	CROSS REFERENCE
APPENDIX II:	SURVEY QUESTIONS PRE-TEST
APPENDIX III:	ONLINE SURVEY PRE-TEST
APPENDIX IV:	SURVEY INSTRUMENT
APPENDIX V:	FACTOR STRUCTURE COMPARISONS
APPENDIX VI:	ASSESSMENT OF NORMALITY
APPENDIX VII:	AMOS REPORTS, CONFIRMATORY FACTOR ANALYSIS
APPENDIX VIII:	AMOS REPORT, FORMATIVE MODEL
APPENDIX IX:	AMOS REPORTS, MULTI-INDICATOR STRUCTURAL MODELS

LIST OF FIGURES

Figure 1-1: <i>The Chronology of the Study</i>	19
Figure 2-1: <i>Spreng and Mackoy's model (Spreng & Mackoy, 1996)</i>	26
Figure 2-2: <i>Iacobucci and Ostrom's model (Iacobucci & Ostrom, 1995)</i>	27
Figure 2-3: <i>The general approach to each of the construct literatures based on table summaries</i>	31
Figure 2-4: <i>ServQual dimensions</i>	32
Figure 2-5: <i>Brady and Cronin's integrated model (Brady & Cronin, 2001)</i>	37
Figure 2-6: <i>Model of e-customer satisfaction</i>	51
Figure 2-7: <i>The iterative relationship between past use and future intention</i>	64
Figure 2-8: <i>The Technology Acceptance Model (TAM)</i>	65
Figure 2-9: <i>Task technology Fit and the Technology to Performance Chain</i>	70
Figure 2-10: <i>Delone and McLean's Information Systems Success Model</i>	71
Figure 3-1: <i>Current beliefs lead to attitudes, intentions and behaviours (Adapted from Fishbein & Ajzen (1975, p. 15))</i>	86
Figure 3-2: <i>The formation of attitudes towards websites</i>	90
Figure 3-3: <i>The Gaps Model (Parasuraman, et al., 1985)</i>	99
Figure 3-4: <i>Reflective indicators</i>	107
Figure 3-5: <i>Example of a MIMIC model from (Diamontopoulos & Winklhofer, 2001)</i>	118
Figure 3-6: <i>Alternative conceptualisation of the model in Figure 3.4 from (Diamontopoulos & Winklhofer, 2001)</i>	119
Figure 3-7: <i>Formative indicators as "short-hand" for a causal model, from (Diamontopoulos, 2006)</i>	123
Figure 3-8: <i>The implied latent variables in Figure 3.6 made explicit, from (Diamontopoulos, 2006)</i>	123
Figure 3-9: <i>Possible common cause of factors, from (Hayduk, 1996)</i>	126
Figure 3-10: <i>The usual assumption of factor models, from (Hayduk, 1996)</i>	127
Figure 4-1: <i>Process model of methodology phases</i>	160
Figure 4-2: <i>Screen shot of Victoria University of Wellington's web portal</i>	166
Figure 5-1: <i>Our approach to instrument development and testing</i>	174
Figure 5-2: <i>Our approach to exploratory factor analysis</i>	189
Figure 5-3: <i>Sample scatter plots of variables</i>	191
Figure 5-4: <i>Catell's Scree Plot</i>	195
Figure 6-1: <i>Our approach to Confirmatory factor Analysis using SEM</i>	233
Figure 6-2: <i>Model 1: Hypothesised 4-factor model from literature review</i>	234
Figure 6-3: <i>Revised 4-factor model following EFA</i>	235
Figure 6-4: <i>Model 3: 3-factor model based on eQual 4.0</i>	236
Figure 6-5: <i>Model 4: 5-factor structure based on eQual 4.0</i>	237
Figure 7-1: <i>A formative latent variable predicts a single reflective latent variable</i> ..	266
Figure 7-2: <i>Formative indicators directly predict the latent variable</i>	266
Figure 7-3: <i>The process of developing and specifying formative models</i>	279
Figure 7-4: <i>Model with all the indicators loading on perceived online service quality</i>	284
Figure 7-5: <i>Formative model 2, with formative latent variables for ease of use, usefulness and trust</i>	290
Figure 7-6: <i>Example of formative latent variables with one indicator</i>	297
Figure 7-7: <i>Formative indicators are used as "shorthand" for single indicator formative latent variables</i>	298

Figure 8-1: <i>MISM-1</i>	311
Figure 8-2: <i>MISM-2</i>	312
Figure 8-3 : <i>MISM-3</i>	313
Figure 8-4: <i>MISM-4</i>	314
Figure 8-5: <i>MISM-5</i>	316
Figure 8-6: <i>MISM-6</i>	317
Figure 8-7: <i>MISM-7</i>	318
Figure 9-1: <i>The formation of attitudes towards websites (revisited)</i>	339
Figure 9-2: <i>Affordance theory allows us to identify the salient features of technologies for inclusion in theory</i>	358
Figure 9-3: <i>Theory development and modelling</i>	359

LIST OF TABLES

Table 2.1: <i>An integrated set of face-to-face Service Quality dimensions</i>	40
Table 2.2: <i>Online service quality whakapapa: Service quality dimensions for an IS department</i>	41
Table 2.3: <i>Online service quality whakapapa: Dimensions of end-user computing satisfaction</i>	43
Table 2.4: <i>Online service quality whakapapa: dimensions of website quality</i>	48
Table 2.5: <i>Online service quality whakapapa: dimensions of e-service quality</i>	54
Table 2.6: <i>Online service quality whakapapa: domain-specific studies</i>	57
Table 2.7: <i>Integrated e-service dimensions framed by eQual dimensions</i>	60
Table 2.8: <i>Use of common indicators in eQual and the Information Systems Success Model</i>	72
Table 2.9: <i>A comparison of definitions of affordances</i>	74
Table 3.1: <i>Kuhnian analysis of eQual</i>	100
Table 3.2: <i>The differences between formative and reflective indicators</i>	106
Table 3.3: <i>Impact of mis-specified models on fit indices-1</i>	112
Table 3.4: <i>Impact of mis-specified models on fit indices-2</i>	113
Table 3.5: <i>Heuristics for Formative Models</i>	115
Table 3.6: <i>The potential implications of factor indeterminacy for psychometric research, adapted from (Schonemann, 1996a)</i>	140
Table 3.7: <i>Measurement issues in ServQual</i>	144
Table 3.8: <i>Studies of the dimensions of e-service quality, adapted from (Alzola & Robaina, 2005).</i>	149
Table 3.9: <i>Analysis of eQual indicators</i>	152
Table 4.1: <i>Example of cross reference for question numbers, full text and short text</i>	167
Table 5.1: <i>Roles held by survey respondents</i>	185
Table 5.2: <i>Frequency of access of survey respondents</i>	186
Table 5.3: <i>Age of survey respondents</i>	186
Table 5.4: <i>Descriptive statistics</i>	187
Table 5.5: <i>KMO measure of sampling adequacy and Bartlett's Test of sphericity</i>	190
Table 5.6: <i>Total variance explained</i>	194
Table 5.7: <i>Parallel Analysis</i>	196
Table 5.8: <i>Comparison between actual eigenvalues and parallel analysis</i>	197
Table 5.9: <i>Sample size and factor loading</i>	200
Table 5.10: <i>Factor loading and practical significance</i>	200
Table 5.11: <i>Initial solution: PCA with Varimax (orthogonal) rotation (values below 0.4 not reported)</i>	201
Table 5.12: <i>Initial solution, Pattern Matrix: PCA with Oblimin (oblique) rotation (values below 0.4 not reported)</i>	202
Table 5.13: <i>Initial solution, Structure Matrix: PCA with Oblimin (oblique) rotation (values below 0.4 not reported)</i>	203
Table 5.14: <i>Initial solution, Component Correlation Matrix</i>	204
Table 5.15: <i>Best solution, Pattern Matrix: PCA with Oblimin (oblique) rotation</i>	205
Table 5.16: <i>Best solution, Structure Matrix: PCA with Oblimin (oblique) rotation</i>	206
Table 5.17: <i>Best solution, Component Correlation Matrix</i>	207
Table 5.18: <i>Communalities</i>	208
Table 5.19: <i>Our instrument development process compared to best practice</i>	209

Table 5.20: <i>Reliability Statistics for Factor 1</i>	212
Table 5.21: <i>Reliability Statistics for Factor 2</i>	212
Table 5.22: <i>Reliability Statistics for Factor 3</i>	213
Table 5.23: <i>Reliability Statistics for factor 3</i>	213
Table 5.24: <i>Mandatory validities applied to our instrument</i>	215
Table 5.25: <i>The evolution of eQual (formerly WebQual) factors</i>	218
Table 6.1: <i>Summary of SEM fit indices and guidelines for use: Overall fit indices</i> ..	242
Table 6.2: <i>Summary of SEM fit indices and guidelines for use: comparing a given model with the null model</i>	243
Table 6.3: <i>Summary of SEM fit indices and guidelines for use: comparing predicted versus observed covariance but penalising for lack of parsimony</i>	244
Table 6.4: <i>Summary of SEM fit indices and guidelines for use: measures based on information theory</i>	245
Table 6.5: <i>Evaluating individual model parameters</i>	247
Table 6.6: <i>Summary of important results from analysis of individual model parameters</i>	249
Table 6.7: <i>Factor Loadings (Model 1)</i>	250
Table 6.8: <i>Factor Loadings (Model 2)</i>	251
Table 6.9: <i>Factor Loadings (Model 3)</i>	252
Table 6.10: <i>Factor Loadings (Model 4)</i>	252
Table 6.11: <i>Squared multiple correlations</i>	253
Table 6.12: <i>Summary of fit indices for the four CFA models presented in this chapter</i>	257
Table 7.1: <i>Issues with formative models and our approach</i>	271
Table 7.2: <i>Impact of error terms of different magnitudes on the model</i>	275
Table 7.3: <i>Simple evaluation criteria for formative models (revisited)</i>	277
Table 7.4: <i>More comprehensive evaluation criteria for formative models</i>	278
Table 7.5: <i>Expert validation of indicators</i>	280
Table 7.6: <i>Formative indicators</i>	282
Table 7.7: <i>Summary of indicators specified reflectively to identify the models</i>	283
Table 7.8: <i>Evaluating collinearity using regression</i>	285
Table 7.9: <i>Evaluating collinearity using eigenvalues and condition indices</i>	286
Table 7.10: <i>Item loadings on dependant variable (online service quality): PLS</i>	288
Table 7.11: <i>The r^2 of perceived online service quality from Formative Model 1 (PLS).</i>	289
Table 7.12: <i>Squared Multiple Correlations: (Group number 1 - Default model)</i>	291
Table 7.13: <i>Variances (error terms): (Group number 1 - Default model)</i>	292
Table 7.14: <i>Regression Weights: (Group number 1 - Default model)</i>	292
Table 7.15: <i>CMIN</i>	292
Table 7.16: <i>Baseline comparisons</i>	293
Table 7.17: <i>Parsimony adjusted measures</i>	293
Table 7.18: <i>NCP</i>	293
Table 7.19: <i>FMIN</i>	293
Table 7.20: <i>RMSEA</i>	293
Table 7.21: <i>AIC</i>	294
Table 7.22: <i>ECVI</i>	294
Table 7.23: <i>HOELTER</i>	294
Table 8.1: <i>Candidate latent construct definitions</i>	305
Table 8.2: <i>The indicators included in the models in this chapter.</i>	309
Table 8.3: <i>Overall fit metrics-1</i>	320

Table 8.4: <i>Overall fit metrics-2</i>	321
Table 8.5: <i>Squared multiple correlations</i>	322
Table 8.6: <i>Overall fit indices compared-1</i>	323
Table 8.7: <i>Overall fit indices compared-2</i>	323
Table 9.1: <i>A comparison of modelling methods</i>	350
Table 10.1: <i>Theme 1: Online service quality research</i>	366
Table 10.2: <i>Theme 2: The structure and antecedents of POLSQ</i>	367
Table 10.3: <i>Theme 3: The ontology of POLSQ</i>	368
Table 10.4: <i>Theme 4: Theoretical and paradigmatic approaches</i>	369
Table 10.5: <i>Theme 5: The measurement of POLSQ</i>	370
Table 10.6: <i>Theme 6: IS research into user attitudes and perceptions towards technology</i>	370

1 Introduction

1.1 Growth in online services and demand for quality measures

As the Internet has matured, there has been a huge growth in the volume and value of online transactions. In Australia, for example, not only has the total number of Internet users increased, but the percentage of all internet users engaged in online transactions has increased steadily, from 41% in 2001 to 64% in 2006. By 2006, 43% of all Australian businesses were conducting transactions over the Internet (Department_of_Broadband_Communications_and_the_Digital_Economy, 2008). In the United States of America, by 2006, approximately 2/3 of Internet use was for online shopping (e-marketer.com, 2006). In the government sector, many countries now offer the ability to access a wide range of public sector services and transactions online (Accenture_Report, 2003).

The rise in online transactions has been accompanied by an increased interest in the quality of customer service delivered online. Managing service quality, and identifying those functions and features that contribute to positive user perceptions, and consequently customer loyalty and retention have long been recognised as essential business skills. This is reflected in the popularity of service quality as an area of academic enquiry. A cursory search on online databases conducted now will find thousands of peer-reviewed studies in the last ten years that include “service quality” in the abstract. There is an old saying that “you can’t manage what you can’t measure”. Concern for effective management of service quality has led to a strong uptake of instruments for measuring service quality by the business community, as well as a large volume of academic research (Sylvester, Tate, & Johnstone, 2007). Many of these measurement instruments are based directly or indirectly on ServQual, a popular scale for measuring perceived service quality (Parasuraman, Zeithaml, & Berry, 1988). Interest in service quality has extended into the management of the information systems function and the increasing number of services delivered via information and communication technologies (ICTs). Academic studies to measure the service delivered by information systems departments and information systems have been popular in information systems (IS) and electronic commerce (EC) fields since the concept of service quality was first appropriated from marketing (Pitt &

Watson, 1994). This research has been enthusiastically adopted for measuring service quality in electronic transactions as well. A recent literature review of online service quality research found eighteen papers, all published between 2000 and 2005 (Alzola & Robaina, 2005). However, despite the large number of studies published, there is little consensus about the nature, dimensions and antecedents of online service quality.

1.2 The scope of online services included in this study

Within the Information Systems field, there are several important distinctions that need to be made when we consider the notion of service, between (1) the service delivered by IS personnel, (2) the service science conceptualisation of software as a service, and (3) an online self-service encounter where a customer utilises technology (primarily a website) to obtain a service without direct human interaction. The original IS-ServQual (Pitt & Watson, 1994) was concerned with the activities of information systems departments and staff. More recently, the notion of software-as-a-service (SaaS) has become popular as part of the emerging phenomena of cloud computing, web-services, and service-oriented architecture. SaaS provides customers with complete, turnkey applications, even complex systems such as CRM or ERP, through web browser (Pitt & Watson, 1994)s via the Internet (Leavitt, 2009). This aspect of IS service delivery is carried out between two organisations (B2B), and is related more to outsourcing management (from the client firm's perspective) and to software development and provision (from the provider firm's perspective). This is distinct from IS mediated self-service which is the focus of this study. The aspect of information systems service that we consider in this discussion is that of individuals utilising an online technology (primarily a website, through any of a growing number of web access devices) to obtain a service such as (for example) banking, accessing online educational materials, or making an appointment, without requiring a human agent. This is conceptually distinct from the other two definitions of IS “service”. It is this third aspect of IS service that is the focus of this study. The first two notions of service are explicitly excluded.

1.3 The state of current research

Despite strong academic interest, and clear salience to practice, there is considerable confusion about what constitutes online service quality. It has been argued that despite extensive research, studies of online service quality have not resulted in a coherent body of knowledge about online service quality. (Tate & Evermann, 2009a). Confusion exists about the dimensions, indicators, and theory net of online service quality.

There is no consensus on the nature or number of dimensions of online service quality. New studies continue to propose new dimensions, or new combinations of existing dimensions, for example (Alzola & Robaina, 2005; Tate, Evermann, Hope, & Barnes, 2007). There is semantic ambiguity, where the same term for a construct is used with different meanings and indicators, or different constructs (and terms) are defined using the same indicators (Tate & Evermann, 2009a). Instruments are not confirmatively revalidated (e.g. using SEM modelling), but rather freely reassembled. This practice of adding, subtracting, modifying and recombining indicators makes meta-analysis and meaningful comparisons between studies impossible (Tate & Evermann, 2009a).

Perhaps as a result of the lack of clarity in definitions and constructs, there is no clear theory net. A recent conference presenter described the frustration of dealing with a body of literature where “the arrows in models are going every which way”. Less informally, Tate and Evermann found constructs named “service quality” included in different information systems theories. For example, information quality has been modelled both as a completely independent construct from online service quality and as a sub-dimension of overall online service quality (Tate & Evermann, 2009a).

The complexities and contradictions in the literature were a major influence on the research presented here. We consider two explanations for the current lack of progress as background to the research questions.

1.4 Explanations for the current lack of progress

Two possibilities exist to explain this lack of consensus and progress. The first is that the phenomenon of online service quality is, in fact, changing, and this explains the inconsistent results over time. The other is that the existing theoretical approaches to measuring online service quality are flawed.

We first consider the possibility that the phenomenon of online service quality is changing. Early measurement instruments concentrated on information provisioning and communication (Barnes, 2001). These were the main function of early websites, and it was not unusual for organisations to have a “brochure ware” Internet presence that fell short of full transactional capability (Lawrence, Corbitt, Fisher, Lawrence, & Tidwell, 1998). More recently, online shopping and full transactional capability has become the standard. Increasingly, as technologies offer new possibilities, such as e-commerce portals, or integration with SMS and mobile phone technologies, customers come to expect these features in their turn (Tate, Evermann, & Hope, 2008b). These changes could contribute to changes over time in user perceptions of what constitutes online service quality, which might result in the instability of the constructs

The second possibility is that the existing theoretical and measurement approaches are flawed. Customer perceptions and attitudes towards technology are often conceptualised as an internal psychological state (perceptions, attitudes, or beliefs) held by individuals. For example, Gefen et al. define trust towards technology as “an *expectation... one's belief that...*” (Gefen, Karahanna, & Straub, 2003, p. 54 our emphasis). Researchers tend to make the implicit assumption that these internal states exist (a realist ontology). For example, trust is described as “*the defining attribute of the relationship, determining its very existence and nature*” (Gefen et al., 2003, p. 54).

The investigation of these internal states is frequently carried out quantitatively using survey research and statistical analysis techniques including exploratory and confirmatory factor analysis and regression and co-variance-based models. Modelling techniques used within the information systems community draw heavily from methods that originated within psychometrics. However, it frequently takes a period

of time for debates and issues in reference disciplines to be adopted by the mainstream of the information systems research community.

Psychometric journals and discussion boards have debated (for example): the use of factor analysis in behavioural research (in a special issue of the *Multivariate Behavioural Research* journal, volume 31, issue 4 in 1996); the theoretical and ontological implications of various methods popularly used for measuring psychological states, for example, work by Denny Boorsboom (Borsboom, 2005; Borsboom, Mellenbergh, & van Heerden, 2003); and the most appropriate approach for developing and testing co-variance-based models (for example, the SEMNET discussion board archive).

It is likely that these debates will offer insights for information systems researchers measuring user attitudes and perceptions towards technology, which may allow us to re-evaluate existing knowledge and open avenues for new research.

1.5 Research Questions

The research questions are:

1. What is the structure of POLSQ?
2. What are the antecedents of POLSQ?
3. What is the ontology of POLSQ?
4. What are the most appropriate theoretical, paradigmatic and measurement approaches for measuring online service quality quantitatively, and what insights can be gained from psychometrics?
5. What insights does this offer IS researchers for the measurement of user attitudes and perceptions towards technologies?

1.6 Background and chronology of the study

The discerning reader may already have some hunches about the background to the research questions. Why select online service quality research for the study, and why suspect that there might be inadequacies in the existing measurement instruments?

The origins of this study lay in another study that did not proceed as expected. Our early assumption was that the structure of online service quality was well established. Our intention was to use a tested instrument for measuring POLSQ as the dependant variable in another model. The chronology of the study is presented in Figure 1.1.

Figure 1-1: *The Chronology of the Study*

It is important to realise that some of the psychometric research that was extremely influential for us was published AFTER the studies that we based the initial models on. We acknowledge that research can only be conducted using the knowledge available at the time, although it behoves later researchers to reinterpret that knowledge in the light of more recent advances. The study aims to contribute to this ongoing discourse.

1.7 The Approach taken in this Study

Broadly, the approach taken in this study is to investigate, quantitatively, the two explanations we posited for the lack of consensus in the measurement of POLSQ. The first assumes that theoretical and measurement approaches popularly used by information systems researchers are correct, but the world has changed, which leads to different perceptions of online service quality. The second is that the theoretical and measurement approaches have limitations, and can be improved by applying insights from psychometric research.

The rest of the thesis is organised as follows. In chapter 2, the literature review, we examine existing research in service quality and related constructs such as customer satisfaction. We also examine affordance theory, from human computer interaction, general theories of attitudes and perceptions from social psychology, and the application of those theories to attitudes and perceptions towards technologies. In Chapter 3 we review theoretical and paradigmatic approaches to measuring service quality and other attitudes and perceptions. We examine and critique existing assumptions, theories and measurement approaches; and suggest alternatives. This discussion forms the basis of the research design and research models. In chapter 4 we describe the research method and sample. This chapter is followed by four chapters that report the models, data analysis and results we obtained using different measurement approaches. Chapter 5 reports the exploratory factor analysis. Chapter 6 reports the confirmatory factor analysis. Chapter 7 reports the formative models, and Chapter 8 reports the multi-indicator-structural-models. In chapter 9, we compare the approaches and discuss the implications for the online service quality construct itself, its measurement, and wider issues of measurement of customer attitudes and

perceptions towards technology. We finish with implications for the future and in Chapter 10 offer some conclusions.

2 Literature review

2.1 Overview

This chapter offers a literature review that supports the first line of enquiry: that the structure of online service quality is unstable because the underlying technology is evolving. As we noted in the chronology of the study, this stage of the research is aiming at a synthesis of previous literature (despite the difficulties). We are aiming to identify what is common from previous research, and what has changed. We investigate the construct of POLSQ and its (extensive) provenance in previous research literature.

This chapter follows the widely used technique of “narrative review”; which is the traditional way of reviewing the literature and is skewed towards a qualitative interpretation of the literature. It is conducted by verbally describing the past studies, focusing on theories and frameworks, elementary factors and their research outcomes, with regard to arriving at hypothesized relationships (King & He, 2005). A risk of this approach is that there is no standardized procedure for narrative review. It is not unusual for “*two reviews to arrive at rather different conclusions from the same general body of literature*” (Guzzo, Jackson, & Katzell, 1987, p. 408).

The remainder of this chapter is presented in four sections. In overview, section 2.2 offers definitions of service quality, and examines the distinctions between service quality and closely-related constructs. Having examined candidate definitions, section 2.3 describes the complex provenance of online service quality and offers a synthesis of relevant literature. We then examine other literature which is used to inform various aspects of the arguments included in this thesis in section 2.4. In section 2.5 we offer some reflections on the previous sections, and the process of developing a narrative literature review in this complex and heterogeneous research area.

In Section 2.3 we offer some leading definitions of service quality, and a discussion of the difference between service quality and related concepts like customer satisfaction and service encounters. Next, the distinction between the process of receiving an on-

line service, and the product used to deliver the service is discussed. Then, the provenance of on-line service quality literature is discussed, in particular service quality literature, and literature on user satisfaction with end-user computing. This is followed by a review of studies that apply service quality literature to information systems to establish the dimensions of IS service quality. Following that, general studies of the dimensions of e-commerce quality are discussed, and literature that considers on-line services specifically is discussed. Finally, an integration of the dimensions identified from each of the literatures is presented.

In section 2.4, we examine the wider theoretical context for the study and measurement of online service quality. We first review expectation-disconfirmation theory, theories of user attitudes and perceptions towards technology, and the use of affordance theory for determining the functional affordances of technologies. We also review other constructs that we found to have potential for developing alternative conceptualisations of online service quality, including ease of use, usefulness, trust and fit.

In section 2.5, we reflect on the process we have conducted to develop a narrative literature review for a phenomenon with extensive research “family tree”, and its strengths and weaknesses.

2.2 Definitions

2.2.1 Service Quality

Although very much studied, Service Quality is a difficult concept to define. Some overall definitions have been proposed by leading theorists, and numerous studies have been done to explore the dimensions of the construct, and to distinguish it from other closely related constructs such as customer satisfaction.

Some examples of definitions include:

“Service quality is a comparison between expectations and performance”
(Parasuraman, Zeithaml, & Berry, 1985, p. 42);

“An evaluative, affective, or emotional response” (Oliver & Swan, 1989, p. 1);

Perception of service quality is determined by “prior expectations of what will and what should transpire...and the actual delivered service”
(Parasuraman, Zeithaml, & Berry, 1994, p. 111);

“Perhaps [service] quality factors are those primarily under the control of management....in contrast, perhaps the [customer] satisfaction factors are those that impact the experiential aspects of the service purchase from the customer’s point of view” (Iacobucci & Ostrom, 1995, pp. 295-296).

We can see some consensus at the highest, most abstract level that customer perceived service quality is a customer judgement based on a comparison between prior expectations and performance. However, even this has been called into question, with Cronin and Taylor finding little support for the expectations construct, at least as an explicit construct that customers could readily quantify (Cronin & Taylor, 1994). This does not necessarily invalidate the conceptual definition, as the customer might be making implicit comparisons when arriving at their assessment of the service they have received.

Despite some loose consensus, a clear, usable definition of service quality that can be applied directly to e-services does not exist in the literature. Service quality is a many-faceted concept, and many studies have illuminated different aspects of service quality, without providing the definitive definition, if such a thing can exist. It is also unclear how existing definitions will transfer to a computer mediated, e-service or self-service environment, where aspects of the service experience previously supplied by the organisation have been transferred to the customer. If a customer connects to a web site over a slow Internet connection, they may not receive the service they desire, but will they perceive the service quality as poor?

Most studies have attempted to define the service quality in terms of its component parts, or constructs. Accordingly, this discussion takes a "bottom-up" approach to building a definition of e-service quality.

2.2.2 Service Quality and Customer Satisfaction

One key area of confusion has been clearly distinguishing perceived service quality from closely related concepts such as product quality and customer satisfaction.

Although there is a lot of similarity between customer satisfaction and service quality, (Spreng & Mackoy, 1996) found that many researchers stated they were different constructs, without being clear what the differences were (Bitner, 1990; Parasuraman et al., 1988; Taylor & Barker, 1994).

Both service quality and customer satisfaction are commonly conceptualised as some form of evaluative response by the customer that compares the performance of the product or service and some standard.

This suggests that one possible difference between service quality and customer satisfaction is a difference in the standard of comparison used by the customer (Bitner, 1990; Oliver & Swan, 1989; Parasuraman et al., 1994). The differences in comparison standard are hypothesised as different antecedents, following the argument that if the antecedents of the two constructs cannot be demonstrated to be different then the concepts are structurally equivalent and cannot be studied independently.

Spreng and Mackoy (1996), and Iacobucci and Ostrom (1995) both developed models of the antecedents service quality and customer satisfaction based on extensive literature reviews.

Spreng and Mackoy's constructs (see Figure 2.1) were tested using a survey methodology. They found considerable support for a modified version of Oliver's model (Oliver & Swan, 1989), which defines the key determinant of customer perceived service quality as "desires congruity" – a comparison between "performance and what the customer feels the firm should provide" quoted in (Spreng & Mackoy, 1996, p. 202). By contrast customer satisfaction assessments are formed based on a comparison between predictive expectations ("what the customer believes will happen"), and what actually occurred. In particular, lower expectations can result

in customer satisfaction, even if their overall assessment of the service quality was poor.

Figure 2-1: *Spreng and Mackoy's model (Spreng & Mackoy, 1996)*

Iacobucci and Ostrom took a different approach. In their study, a qualitative, critical incident method approach was adopted to capture the customer view of the constructs (Iacobucci & Ostrom, 1995). This produced a more concrete set of antecedents that included constructs such as price, "back stage" (supporting back-office systems), expertise, timeliness, service recovery and physical environment. Iacobucci and Ostrom found that "supply side" antecedents such as price, back-stage and expertise were most likely to affect customer perceived service quality, while "demand side" antecedents such as timeliness, service recovery and physical environment were most likely to affect customer satisfaction. The final model produced by Iacobucci and Ostrom is shown as Figure 2.2.

Gronroos proposed another view (Gronroos, 1988). He does not seek to distinguish antecedents, but suggests that the difference relates to the time-frame over which the customer's judgement applies. Gronroos suggests that perceived service quality is generated in the process of delivering a service, the result of the customer's "*here and now*" perception of a service experience (Gronroos, 1988, p. 330). Customer satisfaction is a cumulative, long-term measurement of the extent to which these service experiences are meeting the needs and desires of customers.

Figure 2-2: *Iacobucci and Ostrom's model* (Iacobucci & Ostrom, 1995)

Unfortunately, other authors have suggested that the relationship between service quality and customer satisfaction is the complete converse of that posited by Gronroos. Service quality is defined in terms of “*the discrepancy between consumer’s perceptions of the services offered by a particular firm and their expectations about firms offering such services*” (Parasuraman et al., 1988, p. 14). According to the authors, this is formed *cumulatively* across multiple encounters. By contrast, satisfaction relates to one specific transaction (Oliver, 1981; Parasuraman et al., 1988).

One thing that is immediately obvious is that although all these findings are based on convincing studies, there is almost no basis for comparison between them. “Price” and “desires congruity” are different kinds of constructs. It may be that this reflects the different approaches, with Spreng and Mackoy’s model being grounded in theory and literature, while Iacobucci and Ostrom’s model was built from qualitative customer data. Constructs like “desires congruity” and “expectations disconfirmation” can be measured by researchers but are very unlikely to be articulated explicitly by customers. Neither of these models can easily be compared with Gronroos’ findings, as he makes a distinction based on temporal, rather than antecedent factors.

Iacobucci and Ostrom hypothesised that if it is possible for customers to receive what they perceived as quality service and yet be dissatisfied, this would confirm the separate existence of both constructs and help to illuminate the differences (Iacobucci & Ostrom, 1995). In our view, these studies and the literature they draw on have found convincing evidence that it is possible for customers to distinguish the two concepts, so that they can be “satisfied” but perceive they received poor service quality, or vice versa. This provides justification for the assertion that they are not one and the same construct. However, there is no clear consensus from extant literature about the nature of the differences, and there is also clear evidence for some degree of overlap. This study will not attempt a definitive resolution to this discussion. The approach adopted by this author will be to consider service quality literature primarily, because historically this has been the most cited by studies of e-service quality. However studies in customer satisfaction that use similar constructs may also be considered where relevant.

2.2.3 Service quality and service encounters

Some authors have pointed out that while customer perception of service quality often described as a single measure, it is actually the accumulated sum of individual service encounters (Bitner, Brown, & Meuter, 2000). Bitner et al. note that *“the encounter frequently is the service, from the customer's point of view. ...Service encounters”* she contends *“are moments of truth where promises to customers are kept or broken”* (p. 140).

Gronroos also supports the difference between individual cumulative experiences, finding a difference between the actual experience of receiving the service, which he terms “process consumption” and the longer term value the customer derives from the service, which he terms “outcome consumption” (Gronroos, 1988). Of course a single process consumption encounter may produce outcomes, but outcomes are also cumulative over a range of process encounters.

A related concept is explored by Lemon et al. (Lemon, Rust, & Zeithaml, 2001). This paper develops a taxonomy of types of customer equity that can be developed by an organisation. These are first, value equity, defined as customer assessment of quality, price and convenience; second, brand equity, defined as brand awareness and attitude

to the brand; and finally, relationship equity, defined as stickiness, tendency to stick with the brand. This paper has a wider focus than service quality, and was not written specifically with the dimensions of service quality in mind, but it paints a broadly similar picture. It could be argued that value equity results from the process and outcome consumption from a single encounter, while relationship equity is more cumulative.

We suggest that customers that continue to use a service over time are likely to have positive perceptions. We suggest that in an e-self-service environment, where the range of services available remains fairly constant, there will be more variation in user perceptions of individual use encounters for newer users of the service than for experienced users. New users might find a web site initially difficult to navigate, until they get to know it, and may get frustrated. Some may not persist. Alternatively, newer users may be excited by the possibilities and take some time to encounter the constraints and limitations.

Even in face-to-face service encounters, continuing users frequently have more positive perceptions of service quality. In a longitudinal study, it was found that customer's expectations of future service are strongly influenced by what they have previously received. Further, continuing to expect a service level that experience shows they are unlikely to receive causes cognitive dissonance, so customers tend over time to moderate their expectations to bring them closer to what they expect to receive (Boulding, Ajay, Staelin, & Zeithaml, 1993).

As an e-service evolves, the variation in customer perceptions between individual service encounters will decrease. Applying Bitner et al.'s distinction between overall service and service encounter (Bitner et al., 2000), we suggest that the individual service encounters will become more homogeneous as patterns of use emerge, to the point where there is relatively little gap between customer perceptions of individual service encounters and overall perceptions of e-service quality.

2.3 Service Quality and online Service Quality Literature

2.3.1 The service quality “whakapapa” or family tree

An outline of the structure of this section, and the key tables, is included in Figure 2.3. “Whakapapa” is a New Zealand Maori word encompassing ideas of family tree, inheritance and pedigree. An online service is an information system as well as a service. Both literature focusing on the process of service delivery (based on research in face-to-face services in marketing and consumer behaviour), and literature focussing on the characteristics of technologies that influence perceptions of quality have been included in this review. In fact, many of the studies we consider devote considerable effort to establishing equivalences between aspects of face-to-face service quality and aspects of technology-mediated service quality. A discussion of face-to-face service quality literature is included because of its enormous influence on subsequent research in the service quality area.

The general approach to each of the literatures we discuss is as follows. First, the leading model or models in the area are introduced. The leading model(s) are used to frame a more general discussion, and additional constructs relevant to this, and their supporting literature are identified. A table is presented, using the leading model(s) as a frame. New constructs that are supported by the literature but are not included in the leading model(s) are italicised. If no leading models exist, then a general review of relevant literature is offered. The constructs from each area that are carried forward into the integrated set of constructs and items are summarised in a table at the end of each section, as outlined in Figure 2.3.

Figure 2-3: *The general approach to each of the construct literatures based on table summaries*

This approach is important for two reasons. First, there are so many competing studies already in existence proposing various dimensions and indicators, that it is almost impossible to examine and integrate them all; many are overlapping or incommensurate. The second is that as we intend to study online service quality quantitatively, we aim to re-use, as much as possible, previously tested constructs and indicators.

2.3.2 Face-to-face Service quality literature

Many studies of information systems, web site, and e-commerce quality have their origins in service quality literature. In this discussion, two leading models have been selected to frame the discussion. The first is the ServQual model by (Parasuraman et al., 1985); the second a model which integrates ServQual with other conceptualisations of service quality.

2.3.2.1 *ServQual*

ServQual, a highly influential instrument for measuring service quality was based on the expectations/perceptions gap model of service quality (Parasuraman et al., 1985). The ServQual instrument uses customer measures of service quality expectations, and perceived service quality, and expresses the gap as a difference score. This research began with ten initial determinants of service quality. These determinants were:

Reliability (consistency of performance and dependability); Responsiveness (willingness and timeliness of assistance by employees); the Competence of the organisations employees; the ease with which customers can contact and Access the organisation; the Courtesy demonstrated by the company's employees; the level and appropriateness of Communication about the service; the Credibility of the organisation, as demonstrated by their reputation and the perceived trustworthiness of their contact personnel; the Security of the customer expressed as freedom from physical and financial risk, and appropriate exercise of privacy and confidentiality; the Understanding of the customer's needs and the physical aspects or Tangibles of the organisation, its premises, employees, and other physical evidence of service.

These ten determinants were used as the basis for developing the ServQual instrument. Validation of the SERQUAL instrument resulted in the original ten determinants being consolidated into five dimensions of service quality (Parasuraman, Berry, & Zeithaml, 1991a). These dimensions are shown in Figure 2.4: reliability (the ability to perform the promised service dependably and accurately); tangibles (the appearance of the physical facilities, equipment, personnel, and communication materials); responsiveness (the willingness to help customers and provide prompt service); assurance (the knowledge and courtesy of employees and their ability to convey trust and confidence); and empathy (the caring, individualised attention provided to the customer).

Figure 2-4: *ServQual dimensions*

2.3.2.2 Literature supporting ServQual

Some subsequent literature on service quality has reinforced these five dimensions. Takeuchi and Quelch identify efficiency and attitude as important aspects of customer perceived service quality (Takeuchi & Quelch, 1988). They do not provide

comprehensive definitions, but their concept of efficiency is captured in the reliability and responsiveness dimensions of ServQual, and their concept of attitude is captured in the empathy dimension. Berry discusses the factors required to build a strong service brand (Berry, 2000). This discussion has a branding focus, and many of the factors identified relate to building awareness rather than to evaluating the quality of service received. However Berry's discussion supports importance of quality experiences with the company, noting that "*a presented brand cannot rescue a weak service*" (p. 130). This advocacy for the importance of making an emotional connection with customers is captured in the responsiveness and empathy dimensions of ServQual.

Gronroos (1988), made a distinction between the "totally invisible" (to the customer) aspects of services delivery, which include organisational factors such as management support and systems support, and the visible "interactive" aspects that influence customer perception. ServQual includes the visible aspects of physical resources and equipment as the "tangibles" dimension; however "contact persons" does not have a 1:1 match with the ServQual dimensions but is encompassed within the responsiveness, assurance and empathy dimensions.

2.3.2.3 *Criticisms of ServQual*

Despite the enormous influence of the ServQual model, there have been a number of well-founded criticisms. These have fallen into several areas. An excellent summary of the issues identified with ServQual was offered by Stafford et al. (Stafford, Prybutok, Wells, & Kappelman, 1999). This section provides a brief overview of the issues; for a more complete discussion Stafford et al.'s paper is recommended.

The first criticism relates to unstable dimensionality of the five dimensions. Cronin and Taylor found that ServQual was effectively one-dimensional, and suggested it should only be used as a summed index, because all the items loaded predictably on a single factor (Cronin & Taylor, 1994). Other studies found support for seven to eight dimensions (Carman, 1990). In particular, the tangibles dimension proved unstable, even in subsequent studies by the authors of ServQual (Parasuraman, Zeithaml, & Berry, 1991). The tangibles dimension was modified or dropped by some subsequent researchers (see, for example Pitt, Watson, & Kavan, 1995).

Following this debate, Stafford et al. carried out an extensive study to assess stability and goodness of fit for several alternative measures of service quality and found that five dimensions yielded a better fit to their data than a one-dimensional model (Stafford et al., 1999). We return to the discussion on the lack of consensus about the dimensionality of ServQual in the next chapter, when we examine the nature of latent constructs and their indicators.

The second area of concern frequently expressed about SERQUAL relates to the use of the gap concept and the difference score as a basis for measurement. This aspect of ServQual has been one of the most widely criticised, (e.g. Carman, 1990; Teas, 1994). A high profile debate in the Journal of Marketing was conducted between Cronin and Taylor and the ServQual authors. Cronin and Taylor proposed a perceived performance only instrument, ServPerf (Cronin & Taylor, 1994). They acknowledge the ubiquity of the disconfirmation of expectations paradigm, but argue that the degree of disconfirmation is captured in an overall assessment of satisfaction anyway, *“the perceived summary disconfirmation judgement is sufficient as a causal agent for satisfaction”* (Cronin & Taylor, 1994, p. 126).

Van Dyke et al. conclude that measuring customer perception only has been shown to be more reliable (Van Dyke, Kappelman, & Prybutok, 1997). Subsequent studies (Barnes, 2001; Barnes & Vidgen, 2002) have dropped the explicit "expectations" concept and concentrated on measuring customer perceived service according to the ServQual dimensions, but modified for the web. Stafford et al. offer a detailed comparison of the fit and stability of various measures of service quality (Stafford et al., 1999). They found that

“...despite the considerable discourse on the subject and the current findings, the issue of whether the perceptions-only approach or the difference score approach is superior is still not resolved to the satisfaction of these researchers.” (Stafford et al., 1999, p. 27)

Lacking a clear consensus from the literature, we opted for a perception-only measurement, because this approach has been adopted, as we will discuss shortly, for the best existing instrument we were able to identify for measuring e-service quality.

2.3.2.4 Alternative conceptualisations

One of the most carefully argued recent alternative conceptualisations of service quality is offered by (Brady & Cronin, 2001). These authors present a review and synthesis of the most influential conceptualisations of service quality, including ServQual. We review this discussion in detail, because it is the most comprehensive and integrative study we were able to identify, and the most credible alternative to ServQual.

Brady and Cronin also acknowledge that conceptualisations of service quality originated with the “gap” or disconfirmation of expectations model, and trace the development of this model from physical goods literature published primarily in the 1960’s and 1970’s. Brady and Cronin also acknowledge the early influence of the ServQual model that we have already discussed.

Brady and Cronin acknowledge that there has been little consensus about the dimensionality of the Service Quality construct, with many studies both supporting and refuting variations of the five factor model proposed in ServQual. Another influential conceptualisation by Gronroos (1984, as cited in Brady and Cronin, 2001), also uses the disconfirmation paradigm, but identifies only two service quality dimensions: technical quality and functional quality.

In this conceptualisation, functional quality refers to the customer’s perception of the interactions that take place while the service is being delivered. Technical quality refers to the outcome of receiving the service – the customer’s perception of the value of what they have received. Gronroos’ two-dimensional conceptualisation has also proved enduring. Rust and Oliver added a third dimension to the two proposed in Gronroos’ model, which is the service environment – the surrounding environment in which the service is delivered (Rust & Oliver, 1994). The Rust and Oliver model does not explicitly use a “gap” paradigm.

Brady and Cronin have reviewed variations on the disconfirmation of expectations theme, and several alternative, but equally well-supported conceptualisations of the dimensions of service quality. We noted earlier in the discussion on customer satisfaction and service quality, that this whole area is characterised by well-supported but incommensurable models that provide limited basis for comparison and integration.

Brady and Cronin make a major contribution to unravelling this tangle. First, they review work by Dabholkar that suggests that one of the reasons for the wide variations in extant models is that service quality is in fact multi-layered - a higher order factor that is supported by two additional levels of attributes (Dabholkar, 1995). This approach allows a more complete explanation of a very complex phenomenon. Brady and Cronin note that the multi-layer approach has theoretical support but has had limited application in empirical studies.

Brady and Cronin use these conceptualisations, complemented by a literature review that outlines studies that support various specific dimensions, to develop and test an integrated model. In overview, they conclude from a review of extant models that

*“When assessed collectively, the ServQual model appears to be distinct from the others because it uses terms that describe one or more determinants of a quality service encounter. In other words, the five dimensions of ServQual are terms that might be used to refine some aspect of service quality. However of major concern should be the question as to **what** should be reliable, responsive, empathetic, assured, and tangible....We suggest that identifying this “**something**” is critical in the literature. Specifically, a conceptualisation that recognises the significance of the ServQual factors and defines **what** needs to be reliable and so forth will respond to the call” (Brady & Cronin, 2001, p. 36 my emphasis)*

Brady and Cronin then set out to address this challenge, by developing a multi-layered model of service quality, with ServQual dimensions serving as descriptors of the sub-dimensions. For a complete discussion of the development of their model and their hypotheses, please refer to the original paper by Brady and Cronin. The resultant model is shown as Figure 2.5.

Figure 2-5: *Brady and Cronin's integrated model* (Brady & Cronin, 2001)

The dimensions identified by Brady and Cronin are defined as follows. Interaction Quality is the interpersonal interactions that take place during service delivery. This is made up of the sub-dimensions of the attitude behaviour and expertise of staff. Physical environment quality is very similar to the ServQual “tangibles” dimension, and comprises the aspects of the physical or “built” environment (Brady & Cronin, 2001, p. 38). This dimension is sub-divided into ambient conditions, which encompasses non-visual aspects such as temperature, scent and music; facility design, which refers to aspects of the architecture and layout; and social conditions, which can include factors such as “*the number and type of people evident in a service setting as well as their behaviours*” and other “*signs, symbols and artefacts*” Brady and Cronin, 2001, p. 39). Outcome quality has been described as “*what the customer is left with after the production process is finished*” (Grönroos, 1984, quoted in Brady and Cronin, 2001, p.40). The sub-dimensions proved more problematic to identify from literature. Brady and Cronin propose sub-dimensions of waiting time, tangible evidence of the service outcome, and valence, defined as factors that affect the customer’s overall attitude as to whether the service experience was good or bad. Each construct is supported by an extensive literature review. For a summary of the supporting literature, the original paper by Brady and Cronin may be referenced. The model was built initially from a combination of literature support and extensive qualitative data gathering in four service industries. This was then followed by quantitative testing based on survey research. The combination of extensive empirical

research and theoretical grounding based on extant, well-tested models makes this research particularly persuasive.

Although the overall model has many strong points, we are unhappy with the semantic confusion that has occurred within the “tangibles” sub-dimension. The definition proposed by Brady and Cronin describes tangibles in terms of “*tangible evidence of the service **outcome***” (*our emphasis*), and it is offered as a sub-dimension of a dimension that explicitly refers to “*what the customer is left with*” (Gronroos, 1988). There is some semantic confusion here, because although Parasuraman et al. use the same term, “tangibles” (Parasuraman et al., 1985), and are cited by Brady and Cronin in the supporting literature for their tangibles sub-dimension, the term does not have the same definition in both contexts. Parasuraman et al.’s definition of “tangibles” was reviewed earlier, as “*the appearance of the physical facilities, equipment, personnel, and communication materials.*” (Parasuraman et al., 1985) As we noted earlier, this much more closely aligns with Brady and Cronin’s definition of their physical environment quality construct.

We also find compelling the suggestion that the so-called “Nordic” (based on Gronroos model) and “American” (based on ServQual) strands of research can and should be integrated. However we are not entirely convinced that this has been achieved. Brady and Cronin found good support for their first order and second order dimensions. It was less clear the extent to which their argument that the ServQual dimensions can act as descriptors to the third order dimensions is supported by the research. The suggestion that ambient conditions and design (as defined by Brady and Cronin) can demonstrate or convey reliability, responsiveness and empathy (as defined by Parasuraman et al.) remains a bit of a stretch, both at a semantic level, and based on the findings of Brady and Cronin’s study.

The third order constructs for interaction quality and physical environment quality are somewhat problematic. Brady and Cronin admit that the outcome quality construct was the most problematic, and their discussion is weakened by the semantic confusion over the tangibles construct.

Brady and Cronin have posited a many to many relationship between the ServQual dimensions and their own – each of three ServQual dimensions is related to each of Brady and Cronin’s third order constructs. Since they appear to be alternative conceptualisations at the same level, we have opted to use the ServQual dimensions because they are better known and most widely cited.

2.3.2.5 *Summary*

In a comprehensive position paper on service quality, Zeithaml claimed that we know the dimensions of perceived service quality (Zeithaml, 2000). Unfortunately, despite numerous authoritative studies on the subject spanning several decades, this seems a premature claim.

This discussion has established the following:

- There is considerable support for a multi-dimensional view of the service quality constructs
- There appears to be some support for a many to many relationship between Zeithaml’s five dimensions and the two or three second order dimensions of the Nordic model
- The discussion on the dimensionality of ServQual is not conclusive
- Despite some well researched criticisms and some compelling alternatives, ServQual has been the most dominant and most cited model, and the most widely applied in other industries.

We propose the following integrated set of Service Quality dimensions (Table 2.1). We realise that this conceptualisation has not been subject to empirical research and that even if such research had been conducted; it would be unlikely to represent the final word in this debate. This discussion seeks only to provide an integrated framework for service quality that can be used as a basis for discussion.

Table 2.1: *An integrated set of face-to-face Service Quality dimensions*

Integrated View	Service Quality Dimension	Definition	Source Concepts
Outcome	Features	Objective features of the service under the control of the provider	“Backstage” (Iacobucci & Ostrom, 1995) “Outcomes” (Gronroos, 1988) “Value equity” (Lemon et al., 2001) “Technical quality” (Grönroos, 1984), quoted in (Brady & Cronin, 2001) “Outcome quality” (Brady & Cronin, 2001) “Objective physical aspects [of the product or service]” (Lemon et al., 2001)
	Reliability	The ability to perform the promised service dependably and accurately	“Technical quality” (Grönroos, 1984), quoted in (Brady & Cronin, 2001) “Reliability” (Parasuraman et al., 1988) “Reliability” (Cronin & Taylor, 1994) “Efficiency” (Takeuchi & Quelch, 1988)
Tangibles	Tangibles	The appearance of the physical facilities, equipment, personnel, and communication materials	“Physical resources and equipment” (Grönroos, 1984), quoted in (Brady & Cronin, 2001) “Visible aspects” (Gronroos, 1988) “Tangibles” (Parasuraman et al., 1991) “Physical environment quality” (Brady & Cronin, 2001) “Service environment” (Rust & Oliver, 1994)
Interaction Quality	Responsiveness	The willingness to help customers and provide prompt service	“Contact persons” (Gronroos, 1988) “Efficiency” (Takeuchi & Quelch, 1988) “Emotional connection” (Berry, 2000) “Interactive aspects” (Gronroos, 1988) “Functional quality” (Grönroos, 1984), quoted in (Brady & Cronin, 2001) “Interaction quality” (Brady & Cronin, 2001)
	Assurance	The knowledge and courtesy of employees and their ability to convey trust and confidence	“Contact persons” (Gronroos, 1988) “Efficiency” (Takeuchi & Quelch, 1988) “Emotional connection” (Berry, 2000) “Interactive aspects” (Gronroos, 1988) “Functional quality” (Grönroos, 1984), quoted in (Brady & Cronin, 2001) “Interaction quality” (Brady & Cronin, 2001)
	Empathy	The caring, individualised attention provided to the customer	“Contact persons” (Gronroos, 1988) “Attitude” (Takeuchi & Quelch, 1988) “Emotional connection” (Berry, 2000) “Interactive aspects” (Gronroos, 1988) “Functional quality” (Grönroos, 1984), quoted in (Brady & Cronin, 2001) “Interaction quality (Brady & Cronin, 2001)
“Nordic” dimensions shown in light grey “ServQual” dimensions shown in dark grey			

2.3.3 Service quality in information systems

In the early 1990's IS was increasingly acknowledging the need to incorporate customer expectations and perceptions into its engineering and organisational behaviour lines of inquiry (Kettinger & Lee, 1995; Pitt et al., 1995).

At the same time, customers were acquiring unprecedented access to information on which to form expectations and base their service perceptions, compared with similar customers of even a decade prior (Davenport, Harris, & Kohli, 2001; Jon, 2000). The article; *Service Quality: A Measure of Information Systems Effectiveness* published in *MIS Quarterly/June 1995* is a widely-cited article that heralded the crossover of the ServQual paradigm into the IS domain (Pitt et al., 1995). A summary of Pitt et al.'s operationalisation of the ServQual dimensions for an IS department is provided in Table 2.2. (Pitt et al., 1995).

Table 2.2: *Online service quality whakapapa: Service quality dimensions for an IS department*

Service Dimension	Definition
Reliability	IS departments keep promises, are sincerely interested in solving problems, are dependable, timely, and error free
Tangibles	IS departments have up to date hardware and software, appealing physical facilities, well-dressed personnel.
Responsiveness	IS department are prompt, willing, helpful, are not too busy to respond to requests, and provide accurate information about when services will be performed.
Assurance	IS department employees are courteous and knowledgeable. Their behaviour inspires confidence. Users feel safe in their transactions with IS department employees.
Empathy	IS department employees give individual and personal attention, understand the specific needs of their users, and have the user's best interests at heart? The IS department has convenient operating hours.

IS researchers were quick to adopt the ServQual paradigm as presented by Pitt et al. However, it was not long (almost immediately allowing for the *MIS Quarterly* publishing cycle) before other IS researchers found and took up the concerns previously voiced by Carman, Teas, Cronin and others in the marketing literature. The June issue of *MIS Quarterly* in 1997 was dominated by challenges to, and defences of IS-ServQual (Kettinger & Lee, 1997; Van Dyke et al., 1997). Van Dyke et al began

the debate by calling into question Pitt et al.'s choice of the ServQual paradigm and by drawing attention to the collective concerns of Carman, Cronin and Taylor, and Teas's about the predictive value of ServQual (Van Dyke et al., 1997). Their debate mainly focused on restating Teas's concerns about the instrument development and instantiation, challenging the E-P construct suitability, and describing in detail their concerns about the finer points of the use of statistical techniques. They concluded with a discussion about their opinions on the dimensional instability of ServQual, but stopped short of offering a practical alternative (Van Dyke et al., 1997).

Pitt, Watson and Kavan had the right of reply in the same issue of *MIS Quarterly*. Interestingly, as much as Van Dyke had restated the marketing literatures concerns that had been voiced in the marketing context, Pitt et al. employed the same strategy when it came to the response. They addressed Van Dyke's points by citing the matching PZB responses that had already been given in the marketing literature to the same issues. In fact, they said as much:

"The ServQual debate has certainly not been one sided. Parasuraman et al. have responded in full, with sound and solid arguments regarding the conceptual and empirical aspects of their conceptualization of service quality. The IS community deserves to hear these counter arguments."
(Pitt, Watson, & Kavan, 1997, p. 210)

The impact of this highly influential debate in MISQ is that the SERQUAL stream of research is the most influential services marketing literature cited by information systems and electronic commerce researchers. This has tended to result in a strong sense of consensus within the information systems field about the nature and dimensions of service quality. This might not be the case if the wider discourse on service quality within marketing literature, and some of the alternative conceptualisations of service quality, were more widely cited by information systems researchers (Sylvester & Tate, 2008).

In summary, this literature does not substantially advance discussion of online service quality, as Pitt et al.'s basic dimensions follow ServQual, but are operationalised for IS. However, the debate demonstrates that many of the issues identified in wider service quality literature were equally intractable in an IS context.

2.3.4 User satisfaction with end-user computing

Many influential studies in user satisfaction with Information systems pre-date the widespread availability and use of the Internet. The explosion of end user computing in the 1980s resulted in a huge increase in the number of non-IT professionals using computers, some increase in user choice as to which applications they used, and some shifts in power and control over IT resources from central IT departments to line business units and individuals; in summary, a significant change in the way IT services were delivered and used. One important dimension of this change was a new emphasis on measuring user satisfaction with end-user computing. A seminal paper by Doll et al. identified five dimensions of user satisfaction with end-user computing (Doll, Xia, & Torkzadeh, 1994). These were content, accuracy, format, ease of use, and timeliness. A summary is provided in Table 2.3.

Table 2.3: *Online service quality whakapapa: Dimensions of end-user computing satisfaction*

EUC satisfaction dimension	Definition
Content	Specificity, sufficiency, usefulness, and fit to purpose of the system
Accuracy	User perception of, and satisfaction with, system accuracy
Format	Clarity and usefulness of the way the output was presented
Ease of Use	User perception of system ease of use and friendliness
Timeliness	The currency of the information provided by the system and the timeliness with which the user was able to access it

In more detail, Content covered the specificity, sufficiency, usefulness, and fit to purpose of the system. Accuracy covered user perception of and satisfaction with, system accuracy. Format covered the clarity and usefulness of the way the output was presented. Ease of use covered user perception of system ease of use and friendliness, and Timeliness evaluated the currency of the information provided by the system and the timeliness with which the user was able to access it. These dimensions proved stable through subsequent confirmatory studies.

In summary, studies on end user computing satisfaction represent early attempts to assign quality and satisfaction measure to services delivered by technology. Each of the five dimensions identified by Doll et al. have potential relevance, and proved to have considerable commonality with later studies into the quality dimensions of e-services, as we examine later in this chapter.

2.3.5 Quality in web-sites

When a customer purchases or consumes a product, the production of the product is often transparent to the customer (Gronroos, 1988). The customer consumes and evaluates the outcome of the product production process. By contrast, with services, the customer consumes the delivery, as well as receiving outcomes. In an online service context, the delivery of the service requires the customer to interact with a (possibly complex) software product. This means that quality characteristics of web sites are relevant to this discussion.

Following the explosion of business and personal use of the Internet, and the massive increases in the number of subscribers and the value of transactions, there has been very considerable research interest in quality in web sites. To the extent that online services are delivered via web sites, it may be assumed that quality factors that apply to web sites in general are also relevant to online service quality

There is a plethora of e-commerce quality literature identifying a very large number of candidate dimensions. There is a lack of any existing theoretical framework for integration of these disparate studies; and ambiguity and overlap in the definitions. The following review and synthesis of the main themes from web quality literature aims to arrive at some working definitions of core dimensions identified in web

quality literature that can be compared with the candidate dimensions identified from other disciplines.¹

Web design features that influence users perceptions of quality were integrated from studies in four main areas. First, studies that identified web features associated with user satisfaction; second, instruments measuring web quality from the users' perspective; third, studies examining the effectiveness of web-sites; and finally, research into the characteristics of user-centred interfaces.

In studies of web design features associated with user satisfaction, content presentation, and content quality, is repeatedly mentioned as an essential dimension (Eighmey, 1997; Gehrke & Turban, 1999; Kim, 1999; Loiacono & Taylor, 1999; Rice, 1997; Scharl & Bauer, 1999; Stafford & Stafford, 2001). In addition to content, visual attractiveness is also strongly emphasized (Kim, 1999; Rice, 1997; Stafford & Stafford, 2001). To these, additional web design features of navigation, enjoyment and playfulness were added (Eighmey, 1997; Liu, Arnett, Capella, & Taylor, 2001).

Instruments developed to assess web quality perceived by users also indicate what is important for users. These studies add some features to those already identified, such as reliability (Barnes & Vidgen, 2002), accessibility (Barnes & Vidgen, 2002), privacy (Aladwani & Palvia, 2002), and interactivity (Aladwani & Palvia, 2002). The problem mentioned already, that it is difficult to make one to one comparisons between the terms used by different authors, means that some subjective assessment of the equivalence of constructs has been required.

Several dimensions were derived from an extensive stream of work on development of the eQual instrument (Barnes, 2001; Barnes & Vidgen, 2002). This instrument has three dimensions - site quality, information quality and interaction quality - that are divided into 10 subcategories: aesthetics (included as visual attractiveness), navigation (already included), reliability (added), responsiveness or timeliness (added as interactivity dimension), security (included in my privacy and security dimension),

¹ Some parts of this section have been adapted from a previously published paper by (Baierova, Tate, & Hope, 2003).

access or ease of contact (added), communication (aspects of this are already included in my content presentation dimension).

Three additional eQual dimensions relate more to the customer's experience of the service quality they have received than to specific features of the web-site; understanding of individual (empathy to provide right content, product or service), competence (knowledge), and credibility (trustworthiness). These will be discussed in more detail when eQual is revisited in the context of e-services literature later in this chapter.

Aladwani and Palvia also developed an instrument for measuring user perceived web quality (Aladwani & Palvia, 2002). The instrument measures web quality in four dimensions: content quality (information accuracy, usefulness, clarity, currency, uniqueness, and originality), specific content (privacy policies, customer support, specific details about product/services), appearance (attractiveness, organisation, proper use of colours, fonts, graphics, language, and graphics-text balance) and technical adequacy (navigation, links, reliability, customisation, speed, interactivity, speed, accessibility). The division of categories and sub-categories varies slightly from previous studies, but overall, it does not add any significant new dimensions to the study.

Research on effectiveness of web sites provides additional features for assistance and originality and uniqueness (this concept is also touched on by some previous researchers). According to Bell and Tang, effective web sites comprise of the eight following features: accessibility, content, visual design, and structure (already included), and uniqueness (added) (Bell & Tang, 1998). Additional dimensions identified include usefulness and friendliness. These dimensions, similar to some of the eQual dimensions, relate more closely to the customer's perception of the service they have received, and will be discussed in more detail in a later section of this chapter.

The business domain of the web-site appears to have an influence on service quality dimensions, and some design features identified by previous studies in user preferences for e-commerce sites are e-commerce specific. This section concentrates

on identifying domain independent dimensions that are common across a range of studies. Other studies support the overall classification scheme, but include additional e-commerce dimensions (Srivihok, Ho, & Burstein, 2000; Tilson, Dong, Martin, & Kieke, 1998; Wang, Tang, & Tang, 2001).

Criteria for user-centred Web design were developed by Abels, White, and Hahn (Abels, White, & Hahn, 1997) and Neilson (Nielsen, 1999). Abels et al. identified the dimensions which include use, which covers ease of use and navigation and is included in the navigation and structure category, unique, relevant and credible content (included in the content quality category), visible and well organised structure, linkage, and search capability (included in navigation and structure category), and appearance (visual attractiveness) as the most important features. Neilson also identifies similar features, and adds the dimension of credibility, which will be discussed in a later section of this chapter.

In summary, the dimensions identified from web quality literature were accessibility, assistance, content presentation, content quality, enjoyability, interactivity, navigation and structure, security and privacy, uniqueness and originality, and visual attractiveness. Definitions of each of these dimensions are provided in Table 2.4. This table is based on the definitions included in the survey instrument used by Baierova et al. (Baierova et al., 2003). These dimensions are potentially of relevance to the study of e-service quality.

Table 2.4: *Online service quality whakapapa: dimensions of website quality*

Web Quality Dimension	Definition
Accessibility	Web pages download quickly and completely; Live and relevant links (no dead-end links); Fast and easy to access all parts of a web site
Assistance	Visible instructions; Help is easily available on every web page; Assistance when having trouble finding or using data; Easy recovery from mistakes or errors
Content presentation	Content is communicated in the most appropriate format (effective combination of multimedia - text, pictures, sounds, animations); Information is easy to understand, (e.g. - presentation of information includes also pictures or graphs, the web site about music includes music samples, video); Visible coverage and organisation of content
Content quality	Relevancy, accuracy and credibility of information; Wide coverage and depth of information; Frequent updates of content
Enjoyability	Using a web site is entertaining, exciting, enjoyable; A web site is playful (through interactive features, use of graphics, animations, menus...)
Interactivity	Ability to interact via the web site with the organisation or other users; The web site includes some of the interactive features Chat rooms, Discussion forums or/and Bulletin boards; Virtual community, On-line feedback form, or FAQ; Fast response to e-mail inquiries
Navigation and structure	It is easy to find information on the web site; Clear structure and consistent layout of web pages; Includes search capability and a site map; Clear organisation (hierarchy) of web pages into a web site; Relevant linkage between web pages; Easy navigation through navigation menus and bars; Consistency of navigation menus, bars and buttons
Security and privacy	Encryption of user-entered data (log-on information, registration); Security of data transfers, credit card numbers; Security and Privacy policies
Uniqueness and originality	A web site is innovative, unique; Originality of provided content
Visual attractiveness	The web site as a whole is visually appealing and pleasant to look at (good use of colours, fonts, graphics, animations)

2.3.6 E-commerce and e-service quality literature

2.3.6.1 *Overview*

Previous studies in e-service have come from two main sources, first from researchers in e-commerce investigating how Internet technologies can be usefully applied to improvements in service quality, and secondly, from studies with a marketing and consumer research focus, investigating the impacts of the injection of new technology into customer perceived service quality and service encounters. In this section, we note that while e-commerce research into e-service quality has drawn extensively on marketing literature, there has been relatively little consideration of relevant e-commerce literature by researchers in Marketing. A synthesis of some of the most influential research from both sources is offered.

2.3.6.2 *eQual*

One of the most extensive streams of study into e-services has been the work by Barnes and Vidgen (Barnes, 2001; Barnes & Vidgen, 2001, 2002) on successive versions of the eQual instrument. A brief history of this work, as well as the latest findings, are included here because the combination of qualitative grounding, extensive quantitative analysis, extensive testing across multiple industries and web-sites, and support from literature makes the eQual instrument a potentially promising candidate for constructs to measure end-user perceptions of online service quality. Further, as well as providing a tested instrument for assessing customer perceptions, scores from a sample of users using eQual to rate a site can be consolidated to produce the eQual Index; this is a quantitative overall assessment of the service quality of a web-site, which can be used for comparative purposes.

The first version (1.0) of eQual was developed using qualitative data from customer focus groups, which was consolidated into a pilot questionnaire with thirty-five questions. This was subsequently reduced to twenty-four questions, based on a pilot study and extensive validation of the constructs.

eQual 1.0 concentrated on information quality. This may be attributed partly to the domain in which it was originally tested - business school web sites - which are likely to be information rich rather than transaction rich. Another possibility is that in 1998,

when the initial quantitative data was gathered, a far higher proportion of web sites were still at an early stage of development that emphasised information provisioning; fully interactive e-commerce web sites were less widespread, and therefore user expectations of interaction quality were less sophisticated.

Version 2.0 of eQual was developed based on a study of customer satisfaction with the service provided by on-line bookstores. Customer focus groups indicated that insufficient attention had been paid in eQual 1.0 to interaction quality. To address this gap, service quality literature (primarily work on ServQual, discussed above in the service quality section) and information systems service quality literature, (primarily work in IS ServQual, discussed above in the IS services section) were reviewed to develop additional constructs for interaction quality.

The next version of eQual (3.0) tested the eQual instrument in the on-line auction domain, and grouped the constructs into three high-level dimensions of e-commerce web-site quality - site quality, information quality and service interaction quality.

The current version of eQual (4.0) is based on the analysis of the results of eQual 3.0, complemented by an extensive discussion of the literature that supports each dimension. The main change is an increased emphasis on usability rather than site quality. Site quality still encompasses technical features under the control of the designer, for example use of frames. These features are (or should be) transparent to the user, except to the extent that they enhance or detract from their experience of using the site. The focus on measuring the site as experienced by the user, rather than empirical characteristics of the site is captured better in eQual 4.0. The authors also offer an extensive discussion of the primary and secondary literature sources for the questions in eQual 4.0.

The final high-level categories identified by eQual 4.0 are usability, which encompasses aspects of the user's impression of using and interacting with the site, such as navigation, appearance, and general ease of use; information quality, which encompasses the accuracy, timeliness, relevance, granularity, and general "believability" of the information, and service interaction quality, which encompasses

more general service quality constructs such as security, trust, personalisation and access to the organisation.

2.3.6.3 Other e-commerce studies

Although eQual has been the most comprehensive stream of research into e-service quality, several other studies have been identified that offer additional literature support for eQual constructs, or that offer additional constructs.

A comprehensive model of e-customer satisfaction included in a text by Turban et al. (Turban, Lee, Warkentin, & Chung, 2002, p. 127) is shown as Figure 2.6.

Figure 2-6: Model of e-customer satisfaction

This model offers an interesting comparison with eQual. The transaction safety constructs are more granular than eQual 4.0 - including aspects such as integrity, non-repudiation and authentication. The content quality sub-constructs show a better fit between the two models, several, for example format and timeliness, appear to be directly equivalent. We note also that this model was developed for e-commerce, rather than online service quality. This means that constructs such as logistics are less relevant to our research.

There is also a good fit, at the sub-construct level between the eQual sub-constructs of Interaction Quality and Turban et al.'s model, although in many cases the sub-constructs are grouped into different high-level constructs. This may simply reflect the more extensive empirical research that has been carried out for eQual. In summary, the Turban et al. model does not significantly extend eQual in either overall scope, or at the sub-construct level.

Also worthy of mention in this discussion is a preliminary study of e-service adoption factors by (Sandhu & Corbitt, 2003). Their study concentrates on user attitudes and behaviours in the adoption of e-services, but also offers some insights into user-perceived quality factors. In particular, Sandu and Corbitt found both literature and empirical support for the importance of a "sense of control", "user support" and "self-service" to users of web sites. "User support" is similar to the "customer service" construct in Turban et al.'s model, and the "access to the organisation" construct in eQual. However "self-service" and "sense of control" are new constructs in e-commerce e-services literature.

More recently (and subsequent to most of the research reported in this study), Bressoles conducted a study that compared several instruments designed to measure online service quality, including eQual and eTailQ (discussed below) with their own NetQual instrument (Bressolles, 2008). These studies did not add anything new to the understanding of the overall business domain of service quality, although the overall factor structure and the item set for each factor varied somewhat from instrument to instrument.

A comprehensive literature review of recent studies in e-service quality, and a comparison between e-commerce service quality dimensions and ServQual, was carried out by (Alzola & Robaina, 2005). This paper found the same level of confusion in e-service quality constructs that we have also observed. However, this review provided support for a broader concept of online service quality, to include the quality and range of products and services offered. Three of the eighteen studies they reviewed also included constructs related to the website's product portfolio: "product portfolio" (Yang, Jun, & Peterson, 2004); "product offers" (Syzmanski & Hise, 2000); and "digital products/services" (Wang et al., 2001).

2.3.6.4 Marketing studies

Researchers with a predominantly marketing and consumer behaviour focus have also studied the impact of introducing technology into customer service; replacing or supplementing traditional service by employees with online service. This research suggests that technology has had far-reaching effects on the service relationship that are not yet fully understood. Technology now has a potential impact in all aspects services marketing. Exploring the pervasive impact of technology in services marketing, Parasuraman and Grewal set out a comprehensive research agenda for studying the impact of technology on the quality-value-loyalty chain (Parasuraman & Grewal, 2000). This includes the question "Do the definitions and relative importance of the five [ServQual] service quality dimensions change when customers interact with technology rather than service personnel?" This suggests that within marketing literature there is not yet a consensus the drivers of online service quality.

2.3.6.5 Summary

The research on eQual (Barnes, 2001; Barnes & Vidgen, 2001, 2002) offers the most comprehensive answer we were able to identify to the question raised by Zeithaml (Zeithaml, 2000), although some minor extensions are required. A summary is provided in Table 2.5. As we discussed above, eQual draws extensively on service quality literature, but the eQual research does not seem to be well known or cited within marketing and service quality literature.

Despite our view that on balance, the e-commerce studies are more comprehensive or a better fit to the phenomena, some additional constructs of potential relevance for

this study have been identified from within marketing literature. Research by Dabholkar and Bagozzi has examined aspects of customer satisfaction with self-service technologies, which includes technologies such as ATM machines as well as the Internet (Dabholkar & Bagozzi, 2002). They found a sense of being in control, and minimal waiting time to be important to customers, as well as factors such as ease of use and a sense of fun that have already been identified by previous authors.

Table 2.5: *Online service quality whakapapa: dimensions of e-service quality*

eQual Dimension (Barnes and Vidgen, 2002)	Items
Usability	I find the site easy to learn to operate My interaction with the site is clear and understandable I find the site easy to navigate I find the site easy to use The site has an attractive appearance The design is appropriate to the type of site The site conveys a sense of competency The site creates a positive experience for me
Information Quality	The site provides accurate information The site provides believable information The site provides timely information The site provides relevant information The site provides easy to understand information The site provides Information at the right level of detail The site presents information in the appropriate format
Service Interaction Quality	The site has a good reputation It feels safe to complete transactions on this site My personal information feels secure This site creates a sense of personalisation This site makes it easy to communicate with the organisation I feel confident that goods/services I have requested will be delivered as promised
Additions (from our literature review)	Items
Service Interaction Quality	This site creates a sense of enjoyability or entertainment I feel in control on this site
Usability	Overall, the response time was acceptable to me
Transaction quality	This site offered the range of on-line transactions I expected

A more recently published study by Wolfingerbarger and Gilly reports the development of the eTailQ instrument (Wolfingerbarger & Gilly, 2003). This instrument concentrates on e-tail (online shopping) quality overall, rather than service quality specifically. This study confirmed the importance of aspects of website design, security and privacy, interaction with the organisation, and fulfilment of promises. They found the range (selection) of products available to be an important quality criterion. The equivalent in a services context would be the range and richness of services available.

In summary, eQual appears to cover the domain of e-service quality fairly thoroughly, but our literature review suggests some minor additions. We suggest the addition of items for “a sense of control” and “a sense of enjoyability” as part of service interaction quality; the addition of “response time” as part of usability; and the addition of a new factor for transaction quality, representing the increased availability and importance of online transactions.

2.3.7 Domain-specific studies

Our study aims to measure user perceptions of online service quality in a university web portal. We were able to identify domain-specific service quality studies for e-portals, the public sector, and higher education.

2.3.7.1 *e-Portal Quality*

The studies in e-Portal quality that we were able to identify suggest that perceived service quality in web portals is an extension of perceived service quality in other websites, rather than a radical new departure. Two recent studies on perceived portal quality (Gounaris & Dimitiadis, 2003; Kuo, Lu, Huang, & Wu, 2005) are both based on previous research in service quality and web service quality. The factors identified in these studies are similar to other studies we have reviewed, with slight differences in terminology and emphasis. An empirical study on organisational portals (Kuo et al., 2005) identified four factors of service quality: empathy (similar to service interaction quality in eQual); ease of use (usability); information quality (content quality) and accessibility (similar to aspects of usability and service interaction quality). A study on business to consumer portals in Greece (Gounaris & Dimitiadis, 2003) identified three service quality factors: Customer care and risk reduction benefit (similar to

service interaction quality); information benefit (content quality); and interaction facilitation benefit (usability, extended by adding consideration of technical design and speed). This increased our confidence in the use of the eQual dimensions and items as a starting point for the current research.

2.3.7.2 Measuring e-service quality in the public sector

Many universities are public sector organisations, and it might be expected that e-service quality issues associated with e-government would also apply to universities. However, in our view, there is a limited degree of generalizability between university and other e-government websites and web portals. The university user community is more homogeneous than the community of users of other e-government services. First, much of the user community can be expected to have a high level of educational attainment and literacy, because they will have attained the minimum requirements for university admission. Second, many universities bundle access to computing resources along with access to a range of other services as part of the cluster of benefits available to all students. This means that e-government service quality considerations such as availability, equity, access, and language (Tiecher, Hughes, & Dow, 2002) are not likely to be as important in a university context.

Qualities valued by users of e-government websites include a sense of control, convenience, and time and cost saving (Gilbert, Balestrini, & Littleboy, 2004). We expect that these issues will also be important considerations for users of university websites.

2.3.7.3 Service quality in higher education

There have been a number of previous studies on service quality in higher education. These give confidence that user perceptions of service quality in higher education are similar to those in other domains (Lagrosen, Seyyed-Hashemi, & Leitner, 2004). However the dimensions were operationalized differently, for example, library resources and computer facilities were instances of “tangibles” (Lagrosen et al., 2004). Factor analysis identified three dimensions of service quality in higher education: process (including prompt, willing, error-free service); empathy (including convenience, and attentiveness) and tangibles. This study also suggests that perceived service quality in universities is similar to that in other business domains, and by

implication, perceived e-service quality is also likely to be similar (Lagrosen et al., 2004).

2.3.7.4 Summary

In summary, domain-specific studies have followed a similar path to more generalised studies of online service quality, starting with studies of face-to-face services, particularly ServQual, and extending the concepts to encompass the characteristics of the specific domain of study. The additions to the constructs/items already suggested in eQual are similar to those identified previously in my review of e-commerce studies (Table 2.6). These studies support the addition of a construct for transaction quality, and based on this literature, we suggest the items for this construct should include the range of transactions available, the perceived usefulness (convenience) of the transactions, and the degree to which they save time or money.

Table 2.6: *Online service quality whakapapa: domain-specific studies*

Additions to eQual (from our literature review)	Items
Service Interaction Quality	I feel in control on this site
Usability	Overall, the response time was acceptable to me
Transaction quality	Completing transactions on this site will save me time or money I believe I would be able to complete transactions that are useful to me This site offered the range of on-line transactions I expected

2.3.8 Reflections on service quality and online service quality literature

We used a “narrative” approach to the literature review and examined the whakapapa (family tree) of online service quality literature starting with the early notions of face-to-face service quality from service marketing literature, through the appropriation of service quality measured in information systems literature to measure the information systems function; to the development of satisfaction and service quality measures for technology-based products and services such as end-user computing and e-commerce.

In combination, this is a very large body of literature. Reflecting on the totality of what we have read so far, several interesting trends can be observed. First, we notice the contribution of both face-to-face service quality literature and technology considerations. Second, we notice that the “dimensionality” described is often unclear. Third, we notice that many authors have gone to extensive lengths to synthesise or establish equivalences between technology characteristics and dimensions of face-to-face service quality, particularly those identified in ServQual.

As we mentioned in the introduction to this chapter, online services are technologies and information systems as well as services, and both perspectives have contributed to the literature we reviewed. This has resulted in some confusion between technology features and the user perceptions that arise from them. Several of the studies we reviewed include a mixture of indicators of both types, for example Alzola and Robaina’s literature review paper mixes “mechanisms to protect personal and financial data” (a technical characteristic) with “image” (a perception); both are suggested as indicators of security (Alzola & Robaina, 2005). It is likely that these are not constructs at the same level – the existence of appropriate technical mechanisms may be antecedent to the image produced as a result.

This is symptomatic of a more general issue with many of the models included in the review; although they refer to dimensions, and constructs, in some cases proposing multi-layer models, the directionality of the relationships is often not discussed or is unclear. For example, Brady and Cronin’s multi-layer model, reproduced as Figure 2.5, shows four levels (Brady & Cronin, 2001). At the second level of the model, arrows lead out of “interaction quality”, “physical environment quality” and “outcome quality” into “retail service quality” (suggesting these constructs are antecedent to an overall effect towards retail service quality). At the third level of the model these three dimensions also have arrows leading out into more detailed sub-constructs, for example “physical environment quality” has arrows leading to “ambient conditions”, “design” and “social factors”. This suggests a different kind of relationship, one of decomposition, rather than of antecedent and consequence.

Finally, considerable effort has been expended establishing equivalence between face-to-face dimensions of service quality, primarily those from ServQual, with quality

characteristics that are more relevant to electronic media. Alzola and Robaina found that two thirds of the e-commerce service quality studies they reviewed were based to a greater or lesser degree on ServQual (Alzola & Robaina, 2005). For example, Alzola and Robaina argue for equivalence between the physical site of the service delivery (ServQual's "tangibles") and the virtual interface of the website (a quality characteristic of electronic media). In some cases this becomes a little stretched, with equivalence argued between face-to-face and online constructs even though they have different indicators and different contexts.

2.3.9 Summary and Integration

Overall, changes in technology appear to be reflected in instruments to measure service quality. Studies of end-user computing introduced constructs such as "format" – the clarity of the way output was presented on screen, e-commerce studies introduced technical features such as non-repudiation, response-time to emails, and download time. It therefore seems likely that changes in measures are influenced by changes in technology.

Since one of my goals is to investigate the degree to which instability in the dimensions of online service quality is driven by changes in the underlying technology, we need to re-use an existing instrument in order to be able to make meaningful comparisons.

Bearing in mind the provisos from the previous section, we found eQual to be the most complete and best-supported stream of research into e-services quality. Overall, the constructs provided in online service quality literature from marketing offer only a limited extension to the constructs identified in e-commerce literature. In some cases e-commerce literature provides a better explanation for phenomena. A possible reason for this is that the e-commerce literature is informed by many seminal studies in services marketing, while we were unable to find examples where e-commerce studies in e-services, such as the extensive stream of research by Barnes and Vidgen, were cited in marketing literature. We noticed some differences in emphasis in studies of online service quality within e-commerce; with Li et al. exploring aspects of customer interaction with the organisation (Li, Tan, & Xie, 2002) in more detail than Barnes and Vigden; and suggesting different and more detailed metrics for measuring this

interaction. This made comparison difficult, but overall, we found e-commerce studies showed a difference in emphasis and granularity, rather than entirely different findings.

In summary, we have made some additions to the eQual dimensions from the literature review (shown in Table 2.7, with additions italicised). Most importantly, we suggest the addition of a new construct, which we have called “transaction quality”, defined as the richness and range of transactional functionality provided by the service, reflecting the evolution of website functionality, and the increasing availability of a range of online transactions.

We have also added two new items for service interaction, “control”, defined as a sense of being in control of the service experience, and “enjoyability” defined as a sense of fun and enjoyment. We also found considerable literature support for the importance of minimal waiting time, so this has been proposed as new item for usability.

Table 2.7: *Integrated e-service dimensions framed by eQual dimensions*

eQual dimension	Items	Literature support
Usability	Overall dimension supported, items defined differently/overlapping	“Accessibility” (Baierova et al., 2003)
	I find the site easy to learn to operate	(Barnes, 2001; Barnes & Vidgen, 2001, 2002)
	My interaction with the site is clear and understandable	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) “Clear and simple” (Bressolles, 2008) “Organisational and layout” (Bressolles, 2008) “Ease of Use” (Turban et al., 2002)
	I find the site easy to navigate	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) (Bressolles, 2008) “Ease of Use” (Turban et al., 2002)

Table 2-7 (cont.)

eQual dimension	Items	Literature support
Usability - cont	I find the site easy to use	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) (Bressolles, 2008) "Ease of Use" (Turban et al., 2002) "Ease of Use" (Dabholkar, 1996) "Ease of Use" (Doll et al., 1994)
	The site has an attractive appearance visual appearance	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) Site design (3 items – colourful, creative, attractive) (Bressolles, 2008) "Visual Attractiveness" (Baierova et al., 2003)
	The design is appropriate to the type of site	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) "Content quality/format" (Turban et al., 2002) "Organisation and layout" (Bressolles, 2008)
	The site conveys a sense of competency	(Barnes, 2001; Barnes & Vidgen, 2001, 2002)
	The site creates a positive experience for me	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) "Sense of fun" (Dabholkar, 1996)
	Overall, the response time was acceptable to me	"Minimal waiting time" (Dabholkar, 1996) "Quick and easy to complete a transaction" (Wolfenbarger & Gilly, 2003) Part of "Accessibility" > "web pages download quickly and completely" (Baierova et al., 2003) "technical design and speed" (Gounaris & Dimitiadis, 2003)
Information Quality	Overall dimension supported, items defined differently/overlapping	"Content" and "Content quality" (Baierova et al., 2003) "Content" (Doll et al., 1994)
	The site provides accurate information	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) "Content quality/reliability" (Turban et al., 2002) "Quality of information" (Li et al., 2002) "Accuracy" (Doll et al., 1994)
	The site provides believable information	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) (Bressolles, 2008) "Content quality/reliability" (Turban et al., 2002) "Quality of information" (Li et al., 2002)
	The site provides timely information	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) "Content quality/timeliness" (Turban et al., 2002) "Timeliness" (Doll et al., 1994)
	The site provides relevant information	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) (Bressolles, 2008) Customisation and flexibility (Bitner et al., 2000)
	The site provides easy to understand information	(Barnes, 2001; Barnes & Vidgen, 2001, 2002)
	The site provides information at the right level of detail	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) "In depth information" (Wolfenbarger & Gilly, 2003) "In depth information" (Bressolles, 2008)
	The site presents the information in an appropriate format	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) "Content quality/format" (Turban et al., 2002) "Format" (Doll et al., 1994)

Table 2-7 (cont.)

eQual dimension	Items	Literature support
Service Interaction Quality	This site has a good reputation	(Barnes, 2001; Barnes & Vidgen, 2001, 2002)
	It feels safe to complete transactions on this site	(Barnes, 2001; Barnes & Vidgen, 2001, 2002)
	My personal information feels secure	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) “Security/privacy” (Turban et al., 2002) “Privacy” (Wolfenbarger & Gilly, 2003) “Security” (Wolfenbarger & Gilly, 2003) Security/privacy (3 items) (Bressolles, 2008) “Security and Privacy” (Baierova et al., 2003)
	This site creates a sense of personalization	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) (Wolfenbarger & Gilly, 2003) “Customisation and flexibility” (Bitner et al., 2000) “Empathy” (Li et al., 2002)
	This site makes it easy to communicate with the organization	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) (Wolfenbarger & Gilly, 2003) “Customer Service” (Turban et al., 2002) “Responsiveness” (Li et al., 2002) “Call back systems” (Li et al., 2002) “User support” (Sandhu & Corbitt, 2003) “Customer service” (3 items) (Wolfenbarger & Gilly, 2003)
	I feel confident that goods/services I have requested will be delivered as promised	(Barnes, 2001; Barnes & Vidgen, 2001, 2002) “Logistics support” (Turban et al., 2002) “Fulfilment/reliability” (3 items) (Wolfenbarger & Gilly, 2003) “Reliability” (4 items) (Bressolles, 2008)
	I feel in control on this site	“Sense of control” (Sandhu & Corbitt, 2003) “Sense of control” (Dabholkar, 1996) “Sense of control” (Gilbert et al., 2004)
	This site conveys a sense of enjoyability or entertainment	(Dabholkar, 1996) “Enjoyability” (Baierova et al., 2003)
Transaction quality	This site offered the range of on-line transactions I expected	Statistics of Internet use “Selection” (Wolfenbarger & Gilly, 2003) “Outcome quality” (Gronroos, 2001) “Product portfolio” & “product offers” (Szymanski & Hise, 2000; Wang et al., 2001; Yang et al., 2004)
	I believe I would be able to complete transactions that are useful to me	“Time and cost saving” (Gilbert et al., 2004)
	I believe I would be able to complete transactions that are useful to me	“Convenience” (Gilbert et al., 2004)

2.4 Other relevant theory

In this second part of this chapter, we examine the wider theoretical context for the study and measurement of online service quality, including expectation-disconfirmation theory; theories of user attitudes and perceptions towards technology; other constructs of relevance to online service quality including ease of use, usefulness, trust and fit; and the use of affordance theory for determining the functional affordances of technologies.

As we signalled in the chronology of the study, developing an understanding of the nature of online service quality was a highly iterative process. Although we have presented this literature “at the beginning”, in reality, we went back and forth between modelling and literature as the study progressed.

On-line services are information systems as well as services, yet this aspect is frequently overlooked. Information systems theories can have strong explanatory power for studies of online service quality (Tate, Evermann, & Hope, 2008a); but instead, the service aspect is often emphasised, and the provenance of much online service quality research (for example, Barnes & Vidgen, 2002; Zeithaml, Berry, & Parasuraman, 1996) is services marketing literature. In this section, we examine the contribution that information systems theory can make to the study of online service quality, and the linkages between information systems theory and services marketing theory.

User attitudes and perceptions towards online services are also a subset of user attitudes and perceptions towards technologies in general, so theories that examine user attitudes and perceptions towards technology (and their shared background in expectation disconfirmation theory) may have salience.

In some cases, as with theories of IS adoption and continuance, we argue that some of the constructs included in these models are analogous to some of the dimensions claimed for online service quality. Since the dimensions of online service quality have proved unstable, we considered the possibility that the indicators (survey items) from eQual covered the scope of the domain of online service quality, but the higher order dimensions (content quality, usability and so on) were incorrect. We therefore examined other models that included the same indicators. These include the Task technology Fit Model (Goodhue & Thompson, 1995), and the IS Success Model

(DeLone & McLean, 2003). Finally, since we have argued that the dimensions of POLSQ are potentially affected by changes in the functionality of the underlying technology, we introduce affordance theory as a means of identifying relevant functionality.

2.4.1 Expectation disconfirmation theory

Both consumer behaviour theory and IS theories of user attitudes towards technology draw on expectation (dis)confirmation theory (ECT). This posits that consumers form an initial expectation of a service prior to use, typically from formal marketing communication by the firm and informal communication by peers. Second, they use the service (technology), and form perceptions about that experience. Third, they compare their post-use perception of performance with their expectation to determine whether their expectations (of this encounter) were met (confirmed or disconfirmed). This is the perception of satisfaction. Fourth, they modify and update their overall attitude towards the service ("service quality") based on their satisfaction of this instance of use. Finally, this may lead to an intention to continue to use the service (Oliver, 1981).

IS acceptance and continuance literature also draws on expectation disconfirmation theory. After the first usage experience, past use will inform attitudes towards ongoing interaction with the information system. The iterative relationship between past use and future intentions is described as “the basic concept underlying user acceptance models, and is shown in Figure 2.7 (Venkatesh, Morris, Davis, & Davis, 2003).

Figure 2-7: *The iterative relationship between past use and future intention*

For continuing users of information systems, both confirmation (an experience with an online service that met or exceeded expectations) and perceived usefulness (performance, productivity and effectiveness), contribute to attitude towards the service (Bhattacharjee, 2001).

The existence of this iterative relationship means user expectations of technology (prior to use) are updated by actual use. This means that constructs used in information systems theories of technology acceptance, adoption and continuance can be relevant to the study of POLSQ.

2.4.2 Theories of user acceptance, adoption and continuance of technology use

2.4.2.1 *TAM and UTAUT*

Information systems theory draws heavily on theories of attitudes and behaviour from social psychology. The TAM model (Davis, Bagozzi, & Warshaw, 1989) shown in Figure 2.8, and its more comprehensive successor, the unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003) both cite the Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975) and its successor, the Theory of Planned Behaviour (Ajzen, 1991) as an important part of their provenance.

Figure 2-8: *The Technology Acceptance Model (TAM)*

In the TAM model, intention to use an information system is determined by the person's attitude towards using the system. In determining that attitude, the person makes a judgement as to the subjective probability that using a technology will increase their job performance (perceived usefulness). Perceptions of usefulness are also determined by the degree to which the user expects that using the system will be free from effort (perceived ease of use).

2.4.2.2 TAM extensions in an online context

Extensions to PU and PEOU in an on-line environment have also found support in empirical studies. Since this literature is equally as large, diverse and contradictory as service quality and online service quality literature, we have opted to concentrate on those constructs that are closely related to constructs we have already identified. Gefen et al. found that consumer trust is as important to online commerce as the widely accepted TAM use-antecedents, perceived usefulness, and perceived ease of use (Gefen et al., 2003). “Playfulness” has also been found to be a determinant of user acceptance and attitude towards websites (Moon & Kim, 2001).

2.4.2.3 Playfulness

A sense of fun and enjoyability have been found to be associated with satisfaction with websites (Moon & Kim, 2001), online retailing (Ahn, Ryu, & Han, 2007), and web portals (Lin, Wu, & Tsai, 2005). Fun and enjoyability are considered by these authors to be a dimension of “playfulness”. Playfulness in an IS research context has a slightly different meaning to the vernacular sense of the word; it incorporates notions of concentration (involvement and losing track of time), curiosity (exploration), and enjoyability (Ahn et al., 2007; Lin et al., 2005; Moon & Kim, 2001).

As with many other constructs in this research area; there is little consensus about the causal relationships. Lin et al. modelled playfulness as a consequence of confirmation, and as an antecedent of satisfaction and continuance intention for a web portal (Lin et al., 2005). Ahn et al. modelled playfulness as a consequence of system quality, information quality, service quality and perceived ease of use, and an antecedent of attitude towards online retailing and behavioural intention to use (Ahn et al., 2007). None of these are a good fit to the structure of the eQual constructs that we used as a basis for the study.

Shopping is a more hedonic activity than performing an online self-service task. Online self-service users are task oriented, and are more interested in carrying out their tasks efficiently than in “surfing” and exploring a website (Tate et al., 2008b). We therefore did not include questions associated with curiosity and concentration. The subset of the playfulness construct that related to a sense of fun and enjoyability

has been previously identified in the discussion of website quality, and is further supported by TAM-related literature.

2.4.2.4 Trust

Trust in an online service context relates to the customer's willingness to accept risks (for example, by divulging their personal information) and their expectation that the online service will function as expected and deliver on its promises. Many studies have emphasised the importance of trust in an online context, for example (McKnight, Choudhury, & Kacmar, 2002; McKnight, Kacmar, & Choudhury, 2004; Zeithaml, Parasuraman, & Malhorta, 2002). In TAM-related literature in an online shopping context, trust has been found to be a consequence of both perceived usefulness and perceived ease of use, and an important determinant of intended use (Gefen et al., 2003). This provides additional support, from TAM-related literature, for the inclusion of items relating to trust on an online self-service website.

2.4.2.5 IS continuance theory

IS continuance theory is an extension of widely-used information systems user acceptance models such as the Technology Acceptance Model (TAM) (Davis et al., 1989), and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003). Continuance theory focuses on experienced, continuing users rather than those in the initial stages of technology adoption (Bhattacharjee, 2001; Hsu, Chiu, & Ju, 2004). This theory naturally has more synergy with service quality theory than technology acceptance theory, since a meaningful opinion of online service quality requires that the user has experienced the online service.

The basic concept underlying user continuance models is that actual use updates individual reactions (including confirmation/disconfirmation, or service quality perceptions), which in turn predict further expectations and intentions. Intention to use is preceded by perceived usefulness (PU) and a positive attitude (satisfaction). Perceived ease of use (PEOU) was found to have minimal influence on IS continuance intentions, in contrast to IS adoption models (Davis et al., 1989).

Within marketing literature, the importance of understanding continuance has also been emphasised. *"In planning new strategies to enhance adoption and usage of new*

SST's, it is important to be able to address both the factors leading to the trial as well as those impacting repeated use and commitment" (Bitner, Ostrom, & Meuter, 2002, p. 103).

2.4.3 Aligning expectations and perceptions of online service quality

In an online context, the only real difference between theories of user adoption, acceptance, and continuance of technology use, and user attitudes and perceptions towards online service quality is the point in time at which the measure is taken. Theories of user attitudes and intentions towards technology look forward towards the user's future behaviour, while theories of user perceived service quality look backwards, reflecting on past experience. However, given the feedback loop that operates between past experience and future expectation and intentions, there are no clear antecedents and consequences. This iterative process was modelled in a face-to-face context in the hospitality industry in a longitudinal study (Boulding et al., 1993). These authors found a difference between what customers would ideally like ("should" expectations), and what they actually expected to happen ("will expectations"). Each of these contributed to the perception of the service quality of the encounter (post encounter). Repeated experiences led to an overall attitude towards the organisation.

The authors suggest that maintaining an ongoing gap between "should" and "will" expectations creates cognitive dissonance for customers. This is defined as the discomfort arising from holding two contradictory attitudes or beliefs simultaneously. The theory of cognitive dissonance proposes that people are motivated to reduce dissonance by changing or rationalizing their attitudes, beliefs, and behaviors (Festinger, 1957). Boulding et al. suggest that this explains a tendency for users to align their "should" and "will" expectations of a service over time.

In an online context, the expectation-perception gap is likely to be negligible for an experienced user, because commercial software usually behaves consistently and predictably. This means that "after the event" perceived outcome quality (from service quality theories), and "before the event" perceived usefulness (from information systems adoption theories) are very closely related constructs. Similarly

process quality is similar to perceived ease of use. We can also assume that continuing users are relatively satisfied with the service quality they have so far received.

You can see from this discussion that we can view “perceived usefulness” (from the TAM) as conceptually equivalent to “outcome quality” (from the “Nordic” model), and “perceived ease of use” (from the TAM) as conceptually equivalent to “process quality” (from the “Nordic” model); but measured at a different point in time. This provides a bridge between the “Nordic” model of service quality and TAM-related information systems theory.

2.4.4 Theories of Task Technology Fit (TTF)

Theories based on task-technology fit offer an alternative conceptualisation of online service quality to theories based on expectation-disconfirmation theory. Rather than beliefs and affect towards a technology being based on the gap between expected and perceived service quality, TTF posits that these beliefs are based on the extent of the fit between the individual, the task, and the technology. The better the fit, the more likely the user will have a positive perception of the service quality. The basic TTF suggests that user’s perceptions of whether the information on a website is accurate, relevant, at the right level of detail, in the appropriate format, easy to use, and so on are based on the degree to which it is a good fit to the purpose for which they wish to use it. Figure 2.9 shows the TTF model (in solid lines) (Goodhue & Thompson, 1995). We have also included (in dotted lines) a subset of the wider model, which Goodhue and Thompson called the technology to Performance Chain.

Figure 2-9: *Task technology Fit and the Technology to Performance Chain*

Tasks – the process and actions of an individual turning inputs into outputs

Technology characteristics – the tools used by individuals in carrying out tasks, both computer systems (hardware, software, data) and the user support services (training, help lines etc) provided to assist users in their tasks.

Individual characteristics – including training, computer experience, motivations

Task Technology Fit – the degree to which a technology assists an individual in performing his or her tasks – the degree of correspondence between task requirements, individual abilities and the characteristics of the technology or system used. We note that this includes several constructs that are also used as measures of POLSQ including data (information) quality, ease of use, and learnability.

Precursors of Utilisation - Attitudes and beliefs that affect the decision to use the system.

Utilisation - The uptake of the technology, and the degree to which it is incorporated into work routine.

Performance impact - this is the accomplishment of a portfolio of tasks by an individual.

This shows that Goodhue and Thompson conceptualised beliefs and affect (attitude) towards a technology to be a consequence of perceived fit on a range of factors including several that are used as indicators of POLSQ (Goodhue & Thompson, 1995). This suggests that user perceptions of task technology fit may influence perceptions of online service quality.

2.4.5 E-commerce Success

Another conceptualisation of service quality in e-commerce is provided by the updated Information Systems Success Model (DeLone & McLean, 2003), shown as Figure 2.10. In this paper, the authors do an extensive literature review to update their 1992 model (DeLone & McLean, 1992), including applying it to an e-commerce context. In this model, information quality and system quality are conceptualised as independent constructs at the same level as service quality, rather than as dimensions of service quality. Service quality was added to the 1992 model, and is measured using “ServQual-like” constructs based on the work of (Pitt et al., 1995); assurance, empathy and responsiveness. Note that these authors suggest indicative items to measure their constructs, but do not include a survey instrument.

Figure 2-10: *DeLone and McLean's Information Systems Success Model*

The items suggested by Delone and McLean for Information Quality include many that are also used in eQual, or the revision of eQual, as indicators of different constructs, as shown in Table 2.8.

Table 2.8: *Use of common indicators in eQual and the Information Systems Success Model*

Indicator	Delone and Mclean (1993)	eQual 4.0
Adaptability	Systems quality	
Availability	Systems quality	
Reliability	Systems quality	
Response time	Systems quality	
Usability	Systems quality	Usability
Completeness	Information quality	
Ease of Understanding	Information quality	Content quality
Personalisation	Information quality	Service Interaction Quality
Relevance	Information quality	Content quality
Security	Information quality	Service Interaction Quality

This multiplies the degree of confusion and ambiguity associated with these constructs. Interestingly, however, there is considerable overlap in the indicators, although the dimensions and relationships are conceptualised very differently. This suggests that the indicators identified in the literature review, building on those identified by the eQual authors, may be salient, even if the dimensions and causal relationships can be hypothesised differently.

2.4.6 Software Affordances

A critical difference between technology mediated self-service and face-to-face service is that the technology does not exercise any independent judgement and adaptivity (or can only do so to the extent that it has been pre-programmed). While people can also be constrained by process design and the limits of their authority in a job, they are, to some degree, almost infinitely adaptive. A person can be asked for advice, directions, to call a superior, and so on. Software cannot. The range of

possible options available from most technology-mediated self-service offerings (outside of research laboratories anyway) is finite.

We have also argued that changing functionality over time has the potential to change the dimensions of POLSQ. However, relevant and purposeful functionality does not change as rapidly as individual technologies and applications, and some mechanism is required to identify relevant functional changes.

Affordance theory is a useful way to frame the differences between software and human agents, and to identify salient changes in the functionality of online services. Affordance theory is widely applied in ergonomics, design and human computer interaction, but it is seldom used for the wider domain of software application design. In design theory, affordances involve the interaction between an actor and the properties of an environment (Garver, 1991; Norman, 1999). The key notion of affordances is that human activities are shaped by the material environment they occur in (Garver, 1991). Broadly, an affordance is what is allowed or “afforded” by an object in an environment. For example, something we might normally recognise as a chair (for sitting on) might also afford standing on, or in rare circumstances, being used as a weapon. It would not afford use for a fine motor function such as sewing or tooth cleaning. In software and HCI, various types of affordances have been identified.

As with many other fields of enquiry, affordance theory suffers from semantic ambiguity, where different terms are used to mean similar things, and the same term is used to mean different things. In this section, we are indebted to Rex Hartson for his review and definitions of various types of affordances relevant to software design (Hartson, 2003). Hartson identifies four types of affordance: real or physical; perceived or cognitive; functional and perceptual. A summary of Hartson’s definitions, compared with definitions used by other authors, is provided in Table 2.9.

A real or physical affordance is defined as a design feature that helps, aids, supports, facilitates, or enables physically doing something. Adequate size and easy-to-access location could be physical affordance features of an interface button design enabling users to click easily on the button (Hartson, 2003).

Table 2.9: *A comparison of definitions of affordances*

Hartson (2003)	Physical affordance	Cognitive affordance	Perceptual affordance	Functional Affordance
(Gibson, 1977)	Affordance	Perceptual information about an affordance	Implied	Not addressed
(Norman, 1999)	Real affordance	Perceived affordance	Implied	Not addressed
(McGrenere & Ho, 2000)	Affordance	Perceptual information about an affordance	Indirectly included in perceptibility of an affordance	Affordances in software
(Garver, 1991)	Affordance, also perceptible affordance	Perceptual information about an affordance, also apparent affordance	Indirectly included in perceptibility of an affordance	Implied

Perceived or cognitive affordance is a design feature that helps, aids, supports, facilitates, or enables thinking and/or knowing about something. As a simple example, clear and precise words in a button label could be a cognitive affordance feature enabling users to understand the meaning of the button in terms of the functionality behind the button and the consequences of clicking on it (Hartson, 2003).

A perceptual affordance is a design feature that helps, aids, supports, facilitates, or enables the user in perception (e.g., seeing, hearing, feeling, sensing). Perceptual affordance includes design features or devices associated with visual, auditory, or haptic/tactile sensations. Cognitive affordance and physical affordance are key concepts in interaction design but perceptual affordance supports physical and cognitive affordance. Users must be able to perceive cognitive affordances and physical affordances in order for them to aid the user's cognitive and physical actions (Hartson, 2003).

A functional affordance is a higher-level enablement (than simply the result of a button-click) that is a system function that helps the user to do something in the work domain. This draws on McGrenere & Ho, who refer to the concept of application functionality usefulness, something they call 'affordances in software'. This takes the discussion of affordances beyond user interfaces to the larger context of overall

system design, and the use of software for purposeful actions (McGrenere & Ho, 2000).

This concept of functional affordances allowing purposeful user action combines the concepts of usability and usefulness (Landauer, 1995). Usefulness is created by the utility of functional outcomes. Usability stems from the effectiveness of cognitive affordances for understanding how to use physical affordances, from the physical ease of using the physical affordances, and from the perception of these via perceptual affordances (Hartson, 2003).

In the context of this research, we are mainly interested in functional affordances. Software has a finite number of functions embedded in its code and business rules. For example, if the software behind an online service does not include code and business rules to allow online fee payments, or the display of full-text articles from certain journals, no user action, knowledge, or skill will ever induce it to do carry out this function.

Functional affordances can also work at a wider technology level. Until functions such as hyper linking became available, no one expected software applications to use them. Current industry trends such as integrated portals, and the integration of SMS (short message system) functionality with online services can be regarded as supporting new functional affordances – they support functions such as push notifications that were not previously available.

Defining functional affordances is complex, and may require a degree of negotiation and adaptation between the user and the software. For example, many users are not aware of all the functionality that an online service offers, and may only use a subset of it. So we have a potential gap between perceived and actual functional affordances. In a further complication, people do not always use software for the purpose for which it is designed, so the user's set of perceived functional affordances may be different to the designers. For example, many people anecdotally report using Amazon.com as a research database, since the list of titles it supports is so extensive, while the software was designed as a virtual shop-front for book-selling. The site affords searching book titles for research purposes as well as for book buying. This means you cannot simply

view perceived functional affordances as a subset of designed functional affordances, although there is clearly a relationship.

The growing popularity of off-the-shelf package software, and internet-based services means that increasingly, many people will use, or attempt to use software without receiving any training; and many online services and software packages offer a smorgasbord of functions in the hope that a reasonably wide base of users will find something to meet their needs. This means that there is an adaptation process as users become familiar with the service, which will influence user's perceptions of the quality of the service.

The evolution of online service quality research, described in previous sections of this chapter argues that underlying technologies have changed the nature of POLSQ. Characteristics of face-to-face services are not relevant in a technology context, and characteristics of end-user computing service quality may not be the same as those of internet-based service quality. The functional affordances of web-browser based information systems have evolved from offering primarily information provisioning and communication, to offering full transactional capability.

2.4.7 Summary

One of the reasons we suggested for the instability of the dimensions of online service quality was that the phenomenon itself was changing as technology evolved. There appeared to be evidence for this in the literature we discussed in the previous section. Affordance theory provides a means of framing this evolution in functionality. The widespread availability of increasingly sophisticated self-service transactions is a new affordance which has been added to websites, which previously concentrated on communication and information provisioning.

We have also examined theories of user attitudes and perceptions towards technology; since user perceptions towards online service quality are a subset of this wider body of theory. Although the constructs and theory nets for leading information systems theory are different to those posited for service quality and online service quality, there is some degree of overlap. Information systems theory and services marketing theory have a shared background in social psychology and expectation

disconfirmation theory. We also note that there is a considerable degree of overlap at the indicator (rather than the construct) level between information systems theories and theories of user attitudes and perceptions towards technology. This complicates the picture further, but may also provide new avenues of enquiry.

2.5 Reflection

This chapter takes a conventional “narrative” approach to the literature of online service quality, its extensive family tree, and some related theory from information systems. In summary, we offer some cautious conclusions about online service quality and summarise some of the issues and limitations with this approach.

The richness of web functionality has been steadily increasing. This means that there is an increasing set of functional affordances that form an essential part of a quality online service. We propose a new construct, “transaction quality” to capture the fact that online services increasingly afford a range of transactions as well as information provisioning and communication; and the range and richness of those transactions may be a differentiating factor for POLSQ. This is supported by the review of online service quality literature offered in the first part of this chapter.

Although it is conventional, there are a number of issues with a “narrative” approach to a literature review in this subject area. These include an overuse of this “narrative” approach by previous researchers, resulting in a lack of consensus, confusion, and failure to develop a cumulative literature tradition. These challenges are multiplied by the fact that research in online service quality is based not on a single body of work, but multiple streams of literature from services marketing, information systems, and electronic commerce.

We recollect that a narrative literature review is conducted by verbally describing the past studies, focusing on theories and frameworks, elementary factors and their research outcomes, with regard to arriving at hypothesized relationships (King & He, 2005). This literature is vast, and past information systems and marketing researchers have not been very disciplined or consistent in construct formulation. The same term is used for constructs that are measured with different items, while different terms are

used for constructs that include some of the same items (Barki, 2008); these ambiguous constructs are included in different theory nets. The result, as a presenter said at a conference we attended recently is that “Arrows go every which way. You can show that A causes B and also that B causes A”. Some attempts have been made to offer alternative conceptualisations (Sylvester & Tate, 2008; Sylvester et al., 2007). These have concluded that the perception that key questions have been “answered”, and that there is a cumulative research tradition are largely illusory, and have been created by determined advocacy from influential researchers. Overall, brokering a consensus of previous literature is a highly subjective and almost impossible task.

This means we are in a difficult position with regard to our use of previous literature. Although this study questions the validity of many past efforts to conceptualise and operationalize constructs in research, it is too believed that much research has been done well, at least in part. Wherever the researcher references past research in support of research decisions, it is with the full belief that prior research is sound in those ways pertaining to the citation.

Given this level of confusion, with multiple studies piling up different and competing theories in an unsorted (and unsortable) heap, how should the study of online service quality be advanced? Since we aim to study the phenomenon quantitatively, we need a survey instrument. Since we aim to investigate whether a new affordance has affected the dimensionality of the construct, we needed to be able to compare the results from a study that included items associated with transaction quality with one that did not. This suggested that we needed to use an existing instrument as our starting point.

In the first part of this chapter, we established that the eQual instrument was an appropriate starting point for the quantitative research. The current version of eQual (4.0) does not include items to measure transaction quality, but otherwise provides good coverage of the domain of online service quality. By adding a construct and items for transaction quality we can compare the revision of the instrument with the results of previous studies.

There are a number of risks and challenges in this approach, as we subsequently discovered. These set-backs sent us back to the library to investigate theoretical approaches to modelling and measurement, which are discussed in the next chapter.

3 Theory and Measurement

3.1 Overview

This chapter follows the second line of enquiry: the possibility that the instability in the online service quality construct is because existing theoretical approaches are flawed. This could be the result of the theory net, or way the constructs are modelled and measured. This chapter presents the results of the ongoing reading program; a critique of existing approaches; and a justification for alternative conceptualisations of the online service quality construct.

This chapter is organised as follows: first we consider general issues in scientific and social science theory and then apply them to online service quality research, next we consider general issues in measurement, particularly factor-analytic approaches, the ontological implications of various measurement approaches, and the likely impact of current debates in measurement theory; and apply them to the research area.

3.2 Theoretical foundations

3.2.1 Introduction

In this section, we examine the nature of scientific theory in order to frame a discussion about the theoretical bases of online service quality research. If you were to ask most researchers in user perceived on-line service quality what was the theoretical foundation of their work, they would most likely answer “ServQual” or perhaps “expectation-disconfirmation theory.” A very cursory on-line search provided at least 400 peer-reviewed papers published in scholarly journals from 1988 to the present that included the keyword ServQual in the citation and abstract, and more than 4000 that included the keywords “Service Quality”.

It seems reasonable, at this point, to examine and critique some of the common bases for making truth claims about online service quality. This is a very large area, which cannot be given a comprehensive coverage without writing a philosophy thesis, rather than an information systems thesis. Here, we review briefly some of the important

concepts. We concentrate on those that we believe have the most potential to explain aspects of research in POLSQ. We discuss Karl Popper and logical positivism, Kuhn and “normal science,” critics of Kuhn, and post-positivism.

3.2.2 Philosophy of science

3.2.2.1 *Logical positivism*

Logical positivism was a school of scientific thought that was pioneered in the 1920's by the so-called “Vienna Circle.” Logical positivism assumes that theories can be verified by evidence. In this context, evidence is considered to be a set of statements reporting the outcome of an observation or experiment. The statements are assumed to be directly verifiable by the experiences they describe. In other words, truth statements about the world can be tested by making observations or conducting experiments, and accurately reporting them (Curd & Cover, 1998). Statements other than directly verifiable ones are not admissible.

3.2.2.2 *Karl Popper*

This approach was extensively critiqued by Popper, who observed that many “theories” such as Freudian psycho-analysis and Marxist social theory are endlessly verifiable, but not falsifiable. By way of example, Popper argues that almost any interpersonal phenomenon lends itself to a Freudian explanation, and almost any social phenomenon to a Marxist explanation. Popper proposed as an alternative that scientific theories should make “risky” predictions that can be tested and shown to be false (Popper, 1959). Theorising should involve the formulation of universal statements in a form that enables them to be tested and falsified against observations of what occurs in the real world.

3.2.2.3 *Thomas Kuhn*

This view was critiqued by Kuhn, who argued that the conduct of scientific enquiry was frequently better explained by what he termed “normal science”. Normal science refers to the extended periods of time between “scientific revolutions”, when he proposes that for much of the time, scientists work under the aegis of a paradigm that they all accept. Scientific activity, rather than consisting of theory testing, could better be described as “puzzle-solving”: making more accurate measurements, looking for

entities and processes the paradigm tells us must exist, extending the paradigm to new areas and types of phenomena (Kuhn, 1996).

The central assumptions of normal science are that: (1) science is based on a paradigm which tells scientists what experiments to perform, which observations to make, how to modify their theories, how to make choices between competing theories; (2) science is dogmatic, and involves an attempt to force nature into the pre-formed and relatively inflexible box of the paradigm; (3) science is objectively progressive, and the outcome of scientific enquiry is to accumulate solutions to scientific puzzles; (4) science is holistic, the elements of a paradigm are intertwined, so that you cannot accept some elements and reject others; (5) Observation is theory-laden; (6) meaning is theory-dependant, theories affect what scientists mean by the theoretical terms they employ (Kuhn, 1983).

Kuhn suggests that the quality of normal science is evaluated in a context of values that are widely shared by the scientific community (Kuhn, 1983). During periods of normal science, *“failure to get the right answer to an exercise is regarded as a failure of the student, not of the theory”* (Curd & Cover, 1998, p. 67). Shared values include:

- **Accuracy:** The degree to which theory agrees with experiments and observations that have already been made, and the number and importance of the problems (puzzles) that it solves.
- **Consistency:** Theory should be internally consistent, that is, free from logical contradiction; and externally consistent, that is, it does not contradict any other accepted theory.
- **Scope:** The breadth of the logical consequences of the theory should have consequences that go far beyond the set of observations, laws, and sub-theories it was originally designed to explain.
- **Simplicity:** A theory should have unifying power and bring order to phenomena that would otherwise be considered separate and isolated; and parsimony, all other things being equal, the fewer assumptions, and therefore simpler, the better.

3.2.2.4 *Critiques of Kuhn*

Although extremely influential, Kuhn's arguments in turn have been critiqued by philosophers of science. Longino refutes the notion of value neutrality, and instead, proposes that science should be moderated by criticism, and must permit and encourage transformative criticism (Longino, 1990). In order for this to occur, a number of conditions must exist. These include: avenues for publication of criticism of evidence, methods, assumptions and reasoning; shared standards that critics can invoke; a willingness for the scientific community to respond to such criticism; an environment where intellectual authority is shared amongst qualified practitioners, so that discourse, debate can occur, and a "rational consensus" can be established (Longino, 1990).

3.2.2.5 *Other concepts*

Other key concepts from philosophy of science that are essential to the evaluation of the status of existing research as theory include:

- Hypothesis: a scientific claim that is put forward as a plausible conjecture before there is enough evidence to warrant its acceptance or rejection. A theory, on the other hand, is a well-established scientific claim that is provisionally accepted as true or at least well-confirmed by evidence.
- The assumption of value-neutrality suggests that scientific theories should be governed exclusively by cognitive values, which are linked to the aims of science as a knowledge-seeking enterprise. These include predictive power, explanatory scope, and simplicity. Contextual values should have no place in science – "good" science is objective and free from subjective values.
- Meaning and meaning variance. Establishing the meaning of terms used in scientific evidence is extremely important in order to provide a basis for accumulation of knowledge and shared standards that critics can invoke. Otherwise there is a real risk that "they may utter the same words, but the words have different meanings, so any logical comparison of their utterances is precluded, leaving adherents of rival theories simply talking past one another." (Curd & Cover, 1998, p. 222).

- Scientific realism believes that the world studied by science actually exists and has the properties it does independently of our beliefs, perceptions and theories. The proper aim of science is to describe and explain that world, including aspects that cannot be observed directly. By and large, what science says about the world is true, and by continually replacing existing theories with better ones, science makes progress (Curd & Cover, 1998).

3.2.2.6 *Post-positivism*

The influence of philosophers of science such as Popper, Kuhn and Longino have been credited with the “death of positivism”, although many of the basic assumptions, such as the existence of an independent reality, have continued in what has come to be known as the “scientific” perspective. “Post-positivists”, and in particular, critical realists believe that there is a reality independent of our thinking about it that science can study, but influenced by Kuhn, they recognise that observation is fallible, and that theory is continually revised. Measurement is considered to be fallible, so post-positivism relies on multiple measures and observations to obtain triangulation of observations across these multiple sources in order to obtain a more accurate and complete picture. Following Kuhn, the post-positivist also believes that all observations are theory-laden and that researchers are biased by their cultural experiences and world views.

Positivists believed that objectivity was a characteristic that resided in the individual scientist. The proper aim of science was to put aside personal biases and beliefs and try to uncover truths about the world as it 'really' is. Post-positivists reject the idea that individuals can see the world perfectly as it really is. We approximate truth by triangulating across multiple (individually) fallible perspectives. Therefore, following Longino, objectivity is not the characteristic of an individual; it is inherently a social phenomenon that is approximated when researchers criticize each other's work. Theories that survive are a bit like the species that survive in the evolutionary struggle. This is sometimes called the natural selection theory of knowledge and holds that ideas have 'survival value' and that knowledge evolves through a process of variation, selection and retention. Enduring theories adapt and approximate being objective and providing an understanding of reality.

3.2.3 Theory in social sciences

Many words have been written on social science, its relationship to the physical sciences, so-called “physics envy” and the issues associated with applying quantitative techniques and normal science assumptions to the social sciences. It is not possible to review the entire debate here. Instead, we have selected some key aspects of the debate that have relevance to this research. We first review the foundations of social science theory in social psychology, and Gregor’s taxonomy of the purposes of information systems theory (Gregor, 2006). We then use these insights to evaluate the quality of service quality theories and identify issues with existing theories.

Quantitative research in social science is firmly based in Kuhnian, normal science assumptions. Accuracy, consistency, scope (generalisability) and simplicity (parsimony) are highly valued. Another important point that should be added is originality, which is a key publication criterion for top information systems journals (Seddon, 2006).

These definitions of quality in social science theory are closely modeled on conceptualizations of theory in the physical sciences. Translating social phenomenon into theory of this type is non-trivial. A leading exponent of the application of the scientific, hypothetico-deductive model of theory building in the social sciences is (Dubin, 1978). In overview, Dubin suggests that propositions are a kind of truth statement that can be verified (or falsified) according to the system of logic specified in the researcher’s model. The model is a collection of theory units and their relationships. “Theory units” is a term coined by Dubin as an alternative to the more widely used term “concept”. This is because “*concepts may also mean whole theories or laws of science, or even conceptual frameworks... [this] has led me to employ the more neutral term unit to designate the things out of which theories are built*” (Dubin, 1978, p. 38).

Propositions are predictions of the value of one or more units in the model. If propositions are to be tested, it is necessary to identify empirical indicators that produce values for the units included in the model. Empirical indicators should make clearly visible, and replicable, both the operation (process) of measuring, and the value produced, in order to be considered reliable (Dubin, 1978).

3.3 Theory in information systems and online service quality research

3.3.1 The Foundations of Information Systems Theory in Social Psychology

3.3.1.1 *Attitudes and Behaviours*

Information systems theory draws heavily on theories of attitudes and behaviour from social psychology. The TAM model (Davis et al., 1989) and its more comprehensive successor, the unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003) both cite the theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975) and its successor, the Theory of Planned behaviour (Ajzen, 1991) as an important part of their provenance.

Although the core theory is widely known, some of the more detailed discussion that lies behind these theories is less well cited in information systems research, and has salience for this research. We turn our attention to a review of the definitions of attitudes and beliefs, and the process by which attitudes are formed (Fishbein & Ajzen, 1975). An overview is provided in Figure 3.1.

Figure 3-1: *Current beliefs lead to attitudes, intentions and behaviours (Adapted from Fishbein & Ajzen (1975, p. 15))*

Fishbein & Ajzen reiterate the distinction between affect, cognition and conation. Affect is a person's feelings towards and evaluation of some object (person, issue or event); cognition is his or her knowledge, opinions, beliefs and thoughts about the object; and conation refers to his or her behavioural intentions. Fishbein & Ajzen use "attitude" to denote affect, "belief" to denote cognition, and "intention" for conation

(Fishbein & Ajzen, 1975). To elaborate a little further, belief links an object to some attribute. The object, in this research, would be a website, and the attribute may be any trait, property, characteristic, or outcome associated with that website. The authors also note that attitude is frequently measured by a single indicator of affect for or against an object, and should only be used when it is sensible to place the respondent on a bipolar affective dimension. When the measure places the individual on a dimension of subjective probability relating an object to an attribute, the label “belief” should be applied.

3.3.1.2 *The formation of beliefs*

This is an important argument for the research, so we will step through it in some detail. Fishbein & Ajzen assert:

1. The belief concept is central to understanding attitudes, intentions and behaviours.
2. Beliefs about an object form the basis for the formation of attitudes towards that object
3. Belief formation involves the establishment of a link between two aspects of an individual’s world. One is direct observation (via the senses) of an object (for example, that a table is round, or that Fred is crying), which give rise to descriptive beliefs about the object. Since people rarely doubt the validity of their own senses, descriptive beliefs are usually held with maximum certainty.
4. Beliefs that go beyond directly observable events are inferential beliefs. These are often formed on the basis of descriptive beliefs – for example a person that is crying may be inferred to be feeling sad. Inferential beliefs may be based in the first instance on previous inferences, but most inferential beliefs can eventually be traced back to descriptive beliefs.
5. However, the distinction between descriptive and inferential beliefs can be seen as a continuum, rather than a clear distinction.

“At the descriptive end of the continuum, a person’s beliefs are directly tied to the stimulus situation, and at the inferential end, beliefs are formed on the basis of these stimuli as well as the residues of the person’s past experiences; the continuum may be seen as involving [a range from] maximal to minimal use of such past experiential residues” (Fishbein & Ajzen, 1975, p. 133).

3.3.1.3 Conceptual inclusiveness

A related issue which can also cause confusion in IS research is a tendency to mix constructs at different levels of generality. In social science research this is sometimes described as different levels of “conceptual inclusiveness” (Blalock, 1982).

Blalock observes that “[Social scientists] hope to increase both the scope and inclusiveness of our most important concepts”, while acknowledging there is a trade-off between conceptual inclusiveness and precision (Blalock, 1982).

“The greater the inclusiveness and scope [of indicators] the less likely that those indicators that are common to a diversity of actors and settings will also be identical or equivalent indicators across the full range of applicability of the concept” (Blalock, 1982, p. 74).

There is a greater degree of inference involved in forming beliefs about more highly generalised constructs. This means that a greater proportion of the belief will be based on a person’s individual context and prior experience. In other words, the “meaning” will not be the same across different people. Using eQual as an example, the question “I find the site easy to use” is included (Barnes & Vidgen, 2002). “Ease of use” is a consequence of more specific beliefs such as “I find the site easy to navigate” and “I find the site easy to learn to operate”.

For example, the perceptions of various website users of the navigability of a specific site are likely to be reasonably precise, and to be related to objective features such as the number of clicks required. There may be some variation based on the individual, but the responses are likely to be reasonably similar between individuals, with a substantial portion of the differences in scores explained by objective qualities of the site. However, “ease of use” has a much wider scope, and is much more generalisable. To extend the argument, “is it easy to use?” is a question one could reasonably ask about a diversity of settings ranging between using a lawn-mower and using a cook-book, but the actor’s responses in those settings are unlikely to mean the same thing, so the inclusiveness is gained at the cost of precision.

Previous research has found that individual responses on a highly generalised construct such as effort expectancy (a similar construct to ease of use) may vary based

on a wide range of factors, including tool knowledge and domain knowledge (Tate et al., 2008a), and may be moderated by gender, age, experience and voluntariness (Venkatesh et al., 2003).

3.3.1.4 Salience for this research

This research is concerned with perceptions (beliefs) and overall affect (attitude) towards online service quality. We assert that online service quality research in the past has often failed to adequately distinguish between more descriptive and more inferential beliefs. The terms “perception”, or “user perception” or “customer perception” are often used in information systems and consumer behaviour literature as a catch-all that encompasses both descriptive and inferential beliefs. For example, Watson et al.’s instrument for measuring IS service quality includes both descriptive beliefs “IS’s employees are well-dressed and neat in appearance” and inferential beliefs “IS is dependable” (Watson, Pitt, & Kavan, 1998). Applying theory of beliefs and attitudes to my research questions, we can derive the following principles. An overview of these is provided in Figure 3.2.

1. Descriptive beliefs (perceptions) will (must) change with the objective qualities and affordances of the technology artefact (Fishbein & Ajzen, 1975).
 - 1a. Therefore descriptive belief (perception) measures can be very precise but can never be highly generalizable beyond artefacts of a similar type with similar qualities and affordances.
2. Descriptive beliefs are usually held with a high certainty (Fishbein & Ajzen, 1975).
 - 2a. Therefore it may be reasonable to fix the error term to 0 or a small value on measures of descriptive beliefs.
3. Descriptive beliefs ultimately contribute to attitude or effect towards something, but there may be a series of current and previous inferences based on experience involved; or stated another way, attitude or effect is caused by descriptive beliefs plus previous inferences based on knowledge and experience (Fishbein & Ajzen, 1975).

- 3a. Therefore perceptions cause attitudes not vice versa (although there is a feedback loop into further inferential beliefs). Therefore an affect cannot be modelled as a latent that causes a perception, except in a longitudinal model with a feedback loop. The correct causal structure is from perception to affect.
4. Attitudes (affect) are both more generalised and more conceptually inclusive than perceptions.
- 4a. Therefore measurement of attitudes is less reliable than measurement of descriptive beliefs, since there are more potential unmeasured causes (experiential residue) (Fishbein & Ajzen, 1975).

Figure 3-2: *The formation of attitudes towards websites*

The implications of these principles are as follows:

1. Measurement of perceptions about an artefact (e.g. a website) will be artefact-specific and cannot aim for a high level of generalisability.
2. Affect towards to a website is caused by inferential beliefs held by the person.
3. Affect is a more highly generalised concept than perception and may be meaningfully applied to classes of things (e.g. technology), as well as to a specific instances of technology such as a website.

4. However, the pursuit of generalisability and parsimony in theory has tended to lead to a lack of precision in measuring the constructs that contribute to the formation of affect towards website service quality.
5. One of the roles of IS theory is to explain the process of the formation of affect towards technology artefacts.
6. This can be done with greater flexibility and precision if perceptual (descriptive) beliefs are clearly separated from affects (attitudes), and relevant causal links hypothesised.
7. In practice, this will result in a greater number of hypothesised variables, as many individual indicators are more correctly conceptualised as concepts representing descriptive beliefs (perceptions).
8. Measures of affect (for example usefulness or ease of use) are generalizable, but their meaning is not the same in different theory nets.

These principles and implications will be revisited in the sections on measurement, and the detailed development of the research models.

3.3.2 A Taxonomy of information systems theory

The role of theory in the field of information systems has been controversial almost since the beginning, when in 1980; Keen's highly cited paper laid out some of the issues facing the then new discipline (Keen, 1980). Since 1980, the discipline has flourished but in a state of diversity both with respect to problems addressed and the theories used. The lack of "core theory"; the flourishing of a multitude "frameworks" rather than theories; and the tendency to import theory from reference disciplines, have been advanced as evidence of lack of disciplinary progress (Weber, 1997). The extent to which this is really a problem has been extensively debated (for example, Baskerville & Myers, 2002; Lyytinen & King, 2004), but it is quite widely agreed that the body of "theory" unique to information systems, that is, theory demonstrating accuracy, consistency, scope and simplicity, is fairly limited. It has also been suggested that information systems researchers frequently make exaggerated claims of generalizability at the expense of accuracy, and sacrifice accuracy and salience to achieve parsimony (Seddon & Lyytinen, 2008).

Although the state of theory, or lack thereof, in Information Systems has been lamented, relatively few IS researchers have examined the nature and role of theory in IS research. Recently, influential work by Gregor has addressed this gap. In this section, we present Gregor's taxonomy of information systems theory. This has been published in various forums, including MISQ (Gregor, 2006). We use the longer, e-book version of this taxonomy, available online (Gregor, 2004). Gregor identifies five purposes for information systems theory.

Type I – Theory for analysing and describing. Descriptive theories analyse “what-is”. They describe or classify dimensions and characteristics by summarising the commonalities found over a number of discrete observations. A “naming” theory is simply a description of the dimensions of characteristics. A classification theory goes further, and specifies the structural inter-relationships between these dimensions or characteristics. It is difficult to test a descriptive theory, so other means must be found for evaluating it. It is suggested that the classification system should be useful for aiding analysis, meaningful, natural, logical, and complete (no important elements should be left out). Classification systems can be revised as new entities come to light, or new ways of grouping or naming categories are devised.

Type II – Theory for understanding: These theories that explain ‘how’ and ‘why’ something occurred. Two main types are identified. The first is described as “theory as enlightenment”, which aims to create an original set of categories and assumptions that clear away conventional notions and offer the opportunity for fresh and exciting insights. Structuration theory is offered as an example of this type of theory. The other type of “theory for understanding’ is described as “conjectures” developed from a study of how and why things happened in a real-world situation. Examples can include case studies, surveys and interpretive field studies. These conjectures can form the basis of subsequent theory development. Once again, there is little emphasis on testable prediction, and generalisations need to be made cautiously. This type of theory can be evaluated by its ability to explain something that was poorly or imperfectly understood beforehand, and whether new and interesting insights are provided.

Type III – Theory for predicting. This type of theory explains “what will be”. Outcomes are predicted from a set of explanatory factors, without necessarily understanding the connections between the dependent and independent variables. This type of theory is often associated with statistical techniques such as correlation or regression. The advantages of this type of theory are that statistical methods are well developed for testing the significance of the relationships and the degree of generality with which statements can be made. However, a key limitation is that the existence of correlation does not necessarily imply a causal relationship. These types of theories are most useful if we are able to understand why the two variables are related. Gregor notes that there are few examples of this type of theory in information systems.

Type IV – Theory for predicting and explaining. This type of theory includes both prediction and understanding of underlying causes, and is often considered to be the “real” domain of scientific theory, as expressed by Dubin, and others. This type of theory says “what is”, “how”, “why”, and “what will be”. It is often considered that Information Systems does not have any well-developed theory of this type.

Type V – Theory for Design and Action. This is an emerging area of theory in information systems, widely known as “design science”, aimed at information “how to do” something. There is limited agreement as to the nature of this kind of theory, its contribution to the academic field of information systems, or how to evaluate it.

3.3.3 The quality of service quality theories

Despite the relatively modest predictive claims, service quality research generally follows a hypothetico-deductive model of science enquiry. “*Testing a theory consists of deriving from it consequences that can then be compared with observations and experimental results. If the results agree with the predictions, the theory is inductively confirmed; if the results disagree, the theory is disconfirmed*” (Curd & Cover, 1998, p. 1298). The majority of research papers written about POLSQ assume that truth claims about perceptions of service quality could be verified by quantitative empirical evidence. Of these, most also offered empirical evidence in support of their truth claims, or at least suggested that the outcomes of their research could be used as a basis for further empirical testing (for example, Alzola & Robaina, 2005).

According to Gregor's taxonomy of IS theories, most online service quality research falls into the category of Type 1, descriptive theories that analyse "what-is". Gregor describes these theories as describing or classifying dimensions and characteristics by summarising the commonalities found over a number of discrete observations. This is an extremely good description of most service quality research in electronic commerce, which typically uses factor analysis to identify commonalities in measured variables, i.e. uses data driven methodologies, rather than theory driven ones. Alzola and Robaina conduct a literature review of eighteen scales designed to measure quality in electronic commerce developed between 2000 and 2005. All of them use either exploratory or confirmatory factor analysis and all of them discover a different factor set (Alzola & Robaina, 2005).

According to Dubin's discussion of the components of social science theory, many of these studies can be found lacking from a theoretical perspective (Dubin, 1978). The fact that there is little agreement on the set of constructs and indicators that make up online service quality means that the "units" of the theory are not clearly defined. The fact that many studies use factor analysis means that they do not meet the criteria of falsifiability, as factor analysis is descriptive not predictive.

We suggest that there is evidence that a considerable body of the research into on-line service quality is currently in the "puzzle solving" stage of normal science. The research we examined in the information systems and electronic commerce domain was generally assumed to be "progressive", in that it was adding to or refining our understanding of perceived service quality based on existing knowledge created within the marketing field. We identified at least forty papers that included a discussion of ServQual and subsequent service quality studies in the literature review.

Of these, many fit Kuhn's definition of "puzzle solving", in particular, extending the paradigm to new areas and types of phenomena. In fact, there was a somewhat formulaic approach to many of the reviews of extant service quality literature. The discourse and known issues within the ServQual paradigm were acknowledged, and authors gave summaries of the known issues. Following this, the authors would then conclude that despite some known issues, ServQual is a valid starting point for their research. They then proceed to "puzzle solve" by proposing extensions, modifications,

or reinterpretations of ServQual in a new business domain such as web portals (Kuo et al., 2005), and e-government (Tan, Benbasat, & Cenfetelli, 2008).

There is no clear agreement on the ontology of psychological constructs in service quality research. Many researchers appear to assume a realist ontology (dimensions of POLSQ really exist and can be discovered by research); following Fishbein and Azjen who argue for the existence of psychological phenomena such as attitudes and perceptions. An alternative view; what Kuhn would describe as “theory dependent meaning” is advanced by Grapentine in his reflection of the state of service quality research (Grapentine, 1998). Grapentine suggests the entire phenomenon we call perceived service quality is constructed by research, and has no independent meaning apart from that ascribed to it by researchers.

“The field of marketing needs to develop and agree on a formal language system to define its conceptual definitions; otherwise debates such as the one between Parasuraman and Teas will never be resolved...Marketing is not a hard science...Marketing constructs and terms such as ‘attitude’, ‘perception’, ‘satisfaction’, and ‘loyalty’ do not have objective reality. They cannot be observed by marketing researchers who want to measure them, or by consumers whose minds create them. The reason a construct is called a construct is because marketing scholars and practitioners construct them....Often we make the mistake and think they really exist, a problem called reification.” (Grapentine, 1998, p. 17)

Another possibility is that there is some kind of middle ground between these positions. Some internal psychological constructs do have an independent ontology; some constructs used by service quality researchers may in fact have meaning beyond those whose “minds create them.” Others may be merely convenient labels for a group of indicators that happen to co-vary. The challenge is to define the first type and not become distracted by the second type.

3.3.4 Issues with information systems and service quality theories

3.3.4.1 *Interpretational confounding and meaning variance*

The problem of uttering the same words with different meanings, “so any logical comparison of their utterances is precluded” (Curd & Cover, 1998, p. 222) is unfortunately widespread in information systems. Service quality, and online service quality research are rife with it. In terms of its relationship to other constructs, service

quality sometimes encompasses notions such as usability or content quality and sometimes is considered as a separate factor at the same level as usability or content quality. For example, one of the changes made by DeLone and Mclean when they updated their IS success model was to add service quality as a factor at the same level as system quality and usability (DeLone & McLean, 2003). By contrast, Barnes and Vidgen include usability as a sub-dimension of e-service quality (Barnes & Vidgen, 2002). Perceived service quality is sometimes seen as an antecedent of use and sometimes as a consequence. For example some authors include service quality, information quality and systems quality as separate constructs that all act as antecedents of use and user satisfaction (Cheung & Lee, 2005; Pitt et al., 1995), while Barnes and Vidgen specify information quality as a sub-dimension of POLSQ (Barnes & Vidgen, 2002).

It is also common for the assignment of items to constructs to vary, thereby changing the meaning of the construct. For example, in their 1991 paper, the ServQual authors found different patterns in the organisations they surveyed. Across three organisations, an item measuring the degree to which organisations “understand the specific needs of their customers” was sometimes associated with responsiveness, sometimes with assurance, sometimes with empathy, and sometimes with more than one of these constructs (Parasuraman et al., 1991).

Due to the diversity of social phenomena, it is often expected that rather than being expressed in universal terms, like physical laws, social theories are expressed in statistical terms, as a likelihood; that does not necessarily explain every instance. Put another way “*law applies to **all** cases, a model to **ideal**, that is, to **particular** cases*” (Boudon, 1991, p. 68). This tends to result in a narrower scope of the explanatory power, or a “lower-order generality” for social science theory than for theories in the physical sciences (Gregor, 2006). In information systems research, this tension is evident. We regularly observe research papers which include disclaimers that the model has only been tested in a specific context (for example, Lee & Lin, 2005). Other researchers advocate that their survey instrument should be adapted for different contexts, for example, a critical review of ServQual studies suggested that the instrument was only valid for domains in which it has already been tested: “*recall that our review has suggested that the ServQual dimensions are likely to be industry*

specific” (Asubonteng, MacCleary, & Swan, 1996, p. 75) . The danger is that there may be a limit to the extent to which an instrument can be adapted and still be claimed to be the same instrument.

3.3.4.2 *Parsimony*

The diversity of social phenomena, and the tendency for reality and theory to have varying degrees of correspondence, tends to also create issues with theory parsimony. The quest for parsimony is often the cause of a gap between models and the complexities of the real world.

“[Although social scientists] strive for theories that are simultaneously parsimonious, highly general, and therefore applicable to a wide range of phenomena, yet precise enough to imply rejectable hypotheses, it does not appear possible...to achieve simultaneously all three of these ideal characteristics...my own position is that of the three, parsimony is the most expendable.” (Blalock, 1982, p. 28).

In a recent study, (Tate et al., 2008a) illustrate this point in qualitative service quality research which was conducted with continuing users of a website. Previous studies had found self-efficacy to be influential in forming expectations that led to perceptions of service quality (Hsu et al., 2004). However a qualitative study found that the self-efficacy construct was likely more complex than previous authors had suggested; and included specific knowledge and experience with the service itself, Internet self-efficacy, and general knowledge of the business domain in which the tool was operating. This specificity and preciseness is lost in a single parsimonious self-efficacy construct (Tate et al., 2008a).

3.3.4.3 *Unjustified normal science assumptions*

Overall, research in service quality and online service quality tends to make Kuhnian normal science assumptions, which are not justifiable. The relatively weak status of the gaps model and ServQual as a scientific theory (Tate & Evermann, 2009a), the issues with accuracy, consistency and meaning variance mean that there is not a cumulative research tradition, which extends existing theory to solve new puzzles; although literature reviews in service quality research papers are often presented as though that is the case (for example, Alzola & Robaina, 2005). There is a widespread

tendency to inventory the propositions of the field, and try to give order to the totality, which commits the sin of adding non-comparable things (Dubin, 1978).

We suggest that ServQual has been accorded de facto status as a theory that may not be deserved and subsequent research into online service quality has “built” on it based on normal science assumptions. eQual is part of this tradition; “puzzle-solving” in the area of e-services, based on Kuhnian assumptions about the ServQual tradition of research. A more detailed examination of issues with the ServQual tradition is provided in Tate and Evermann (2009).

The reality of service quality research is much closer to a social constructivist approach than a normal science or even a critical realist approach. In many cases, the constructs reported in studies do not stand up to scrutiny either as meaningful theory units or as real entities, although they are often tacitly assumed by the study authors to be both. Instead, we suggest that researchers are actively involved in the construction of constructs such as POLSQ. Multiple studies report sets of definitions, observable indicators, and relationships that are fully or partly unique to that study and not interchangeable with those used by other researchers. We add to this the fact that replication studies, which are necessary to the activity of negotiating, are widely disparaged in the IS field. Moreover, the difficulty of recruiting an adequate sample, evidenced by the low response rates, makes it impractical to use comparative or “replicate and extend” studies in many cases.

3.3.4.4 The theoretical provenance of eQual

The research provenance of eQual includes services marketing literature, primarily ServQual, supported by a literature review and focus groups (Barnes & Vidgen, 2002). No specific theory net is cited. So to what extent is eQual based on theory? Can ServQual lay any claims to form part of a Type III or IV theory? The theory net that was originally proposed for the ServQual instrument is the gaps model (Parasuraman et al., 1985) shown as Figure 3.3.

Figure 3-3: *The Gaps Model (Parasuraman et al., 1985)*

It is hypothesised that gaps one through four cause gap five. The ServQual instrument was developed to measure gap five. The hypothesis that gap 5 is caused by gaps 1 through 4 is explicitly stated by the original ServQual authors (Parasuraman et al., 1985), who even made an attempt to test it (Parasuraman et al., 1991). This attempt was inconclusive. So to the extent that eQual builds on existing theory, that theory is inconclusive and unproven. The ServQual instrument, (and subsequent service quality measurement instruments such as eQual that look back to ServQual for their provenance) is frequently disconnected from its original theory net, and has taken on something of a life of its own, as a stand-alone measurement instrument.

eQual, while not as widely adopted as ServQual, has achieved a reasonable level of uptake, and has been used to measure service quality in a range of domains, including auction sites (Barnes & Vidgen, 2001), WAP news sites (Barnes, Liu, & Vidgen, 2001), and Russian B2C cosmetic sites (Durova & Amin, 2009). However, exactly what is being measured, and what truth claims can be made on the basis of those measures, is not clear, since the theoretical status of the instrument is not clearly established.

We apply Kuhnian principles and quality dimensions to evaluate the eQual instrument and the theory that tacitly underlies it. A summary is provided in table 3.1.

Table 3.1: *Kuhnian analysis of eQual*

Quality dimension	Evaluation of eQual
Accuracy Does it agree with experiments and observations that have already been made and solve important puzzles (problems)?	No. There are many different versions of the dimensions of online service quality from different authors – for a more extended discussion refer to (Alzola & Robaina, 2005).
Consistency Is it consistent, free from logical contradiction; and externally consistent in that it does not contradict any other accepted theory?	No. It contradicts the widely cited IS success model. DeLone and McLean specify information quality, system quality and service quality as independent antecedents of intention to use and use satisfaction (DeLone & McLean, 2003), while Barnes and Vidgen include information quality and aspects of system quality as sub-dimensions of online service quality (Barnes & Vidgen, 2002).
Scope Does it have consequences that go far beyond the set of observations, laws, and sub-theories it was originally designed to explain?	To some degree. The instrument has proved useful for benchmarking online service quality in a variety of business and public sector domains For example, the instrument has been used to benchmark online bookshops (Barnes, 2001), and auction sites (Barnes & Vidgen, 2001) However, it does not appear to have any predictive value, and no claims for predictive value have been made by the authors.
Simplicity and parsimony Does it bring order to multiple, apparently different phenomena, and contain as few assumptions as possible?	Yes. The factor analysis is a reduction technique aimed at improving parsimony and simplicity. “We need to move beyond the scores and indices on individual questions towards a set of meaningful and reliable sub-groupings. ...a factor analysis was conducted on the data” (Barnes & Vidgen, 2002, p. 121). “We were able to build a profile of the qualities of an individual Web site that makes it easy to compare with its rivals.” (Barnes & Vidgen, 2002, p. 122).
Meaning variance Are the same words used, but the words have different meanings?	Yes. For example “usability” is used to confusingly describe both a second order construct which encompasses both “usability” and “design” as sub-constructs, and for a first order construct titled “usability” which is similar to ease of use.

Based on this analysis, we can see that the eQual instrument does not possess the range of characteristics to constitute a scientific theory, according to criteria defined by influential philosophers of science. The main strength of eQual is as a data reduction tool, culminating in summative measures of service quality. Dubin notes that summative measures are “legion” in social science and often imply much more than they state (Dubin, 1978). The laws of interaction between theory components are not well defined, and many of the explicit or implicit hypotheses have never been tested or are difficult to test. This study addresses this issue for online service quality research by making the hypotheses implicit in eQual explicit, and testing them using structural equation modelling.

3.3.5 Summary

In summary, quantitative information systems research, including research on online service quality tends to adopt a rhetoric of realism, positivism or post-positivism, normal science assumptions, and valuing accuracy, consistency, generalisability, and parsimony in theory. Research on attitudes and perceptions towards technology is nominally based in social psychology, but often fails to adequately distinguish between attitudes and beliefs at different levels of generalisability and conceptual inclusiveness.

Online service quality research is dominated by “type I” theories (theory for analysing and describing). eQual fits this description, in that the main contribution is to *“describe or classify dimensions and characteristics by summarising the commonalities found over a number of discrete observations”* (Gregor, 2004, p. 6). To the extent that current online service quality research can be considered as theory, it frequently includes fallacies such as widespread meaning variance.

An excessive focus on parsimony and generalisability tends to lead to meaning variance in key constructs when they are applied in different contexts; in different web quality studies, constructs such as “usability” are represented by different items with different strengths of relationship, which implies a different meaning for the construct.

A related issue is that items with different levels of conceptual inclusiveness are often grouped together, for example, “ease of use” has high conceptual inclusiveness, while “ease of navigation” is more specific. This leads to measurement issues, which will be discussed in more detail in subsequent chapters.

Literature reviews in the online service quality area frequently “inventory” the knowledge in the area, and commit the sin of adding non-comparable things (Dubin, 1978). This leads to theoretical confusion, which flows into measurement and modelling issues.

We propose by way of alternative, that rather than striving for excessive parsimony and generalisability, with the risk of meaning variance and varying levels of conceptual inclusiveness, we explore the important distinction between descriptive and inferential beliefs. By definition, descriptive beliefs must vary with the cognitive affordances of the specific technology (an experienced user is highly unlikely to have a descriptive belief about an affordance to technology does not, in fact, possess). These are likely antecedent to inferential beliefs, are more generalised but less reliable, and more likely to have unmeasured causes. In later sections of this study, we offer some alternative measurement approaches that address some of these issues.

3.4 Issues in Measurement

3.4.1 Introduction

Quantitative research in social science tends to draw heavily on psychometrics as a reference discipline for the measurement of human perceptions, attitudes and behaviours. Information Systems is no exception. A cynic might say that this importing is often done uncritically and superficially. While information systems researchers are conducting exploratory factor analysis using SPSS, selecting the default settings of Principle Factor Analysis, and Varimax rotation, to reveal latent variables hidden in their data; in the meantime measurement theorists are debating whether latent variables actually exist.

In this section we offer a brief summary of some key aspects of measurement theory to frame some of the current debates in the psychometrics, measurement theory, and SEM (Structural equation Modelling) communities that are relevant to this research.

3.4.2 Measurement Theory

The foundations of structural equation modelling are in measurement theory. We briefly review the background to the development of SEMs, with a particular focus on why they represent the most appropriate form of model for this research design.

The beginnings of classical measurement theory are usually considered to be in the 1930's, with founding work by Karl Pearson 1857-1936, who is sometimes known as "The founder of Statistics", who pioneered techniques that are still used today, such as the chi-squared goodness of fit test (Allen & Yen, 1979). The widespread adoption of measurement theory for psychological testing and pioneering work on factor analysis and test reliability is usually credited to Charles Spearman (1863-1945). The journal *Psychometrika* was founded in 1935 (Allen & Yen, 1979).

3.4.2.1 *Latent variable theory*

A popular approach for measuring psychological phenomena is latent variable theory. A latent variable is an internal, unobserved and unobservable state (for example, an attitude) which cannot be measured directly, but which is assumed to cause responses that can be measured (for example, responses to questions on a survey). The basic assumption of latent variable theory is that a survey response on an item will be determined by an examinee's standing on one or a set of (unobserved and unobservable) latent traits. For example, a respondent that has (internally) a positive view of the ease of use of a technology (which is unobservable) will give a high score on a survey question such as "I find this technology easy to use". The survey question response can be observed and is assumed to be caused by the internal state. So latent variable models are a representation of how a hypothesised, internal latent variable influences a subject's performance on each survey item response (Allen & Yen, 1979).

In measurement theory terms, structural equation models (SEMs) are a statistical implementation of latent variable theory. SEMs are widely used for psychological

studies, and are therefore very suitable for studies of user attitudes and perceptions towards technology.

3.4.2.2 Structural equation modelling

Structural equation modelling is a statistical methodology based on latent variable theory. Structural equation modelling (SEM) is not a single technique, but a family of related procedures, with a number of important characteristics in common (Kline, 2005). The key characteristics of SEM models are:

1. They are confirmatory and support an hypothesis testing approach to the phenomenon being investigated. SEMs require an a-prior statistical model (the model must be specified in advance). However, the ability to respecify models means that they can be used for a combination of exploratory and confirmatory purposes.
2. Typically, this theory represents “causal” processes that generate responses on multiple variables (Bentler, 1988).
3. The causal processes under study are represented by a series of regression equations.
4. The entire model (system of variables and hypothesised relationships) can be tested simultaneously to determine the extent to which it is consistent with the data.
5. SEMs explicitly specify the relationships between latent (unobserved) variables and observed variables (such as a survey question response).
6. The basic statistic in SEMs is covariance. Within the information systems community, partial least squares models are sometimes also described as SEMs, but this use of the term is an exception within the wider SEM community (Rouse & Corbitt, 2008).

3.4.3 Factor analytic approaches: Overview of formative and reflective variables

SEMs provide a method for modelling the causality between latent variables and their indicators; these are described as reflective indicators. It is also possible to model relationships in the other direction, from indicators to latent variables; these are

described as formative indicators. Recently, there has been considerable interest in research literature in the use of formative variables, for example, a paper on this issue was recently published in MISQ (Petter, Straub, & Rai, 2007).

Psychometric theory assumes that the response by a subject on a reflective indicator is caused by the psychological state of the subject. The state of the subject is captured in the unobservable latent variable. Reflective indicators are caused by the underlying latent variable, which is considered to give rise to the indicators. If the value of the underlying latent variable changes, the values of the reflective indicators will also change. Formative indicators, by comparison, cause the latent variable they are measuring (Diamontopoulos & Winklhofer, 2001). In that sense, the terms "indicate" and "measure" are ill-chosen for formative indicators.

In this section, we discuss both reflective and formative indicators, but we concentrate in more detail in explicating aspects of specifying, modelling and analysing formative indicators. This is because the use of reflective indicators is the norm, and is supported by an extensive literature. By contrast, the use of formative indicators has a more limited and more contemporary literature, and some aspects relating to the use of formative indicators in structural equation models are still being debated. These issues are highly pertinent to the research question, and are therefore covered in more detail.

The differences between formative and reflective indicators are summarised in Table 3.2.

Table 3.2: *The differences between formative and reflective indicators*

Reflective indicators	Formative indicators
Indicators interchangeable and reflect the same latent construct	Indicators not interchangeable – omitting an item is omitting part of the construct. Indicators do not necessarily share a common theme
It is often stated that the indicators reflect the underlying latent construct and are selected from a large pool of potential indicators (Jarvis et al.), or that items are “chosen randomly from the universe of items relating to the construct of interest” (DeVellis, 1991, p. 55).	The indicators are the set of defining characteristics of the construct (Jarvis, MacKenzie, & Podsakoff, 2003). A complete census of indicators is required. Items must cover the entire set of antecedents of the latent variable (DeVellis, 1991, p. 55).
Eliminating or changing an indicator will not change the meaning of the latent variable (Jarvis et al., 2003).	Eliminating or changing an indicator may change the conceptual domain of the construct (Jarvis et al., 2003).
Causality is from construct to measure (Jarvis et al., 2003).	Causality is from measure to construct
Correlations explained by the measurement model	Correlations not explained by the measurement model. Indicators are exogenously determined
Construct possesses “surplus” meaning in that it has independent ontological existence that “causes” the response on the indicators (Jarvis et al., 2003).	Construct possesses “surplus” meaning in the form of an error term that represents the portion not caused by the indicators (Jarvis et al., 2003).
A change in the value of one indicator should be accompanied by a change in the value of all indicators (Jarvis et al., 2003). Specific pattern of signs and magnitude should characterise correlation amongst indicators.	Indicators may vary independently (Jarvis et al., 2003). There is no reasons to expect a specific pattern of signs and magnitude of correlation – internal consistency is of minimal importance – even negatively related variables can be meaningful indicators of a construct They should be maximally independent, or we will get multi-co linearity issues
Indicators have error terms	Indicators do not have error terms; error variance is represented by the disturbance, which is uncorrelated with the X’s. The error (disturbance) is at the construct level, not the item level

3.4.4 Statistical differences

The statistical representation of formative and reflective variables needs to be reviewed briefly, since different statistical models imply different ontological and epistemological assumptions.

3.4.4.1 *Reflective indicators*

Reflective indicators are the most commonly used in structural equation modelling. Reflective measures are assumed, as we discussed above, to all measure the same construct in slightly different ways. Therefore we expect them to be correlated, and we expect changes on the latent to be reflected in all its indicators, and changes in one reflective indicator to be similarly reflected in the others (Figure 3.4).

Figure 3-4: *Reflective indicators*

The model in Figure 3.4 can be represented as follows, where the variance in each indicator is shown as a (linear) function of the underlying latent variable plus error.

$$x_1 = \lambda_1 P + \varepsilon_1$$

$$x_2 = \lambda_2 P + \varepsilon_2$$

$$x_3 = \lambda_3 P + \varepsilon_3$$

Where λ_i is the expected effect of P on x_i and ε_i is the measurement error for the i th indicator ($i = 1, 2, 3$). It is typically assumed that $\text{Cov}(P, \varepsilon_i) = 0$, $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$, where $i \neq j$ and $E(\varepsilon_i) = 0$.

Implications of a reflective model are that $\text{Cov}(x_i, x_j) > 0$ and that $\text{Cov}(x_i, x_j) > 0$ for $i \neq j$ (Bollen, 1984). This means high correlations among the indicators enhance internal consistency (Bollen & Lennox, 1991), and that dimensionality, reliability, and convergent/discriminant validity assessments are all meaningful (Diamontopoulos, 1996).

3.4.4.2 Formative indicators

Statistically, there are two slightly different ways of expressing the equations for formative indicator models, where γ_1 is a parameter reflecting the contribution of x_1 to the latent variable η , and ζ is the error term.

$$(1) \eta = \gamma_1 x_1 + \gamma_2 x_2 + \dots \gamma_n x_n$$

$$(2) \eta = \gamma_1 x_1 + \gamma_2 x_2 + \dots \gamma_n x_n + \zeta$$

Although both of these are described as formative variables, they are not the same. Equation 1 is a simple linear regression equation with no error term. Mathematically, it is a sum or composite of its indicators. Ontologically η is also defined as the sum of its measures. This is consistent with a principal components model of factor analysis, and with the partial least squares approach to statistical modelling. In theoretical terms η can be interpreted as a dependent variable, but this explanation assumes it is fully explained by its formative measures and is “no more than” the sum of its measures.

Equation 2 includes a disturbance term. In this equation, unless all y 's are equal to 1, and the error term is equal to 0, a linear composite based on the sum of equivalently weighted indicators (equation 1) will not be equivalent to the latent variable in equation 2. In the second type of equation η can be explained as a dependent variable with an error term. This is more consistent with a claim of a realist ontology, and allows the explanation that η is “caused by” its indicators because the disturbance term allows an explanation of the causes that are not fully explained by the indicators (depending on how you treat the explanation of error terms) (Borsboom et al., 2003). Co-variance-based SEMs permit either treatment of formative indicators.

3.4.5 Model fit

Assessing model “fit” in co-variance-based SEMs is the process of comparing an observed covariance matrix (based on the actual covariances of the observed data) with the implied covariance of the model, and evaluating the degree of difference between the two covariance matrices (Evermann & Tate, 2009a)

In a SEM model, the researcher constrains some of the parameters (makes an assumption or approximation of their value), and leaves others free to vary. This, and the number of unique covariances, determines the degrees of freedom of the model. In statistics, the number of degrees of freedom (df) is the number of independent pieces of data being used to make a calculation. The df can be viewed as the number of independent parameters available to fit a model to data. Generally, the more parameters you have, the more accurate your fit will be. However, for each estimate made in a calculation, you remove one degree of freedom. This is because each assumption or approximation you make puts one more restriction on how many parameters are used to generate the model.

Model fit is subject to extensive discussion, like many other areas in SEM modelling literature, most recently in SEMNET, and in a special issue of the journal “Personality and Individual Differences” in 2007 (volume 42 issue 5). Understandably, researchers who want to use SEMs without becoming advanced SEM theorists, seek best practice guidelines that can be used to evaluate model fit.

Key issues in the debate have been the use of the χ^2 test versus approximate fit indices, the degree to which the χ^2 test is sensitive to sample size, the degree to which “mis-fit” and “mis-specification” are related, and whether there are “golden rules” for fit that can be applied by researchers and editors. We will address these one at a time.

There appears to be a consensus among SEM specialists that the χ^2 statistic should always be reported, and is agreed to be the only “true” goodness of fit test, in the sense that it compares the model implied covariance with the actual observed sample covariances to determine whether the discrepancies (residuals) are greater than would have occurred by chance alone. Very little discrepancy suggests a good-fitting model (Barret, 2007; Evermann & Tate, 2009a; Marsh, Hau, & Wen, 2004). However, if the χ^2 test were perfect, there would be no requirement for other fit indices. It is suggested that the likelihood of a “perfectly-fitting” model is extremely low, particularly in models with a large number of degrees of freedom; given that any model is only an approximation of reality (Bentler, 2007; Goffin, 2007; Millsap, 2007). Since all models are to some degree approximations of reality, then approximate fit indices may be appropriate.

It also claimed that the χ^2 statistic is sensitive to sample size, and in particular, that very large sample sizes tend to result in a high χ^2 that would result in the rejection of good-fitting models (Barret, 2007; Byrne, 2001; Kline, 2005). However, a study by Bollen, cited 93 times, suggests this view is simplistic, and only applies for all forms of mis-specified models (Bollen, 1990).

Hayduk et al. are critical of approximate fit indices on the grounds that it can be demonstrated that there is very little relationship between degree of mis-fit (the size of the covariance residuals) and degree of mis-specification (Hayduk, Cummings, Boadu, Pazderka-Robinson, & Boulianne, 2007). Relatively low residuals can be associated with a badly mis-specified model, and vice versa (Fan, Thompson, & Wang, 1999; Marsh et al., 2004).

Despite this, there are still many proponents of approximate fit indices. Over time, two things have happened. First, the guidelines for “acceptable fit” have tended to become more restrictive, and second, guidelines have tended to be adopted as “golden rules” rules of model fit, and have become separated from the contexts and caveats specified by the original authors.

Research by Hu and Bentler appeared to support adopting .95 rather than .9 as the criteria for good fit for many approximate fit indices and this has been widely adopted by researchers (Hu & Bentler, 1999). Many researchers use this research as a “golden rule” that has broad generalisability across different conditions and sample sizes, when in fact most criteria for evaluating fit indices are rules of thumb. A high goodness of fit (GOF) is often treated as though it is both necessary and sufficient. However Marsh et al. recommend caution in the interpretation of Hu and Bentler’s findings, suggesting that more stringent criteria may result in the rejection of acceptable models and recommend that more attention is paid to Hu and Bentler’s own caveats (Marsh et al., 2004). They suggest there are two possible approaches to interpreting fit indices, the normed reference approach, and the criterion reference approach.

Normed reference approach (intuition and received wisdom) mandates “higher fit indexes are better”. Marsh et al. suggest that it is very difficult to get good fit indexes when analysis is done at the item level and there are multiple factors (say 5-10) each measured with say 5-10 items, resulting in an instrument with at least 50 items (Marsh et al., 2004).

Responses to this discussion on the SEMNET forum mostly argued that more stringent cut-off criteria excluded many existing models and instruments. Some suggested that factors should not be represented by more than 2-3 items, or maybe even a single item (see for example, post by Les Hayduk, 17 February 2009). Marsh et al. suggest that a smaller number of items were likely to improve the fit indices, but result in a measure with poorer construct validity (Marsh et al., 2004). Marsh observed that researchers sometimes resort to dubious practices such as smaller sample sizes or a lower number of items to improve their fit metrics (Marsh et al., 2004).

We suggest that the problem with fit indices and larger numbers of items occurs when the items are not fully reflective of the same factor. When this is the case, with more items, you potentially have more interdependencies between the observed variables than can be explained by a single common cause. This is the reason why 5-10 factors with 5-10 items tend to yield bad fit. There may be structures between the items that the factors don't explain as they are too generalised. In the models reported in this study, models with fewer indicators were better-fitting.

Criterion reference approach (using some external criterion of model fit). Marsh et al. (Marsh et al., 2004) created nested models that systematically varied in terms of mis-specification (over or under parameterisation), and used them to apply different fit indices. The authors concluded that most GOF indices were better able to distinguish between the degrees of fit of alternative models of the same data than to establish criteria of acceptable overall model fit.

The authors report on Hu and Bentler's (Hu & Bentler, 1999) study of mis-specified models, to see how frequently type 1 errors (rejecting the null hypothesis when it is true – considering a model good-fitting when it is not) or type 2 errors (accepting the null hypothesis when it is false – rejecting a model as poorly fitting when it is acceptable) occurred. This is summarised in Table 3.3.

Table 3.3: *Impact of mis-specified models on fit indices-1*

Fit Index	Simple mis-specified	Complex mis-specified	Comments
TLI	Acceptable (pop est of TLI >.95)	Not acceptable (pop est of TLI <.95)	Large % of type 1 errors and lower type 2 errors than other fit indices.
RNI		Only most mis-specified complex model was unacceptable	
SRMR	Higher rejection rates	Lower rejection rates	Sensitive to different sorts of mis-specification.

They experimented with “simple” mis-specification (one or two factor loadings fixed to zero), and “complex” mis-specified models (with one or two cross-loading items fixed to zero). However, the probability of correctly rejecting a false model decreased with sample size rather than increasing. The reason for this is that some of their “mis-specified” models should be considered acceptable, and were only marginally mis-specified. In their study, ML chi-squared tests outperformed all seven GOF indices in correctly rejecting mis-specified models.

Overall, the authors concluded that fit indices were better at distinguishing degrees of mis-specification than providing overall guidelines for goodness of fit.

Another study by Fan et al., summarised in Table 3.4, also addresses the effect of mis-specification on fit indices (Fan et al., 1999). These authors concluded that too much work has been concentrated on true models, and found that the estimation method and sample size influenced fit indices.

Table 3.4: *Impact of mis-specified models on fit indices-2*

Indicator	Comment
RMSEA	Sensitive to model mis-specification. Less influenced by sample size. Less influenced by estimation method.
CI	Sensitive to model mis-specification. Less influenced by sample size.
GFI, AGFI	Sensitive to model mis-specification. Downwards bias with smaller sample size. Less influenced by estimation method.
CFI	Less sensitive to model mis-specification. Found to be influenced by sample size in a study by Tanaka (1993) but not in this study.
NNFI, NFI, RH01, Delta2	Less sensitive to model mis-specification.

The current consensus from the Personality and Individual Differences special issue in 2007 (42, 5) appears to be a test of overall model fit, such as the χ^2 statistic should always be reported, but with the understanding that it is a fairly “blunt” measure, especially for large and complex models. In addition, other measures of fit should be reported as appropriate (Bentler, 2007).

3.4.5.1 Assessing the fit of Formative Models

Collinearity

Formative measurement is similar to multiple regression, therefore the stability of the indicator coefficients (the items) is affected by the sample size and the strength of indicator intercorrelations. Excessive collinearity means it is difficult to assess the impact of an individual x (item) on the latent variable. Since each y is a direct structural relation between x and η , the magnitude of the y 's can be interpreted as validity coefficients, so high multi-collinearity makes this problematic. Further, since each item (x) is supposed to contribute directly to the formation of the latent dependent variable η , then an item that is a perfect linear combination of others is redundant.

In formative measure each variable is constructed from a set of indicators so it can be interpreted as a multiple regression model. For example, in a simple model it means that:

$$\eta = y_1 x_1 + y_2 x_2 + \dots y_n x_n + \zeta$$

Where y_1 and y_2 are regression coefficients and the x_1 and x_2 are two formative "indicators". In a path diagram this would look like:

$$x1 \rightarrow \eta \leftarrow x2$$

If $x1$ and $x2$ are collinear, i.e. have non-zero covariance, multiple explanations are possible (because $x1$ might cause $x2$ or vice versa), so there might be indirect effects as well as direct effects, as follows.

$$x1 \rightarrow x2 \rightarrow \eta$$

$$x2 \rightarrow x1 \rightarrow \eta$$

$$x2 \rightarrow x1 \rightarrow \eta \leftarrow x2$$

$$x1 \rightarrow x2 \rightarrow \eta \leftarrow x1$$

Error term

The error term of a formative latent variable indicates the amount of variance that is surplus to that explained by the indicators. A high error term indicates that the domain is not fully captured by the existing indicators (Diamontopoulos, 2006).

Indicator coefficients

The indicator coefficients explain the amount of variance in the formative construct that is explained by the indicator. A high coefficient suggests that the indicator is making a significant contribution to the formative latent variable (Diamontopoulos, 2006).

Heuristics for formative models

In combination, these lead to the following heuristic for formative models, shown in Table 3.5.

Table 3.5: *Heuristics for Formative Models*

Error term	Coefficients	Comments
Small	High	Good-fitting formative model
Small	Some Low	Test co-linearity, some indicators may be redundant. Delete some indicators?
Large	High	Check if the domain is fully covered – may need additional indicators?
Large	Low	Re-conceptualisation required

3.4.6 Measurement: Scale and index construction

3.4.6.1 *Scale construction*

A scale is a set of reflective measurement indicators for a latent (unobserved) variable. A scale is a set of items that are conceptualized as being sampled from an infinite pool of items that reflect the unobserved latent construct that they are measuring. Scale items are considered to have the same common cause, and to conform to the heuristics in table 3.2 for reflective variables.

3.4.6.2 Index construction

An index is a set of formative measurement indicators for an unobserved construct. We have avoided describing the construct as a latent variable, because the meaning of an unobserved variable described by formative indicators requires some further discussion. For the meantime, we will examine the process of developing an index, which is a set of formative measurement indicators.

First, content specification: the domain of the content is extremely important – failure to capture all facets of the construct will exclude part of the construct itself. Items must cover the entire scope of the latent construct. Second, indicator specification: a census of indicators is required. Specification must be inclusive enough to capture the construct's domain of content. Third, indicator collinearity must be considered, and should be avoided.

Scale development places major emphasis on the intercorrelations among the items, focuses on common variance, and emphasizes unidimensionality and internal consistency. Index construction, emphasizes the role of indicators as predictor rather than predicted variables (Diamontopoulos & Siguaw, 2006).

3.4.6.3 Differences in the item sets for scales and indexes

The obvious question that needs to be asked when considering these issues is “Can one use the same item set, but simply reverse the arrows to respecify reflective indicators as formative?” In software packages such as AMOS, this is the work of a few moments, and it is widely considered to be a viable option (Diamontopoulos & Siguaw, 2006, p. 266). However, it does not appear to be a sound approach. In their recent paper, Diamontopoulos and Siguaw examined the implications of reversing the direction of the indicators by following a parallel process of item generation, measure purification, and measure validation for formative and reflective indicators in the same business domain to determine to what extent the item pool would be the same (Diamontopoulos & Siguaw, 2006). They note that literature on both approaches emphasises the need for a comprehensive and inclusive item pool at the beginning. The two approaches diverge when considering measure purification.

High intercorrelations between items are desirable and expected in scale development, but in index construction, items should only be retained so long as they have a distinct influence on the latent variable. High correlation between items suggests unnecessary redundancy. This means that the final set of items following purification is likely to be different and that making incorrect assumptions as to whether the unobserved variable should be measured formatively or reflectively is likely to significantly affect the resulting item pool at the point when the item set is purified. In their study, a comparison of scale and index development for the same construct yielded only two items that were common to both measures after measure purification. They conclude that the different approaches will likely result in different item sets rather than respecifying the same set of items, because the formative indicators are the causes of the construct, and the reflective indicators are resultants of the construct.

There are also obvious problems if we examine this issue causally: It is absurd to think that a person's response on a survey item could be the cause of a psychological state. This does not make much sense under any circumstances, but it becomes even more unlikely if what we are trying to explain in the survey happened in the past – i.e. we are asking subjects about past experiences.

3.4.6.4 Model specification and identification

In order to evaluate a SEM model, it must be identified. This means that estimates for the free parameters can be unambiguously derived from the known covariances. Amongst other criteria, this requires the model to have at least as many observed values to be fitted to the model, as there are free model parameters (to be estimated by the analysis), and every latent variable must be assigned a scale. This is usually not a problem for reflectively specified models, but it can be an issue for formative and mixed models.

3.4.6.5 Formative and mixed models

The correct method for achieving model identification for formative models is the subject of ongoing discussion amongst SEM theorists at present. Although there are some preferred approaches, in my view, many of these have weaknesses, and the question is not adequately resolved. Some of the key issues in the debate are summarised in this section.

There are several ways to handle this. First, each indicator can be correlated to another variable external to the index, and only indicators that are significantly correlated with this additional external variable are retained. This requires a strong theoretical reason why the index item should be correlated to the external variable, and why this correlation should not be the result of a common cause (Diamontopoulos & Winklhofer, 2001; Jarvis et al., 2003).

Second, you can add some reflective indicators to each dependent variable, to give a MIMIC (multiple indicators and multiple causes) model. This resolves some of the statistical issues, but in my view, it causes ontological issues, which are explored later in this chapter, as the different types of measures are based on different ontological assumptions. It is difficult to see how a construct could be formed by an index of formative variables (because then it is just a sum-score based on its indicators, and at the same time be the common cause of reflective variables (because that would presuppose the separate ontological existence of a latent variable).

In the example provided in Figure 3.5, the index indicators x_j , ($j = 1, 2, 3$) are the direct causes of the latent variable η , which in turn is indicated by the reflective measures y_j ($j = 1, 2, \dots, m$).

Figure 3-5: Example of a MIMIC model from (Diamontopoulos & Winklhofer, 2001)

Finally, you can focus on the nomological aspects and link the index to other constructs with which it could be expected to be linked. When adding an index item in this way, you need to establish that the new version functions in a predictable way. This requires that information is gathered for one more construct than the one directly measured by the index item that the additional construct is measured reflectively, and there is some theoretical reason for arguing that a relationship exists between the new index item and the additional construct (Diamontopoulos & Winklhofer, 2001).

In the example in Figure 3.6, reproduced from D&W, (Diamontopoulos & Winklhofer, 2001) the variable η_1 is constructed formatively, and is hypothesised as a direct or indirect cause of η_2 which is measured reflectively. Note that this is statistically the same MIMIC as above (Figure 3.4), as the β_{21} and the $\text{var}(\text{zetaeta2})$ are aliased parameters with $\text{var}(\text{zetaeta1})$

Figure 3-6: *Alternative conceptualisation of the model in Figure 3.4 from (Diamontopoulos & Winklhofer, 2001)*

Edwards and Bagozzi support the third approach, and note that for formative measures the relation between the measure and the construct can be estimated provided the construct is specified as the direct or indirect cause of at least two reflective measures (Edwards & Bagozzi, 2000). The association between a construct and its measures, i.e. the regression coefficients or factor pattern γ_{11} , γ_{12} and γ_{13} , should remain stable regardless of the larger causal

model in which the construct and measures are embedded. If these values vary, then no unique meaning can be assigned to the construct. But, associations of formative measures with their constructs are determined by the relationships between these measures and measures of constructs that are dependent on the construct of interest. Thus, placing it in a different context is likely to change the regression coefficients/factor patterns and therefore the meaning of the construct. Thus, the meaning of the latent construct is as much a function of the dependent variable as it is a function of its indicators, and the relationship between a formative construct and its measures changes with the choice of dependent variable.

Jarvis et al. (2003) note that discussion of validity in the measurement model relies on classical test theory, which assumes that the variation in the scores is a function of the true score, plus error (Jarvis et al., 2003). The underlying latent construct causes the observed variation in the measures.

By contrast, a construct measured formatively may be considered to be a perfect linear combination of its measures (equation 1 in section 3.4.4.2). With a composite latent construct model, the construct is a linear combination of its measures, plus error as shown in equation 2 above (Diamontopoulos & Winklhofer, 2001, p. 202).

3.4.6.6 *The error term*

The “meaning” of the error term in equation 2,

$\eta = y_1, x_1 + y_2, x_2 + \dots y_n, x_n + \zeta$, is at the heart of the issues with formative measures (Diamontopoulos, 2006).

In reflective models, the error term relates to the observed score, and therefore represents measurement error. The observed score equals the true score plus error. By contrast, for formative models, the error terms shows all remaining causes of the index score other than the observed scores. It is uncorrelated with the observed measures. This works statistically, but it does imply a realist ontology for the latent variable – there has to exist some latent construct that can have “causes”. The alternative explanation of a formative variable is that it is a form of data reduction and simplification - a mathematical combination of the indicators, in which case an error term is not required.

In a formative model, the error term is NOT an estimate of the overall measurement error contained in all the X's. The error term is the residual variance in the construct after the influence of all the items has been accounted for. In some cases, it is possible to simply exclude the error term – for example, if you can have a clearly exhaustive set of items that covers all possible cases.

It needs to be remembered that in every model, the "error" term shows unexplained causes. The "disturbance" term on an endogenous latent variable is not called measurement error, because the variable is not measured. However, it still represents unexplained causes. Because a formative construct is an endogenous latent, this applies here as well.

If the magnitude of the error term is small and all indicator coefficients are significant, then it could be concluded that the formative measure is appropriate *in that context*. If the magnitude of the error term is small, but several of the indicator coefficients are non-significant, then an investigation of multicollinearity should be done, because some items might be mediators of others. Low collinearity suggests that irrelevant indicators have been included in the set of items. Presence of collinearity would suggest item overlap. If the magnitude of the error term is large, but all the indicator coefficients are all significant, then the constructs domain of content may not be fully captured by the set of items. If the magnitude of the error term is large, but all the indicator coefficients are non-significant, then you have made a mess and should start again.

Diamontopoulus suggests that heuristics from multiple regression can be used to evaluate the effect of error terms in formative models (Diamontopoulos, 2006).

Guidelines: small effect

$f^2 = 0.02$ corresponds to an R^2 of 0.0196

Moderate effect

$f^2 = 0.15$ corresponds to an R^2 of 0.15

Large effect

$f^2 = 0.35$ corresponds to an R^2 of 0.26

Where f^2 is Cohen's f^2 used to measure the effect size in multiple regression relationships (Cohen, 1988), and R^2 is the measure of the proportion of variance explained by the variables.

In conclusion, for formative constructs, variation in the indicators precedes variation in the latent variable (Borsboom et al., 2003). By definition, formative constructs do not exist independently of measurement and are inextricably tied to their measures. The surplus meaning then, is directly associated with the error term in the formative model specification. The surplus meaning relates to un-measured causes – indicators not included in the model.

$$\text{VAR}(T) = \text{VAR}(O) + \text{VAR}(e)$$

Therefore

$$\text{VAR}(T) > \text{VAR}(O)$$

A formative model can be seen as “short-hand” for a causal model, where the observed variables do double-duty for implied latents that are not specified. Diamontopoulus demonstrates this in the models reproduced in Figures 3.7 and 3.8 (Diamontopoulos, 2006).

Figure 3-7: *Formative indicators as “short-hand” for a causal model, from (Diamontopoulos, 2006)*

Figure 3.7 is a special case of figure 3.6, if you set the loadings of the indicators $x_1 \cdots x_3$ on their latents to 1.0, and set the error terms to zero (assume that the measures are error free).

Figure 3-8: *The implied latent variables in Figure 3.6 made explicit, from (Diamontopoulos, 2006)*

This demonstrated equivalence means that the error term on η cannot be the result of any combination of the error terms for $x_1 \cdots x_3$, because the formative model assumes that these are error free.

3.4.7 Non-factor analytic approaches

Factor analysis has frequently been justified in pursuit of parsimony, but some theorists have questioned whether *theory* parsimony is always the major justification for SEM modellers (Hayduk, 1996). In many cases, parsimony is sought to increase ease of modelling and analysis. As the number of factors increases, the number of covariances between those factors increases more rapidly, so a major justification for a minimum number of factors is often to avoid an under-identified SEM model (Hayduk, 1996).

Hayduk makes a distinction between factor models, and multiple indicator structural models (Hayduk, 1996). What is a multiple indicator structural model? In simple terms, Hayduk posits that **all** arrows in a SEM model indicate causal relationships that need to be supported by theory; and there are no important theoretical differences between measurement and structural relationships.

This approach leads to a different approach to structural equation modelling that offers an alternative to the conventional wisdom of the “two-step” approach, which separates the “measurement model” from the “structural model”. The measurement model represents latent variables and their indicators, and is intended to confirm the validity of the measures. Once these are confirmed the structural relationships between latent variables can then be evaluated (Hayduk, 1996).

Hayduk argues that the loading of an indicator on a latent variable (the λ value) is also a truth claim that needs to be supported by theory (Hayduk, 1996). For example, consider two studies that examined a latent construct “content quality”, with indicators accuracy, believability, timeliness and relevance. The first study had loadings of 0.87, 0.83, 0.67, and 0.55 respectively, while the second had loadings of 0.56, 0.62, 0.89, and 0.71. Hayduk would suggest that the meaning of content quality was different from one study to another – we were using the same term, and indeed, the same indicators, but the two latent constructs were not in fact the same (Hayduk, 1996). The second was best represented by timeliness, while the first was best represented by accuracy and believability. In each case, the internal state we claim to

be measuring is having a different effect on the subject's responses to questions, so they must be different things.

Changing a λ value changes the correlation between the latent variable and its indicator. "Since the data has not changed, the only alternative is to admit that the change in the correlation arises because the identity of the concept (and hence its meaning) has changed. Even changing the variance of the concept with a fixed λ , changes the meaning of the concept, because the correlation between the concept and indicator is again altered, and the meaning of the data is unchanged" (Hayduk, 1996, p. 32).

Hayduk notes that survey design and modelling methods texts frequently advise is that every latent construct in the model should have three or four indicators (Hayduk, 1996). This nearly always guarantees that the model will be identifiable, even with all the covariances between latent constructs freed. However, Hayduk suggests that freeing all covariances lacks theoretical rigour, and if you specify the expected effects among the constructs, instead of permitting all possible correlations, you will not have so many issues with model identification, and therefore will not (necessarily) need so many indicators. A further argument against the use of multiple indicators is that some latent variables have only one sensible indicator (e.g. age, sex) (Hayduk, 1996).

The argument continues by noting that there is considerable similarity in how factor models and multiple indicator structural models account for the observed covariances among items. The key difference is that the factor model's primary explanatory mode is "spurious item correlations (covariances) arise from a dependence on the common-factor cause", while SEM permits explanations based on hypothesised causal effects, and may or may not include multiple indicators. The researcher has more freedom to explain the observed item correlations other than through common cause (factor) or through labelling them as spurious (i.e. error). This allows the researcher to reduce the error term. Structural models incorporating multiple indicator models are similar to factor models in that they imply a strict proportionality between factor loadings (λ 's) and the indicator covaries with each and every other modelled variable (Hayduk, 1996).

Hayduk then continues with a proof that a common cause of an effect does not necessarily need to be a common factor, and concludes that “even though a specific variable is named and justifiably identified as a common cause of a set of indicators, this is insufficient to conclude that “this” is the variable that will be located as the common factor in factor analysis. Factor analysis should locate the common cause precisely at the branching that sends unique impacts to each of the indicators. This means we need to be able to provide some justification as to why the posited identity of a variable should correspond to the most proximal of the potentially many common causes of the items. “The researcher must acknowledge, justify, and account for the uniqueness of the causal forces emerging from the common factor and taking separate routes to the indicators.” (Hayduk, 1996, p. 23).

What do we mean by this? Hayduk suggests that we are often misled by factor analysis to assume that a hypothesised latent variable that appears from analysis to be “one” common cause of the indicators is “the” (only) common cause. Factor analysis tends to assume a model like that in Figure 3.10. However, it is possible that the correct model is shown in Figure 3.9, and the common cause is not recognised. Further models, that are also not equivalent to Figure 3.10, will also produce similar item-factor loadings. For a detailed discussion and proof, you are invited to read Hayduk’s book (Hayduk, 1996, pp. 20-25).

Figure 3-9: *Possible common cause of factors, from (Hayduk, 1996)*

Figure 3-10: *The usual assumption of factor models, from (Hayduk, 1996)*

Instead of the “two-step”, Hayduk suggests the following four steps for developing multiple indicator structural models.

1. Chose the indicator you believe is the best single available indicator of the concept in question.
2. Fix the λ for that indicator at 1.0
3. Fix the error variance (θ) for that indicator at a specific value based on theory or prior work.
4. Enter free λ and θ for the second and third best indicators of the concept, unless you are confident that you can fix these also.

Step 1 selects which part of the shared external world (the indicators) is most similar to your hypothesised⁷. Step 2 sets the scale for that concept to correspond to the scale for the observed indicator, by making each real unit of change in the concept correspond exactly to one unit of change in the indicator. Step 3 quantifies your assessment of how similar or dissimilar your concept is to the best indicator. Step 4 is only possible if there are two or more indicators of a concept, and it provides a test for the conceptualisation you have asserted in steps 1-3 (Hayduk, 1996).

3.4.8 Summary

In summary, measurement theory, particularly latent variable theory, in combination with increased computer power, have given rise to a family of powerful statistical techniques broadly known as structural equation modelling. The “received wisdom” for social sciences including information systems is to use reflectively modelled indicators, and to test the “measurement model” (the loading of indicators on their latent variables) before the structural model (the influence of one latent on another). However, there are other ways of specifying and analysing SEMs, including formative models and multi-indicator structural models.

Although some of these methods are slowly gaining popularity, they are not widely used beyond psychology and psychometrics. As a result, it is not uncommon for the models implied by popular survey instruments such as ServQual and eQual to be mis-specified due to a lack of awareness of the differences between formative and reflective indicators. It is also common for latent constructs to include indicators at differing levels of generality and conceptual inclusiveness, which would probably be more accurately modelled as separate latent constructs, some that are more specific and descriptive, leading to others that are more generalised and inferential.

The overall conclusion from this analysis of different modelling methods is that every arrow or potential “arrow” represents a hypothesis. In co-variance-based SEMs constraints are tested, which includes missing arrows (relationships between variables which could exist but are hypothesised not to exist); and other constraints specified by the researcher. Relationships that can be specified and constrained, or left free to be estimated include the relationships between latent constructs and their indicators, as well as the relationships between latent constructs. At an even more detailed level, other research decisions, such as the magnitude of the error term, also represent assumptions about the world which can be specified, for example, descriptive beliefs tend to have lower error terms, while inferential beliefs will have higher error values. This approach has not been widely adopted by information systems researchers, and may offer opportunities for fresh insights when modelling POLSQ.

3.5 Ontological issues and assumptions of modelling methods

Various types of structural equation models are frequently compared from a statistical point of view, and there is extensive literature about how to ensure models are identified, what are appropriate fit indices, and the meaning of the error term. Less frequently discussed are the different truth claims inherent in different types of models. We discuss these briefly in this section.

The problem can be summarised as follows. As soon as we introduce the notion of a latent (unobserved) variable that is different in some way from the observed variable(s) that purportedly measure it, we need to create an explanation of the nature of the relationship between the two. This relationship is usually assumed to be causal. This introduces a number of ontological issues and assumptions.

The problem is that many data generating mechanisms can produce the same structure in data as the hypothesised model, and many different models can be developed to fit the same data (Hayduk & Glaser, 2000a, 2000b). In other words there are multiple models that are covariance-equivalent. The latent variable is some kind of mathematical function of the observed variables, and numerous proofs exist that the connection between the observed and unobserved variables is not self-evident, and lends itself to multiple explanations (Hayduk & Glaser, 2000a, 2000b). There are three possible positions (Borsboom, 2005):

1. The latent variable is a numerical means to simplify our observations – a sum-score or weighted sum-score and no more. This does not make sense if the latent variable is supposed to cause or precede responses on survey items. A sum-score cannot underlie item responses. Factor-analytic techniques such as EFA are often used in this spirit. The problem with this is that there is no way to evaluate equivalent or alternative explanations. And in fact, jumping from an EFA which produces a data reduction to then claiming this data reduction represents something real, is quite a big jump (but often done and assumed unproblematic in research). In fact, users of factor analysis sometimes seem unclear about their ontological assumptions; and seem at times to assume they are carrying out data reduction, and at other times that they have discovered or identified latent variables; or simply fail to make the distinction.

2. The latent variable is merely a means for prediction.
3. The latent variable represents a real entity. The associated theory is either true or false, the entities in the theory exist, and the theoretical entities are responsible for observed phenomena.

3.5.1 The explanations compatible with reflective models

A reflective model is compatible with a realist assumption about the existence of a latent variable, which, in psychologically based models, is some sort of unobservable, internal state held by the respondent. The response on the questionnaire varies as a function of the latent variable. So people with a positive internal state towards “content quality” will score questions about the indicators of content quality highly. Variation in the latent variable precedes variation in the indicators (Borsboom, 2005).

Reflective models are also compatible with an instrumentalist viewpoint, which says that the real existence of the phenomenon being measured is not important, simply whether the theory is useful in predicting or describing observed data. However, we are concerned with issues that arise when models are used as the basis for truth claims to having “discovered” some aspects or characteristics of user perceptions. These issues affect the assumed ontology of various forms of models for perceived on-line service quality.

If we assume that perceived on-line service quality is a construct that has independent existence (realist assumptions), this has a number of implications. If we assume that perceived on-line service quality does not have independent existence (anti-realist assumptions), this has another set of implications for how this it should be modelled.

Reflective models often implicitly assume entity and theory realism, and a causal relationship between an (existing) latent variable and the results for the observed item set. On the other hand, formative variables do not require realist assumptions, and posit that the dependent variable is constructed out of its indicators. The problem is that we cannot claim that latent A causes indicator B and at the same time maintain that latent A is constructed out of indicator B; or claim that latent A must exist but is

unobservable; and at the same time that latent A does not exist except as a sum of its indicators (Borsboom, 2005). These different conceptualisations of POLSQ have different philosophical underpinnings and implications for future research approaches.

Theory realism assumes that truth is a match between the state of affairs posited by the theory and the state of affairs in reality. Relations in reality need to exist independent of what we say about them. This has several implications for latent variable SEMs. 1) For individual subjects, when evaluating the position of the subject on the latent variable, if it is possible that you have made a mistake about the relative position of two subjects (and if there is an error score specifying the probability of such a mistake) then there must be a “true” score or it is impossible to be wrong or to interpret (ascribe meaning to) the unexplained variance as error. 2) The estimation of the parameters. Estimation similarly implies that there is something to be wrong about, and it is impossible to be wrong if it is impossible to be right. In parameter estimation, this requires the existence of a true value. You estimate the strength of your belief that a parameter actually has a value of one. This implies that you can be wrong about your belief.

Latent variable theory often assumes that unobserved latent variables are frequently supposed to be the common cause of the observed variables, although there are dissenting voices that are discussed later. A formal structure relates test scores to the hypothesised attribute, deduces the empirical implications of the model, and evaluates the adequacy of the model by examining the goodness of fit with respect to empirical data. This originated with Spearman (1904) who developed factor analytic models for continuous variables in the context of intelligence testing. Subsequently Confirmatory Factor Analysis (CFA) has become popular, based on software products like LISREL and AMOS, and is the most widely used statistical modelling method in psychological measurement. Latent variable theory is aimed at constructing an explanatory model to account for relations in the data. The model is frequently assumed to be a representation of real, but unobserved relationships and processes “in the world”.

There is a considerable literature in psychometrics arguing that latent variables are merely mathematical constructions that fit data. For an extended discussion of this issue, see for example Schonemann’s work (Schonemann, 1996a, 1996b) and other

papers in the special issue of Multivariate Behavioural Research (Volume 31, issue 4) devoted to factor indeterminacy. Although this does not make any difference to the mathematics of the problems, factor theorists such as Boorsboom and Schonemann suggest that this set of assumptions tends to lead to rather muddled truth claims. It is impossible to make any truth claims to have discovered anything “real” about the state of the world on the basis of a set of constructed equations and correspondences unless they were specified as hypotheses in considerable detail in advance. Further, as Boorsboom pointed out, the notion of the true score requires the assumption of the existence of “true” values that are separate from the observed data (Boorsboom, 2005).

The key point in both arguments is that you cannot have it both ways – you cannot both construct your latent variables out of mathematical transformations of observed data, AND then ascribe “real” properties to them.

3.5.2 The explanations compatible with formative models

Formative models, by contrast, do not require realist assumptions. Formative models are compatible with various anti-realist positions, including instrumentalism, and operationalism (Boorsboom, 2005).

For instrumentalists, theories are instruments that allow us to predict future events and exercise control over our environment. They do not need to refer to structures in the world. The only criterion is whether they are useful. Whether the theoretical entities referred to actually exist or not is scientifically and philosophically unimportant (Boorsboom, 2005). The instrumentalist would suggest that if we were able to explain (for example) a score on overall perceptions of online service quality by measuring scores on (for example) ease of navigation and timeliness of information, then whether there is a real mental state underpinning these measures is not important.

For strict operationalists, scientific concepts can only be defined in terms of the operations by which they are measured (Bridgman, 1927). In science, this has been interpreted to mean that there is not a single concept of (for example) length, but different methods of measuring length measure slightly different things. The empirical methods for measurement constitute meaning; and a scientific test is constituted as the carrying out of such an operationally defined measurement (Bickhard, 2001).

In psychology, it has been suggested that in the process of “operationalising” theoretical, unobserved constructs, there is a tendency to conflate meaning, measurement, and testing (Bickhard, 2001). For example, a person talking loudly, in a certain tone, who appears excited, and chooses a certain vocabulary to express themselves “is” angry, or a person who exhibits a certain pattern of scores “has” a high IQ, or a good ear for music.

A formative model does not require that we postulate the independent existence of the latent variable. If you suggest that the “meaning” of a construct changes in different (quantitatively measured) theory nets, this suggests an operationalist view of measurement. For example, if in one formative model, 75% of the variance in a measure of POLSQ is predicted by content quality, 20% by usability and only 5% by transaction quality; while in another 50% of the variance in our measure is predicted by transaction quality, 45% by usability, and only 5% by content quality, then what is meant by “content quality” is different in each survey. In Information Systems, we tend to accept this sort of variability. This causes issues with both the realist and the instrumentalist account. The realist would say that it shows the underlying thing doesn't exist (or exists as two separate things). The instrumentalist would say this suggests the theory is not as parsimonious as we would like: its explanatory power is less because it is variable.

This ontological discussion assists with the decision about the retention (or otherwise) of formative items that do not load strongly on the formative latent variable. As we noted earlier, in Table 3.2 and sections 3.4.6.2, 3.4.6.3 and 3.4.6.5, there is debate about the nature of the relationship between formative items and formative latent variables, the importance of covering the whole scope of the formative latent, and the contribution of items to the over all meaning. This is hard to resolve without taking a position on the ontology of the formative latent variable. We have suggested that a realist ontology is hard to justify for formative latents. If we do not adopt a realist position with respect to the formative latent, but assume that it is a point variable or index with predictive power for other constructs, then it is not so important that it always “means” the same thing. A different basket of formative indicators; or different loadings by the same formative indicators; result in a slightly different index.

In this formulation, indicators that increase the predictive power (even slightly) should be retained.

A formative model is compatible with explanations of online service quality as an index that is constructed from other factors, and can be used as a bench-marking tool to compare the performance of two or more organisations on different dimensions or to compare organisations from one year to another.

A simple summary of the ontological assumptions of modelling methods is that exogenous latent variables may be (but need not be) assumed to be real (realist assumption) while endogenous latents are necessarily constructed. Thus, if one claims that *all* entities are real and make discoveries about the world, one would have only reflectively modelled latent variables. On the other hand, the presence of reflective latents in a model does not necessarily imply a realist ontology, as reflective latent variables are also compatible with an instrumentalist ontology. This comes down to the position adopted by the researcher and the truth claims they make.

3.5.3 The explanations compatible with multiple indicator structural models

Unfortunately, philosophers of social science like Denny Borsboom (Borsboom, 2005; Borsboom et al., 2003) have not yet addressed the philosophical underpinnings of multiple indicator structural models, so it is left to the author to make some observations. Multi-indicator structural models specify a hypothesised set of covariance based mathematical relations in data, and can be specified formatively, reflectively, and as MIMIC models. They are a technique for representing a theory, and do not embed any inherent ontological assumptions.

In my view, one of the key strengths of these models is that they reduce the distance, and the number of metaphysical assumptions that need to be made when linking the observed and unobserved variables in the model. These models open up the “black box” of theory. By having fewer indicators, with stronger loadings and lower error terms, the hypothesised latent variable is much more closely identified with its indicator(s).

These models are still compatible with both a realist and instrumentalist ontology – that there are internal states with independent existence that cause the respondents to hold a certain position on a scale. However, introducing more structural relationships also introduces more regular checks for correspondence between theorised and observed relationships, which is compatible with an instrumentalist interpretation.

Hayduk himself appears to argue for an operational interpretation of his models, where differences in observed scores imply differences in meaning. Hayduk argues that the loading of an indicator on a latent variable (the λ value) is also a truth claim that needs to be supported by theory. Note that this type of model also lends itself to an instrumental interpretation, where it does not matter whether the constructs actually exist, but only that they appear to have predictive value and salience.

We still have some potential issues, if we want to develop a non-factor-analytic model based on research into the formation of attitudes and behaviour from social psychology. Fishbein and Ajzen have a clearly realist position that the objects that people perceive via direct observation, and the descriptive and inferential beliefs they hold about the object have independent existence that it is possible to measure; while Hayduk appears to want to take an operationalist perspective of imputing meaning from characteristics of data, for example changes on the variance or the strength of a factor loading (Hayduk, 1996). This becomes compatible with a realist ontology if researchers are careful to match their theoretical definitions and assumptions closely to their operationalization, and reduce the distance between observed and unobserved variables.

Although it does not solve all the metaphysical issues, our interpretation of the salience of Fishbein and Ajzen's theory of attitudes and beliefs to this research lends itself well to a non-factor analytic modelling approach. We were able to heed Hayduk's call to be very precise, in advance, about our theoretical assertions. We have a theoretical basis for fixing the (standardized) error terms for descriptive beliefs to 0.1, since descriptive beliefs are held with a high degree of certainty (Fishbein & Ajzen, 1975). We can specify a more detailed theoretical net, which shows the progression from more precise descriptive beliefs to more error prone (due to

unmeasured causes) inferential beliefs. Overall, we can establish a clearer correspondence between our theoretical assertions and our empirical observations. This does not fully resolve the ontological issues we discussed earlier in this chapter, but it reduces the distance between theory and observation and does not require factors to do double duty as mathematical constructions and real world entities. It also provides a cogent real-world explanation for the error term, namely that it represents real-world meaning not captured in the model.

We note that this explanation is also compatible with the explanation of the error term in formative models. As we suggested earlier, many formative models could easily be reconceptualised as multiple indicator structural models. If we consider the “indicators” of formative variables as candidate latents in their own right, and the relationships between formative variables and their latents as structural, then we arrive at a model very similar to a multi-indicator structural model. Consider (for example) perceived information timeliness (or any indicator claiming to measure user perceptions of some aspect of a technology artefact). This indicator implies the separate (and prior) existence of the internal state of perception that it purports to measure. A true formative indicator would need to be something that did not require any further psychological explanation than the measurement. For example, if we were to hypothesise that information no later than 24 hours old, with no factual errors “caused” a perception of quality information, this would be compatible with the assumptions of a formative model.

The theory should determine the statistical model, but there may be more than one way of implementing a theory. Reflective factor models could also be used effectively, provided they are specified in advance (for example, confirmatory factor models), and represent detailed and specific constructs with a relatively small distance between the measure and the hypothesised latent construct. Hayduk’s technique is not incompatible with reflective models as such, just with theory building technique (such as what he describes as the “two step”) that tends to move from “constructing” factors out of data (using EFA) to making truth claims about the constructed factors.

3.5.4 Debates in psychometrics and their relevance to the measurement of online service quality

Knowledge sometimes flows slowly from one field to another. There has been a great deal of discussion about the place of Information Systems with regard to reference disciplines (for example, Baskerville & Myers, 2002). The downside of the appropriation of knowledge into information systems from reference disciplines is that new knowledge sometimes permeates disciplinary boundaries rather slowly. While information systems researchers are carrying out exploratory factor analyses and enthusiastically embracing structural equation models; within psychology, debates are raging about whether quantitative methods are even appropriate for measuring attitudes and perceptions, and whether latent variable theory has reached the end of its useful life.

We cannot resolve these objections with respect to the measurement of online service quality; merely report their existence and some of their potential implications. There is, however, a powerful motivation for examining our assumptions and processes for measuring online service quality. Despite thirty years of measurement using increasingly sophisticated techniques, we have achieved very little by way of consensus or predictive value. This suggests that either the constructs we purport to be measuring do not exist, or that we are doing a poor job of measuring them, or there are common causes that we are not measuring.

3.5.4.1 The factor indeterminacy debate and the operationalist fallacy

The historical problem of the fact that there can potentially be a large number of different factor structures that fit the same data is called factor indeterminacy. Even modern, introductory texts still discuss the question of how to determine if you have extracted the correct number of factors, as we will discuss later. This issue has been debated in psychometric literature since the 1920's (Blinkhorn, 1997). Two major explanations have been advances for factor indeterminacy – broadly the “alternative solution” position (ASP) which considers that the latent common factor is determined by constraints inherent in the model (in other words, it is a mathematical construction) and the “posterior moment” theory, which considers the latent factor to be a single (separately existing) entity.

The issue with realist assumptions about the separate existence of latent factors, if they are considered only in the context of a latent factor model, is that a good-fitting model is expected to simultaneously prove the existence of the factor by acting as a criterion (test) for the existence of the factor, AND to measure the effect of the factor, such that when it is present, its indicators will behave in a certain way.

*“When one analyses data with factor analysis, what one has are manifest [observed] variates, and **constructed** [unobserved] variates called latent factors.” (Maraun, 1996, p. 529).*

This clearly states a view that the latent is an operational construction of the researcher, defined only by its measurement. Other researchers have suggested that if one is convinced by these arguments about the indeterminacy of factors, this has serious implications for psychometrics, as in latent variable theory, the meaning and measurement of constructs tend to be conflated (Bickhard, 2001).

Bickhard makes similar arguments to Maraun, but claims they lead to “pathology” in psychology, that conflates meaning, measurement, and testing. For example, a person talking loudly, in a certain tone, who appears excited, and chooses a certain vocabulary to express themselves is assumed to be angry (although we cannot observe his internal state), or a person who exhibits a certain pattern of scores is assumed to have a high IQ. Bickhard asserts that the statistical testing of a nomological net must not be the same as operationally defining unobserved concepts in the same net (Bickhard, 2001).

When we measure perceptions or affects towards technology without carefully considering the underlying assumptions we are making, we are at risk of similar issues. For example, if we identify a number of indicators, such as information timeliness, information level of detail, information relevance, and so on, that in combination exhibit certain behaviours (for example, they explain the variance of another factor) then if we believe we have discovered the “meaning” of content quality, we have conflated the measurement and meaning of content quality, and defined it in terms of its measures – potentially an operationalist fallacy. If we are not careful, we then carry on to reify the concept we have discovered, by allowing it a separate existence that can “cause” subject’s responses on indicators.

This is a particular problem for exploratory factor analysis using principal components analysis, which is a data reduction technique that constructs factors from a dataset. However, to the extent that any factor model is indeterminate (there may be mathematically equivalent alternate specifications) then it cannot be used simultaneously for both definition and specification of factors and evidence of their truth and existence.

3.5.4.2 *Has modern psychometrics run its course?*

Serious questions are being raised as to whether latent-variable type models are continuing to add value to psychological research, and by implication, to research on attitudes and perceptions in other fields such as information systems that take their lead from psychometrics.

We cannot resolve these issues here, or even report them in detail, as the arguments are extremely subtle and detailed and only of peripheral relevance to our argument. However, we would like to report their existence, so that information systems researchers have the opportunity to become familiar with them. In particular, these issues have been debated in recent special issues of *Measurement* (volume 6, issues 1&2, 2008) and *Multivariate Behavioural Research* (volume 31, issue 4, 1996). We conclude with some salient comments from psychometric researchers.

“A more radical view is that the current styles of test theory simply have no more practical value to offer, that the implicit assumptions that have guided research for nearly a hundred years place constraints on what can be achieved...a new start is necessary....For too long test theory has concerned itself with ever cleverer accounts of unimproved practical effectiveness.”
(Blinkhorn, 1997, p. 183)

“The obvious conclusion is that if we are to generate progress in our science, increasingly sophisticated methods of statistical data-model analysis are not going to help. Something about the way we go about constructing psychological variables, their measurement, and their presumed causal relations is wrong. [We should not concentrate on] yet more questionnaire item test theory and assumption laden structural latent variable models.”
(Barret, 2008, p. 82)

In summary, it is possible that there are serious structural weaknesses in latent factor analysis and latent variable theory, particularly the problem of factor indeterminacy.

This has serious; some might say catastrophic, consequences for psychometrics and other fields such as information systems that use it as a reference discipline. This point was made very succinctly, if somewhat humorously, by Schonemann (Schonemann, 1996a). His argument is summarised in table 3.6.

Table 3.6: *The potential implications of factor indeterminacy for psychometric research, adapted from (Schonemann, 1996a)*

Key points	Argument	Implications
Major premise	P implies Q	Factor indeterminacy (P) implies that the basis of many standard psychometric approaches to the measurement of attitudes and perceptions is incorrect (Q)
Minor premise	Q is too horrible to contemplate	Psychometric research is in denial about this (Q), because the consequences are serious and far-reaching for the discipline
Conclusion	P must be false	<i>So researchers tend to reject factor indeterminacy (P), not because it is wrong but because it is so uncomfortable.</i>

Should we be worried? It's hard to say. At the very least, we suggest, we should understand, and carefully examine the ontological positions and truth claims implicit in our models of user perceptions of, and attitudes towards technologies.

3.6 Modelling and Measurement in Online Service Quality Research

Modelling and measurement in online service quality research is firmly based on latent variable theory and factor analytic approaches. This has generally been applied quite uncritically. Overall, service quality and online service quality researchers have tended to reify factors identified by statistical analyses, and to apply exploratory and confirmatory factor analysis techniques without a careful examination of the underlying assumptions of the techniques that have been applied.

As well as eQual, we give extended attention to ServQual, because as we discussed in our literature review, it is widely cited as the provenance for subsequent service quality research in information systems and electronic commerce. We previously asserted that ServQual has been accorded a de-facto status as theory, with subsequent

research adopting normal-science assumptions of “building” on this body of knowledge. Many of the explicit or implied assumptions of ServQual have been adopted uncritically by subsequent researchers on online service quality. It is therefore important to establish whether these assumptions are sound.

In this section, we apply the theory that we discussed in the previous section to the measurement of service quality and online service quality. We first examine briefly the measurement approaches used for the ServQual instrument and then the approaches used for the eQual instrument that is used as the foundation for the current study. We consider briefly the provenance of the instruments and their items (from a measurement perspective), the nature of the constructs and indicators, and the stated assumptions of the authors.

3.6.1 ServQual

3.6.1.1 *Provenance of the constructs*

The original ServQual constructs were derived from qualitative analysis of focus group and individual interview data across several industries; in other words they were empirically derived, rather than based on a priori hypotheses. The original constructs were: reliability (consistency of performance), responsiveness (readiness and willingness of employees), competence (possession of required skills and knowledge), access (approachability and ease of contact), courtesy (politeness, respect, consideration), communication (keeping customers informed), credibility (trustworthiness and honesty), security (freedom from danger, risk and doubt), understanding (making an effort to understand customer’s needs), tangibles (physical elements) (Parasuraman et al., 1985).

The authors describe their early identification of these constructs: “Specifically, the research revealed 10 dimensions that customers *use in forming* expectations about and perceptions of services, dimensions that transcend different types of services” (Parasuraman et al., 1985, p. 49, our emphasis). This appears to describe a set of latent variables, some kind of unobserved internal concepts held by the customer that are measured by the expectations and perceptions that are measured on the ServQual scale. This would suggest that the indicators in the scale should be reflective of these

constructs. If the constructs are “*used in forming*” an overall affect towards service quality, then this would suggest that the dimensions are antecedent to perceived service quality, rather than consequents of it.

The ServQual instrument was subsequently refined. The original ServQual instrument included 97 statements to reflect those 10 dimensions. Further analysis, including item to total correlation analysis, factor analysis, and assessment of internal consistency, was used to refine the instrument (Parasuraman et al., 1985). It was subsequently reduced to 34 items and seven dimensions (Parasuraman et al., 1988), and then to 22 items and the current five “general” dimensions (Parasuraman et al., 1988).

These five dimensions are: assurance (knowledge and courtesy of the employees and their ability to inspire trust and confidence), empathy (caring individualised attention), reliability (ability to perform the service dependably and accurately), responsiveness (willingness to help and promptness), and tangibles (appearance) (Parasuraman & Zeithaml, 2001). According to the authors “*Assurance is basically a combination of the original dimensions of competence, courtesy, credibility and security. Empathy represents the remaining dimensions of access, communication, and understanding the customer*” (Parasuraman & Zeithaml, 2001, p. 341).

However, as we discussed in our literature review, unstable dimensionality has plagued the ServQual discussion. We suggest that a major reason for that is that the “five” factors are constructions of factor analysis rather than latent constructs for which any realist argument can be made. There is little support from other theories (triangulation) for the existence of the five dimensions – they do not exist in any comparable form in other nomological nets that we were able to identify. Several of the definitions contain “and” statements and are acknowledged by the authors to be composites. For example, communication, credibility, security, competence, and courtesy were collapsed by the authors into “assurance”, which is defined as “knowledge *and* courtesy of employees *and* their ability to inspire trust *and* confidence” (Parasuraman et al., 1988, p. 23, our emphasis). “Assurance” cannot, therefore be a latent variable defined by this set of indicators, as it is not (according to the authors) a single common cause of the indicators that purport to measure it, but a composite of several constructs.

Overall, we suggest that the dimensionality issues with ServQual are a direct result of the over-use of factor analysis and data driven discovery without an adequate discussion of the underlying assumptions of the theory. On close examination, it is clear that the “five” ServQual “dimensions” as currently defined by their indicators are not single constructs that represent unobserved internal states (latent variables) that customers use in forming the expectations and perceptions that ServQual purports to measure.

3.6.1.2 The nature of the constructs and indicators

When the constructs of ServQual, and their indicators, are examined according to the principles of latent variable theory and factor analysis that we discuss above, it becomes clear that there are measurement issues. These are summarised in Table 3.7. The survey items are a mixture of formative and reflective indicators, and in some cases, appear interchangeable with indicators for other, supposedly distinct constructs.

Table 3.7: *Measurement issues in ServQual*

Construct (latent variable)	Indicator	Comments
Reliability	1. When excellent companies promise to do something by a certain time, they will do so	Reflective of expectation of reliability and interchangeable with 4
	2. When customers have a problem, excellent companies will show a sincere interest in solving it	It is not entirely clear that this reflects reliability in a way that has discriminant validity with other constructs. This could also be argued to reflect expectation of empathy (understanding the customer).
	3. Excellent companies will perform the service right first time	More specific than reliability, therefore likely to be a formative indicator. It does not encompass all the meanings of reliability.
	4. Excellent companies will perform their services at the time they promise to do it	Reflective of expectation of reliability and interchangeable with 1
	5. Excellent companies will insist on error free records	More specific than reliability, therefore likely to be a formative indicator. It does not encompass all the meanings of reliability. Could be related to 3, to the extent that “right first time” implies absence of error
Assurance	1. The behaviour of employees in excellent companies will instil confidence in customers	Reflective of expectation of assurance
	2. Customers of excellent companies will feel safe in their transactions	Reflective of trust (security) which is one of the original 10 dimensions that is now included in assurance. Does not encompass the full meaning of assurance and is not interchangeable with other indicators. If included as an indicator of assurance it would need to be specified formatively.
	3. Employees of excellent companies will be consistently courteous with customers	Reflective of courtesy, one of the original 10 dimensions that is now included in assurance. Does not encompass the full meaning of assurance and is not interchangeable with other indicators. If included as an indicator of assurance it would need to be specified formatively.
	4. Employees of excellent companies will have the knowledge to answer customer questions	Reflective of competence, one of the original 10 dimensions that is now included in assurance. Does not encompass the full meaning of assurance and is not interchangeable with other indicators. If included as an indicator of assurance it would need to be specified formatively.

Tangibles	1. Excellent companies will have modern-looking equipment	Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
	2. Physical facilities at excellent companies will be visually appealing	Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
	3. Employees of excellent companies will be neat-appearing	Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
	4. Materials associated with the service will be visually appealing in excellent companies	Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
Empathy	1. Excellent companies will give customers individual attention	Reflective of empathy. Interchangeable with 3
	2. Excellent companies will have operating hours convenient to all their customers	Formative. Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
	3. Excellent companies will have employees who give their customers personal attention	Reflective of empathy. Interchangeable with 1
	4. Excellent companies will have the customer's best interests at heart	Formative. Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
	5. Employees of excellent companies will understand the specific needs of their customers	Formative. Can vary independently from other indicators, not interchangeable, and does not necessarily have the same antecedents and consequences
Responsiveness	1. Excellent companies will tell customers exactly when services will be performed	Unclear. Appears to be interchangeable with items 1 and 4 for reliability
	2. Employees of excellent companies will give prompt service to customers	Hyponym of responsiveness, does not encompass the full meaning, and could vary independently with 1 and 3 (it is possible to be prompt but not willing).
	3. Employees of excellent companies will always be willing to help customers	Unclear. Appears interchangeable with 4 Appears to reflect a similar construct to questions 1 and 3 of empathy
	4. Employees of excellent companies will never be too busy to respond to customer requests	Unclear. Appears interchangeable with 3.

3.6.1.3 Assumptions of the authors: Theoretical basis of ServQual

According to the authors, ServQual was originally developed as a scale to measure the dependent variable for the gaps model. *"We previously developed a multiple-item scale called ServQual to measure service quality as perceived by consumers (Gap 5 in figures 1 and 2)....Therefore, it is possible to recast the conceptual service quality model (the gaps model) as a structural equations model wherein perceived service quality (gap 5) is the unobservable dependent variable, and the four gaps on the marketers side are the unobservable independent variables. This model can be tested by collecting data on the indicators of the five gaps.."* (Zeithaml, Parasuraman, & Berry, 1988, p. 44)

Already, we have difficulty, since service quality is proposed as a multi-dimensional construct. The authors' narrative above includes the implicit assumption that the five dimensions they claim to have identified cause responses on their own indicators, and in turn cause the user's perceptions of overall service quality. This particular conceptualisation has never been tested, and does not appear (on the basis of our discussion above) to be defensible. Nevertheless, over time, ServQual became detached from its original theory net and took on a life of its own.

3.6.1.4 Service Quality as a latent construct or constructs

The ServQual authors themselves have acknowledged that there are issues with discriminant validity of the "dimensions" while continuing to support and promote their scale. Further studies, including the use of confirmatory factor analysis, found cross-loading between items for responsiveness with assurance and reliability, and items for responsiveness and assurance loading on the same factor (Parasuraman et al., 1991). They also acknowledge that it is possible there are unidentified structural relationships between the dimensions. We do not find this at all surprising, given our discussion above. It is difficult to make a convincing realist argument for the separate existence of a latent construct of "service quality" or any of its "dimensions". Both the empirical issues (problems with dimensionality), and the theoretical issues (mis-specification of the ServQual instrument) preclude this.

There has been extensive discussion of the reliability and validity of the ServQual scale using Cronbach's alpha, and factor analysis (Asubonteng et al., 1996). Different explanations have been advanced for the variations in findings, particularly in discriminant validity, including differences in analysis techniques, and differences in respondents' perceptions of the dimensions (Asubonteng et al., 1996). However it is not usually suggested that these differences may have arisen because the model was mis-specified, with a mixture of formative and reflective indicators. "Traditional" techniques for assessing reliability and validity are not appropriate for a formative or mixed model (Diamontopoulos, 1996).

3.6.1.5 *Mis-specification by the authors in online service quality research*

We have suggested in our discussion that the ServQual instrument is mis-specified. It is reasonable to ask how this could be true of such a well-established and highly cited instrument. Our intention is not to attack the SevQual authors as such, but to demonstrate first, the relatively low level of knowledge of key debates in psychometric research in the wider social science research community; and second to show how the uncritical application of factor analytic techniques, and an uncritical adoption of "knowledge" from reference disciplines can have serious and far-reaching consequences.

The ServQual authors appear to have applied statistical techniques fairly uncritically, as in this example:

"The CFA results in table 5 are for a first order factor model specifying the scale items as reflective indicators of their corresponding E-S-Qual dimensions and allowing the four dimensions to inter-correlate. We also conducted a second order CFA with the four first-order dimensions modelled as reflective indicators of a second order e-SQ construct. The results were similar. However, we were not able to conduct a second order CFA with the first order dimensions modelled as formative indicators, because we did not have the additional reflective indicators of the second order construct that are needed for model estimation." (Parasuraman, Zeithaml, & Malhotra, 2005, p. 225 footnote).

This appears to suggest that the authors believe it is possible to simply "reverse the arrows". The differences in meaning between a reflective and a formative model are not discussed.

3.6.1.6 *Summary*

ServQual is not a sound foundation for subsequent research. Its theory net has never been fully tested or proven, so to some extent, it is a dependent variable looking for a model. The ServQual “dimensions” appear to be reified from factor analysis and are difficult to defend as latent constructs with independent existence (with their current set of measures), especially since several of them are acknowledged by the authors to be composites of several other latent variables. These could be decomposed, and respecified. An alternative conceptualisation might model (for example) a sense of assurance as a consequence of the perceived knowledge and competence of employees. The instrument itself is mis-specified as a mixture of formative and reflective variables, and the authors themselves acknowledge that there are likely to be structural relationships between the “dimensions” that have not been identified or tested. We suggest that there are likely relationships between the indicators as well.

The enthusiasm with which ServQual has been adopted by academics and practitioners in a range of business domains suggests that the indicators (question items) included in ServQual have relevance to practice. This gives some legitimacy to the set of twenty-two questions included in ServQual as a starting point for measuring perceived service quality. However, the issues with the ServQual stream of research suggest that it is long overdue for re-examination, to better understand the structure of this important phenomenon.

3.6.2 E-commerce service quality

Research into latent constructs (dimensions) of e-commerce service quality has tended to rely heavily on exploratory factor analysis. If we examine the findings of a recent literature review paper (Alzola & Robaina, 2005), summarised in Table 3.8, the following generalisations can be made.

Table 3.8: *Studies of the dimensions of e-service quality, adapted from (Alzola & Robaina, 2005).*

Authors	Principal Analysis Applied	Dimensions
(Syzmanski & Hise, 2000)	Principal components analysis with varimax rotation	Convenience; product offers, product information, web site design, financial security
(Wang et al., 2001)	Principal components analysis with varimax rotation	Customer support; security; ease of use; digital products/services; transaction and payment; information content; innovation
Yoo and Danthu (2001)	Factor analysis with oblique rotation and confirmatory factor analysis	Ease of use; attractive design; speed of processing; security
(Aladwani & Palvia, 2002)	Delphi method, principal components factor analysis with varimax rotation	Content; quality of content; appearance; technical sufficiency
(Barnes & Vidgen, 2002)	Principal components analysis with varimax rotation	Usability; design; quality of information; quality of interaction with service; trust; empathy
Janda, Trocchia, and Gwinner (2002)	Confirmatory factor analysis	Performance; access; security; sensation; information
Ranganathan and Ganapathy (2002)	Exploratory factor analysis	Content of information; design; security; privacy
Young and Jun (2002)	Principal components analysis with varimax rotation	Reliability; access; ease of use; personalisation; security; credibility; response capacity
Cai and Jun (2003)	Principal components analysis with varimax rotation	Web site design; trust; reliable service; communication
Keating, Rugimbana, and Quazi (2003)	Principal components factor analysis with equimax rotation and confirmatory factor analysis	Quality of service; physical aspects; reliability etc. Relational quality: trust; effort; value; understanding etc
(Wolfenbarger & Gilly, 2003)	Principal components analysis with varimax rotation and confirmatory factor analysis	Reliability; tangibles; empathy; response capacity; security
Jun, Yang, and Kim (2004)	Principal components analysis with varimax rotation	Reliable/prompt responses; access; ease of use; attentiveness; security; credibility
Lim and Dubinsky (2004)	Principal components analysis with varimax rotation and confirmatory factor analysis	Merchandise; interactivity; reliability; navigation
Long and McMellon (2004)	Principal components analysis with varimax rotation	Tangibility; reliability; response capacity; security; purchase process
Muyelle, Moenaert and Despontin (2004)	Critical incident technique and confirmatory factor analysis	Information relevance; precision of information; understanding of information, etc
(Yang et al., 2004)	Confirmatory factor analysis	Reliability; response capacity; competence; ease of use; security; product portfolio
(Lee & Lin, 2005)	Confirmatory factor analysis	Website design; reliability; response capacity; trust; personalisation

The vast majority of the studies (10) use exploratory factor analysis, primarily principal components analysis, in conjunction with orthogonal rotation, in particular varimax (8 studies) and equamax (1 study). A smaller number of studies use oblique rotation (2). Confirmatory factor analysis is also quite widely used (7 studies), likely specifying correlated latents. This means that approximately half of the studies in Alzola and Robaina's review think the dimensions should be orthogonal, while the other half thinks they should be correlated.

A cynic might think that this pattern partly results from the fact that principal components analysis with varimax rotation is the default setting in SPSS™. The question is whether this approach is the most appropriate.

We suggest that it is not. The low level of convergence across these multiple studies suggests that this approach is not successful in building generalisable theory. Further, a brief review of the “dimensions” posited by these studies suggests that they are not all the same type of construct – they are “incommensurable things”. Some are attitudes (affects) towards websites, for example ease of use, trust, empathy, reliability. Others are descriptive beliefs about website affordances, for example purchase process, digital products, and technical sufficiency – effectively inventories of desirable and salient affordances. Our previous discussion suggests that these are not equivalent dimensions. Descriptive beliefs (about a technology) are specific to the affordances of that technology (and therefore extremely unlikely to be generalisable). Attitudes (affects) towards technology are caused by descriptive beliefs and other accumulated experience.

3.6.3 eQual

3.6.3.1 *Provenance of the constructs*

One of the strengths of eQual is the multi-methods approach which has been taken to deriving the constructs and indicators. This included focus groups, several rounds of empirical data gathering in different business domains, and extensive literature support (Barnes & Vidgen, 2002). Subsequent research aimed at identifying additional items/dimensions suggests that the domain is fairly well covered, with the exception of transaction quality and richness, which has emerged as a quality criterion

as websites increasingly offer self-service transactions (Tate et al., 2008b; Tate et al., 2007).

3.6.3.2 The nature of the constructs and indicators

To the extent that there are weaknesses in the eQual instrument that need correction, they appear to be in the technical, measurement theory aspects, rather than in the literature support or domain coverage. The authors themselves are ambiguous in their claims as to the number of constructs. At times three constructs are referred to: information quality, usability (which is a composite of design and usability) and interaction quality (which is a composite of trust and communication). At other times, five constructs (information quality, usability, design, trust and communication) are used (Barnes & Vidgen, 2002).

Our analysis (summarised in Table 3.9) suggests that survey items are a mixture of formative and reflective indicators. Content quality in particular does not appear to have any claims to realist existence as a latent variable which causes responses on its items, as the items can easily be seen to be able to vary independently (it is possible for information to be highly accurate but not relevant, timely, but not easy to understand, and so on). The usability items seem to be the most “reflective” of a latent construct, while many of the others are a mixture of items of different types, which in some cases, do not have strong face validity for their posited constructs (for example positive experience and sense of competency to design).

Table 3.9: *Analysis of eQual indicators*

Construct	Indicator	Comments
Information Quality (IQ)	This site provides accurate information	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
	This site provides believable information	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
	This site provides timely information	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
	This site provides relevant information	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
	This site provides easy to understand information	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
	This site provides information at the right level of detail	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
	This site presents information in the appropriate format	Formative indicator. All of the IQ indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences.
Usability > Usability	I find the site easy to learn to operate	More specific than usability and therefore likely to be a formative indicator. Does not encompass all the meanings of usability.
	My interaction with the site is clear and understandable	Reflective of usability
	I find the site easy to navigate	Reflective of usability
	I find the site easy to use	Reflective of usability
Usability > Design	The site has an attractive visual appearance	Reflective of design
	The design is appropriate design for the type of site	Reflective of design but appears to have another causal or moderating factor. This indicator also requires an evaluation of the fit between the design and the site purpose.
	The site conveys a sense of competency	Not clearly either formative or reflective for “design”. Could be argued to be a summative indicator of overall quality
	The site creates a positive experience for me	Not clearly either formative or reflective for “design”. Could be argued to be a summative indicator of overall quality

Interaction Quality > Trust	The organisation has a good reputation	More specific than trust and therefore likely to be a formative indicator. Does not encompass all the meanings of trust.
	It feels safe to complete transactions	More specific than trust and therefore likely to be a formative indicator. Does not encompass all the meanings of trust.
	My personal information feels secure	More specific than trust and therefore likely to be a formative indicator. Does not encompass all the meanings of trust.
	I feel confident that goods/services will be delivered as promised	Reflective of trust
Interaction Quality > Communication	The site conveys a sense of personalisation	Formative indicator. All of the communication indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences
	The site conveys a sense of community	Formative indicator. All of the communication indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences
	The site makes it easy to communicate with the organisation	Formative indicator. All of the communication indicators can vary independently, are not interchangeable, and do not necessarily have the same antecedents and consequences

3.6.3.3 Assumptions of the authors

The eQual authors do not clearly state a theory net for their instrument. At some times, they appear to be aiming to identify latent constructs, while at other times, they appear to be aiming to develop a sum-score that can be used for data reduction and bench-marking.

“Perhaps more interesting is some conceptual assessment of how the Web sites differ in quality...we need to move beyond the scores and indices on individual questions towards a set of meaningful and reliable sub-groupings....The generation of subcategories is relatively similar to the work associated with ServQual (Zeithaml, Parasaman, & Berry, 1990). As a starting point, and to establish that the qualities can be disentangled and are not part of a single scale, a factor analysis was conducted on the data” (Barnes & Vidgen, 2002, p. 121).

References to “meaningful and reliable sub-groupings”, similar to those identified by the ServQual authors suggests that the authors believe they are identifying generalizable latent constructs with a realist ontology. Strictly, in this world view, this should be specified as a second order model, with “customer perceptions of online

service quality” acting as a dependent variable in some (unspecified) theory net, causing a response on its sub-constructs, which in turn caused a response on their indicators.

At other times, the authors’ description of their instrument suggests that they are developing an index (sum-score) for data reduction and bench-marking purposes, compatible with an anti-realist explanation.

“By utilising the framework of categories examined in the last section, we were able to build a profile of the qualities of an individual Web site that makes it easy to compare with its rivals. Thus, we may examine why some sites fared better than others on the eQual Index” (Barnes & Vidgen, 2002, p. 122).

The “classic” tests of survey reliability and validity are not applicable to formative models (Diamontopoulos, 1996). The eQual authors have used factor analysis to check validity and Cronbach’s alpha to check reliability. These imply a reflective model, and may not be appropriate for the nature of the eQual indicators, as we discussed above, or for the intended purpose of the instrument, which appears to be to develop a calculated index for data reduction and bench-marking purposes.

The nature of the constructs and items suggests that these two goals are in fact rather muddled in the eQual research. Our discussion suggests that these two different types of constructs are based on different ontological assumptions, and require different modelling methods.

3.7 Synthesis

This chapter presents the results of our reading program in philosophy of science, social science theory; and measurement issues, particularly from contemporary psychometrics; and our application of these concepts to the research and measurement of service quality and online service quality. This was initially carried out in an attempt to explain the weak results of our initial data analysis (presented in Chapter 4). This also forms our second line of inquiry to explain the instability of the online service quality construct: the possibility that our existing theoretical approaches to this research are flawed. This possibility received some support from the conclusions

of Chapter 2, where we noted that despite a large volume of literature, there is little consensus, and the existing literature is so confused and contradictory that carrying out a conventional literature review is almost impossible.

Our conclusions from this discussion cover theoretical issues from: social science theory and its applicability to information systems; the theoretical foundations of theories of attitudes and behaviours towards technology in social psychology; types of information systems theory; and an evaluation of the quality of service quality theories. Our conclusions with regard to measurement cover the differences between formative and reflective models, and alternative approaches such as non-factor analytic models (also known as multi-indicator structural models). We explore the underlying ontological assumptions of different modelling and measurement methods, and find issues with the truth claims and theory nets implicit in previous studies.

In overview, the quality of existing theory in service quality and online service quality seems weak; based on unjustified normal science assumptions, unjustified realist claims for the existence of factors reified from exploratory factor analysis, and the illusion of a cumulative research tradition. From a measurement perspective, it seems likely that researchers in business schools (including information systems researchers) have imported measurement techniques such as factor analysis from psychometrics uncritically, without unpacking their underlying assumptions, or understanding their weaknesses and limitations. This supports the findings of a study by Conway and Huffcut on the use of factor analysis in management journals (Conway & Huffcut, 2003). It also seems likely that information systems researchers in particular have continued to use techniques such as multiple regression that have been largely superseded by co-variance-based SEMs within the psychometric community. This supports the assertions of Rouse and Corbitt in a recently published paper that compares PLS and co-variance-based SEMs (Rouse & Corbitt, 2008).

Our overall paradigmatic position is a post-positivist, critical realist view of the phenomenon of online service quality. We assert that there is a reality of IT artefacts and user perceptions thereof, and more specifically, online services and user evaluations thereof; that is independent of our thinking about it. However, following Longino (Longino, 1990) we believe that our explanations of these phenomena are

fallible “best efforts” constructions that are negotiated between researchers. We further suggest that it is possible that some of the theories that have been used to explain online service quality so far may be due for a rethink, if not a replacement.

The existing theory nets for online service quality, when measured against quality standards for social science theory are weak. Dubin called for meaningful “building blocks” of social science theory made up of propositions that predict the value of one or more units in the model, and empirical indicators that make visible and replicable the process of measuring and the value produced. Gregor bemoaned the lack of information systems theories that exhibit these qualities. We demonstrated that the theoretical family tree for this research is largely composed of descriptive,” Type 1” theory (Gregor, 2004).

ServQual, and its successors such as eQual, have been afforded the defacto status of a research paradigm that researchers can accept (Kuhn, 1996), and are used as the basis of a “cumulative” research tradition within information systems. This assumes that the outcome of scientific enquiry is to *accumulate* solutions to scientific puzzles within a given theoretical frame, rather than piling up different and competing theories in an unsorted (and unsortable) heap. Research in online service quality has not succeeded in producing reliable, replicable results, or a consensus of researchers in the field. Further, issues such as causal ambiguity, interpretational confounding, incommensurability, and a pre-occupation with parsimony have dogged some of the critical studies in the research “family tree”, which suggests that previous research does not provide a solid foundation to build on.

Despite early foundations in social psychology, key insights relating to the formation of attitudes and intentions towards technology appear to have been lost by information systems researchers. We suggest that user attitudes and perceptions towards a website are built up from initial descriptive beliefs about the characteristics of the technology, to more generalised inferential beliefs that in turn lead to attitudes and affect. This basic approach provides a solidly grounded theoretical template which can be used as the basis for developing specific theory and models. A more extensive development and illustration of this approach is provided by Evermann and Tate (Evermann & Tate, 2009b; Tate & Evermann, 2009b).

From a measurement perspective, we conclude that the ongoing use of exploratory factor analysis is inappropriate for this research area. In fact, factor analysis appears to be the subject of considerable “angst” among contemporary psychometric researchers, and may be going out of favour as a preferred approach for measuring psychological latent variables.

eQual and ServQual are specified as a set of latent constructs with reflective indicators. A careful examination of the indicators used for eQual supports an argument for reformulating the model as formative. It is difficult to make an intuitive argument for the separate ontological existence of an internal affect towards (for example) “content quality” which causes responses on other items. Further, the items self-evidently can vary independently (it is possible for information to be timely but not accurate, and vice versa), so they cannot be argued to be interchangeable measures that “reflect” the same latent, or to have the same common cause.

A possible option would be to respecify the constructs for POLSQ formatively. However, formative models (particularly using partial least squares) make relatively weak truth claims from an ontological point of view. They do not require realist assumptions. They simply assert “if you measure *a*, *b*, and *c*, and apply these mathematical functions, you will be able to predict the value of *d*. If you simply want measures that appears to have some sort of real-world salience in the business domain, (and the popularity of online service quality studies suggests that this measure does), then perhaps a formative model will suffice. It allows organisations to “benchmark” consistently from year to year, or between organisation and organisation.

On the other hand, a new form of modelling emerging in contemporary psychometrics, multi-indicator structural models; appear to offer the opportunity to create and test clearly articulated hypotheses and predictive theory nets. They are compatible with a critical realist interpretation that a latent variable causes responses on an indicator, with a low error rate, and that unexplained causes are responsible for the error term.

4 Methodology

4.1 Overview

This study is presented as a series of phases, but as we described in the chronology, the work was a highly iterative process, as we moved back and forth between literature and data analysis. This chapter provides an overview of the methodology and modelling approaches we used. Since we used the same context and sample throughout the study, characteristics of the research context and sample are also described in this chapter.

In the first part of this chapter, we recollect our research questions and describe the methodology and modelling methods that will be used to evaluate these questions in the following chapters. In the second part of the chapter, we describe the context and sample used for the study.

4.2 Research questions and modelling methods

We recollect that we established that despite extensive research, there is very little consensus about the structure of POLSQ. Our research questions are as follows:

1. What is the structure of POLSQ?
2. What are the antecedents of POLSQ?
3. What is the ontology of POLSQ?
4. What are the most appropriate modelling and measurement methods for measuring online service quality quantitatively, and what insights can be gained from psychometrics?
5. What insights does this offer IS researchers for the measurement of user attitudes and perceptions towards technologies?

Our investigation follows two lines of enquiry to explain the unstable dimensionality of POLSQ: 1) that the structure of POLSQ has been unstable because the domain itself has changed, and 2) our modelling and measurement approaches are flawed.

Our research investigates both possibilities. Following the literature review in chapter 2, we selected a leading instrument and suggested a (likely) additional factor and some additional indicators based on research literature. Following the discussion of theory and measurement in chapter 3, we identified a number of potential weaknesses in the online service quality literature “family tree”, and some alternative approaches for measuring POLSQ

An overall process diagram for the methodology phases is included as Figure 4.1. Our first set of models, presented in Chapter 5, continues the approach taken initially by the ServQual authors, and more recently by the eQual authors. We apply “conventional” techniques of a narrative literature review and exploratory factor analysis to develop and pre-test our online service quality instrument, based on eQual. We collect data using an online survey, and conduct an exploratory factor analysis (EFA). The results are compared with the eQual factor structure.

As we described in the chronology in our introduction, our results were unsatisfactory, and we went back to literature for a deeper understanding of theory and measurement. The remaining models present various alternative approaches we used for modelling the data.

In Chapter 6, we apply more “traditional” modelling techniques which are widely recommended in SEM methods textbooks (for example, Byrne, 2001; Kline, 2005) which involved developing and testing a reflective “measurement model” using confirmatory factor analysis (CFA) before continuing to test the “structural model”. We evaluate several alternate models of the structure of online service quality based on the dimensions posited by (Barnes & Vidgen, 2002), the dimensions based on a hypothesised a priori factor structure from our literature review, and the emergent factor structure from our EFA.

In Chapter 7 we investigate the possibility we discussed in Chapter 3, that previous reflective models of service quality and online service quality are mis-specified, and many of the indicators used by previous researchers are formative rather than

reflective. We develop a formative model of online service quality based on literature and recent insights from structural equation modelling theorists.

In chapter 8 we apply recent insights from psychometric literature to develop multiple indicator structural models of POLSQ. These are based primarily on recent work by Hayduk (Hayduk, 1996; Hayduk & Glaser, 2000a, 2000b), and recent debates within the structural equation modelling community (for example, in Structural equation Modelling, volume 7 issue 1), and on the SEMNET discussion forum). The models were informed by information systems theory, marketing theory and social psychology.

Figure 4-1: *Process model of methodology phases*

The detailed design of the models is included at the beginning of each of chapters five, six, seven and eight. These chapters describe the detailed, step by step research design for each of these modelling approaches, and then present the results of each step. We believe this provides a more logical flow for the reader than separating the

detailed research steps from the results obtained by several chapters! Where appropriate, detailed explanations of the various technical terms used are included as and when they arise. Some technical terms are included, without a detailed explanation, in this overview of our research phases for the benefit of expert readers who wish to obtain a high-level view of our method.

4.3 Instrument development and exploratory factor analysis (presented in Chapter 5)

This was conducted in two steps: first, developing and pretesting, and administering the survey instrument; followed by exploratory factor analysis to validate the dimensions of online service quality. The survey was conducted online by placing a link on the Victoria University Website.

4.3.1 Instrument development, pre-testing, and survey administration

We describe the development of our survey instrument. This includes specifying the domain of the research constructs, generating scale items based on existing scales and research literature, initial assessments of content validity, and pre-testing.

4.3.2 Exploratory factor analysis

Description: Exploratory factor analysis of the revised model of e-service quality, based in eQual, and incorporating the changes suggested in our literature review in Chapter 2.

Purpose: Apply factor analysis based on previous research and literature. Conduct exploratory analysis on the revised instrument for e-service quality. Perform factor analysis and error variance to identify constructs and indicators and give confidence for their use in subsequent structural equation models.

Procedure: This will be conducted in six steps: determine the suitability of the data for factor analysis, to confirm that factor analysis is appropriate; perform statistical tests on each indicator to confirm that the indicator is suitable for inclusion; determine the number of factors to extract, using a consensus from a range of statistical tests;

determine the most appropriate method of factor analysis and rotation; consider indicator-factor loadings; and iterate and refine the analysis.

Analysis: Discriminant validity will be assessed using scree tests, and testing for eigenvalues greater than 1. Indicator loadings on factors will be examined, using a cut-off value of 0.4 (Gefen, Straub, & Boudreau, 2000). Internal consistency will be measured using Cronbach's alpha and heuristics for exploratory research which suggest a target value above 0.6 (Gefen et al., 2000).

4.4 Confirmatory factor analysis (presented in Chapter 6)

Description: Four alternative confirmatory factor analysis models of online service quality are developed based on theory. We present and compare two models based eQual, a model based on the factor structure we hypothesised from literature, and a model based on our exploratory factor analysis. The survey items are modelled as reflective indicators of latent constructs such as content quality, transaction quality, and usability.

Purpose: Apply the “two-step” SEM modelling process of evaluating the measurement model followed by evaluating the structural model. This step continues traditional, factor-analytic reflective models to represent the online service quality construct. Since the model fit indices were poor, we compare the models, and discuss the possible explanations for the poor fit.

Procedure: This will be conducted in five steps: Specify alternative models based on theory, evaluate the suitability of the data for CFA, select the most appropriate metrics for evaluating the model, evaluate the individual model parameters, and evaluate the overall measurement model using CFA.

Analysis: The data will be evaluated for normality and skew, and appropriate procedures used in the event of non-normal data. List wise deletion will be used for missing data, meaning any response (“case” in SPSS) with missing values was deleted. This will enable us to apply bootstrapping, which is a procedure which is particularly well supported by AMOS for managing non-normal data (Byrne, 2001).

Following this we select the most appropriate fit indices. These are discussed in detail in chapter 6. We evaluate the modification indices and regression weights for individual model parameters to ensure they are suitable for inclusion. Following this, we calculate the estimates for our alternative models using AMOS.

4.5 Formative models (presented in Chapter 7)

Description: Following our review of structural equation modelling literature, we came to believe that many of the indicators of online service quality that had been modelled reflectively were mis-specified and would be more correctly modelled formatively. Applying theory from service quality and information systems literature, and the modelling guidelines discussed earlier in this chapter, we develop two formative models of the online service quality domain. This includes a range of indicators that collectively form perceptions of three formative constructs, ease of use, usefulness and trust. These in turn predict overall perceptions of POLSQ.

Purpose: Apply insights from recent debates within the SEM modelling community to the measurement of online service quality. Apply recent knowledge about the distinctions between formative and reflective indicators to the survey items, and correct previous mis-specification. Determine whether the item set included in our instrument covers the full scope of the domain of user perceptions of online service quality.

Procedure: This will be carried out in four steps. First, do a census of the indicators to ensure that they cover the entire scope of the constructs. Next analyse the direction of the causality, and other characteristics, to identify those indicators that should more correctly be modelled as formative. Then evaluate the model to ensure that appropriate reflective indicators were included to allow the model to be identified. Finally, we evaluate the models.

Analysis: Traditional measures of model fit are not appropriate for formative models. A low error variance on the formative construct and a high value for the squared multiple correlation of the formative construct suggests that a high proportion of the variance in that construct is explained by the hypothesised formative indicators.

4.6 Multi-indicator structural models (presented in Chapter 8)

Description: Develop and fit alternative multi-indicator structural models based on our dataset. As discussed earlier in this chapter, there has been extensive debate in the SEM modelling community about the role of factor analysis as a basis for SEM models. Some methodologists, in particular, Leslie Hayduk (Hayduk, 1996; Hayduk & Glaser, 2000a, 2000b) have advocated strongly for multiple cause structural models, which are not based on the principles of factor analysis. In this phase, we apply this approach to modelling the structural relationships between our indicators.

Purpose: Apply insights from recent debates within the SEM modelling community about “two-step” versus “four-step” modelling to the measurement of online service quality. Apply recent knowledge about multiple cause structural models to our research, and explore the possibility of specifying more granular relationships between our constructs that more closely reflect the evolution from descriptive to inferential beliefs as described in social psychology.

Procedure: This phase will be carried out in three steps. First we develop alternate structural models, based on literature, expert evaluations, vernacular definitions and nomological nets. Then we specify the structural models in detail, including setting error terms. Finally, we evaluated the fit indices.

Analysis: For each model, analysis of a range of appropriate SEM fit indices will be carried out.

4.7 The context of the study

The context of the study was a university web portal. The Victoria University home-page acts as the main portal for information about the university and access to the university’s range of on-line services. An image of the home page is included as Figure 4.2. Information available includes an introduction to the university and a wide range of academic information relating to the degrees available, the subjects that make them up, and the faculties and schools that offer them. This is offered in a variety of different forms, as a structured information hierarchy on the website, and in

the form of various on-line directories. The web-site offers a customised experience for different user segments, the side-bar on the left provides information, navigation, and links targeted to the needs of site users segmented into various categories, for example students of various ethnicities, or students and staff at various stages of their relationship with the university. In addition, the home site provides access to information about the research activities of the university; and overview of student life; links and information about library resources; and other information aimed mainly at current and prospective students, covering topics such as enrolment, distance education offerings, scholarships, and fees.

According to the site owners, the VUW home site is designed for clarity and functionality rather than for entertainment or affect. It is information rich, limited graphics are used, and little or no multi-media. Some on-line transactions are available; these are mostly confined to requesting information packs and forms, and booking resources. However, the portal serves as the main entry point to a variety of other services, such as enrolment, library services, student services, research support, domestic and international student admission.

Figure 4-2: Screen shot of Victoria University of Wellington's web portal

4.8 The dataset used for the study

The detailed process of developing the survey and collecting the data is described in the next chapter (Chapter 5). An overview is given here since the same dataset is used for all of the analysis.

The survey was made available on-line, in the public domain with a link from the VUW home-page (shown as Figure 4.1). This means that anybody, regardless of whether they are a current user of the VUW website, could complete the survey. Participation was voluntary and anonymous. The survey yielded a single, cross-sectional, non-stratified sample. Participants included users of university e-services, including academic and general staff, prospective students, undergraduate and post-experience students.

4.9 References to the dataset in the text

To assist with clarity and readability at different points in the document, a number of abbreviations are used when referring to the observed variables in the various models (which correspond to the questions on the survey instrument. Sometimes the question number (for example “x2”) is used. Sometimes the full text of the question (for example “The site provides accurate information”) is used. Sometimes a short text is used (for example “accuracy”). A table of question numbers, the full text of the questions, and the short text is provided as Appendix I. An example is provided below as Table 4.1.

Table 4.1: *Example of cross reference for question numbers, full text and short text*

Question- Variable number	Full text	Short Text
x2	The site provides accurate information	Accuracy

4.10 The adequacy of the sample

4.10.1 Overview

The next chapters present and compare several different modelling approaches. The same survey, the same set of indicators, and the same data-set are used as the basis for all the models. This means that variation in the results is completely attributable to the various modelling approaches, rather than being potentially explained by differences between the samples or different interaction effects between indicators.

The population from which our sample was drawn was restricted to current, past and potential members of the Victoria University community, and then within that to that subset of those members who might be expected to use the university's online information sources and services at the time the survey was taken. This is difficult to estimate but smaller than the total possible population as no online services were mandatory at the time. The Victoria University website indicates that the university had approximately 21,000 students enrolled in 2008, and approximately 1900 staff (<http://www.victoria.ac.nz/home/about/snapshot.aspx> retrieved 11 June 2009). This means that the survey respondents represent between 0.5 and 1% of the total possible population. More importantly, by placing the survey on the website, respondents were self-selected to be representative of the population of interest (users of online information and services).

We opted not to require the respondents to answer every question, as that would have introduced a response bias. If a respondent did not have an opinion on a question they were forced to answer, then there is a strong likelihood of a meaningless answer. We could have included a not-applicable option, but we chose not to as this option had not been offered in the previous researchers that we intended to compare our results with. Allowing respondents the option of not answering all questions, rather than offering a "no opinion" or "not applicable" option, or simply requiring them to take a position on the scale, is consistent with the approach taken by Barnes and Vidgen in their eQual research, which we used as the starting point for our instrument (Barnes, 2001; Barnes & Vidgen, 2001, 2002).

A total of 250 responses were received. Our approach to data analysis (described in a later section) required us to use complete data without missing values. After data scrubbing involving list-wise deletion of subjects with missing values (the whole response is deleted if any values are missing) we ended up with 147 responses. We note that some statisticians recommend data imputation (using an algorithm to estimate missing values) as preferable to list-wise deletion to minimise the loss of statistical power (Howell, 2009). However, AMOS does not support data imputation for non-normal data, so we were unable to adopt this approach. The data was missing completely at random (MCAR), meaning that there was no consistent pattern in the missing data that might distort the results. This was evaluated using Little's MCAR test (Little, 1988). The results for our dataset (calculated in SPSS 16.0) were: chi-square = 1674.440; $df = 1625$; Sig. = 0.192. Significance greater than 0.05 indicates that the missing data is missing completely at random (Little, 1988; SPSS_Inc, 2007). This means that the cases with missing values (which were deleted) were not systematically different to the cases that were retained, and therefore listwise deletion is acceptable (Howell, 2009).

We note also that research in the related field of human computer interaction (HCI) suggests that broad agreement about the usability of a website (usability is a subset of POLSQ, as we established in Chapter 2) can be achieved with a relatively small number of participants. Jacob Neilson recommends eight different methods for testing the usability of websites with users, and the recommended number of users participating in the study ranges from three (for methods involving observation of actual use) to “at least 30” (for questionnaires) (Nielson, 1993).

The scrubbed dataset was used for all the analysis apart from the initial descriptive statistics, which used the complete dataset. At that stage we had not made a decision as to the approach towards missing data.

4.10.2 Sample size and exploratory quantitative research

Heuristics of exploratory quantitative research suggest that the sample should be 5-10 times the number of questions included in the instrument (Straub & Carlson, 1989). For our study this would require a minimum of approximately 125 participants. Other studies have suggested that data sets used for factor analysis should include at least

five times as many participants as variables (Field, 2005), and up to 10 times as many participants as variables (Nunnally & Bernstein, 1994). Our study, with 147 usable responses and 27 variables, has a ratio 1:5, which is within the accepted guidelines.

4.10.3 Sample size and SEMs

More recently, the issue of sample size has been debated extensively by SEM specialists notably in a special issue on SEM in the journal *Personality and Individual Differences* (volume 42, issue 5). Much of the debate hinges on the representativeness of the sample for the population of interest.

Barret argues strongly against publication of SEM studies with smaller sample sizes. *“SEM analyses based on samples of less than 200 should simply be rejected outright for publication”* but makes an exception if the population from which the population is drawn is itself limited in size *“unless the population from which the sample is hypothesised to be drawn is itself small or restricted in size (measured in the hundreds or thousands).”* (Barret, 2007, p. 820).

Other authors commenting on Barret’s assertions reject the application of blanket rules, and also emphasize the importance of the representativeness of the population. *“In some circumstances, depending on the representativeness of the sample for the population of interest, relatively small samples might be permissible”* (Markland, 2007, p. 855).

Our dataset is highly representative. Research in HCI suggests that relatively small numbers of research participants are required to obtain a consensus about perceptions of the usability of a website (which is one of our quality criteria). In our study, the overall size of the sample frame is unknown but not more than 10,000-20,000. Of these, all the survey respondents are representative of the population of interest by definition. By choosing to place a link on the website, we ensured that only active users of the online service will respond.

A further consideration is that an influential study on determination of sample size required for covariance-based SEM modelling found that the minimum requirements for sample size *decreased* as the degrees of freedom in the model increased.

“Adequate power for the recommended tests can be achieved with relatively moderate levels of N [sample size] when d [degrees of freedom] is not small” (MacCallum, Browne, & Sugawara, 1996, p. 144). This study suggests that a sample size of 132 is adequate for a model with 100 df . Our models typically have in excess of this.

Overall, there are various opinions on sample size in SEM modelling. However, there is not universal agreement that “more is better”. The representativeness of our sample, and the findings of MacCallum et al. that the required sample size for adequate power in model testing decreases with the number of degrees of freedom, suggest that our sample size is acceptable for our study (MacCallum et al., 1996).

4.11 Issues and limitations of the approach

Our approach has a number of issues and limitations. As a single, relatively small sample, there are clearly potential issues with generalisability. However, as we discuss initially in chapter 3, and further in chapter 9, we are less interested in generalisability (which has perhaps been overvalued in information systems research) than in accuracy and salience.

By choosing to use the same data set for all of our analysis we were constrained to using the same set of indicators for every model. This has a number of disadvantages. As we discussed in Chapter 3, many of the indicators that have been modelled reflectively in previous research were probably mis-specified, so it would not be unexpected if we had issues with the confirmatory factor analysis. In developing the formative models (Chapter 7) we run the risk that the indicators would not fully represent the scope of the domain, as we did not follow a process explicitly aimed at identification of formative indicators. However, we identified in Chapter 3 that many of the indicators were formative rather than reflective, and we conducted a literature review in Chapter 2 to identify additional items, so we can have a reasonable degree of confidence in the use of the item-set for developing a formative model. We need to add some reflective items to our multi-indicator structural models to enable them to be identified. Once again, we are somewhat limited by the item set available.

Despite these disadvantages, our approach has a number of strengths. By using the same sample and item set, we can be sure when making comparisons between the models that all the differences in results can be attributed to the modelling method, and are not the result of differences in the sample, the context of the study, or the effect of different interactions between different survey items.

Overall, the approach we have outlined is appropriate for a comparative study of quantitative measurement methods when measuring attitudes and perceptions towards service quality in a specific context (in the case, a university website). However, we will likely need to make some sacrifices with respect to the use of specific modelling methods, particularly in our selection of indicators.

5 Exploratory Factor analysis

5.1 Overview

This chapter describes the first phase of data analysis for this dissertation. In this phase we carried out exploratory quantitative research aimed at establishing a reliable scale for measuring perceived on-line service quality in a university environment. In this chapter, we follow the assumptions of previous researchers from the ServQual and eQual research streams that the constructs should be measured reflectively. The chronology of the research, presented in the introduction, notes that at the time this analysis was performed, we had no particular reason to doubt the appropriateness of the approach. Readers, having read chapter 3, have the benefit of hindsight that we did not have at the time, and will no doubt be able to see the deficiencies. This chapter presents our best attempts to obtain meaningful results from our exploratory factor analysis at the time.

This chapter is structured as follows. We introduce the development of the instrument, questions of validity and reliability, and the process for conducting the survey, the sample used for the study, and the exploratory factor analysis conducted on the survey data. Finally we discuss the results from this phase of the research.

5.2 Instrument Development

5.2.1 Overview of Our Approach to Instrument Development

In this research, we were guided primarily by the approaches recommended by Grover and Hinkin (Grover, 1997; Hinkin, 1998), supplemented by Moore and Benbasat (Moore & Benbasat, 1991). These authors suggest a broadly similar sequence of steps for developing and validating a survey instrument. An integrated view is provided in Figure 5.1. The captions on the left hand side represent our titles for the six-step approach to survey instrument development that we used to guide the process.

This phase of the research covers the specification, scale development, pre-testing and initial survey administration, exploratory analysis, scale purification and reduction steps.

Figure 5-1: *Our approach to instrument development and testing*

5.2.2 Specification

This step involves initial decisions about the type of research, the research design, and unit of analysis, and target respondents for the study (Grover, 1997).

A distinction is frequently made between two types of research, exploratory (theory building) and confirmatory (theory testing) research (Grover, 1997). Between these, there is a middle ground where conceptual refinements are made to existing theory (Straub & Carlson, 1989). This research occupies that middle ground. There is a very extensive body of research for both customer-perceived service quality and customer-perceived on-line service quality. However our literature review in Chapter 2 suggests that new affordances need to be included in measures on POLSQ. This suggests that exploratory research, aimed at making conceptual refinements to existing measures, is appropriate.

The unit of analysis, e.g. an individual, a work group, or a project, needs to be clearly defined from the outset (Grover, 1997). This research measures latent variables (customer perceptions) that are formed as a result of interacting with an on-line service. The unit of analysis is individual customer perceptions of on-line service quality. A possible source of confusion for this research can arise between measuring aspects of the customer's perception, and characteristics of the website itself. The survey instrument for this study needs to keep a clear focus on the customer's internal perceptions of the service quality delivered as the unit of analysis, rather than the objective functionality or usability of the website.

It is important that the respondents chosen are appropriate for the research question (Grover, 1997). The persons most knowledgeable about the construct of interest should be chosen (Grover, 1997). This research requires persons that have interacted with the VUW website. Placing the link to the survey on the website will ensure that respondents that are appropriate to the research question and have knowledge of the constructs of interest are chosen.

Finally, Grover suggests that in this initial step, triangulation (use of multiple methods) is used to assess the variables of interest (Grover, 1997). The author conducted a separate qualitative study of user expectations of service quality in a university web portal, which informed aspects of this study (Tate et al., 2008b; Tate et al., 2007).

5.2.3 Scale Development

This step in the process involves generating and developing items for inclusion in the survey, doing an initial assessment of content validity, and selecting the scale that will be used to measure the responses (Hinkin, 1998). Existing scale items were retained, and scale development literature was used to inform the development of one new construct (transaction quality) and additional items identified from literature.

5.2.3.1 *Generating items*

Useful methods for generating candidate survey items include the use of existing scales (Grover, 1997; Moore & Benbasat, 1991), review of relevant literature (Grover, 1997), and using a panel of experts (Grover, 1997).

There was a considerable body of existing research, and several existing scales available for measuring our dependant variable, perceived on-line service quality. We therefore opted to use an existing scale as the starting point for this research (Boudreau, Gefen, & Straub, 2001; Grover, 1997; Hinkin, 1998). The scale we selected was eQual (formerly WebQual) scale, developed by Barnes and Vidgen. This scale was chosen because it had been developed using multiple rounds of qualitative and quantitative data gathering across multiple business domains and was based on an extensive literature review (Barnes, 2001; Barnes & Vidgen, 2001, 2002). In addition, the eQual authors assert that their scale development process followed the best practices identified from literature. The scale was triangulated using multiple methods. Both deductive (based on existing theory) and inductive (based on input from respondents) methods were used (Barnes, 2001; Barnes & Vidgen, 2001, 2002). Multi-item measures are used for each factor.

In addition, as part of this study, we carried out an extensive, independent literature review (Grover, 1997), which forms Chapter 2 of this thesis. This identified some additional items for existing factors, and hypothesised a new factor, transaction quality.

A formal panel of experts was not used (Grover, 1997), although both of the advisors for this phase of the research (Dr Stuart Barnes and Dr. Beverley Hope) are academics with research experience in the area. In addition, the author has previous experience managing an online channel for a major bank, and consulting to public and private organisations in the area of online service quality.

5.2.3.2 Developing items

Once a set of candidate items has been generated, the actual question wording needs to be developed (Hinkin, 1998). Items should be short, use language familiar to the target respondents, be consistent in terms of perspective, address a single issue, and not be negatively worded (Hinkin, 1998). Each of our new questions related to a single candidate item identified from our literature review. A focus on the user's perceptions, as opposed to the objective characteristics of the site, was maintained, for example, we asked "Overall, the response time was acceptable to me", rather than (say) "The response time was less than six seconds". The target respondents were all web literate and current or prospective university students or staff. The initial questions generated were considered by the advisors for this research, (in their capacity as subject matter experts) to use appropriate language.

5.2.3.3 Initial assessment of content validity

Once the items have been developed, an initial assessment of content validity should be carried out. Various methods have been suggested for assessing content validity. Many involve variations on card sorting techniques; where the items are given to respondents individually, and then sorted into categories (see, for example, Moore & Benbasat, 1991). These categories may be either pre-defined or open and defined by the respondents or can evolve from open to predefined categories over multiple rounds of card sorting (Moore & Benbasat, 1991). We opted not to carry out an extensive card sorting process, as the majority of our items had been confirmed into pre-existing categories in previous studies carried out by Barnes and Vidgen.

An alternative approach to repeated rounds of card sorting is to use a process where respondents match items to definitions (Grover, 1997; Hinkin, 1998). Since our items were relatively short and self-explanatory, it proved difficult to generate definitions that were sufficiently different to the questions themselves to make this process worthwhile. Instead, we opted for a variation on this approach. A panel of six respondents, drawn from the target group of VUW students and staff, was selected for the content validity test. Respondents were given a brief context for the study. Initially, each NEW question was asked to three respondents, who were asked to describe in their own words what the question meant. The answers were recorded. As a further measure of content validity, respondents were asked what variations on the new latent construct “transaction” meant to them. Variations included “transaction” and “task”. Some minor revisions were made based on initial feedback, and the revised questions were trialled with the remaining three respondents. The protocol for the process, and the responses from the participants, are included as Appendix II

A further decision related to the number of items needed to measure each construct. Various heuristics exist for the number of items required to test the homogeneity of items within a construct. It has been suggested that as few as three or four items can provide adequate internal consistency (Harvey, Billings, & Nilan, 1985). Hinkin suggests an eventual goal of between 4 and 6 items for most constructs. It has been suggested that a smaller number of items provides greater confidence, as scales with a large number of items tend to have high internal consistency reliabilities even if item correlations are quite low (Hinkin, 1998). Some researchers suggest starting with approximately twice as many items as you will eventually need and reducing the number of items, particularly for exploratory research (Hinkin, 1998; Moore & Benbasat, 1991). We needed to exercise some judgement with regard to number of new items we generated and the extent of the changes we were making to eQual instrument. Although starting with a larger number of items and reducing them is frequently recommended, we noted that we already had a reasonably lengthy instrument, which we were anxious to keep as short as possible to minimise response bias caused by boredom and fatigue (Hinkin, 1998). We were also keen to minimise the changes to eQual, in order to increase the validity of meta-analysis between our findings and previous eQual studies (Keller & Dansereau, 2001). Accordingly, we

adopted a parsimonious approach. We added three new questions for our proposed new latent construct “transaction quality”, and a handful of new questions suggested by literature to enhance the content validity of existing factors. These were: for service interaction quality, “This site creates a sense of enjoyability or entertainment”, and “I feel in control on this site”; and for Usability “Overall, the response time was acceptable to me”.

5.2.3.4 Selecting the scale

Once the questions have been developed, a decision needs to be made about the scale that will be used for the responses. A likert scale is the most commonly used in behavioural research, and has been found to be suitable for factor analysis (Hinkin, 1998). In order to be consistent with previous eQual research, a seven-point likert scale was adopted.

5.2.3.5 Managing Common Method Variance

Common method variance (CMV) is variance that is attributable to the measurement method or source rather than the constructs the method measures. CMV has been widely debated by organisational researchers (Brannick, Chan, Conway, Lance, & Spector, 2010; Lance, Dawson, Birkelbach, & Hoffman, 2010; Pace, 2010; Podsakoff, Podsakoff, MacKenzie, & Lee, 2003; Richardson, Simmering, & Sturman, 2009). There is no clear agreement about the impact of CMV in organisational research such as ours, which involves “self-reports” (where subjects provide their own evaluation of an organisational phenomenon, without collecting further data on the same phenomenon from another perspective). Some organisational researchers consider all “self-report”, or “same-source, same-method” studies to be highly likely to suffer from common method bias (Cote, 1987; Podsakoff et al., 2003) others consider the existence of, or influence of common method variance to be over-rated (Spector, 2006). Some researchers have further suggested that CMV likely exists in all same-source same-method studies, and that the method effects vary depending on the nature of characteristics of the rater, the item, the construct and the context of the study (Cote, 1987)

There is debate over the extent to which common method effects should be considered a problem. Lance et al. point out that “method” is a broad term; and argue

that some of the *variance* arising from common measurement “method” (which as we recall can include the rater, the item, the construct and the context) is not *bias* (that needs to be fixed or controlled), but a theoretical finding of interest (Lance et al., 2010). For example, if perceptions of online service quality were different between evaluations in an online survey from regular users of a website and evaluations in a paper-based survey from people that seldom or never used it; the variance potentially arising from different raters would be a legitimate source of variance, not a rater and context bias that needed to be controlled by conducting (for example) the survey using both online and paper-based collection methods

Given that the research literature is inconclusive, if we allow that CMV may be a potential issue, there are a number of possible ways of managing it. These include procedural and statistical techniques. The efficacy of statistical techniques has been found to be limited in a study by Richardson et al. (2009) “*applying a statistical correction does not necessarily produce more accurate estimations of the relationships that doing nothing. Overall...when CMV was present in the data, the absolute correction accuracy of all [statistical] techniques tended to be low*” (Richardson et al., 2009, p. 793). We opted not to adopt statistical correction techniques.

Procedural techniques for managing the effects of CMV arising from the rater, the item, the construct, and the context are explored in a paper by Podsakoff et al. (Podsakoff et al., 2003). Common rater effects relate to the fact that the respondent providing the measure of the independent and dependent variables is the same person (Podsakoff et al., 2003). Item effects relate to the ways the items are presented, such as the extent to which they appear ambiguous, complex, or socially acceptable (Podsakoff et al., 2003). Item context effects arise from the way in which an item is presented in relation to other items; if a set of items appears to be related, the respondent may give them a similar rating (Podsakoff et al., 2003). Finally, measurement context is the time, location and media used to measure the constructs (Podsakoff et al., 2003).

It has been noted that self-report measures are necessary for certain types of research. “*The use of self-report measures is not only justifiable but probably necessary when*

assessing constructs that are self-referential perceptions such as job satisfaction, perceived organisational support, and perceived fairness. To find out about the perception of an individual, it is often best to ask the individual about his or her perception rather than infer it directly or indirectly from what others observe about the individual's behaviours." (Brannick et al., 2010, p. 416).

Our use of self-report measures is justified given that we are conducting a study involving self-referential perceptions of online service quality. Nevertheless, since our study is conducted using a same-source same-method approach, the possibility of CMV should be considered. We considered potential issues of common method bias in our survey design.

In our view, some of the potential common rater effects, such as a sense of the social desirability of certain answers, are unlikely to be applicable to our study, as the context is socially neutral and uncontroversial (Podsakoff et al., 2003). Others, such as the influence of the rater's mood, could potentially apply (Podsakoff et al., 2003). However, over a representative sample of the population, it is expected that this might result in some standard deviation around the mean for each item, rather than affecting the overall pattern of responses. Item characteristic effects were minimised by reusing, where possible, exhaustively pre-tested and validated items from other studies. By permitting non-response, we allowed respondents the option of not responding to questions they found to be unclear or had no opinion on. Item context effects were minimised by presenting the questions in a random order that did not represent the construct structure hypothesised by the researchers (Podsakoff et al., 2003). We do not consider measurement context to be an issue for our study. In fact, it is important that both the assessment of detailed items, and the evaluation of the overall perceived quality are made at the same time, otherwise there is a possibility that either the web-site or the subject's perception on some items could have changed in the meantime.

Overall, the use of a same-source, same-method approach is appropriate for our study. The characteristics of the study require a consistent context and a point-in-time measure, and the context is socially neutral. Potential item and item context CMV effects were minimised as far as possible using procedural techniques.

5.2.3.6 Discussion

Some difficult decisions needed to be made during this stage of the research. Since it was based extensively on the existing body of eQual research, it could not be considered to be fully exploratory. However, since research in perceived service quality (both on-line and face-to-face) has been characterised by unstable dimensionality, combined with rapid changes on the objective qualities of the phenomenon of interest (on-line service websites) it did not seem appropriate to carry out further research without considering whether revisions to the scale were required. Conceptually, this places our research as a conceptual refinement of existing theory (Straub & Carlson, 1989). Heuristics exist for both exploratory and confirmatory survey research, although these are not cast in concrete. A grey area where judgement is required applies in between.

We opted to make minimal changes to the existing eQual instrument, in order to minimise the risk to the existing, tested scales, and increase the possibility of meta-analysis between this study and the previous eQual studies (Keller & Dansereau, 2001). However, since we had made changes to the instrument, we also followed recommended best-practice, by conducting exploratory analysis and revalidating the instrument (Boudreau et al., 2001; Straub, Boudreau, & Gefen, 2004).

5.2.4 Pre-testing and initial survey administration

Once initial content validity has been established, the next step is to carry out a pre-test to establish length and wording (Moore & Benbasat, 1991). In this step, we also piloted our test for manipulation validity (Straub et al., 2004). Manipulation validity is concerned with assurance that the subjects responding to the survey had actually had the experience they were being asked to report on.

The pre-test was conducted by placing the survey on-line but not linked to the VUW home-page. Pre-test participants were supplied with the URL. The on-line survey was pre-tested by 8 experts, predominantly information systems faculty and post-graduate students. This group is representative of the sample population. The respondents were

asked to complete the survey and to record their experiences on a hard-copy form that was provided to them. The results for the pre-test are included in Appendix III.

The pre-test identified some minor issues with presentation and wording that were corrected. The majority of the feedback, and the biggest issue for the researchers, related to a trade-off between length and manipulation validity.

To provide manipulation validity, it was essential to ensure that participants had interacted with the website before recording their perceptions of the service quality. Participants were required to carry out some on-line tasks before completing the survey. Initially four tasks were selected, including obtaining fee and timetable information. These tasks were designed in consultation with the VUW Marketing group, and were intended to provide a representative experience of interacting with the website. We selected tasks for which there was a high degree of demand, and with varying levels of objective usability, according the Victoria University Marketing department. Three of the tasks were considered reasonably “simple”. The fourth task, obtaining fee information, was known to be difficult to carry out successfully on the site at the time of the test.

Participants felt that the tasks required were too time-consuming and onerous. One pre-test participant commented,

“The tasks to perform weren't that simple for me. Well, the truth is that I haven't used the VUW site much and I didn't know where to look at to start with. There are a few places where the information could've been located and getting to the right place took me a few minutes. I completed all four tasks because I knew it was for you; if the survey was from someone I didn't know, I don't honestly know if I would've gone through all the tasks.”

We were initially unwilling to make the tasks simpler. We felt it was important to the study that the participants had a representative experience of trying to obtain on-line services on the site. On the other hand, we felt that the risk of deterring potential participants was quite high if the tasks were retained in their original form. Eventually a decision was taken to reduce the number of tasks to two. Both tasks could be completed in a reasonable time-frame, but would require some persistence and were unlikely to be able to be answered from general knowledge of the institution without

actually interacting with the site. Based on this feedback and the subsequent modifications, we expected that the tasks and questionnaire could be completed in 15-20 minutes and would achieve a reasonable completion rate without abandonment. We ensured that participants had completed the manipulation tasks by requiring them to enter the answers to the manipulation questions as part of the survey response. Responses that did not include answers for the manipulation tasks were discarded.

The completed survey instrument, supporting information, and instructions for participants are included in Appendix IV.

5.3 Our Approach to Exploratory Factor Analysis

5.3.1 Conducting the Survey

The survey was made available on-line, in the public domain with a link from the VUW home-page (shown as Figure 4.1). This means that anybody, regardless of whether they are a current user of the VUW website, could complete the survey. Participation was voluntary and anonymous. A prize draw of \$200 worth of entertainment (dinner, book vouchers, or CD vouchers) was provided as an incentive for participants. Participants needed to supply their e-mail address in order to be eligible for the prize draw, but this was stored separately to the survey responses. The survey was available for 4 weeks on the VUW home site and initially attracted about 30 responses. After four weeks, the survey was also placed on the student portal. This is an authenticated environment for current students only, and is used to deliver VUW on-line services that are restricted to currently enrolled students, for example the VUW on-line learning environment. This attracted a much better response. After another three weeks, 250 entries had been received, 85% of them from current students. As we noted earlier, after analysing our data for missing values, we ended up with 147 usable responses. As the survey was an anonymous, on-line survey, it was not considered practical to offer the option to individual participants of requesting feedback about the findings of the survey. Instead, summarised results were posted on the home-site, linked from the home page, for a two-week period shortly after the survey closed.

5.3.2 Descriptive statistics

Most respondents were currently members of the VUW community, either students or staff (Table 5.1). A further 1.7% of respondents were former students. Only 3.7% were prospective students, who might therefore be expected to have limited experience with the site. A further 4.9% of respondents did not specify their role, or selected “other” rather than one of the options offered. This means that the vast majority of the respondents were likely to have some level of familiarity with the organisation, the web-site, and the kind of information and services available, and would therefore fit the research objective of identifying genuine consumers of the service being evaluated.

Table 5.1: *Roles held by survey respondents*

Role		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Prospective Student	9	3.7	3.8	3.8
	Current Student	207	84.1	88.5	92.3
	Current Staff Member	14	5.7	6.0	98.3
	Former Student	4	1.6	1.7	100.0
	Total	234	95.1	100.0	
Missing	System	12	4.9		
TOTAL		246	100.0		

This is supported by Table 5.2, which shows the results for frequency of access. This shows that 62.1% of all respondents were very regular users of the site, with a further 25.1% using it at least once a month. Only 12.8% of respondents used the site less than once per month. This suggests a reasonably high degree of familiarity with the site. In terms of the research objectives, this is a good result, and it mitigates our earlier concerns about reducing the requirement for respondents to interact with the site before completing the survey.

Table 5.2: *Frequency of access of survey respondents*

Access Frequency		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	At least once a week	151	61.4	62.1	62.1
	At least once a month	61	24.8	25.1	87.2
	Less than once a month	25	10.2	10.3	97.5
	This is my first visit	6	2.4	2.5	100.0
	Total	243	98.8	100.0	
Missing	System	3	1.2		
TOTAL		246	100.0		

The age-range of the respondents was concentrated on under 30-year olds (89.3% of the total), with 59.3% being between 20 and 25 (Table 5.4). This reflects the typical demographic of current university students – the largest group of respondents.

Table 5.3: *Age of survey respondents*

Age		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under 20	48	19.5	19.8	19.8
	20-25	146	59.3	60.1	79.8
	26-30	23	9.3	9.5	89.3
	31-40	14	5.7	5.8	95.1
	41-50	9	3.7	3.7	98.8
	Over 50	3	1.2	1.2	100.0
	Total	243	98.8	100.0	
Missing	System	3	1.2		
TOTAL		246	100.0		

5.3.3 Descriptive Statistics

The descriptive statistics for the exploratory survey are included as Table 5.5. Several interesting observations can be made about these responses. Firstly, the standard deviations were low overall, suggesting that there is a relatively high level of agreement from the respondents about both the importance of the website qualities being evaluated, and the score for the VUW website. Since we have already established that 87% of respondents were very regular (more than once per week) or fairly regular (more than once per month) users of the site, it can be inferred that there

is a reasonably high degree of consensus, from a group of informed and experienced users, about the quality of the VUW website with respect to this criteria.

Table 5.4: *Descriptive statistics*

Descriptive Statistics	Score	
	Mean	Std. Dev
1. I find this site easy to learn to operate	3.00	1.357
2. The site provides accurate information	2.29	1.095
3. This site has a good reputation	3.06	1.305
4. I believe I would be able to complete transactions that are useful to me	2.67	1.323
5. My interaction with the site is clear and understandable	2.86	1.430
6. It feels safe to complete transactions	2.28	1.087
7. The site provides believable information	1.95	1.044
8. Completing transactions on this site will save me time or money	2.68	1.502
9. I find the site easy to navigate	3.27	1.624
10. The site provides timely information	2.86	1.357
11. My personal information feels secure	2.41	1.310
12. This site offered the range on on-line transactions I expected	2.74	1.281
13. I find the site easy to use	3.05	1.475
14. The site provides relevant information	2.39	1.154
15. This site creates a sense of personalisation	4.02	1.663
16. This site has an attractive appearance	3.13	1.267
17. The site provides easy to understand information	2.66	1.095
18. The design is appropriate to the type of site	2.45	1.141
19. I feel in control on this site	3.10	1.420
20. The site conveys and sense of competency	2.74	1.190
21. This site conveys a sense of enjoyability or entertainment	4.32	1.531
22. The site creates a positive experience for me	3.32	1.344
23. The site provides information at the right level of detail	2.84	1.351
24. This site makes it easy to communicate with the organisation	3.01	1.446
25. Overall, the response time was acceptable to me	2.68	1.306
26. The site presents the information in an appropriate format	2.49	1.186
27. I feel confident that the goods/services I have requested will be delivered as promised	2.80	1.297
Average	2.64	1.189

5.3.4 Exploratory Factor analysis (EFA)

In essence, exploratory factor analysis is a data reduction technique that aims to explain the scores on observed variables (survey items). As we discussed earlier, exploratory factor analysis is widely used in information systems and other social sciences.

5.3.5 Overview of our approach to EFA

In this section, we describe our approach to exploratory factor analysis of our survey data. In overview, this included the following steps: determine the overall suitability of the data for factor analysis, ensure that each variable is suitable for inclusion in the factor analysis, determine the number of factors to extract, determine the most appropriate method of factor analysis and factor rotation, consider the item-factor loadings, and refine the analysis (for example, by removing cross-loading items). An overview of our approach is provided in Figure 5.2.

Guidelines for conducting EFA emphasize that it is an inexact process, and suggest using a range of different tests, if possible, to establish a consensus about the patterns observed in the data. For example, when determining the number of factors to extract, Habing notes "*There is no best rule to use, so use all three and use the closest thing to a majority decision*" (Habing, 2003, p. 4). Choosing the best method of rotation is similarly inexact; "*many researchers conduct both orthogonal and oblique rotations and then report the clearest and easiest to interpret. You are hoping for...a simple structure*" (Pallant, 2005, p. 176). Due to the wide range of criteria and heuristics available for EFA, a very large number of assessments were conducted in the course of this analysis. Reporting the full range of statistical assessments conducted would be prohibitively time consuming and in many cases, irrelevant. We have chosen to report the assessments most commonly used for the various stages of EFA, as recommended by Pallant and Field (Field, 2005; Pallant, 2005). Where these assessments do not provide a clear consensus, we provide a justification for our decision.

Figure 5-2: *Our approach to exploratory factor analysis*

5.3.6 Determine the overall suitability of the data for factor analysis

First, it is important to establish that your data set is suitable for factor analysis. We previously established (in Chapter 4) that the sample size fell within recommended guidelines.

5.3.6.1 *Degree of correlation between variables*

Common sense tells us that variables that are supposed to be measuring the same underlying factor should be correlated. The degree of correlation is usually suggested to be reasonably high, but not perfect, or this would suggest that the two perfectly correlated variables are measuring exactly the same thing. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is a commonly-used measure of the pattern of correlation between variables. A value of 0 indicates a high degree of diffusion in the pattern of correlations, suggesting that factor analysis is inappropriate. A value close to 1 suggests that the patterns of correlations are quite compact and therefore factor analysis is appropriate. A KMO value above .9 is considered ideal (Field, 2005). The KMO measure for this sample is .93.

The null hypothesis, that the variables are not co-related, can be tested using Bartlett's test of sphericity. An insignificant result would suggest that there is no correlation between the variables, and therefore factor analysis is not appropriate. A significant result ($<.05$) tells us that there are some relationships between the variables (Field, 2005). In this sample Bartlett's test is highly significant ($p<.01$) suggesting factor analysis is appropriate. Results for KMO measure of sampling adequacy and Bartlett's test of sphericity are shown in table 5.6

Table 5.5: *KMO measure of sampling adequacy and Bartlett's Test of sphericity*

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.930
Bartlett's Test of Sphericity	Approx. Chi-Square	2934.374
	Degrees of freedom	351
	Significance	.000

5.3.6.2 “Spot check” for linearity

Factor analysis is based on correlation, and it assumes a linear relationship between the variables. It is not practical to carry out a scatter plot of every variable against every other variable, but a spot check can be carried out on selected variables (Field, 2005; Pallant, 2005). These should show no evidence of a curvilinear relationship. Two sample scatter plots were generated between variables that were hypothesised to load to the same factor (information believability and information accuracy, hypothesised to load to content quality, and transaction importance and transaction safety, hypothesised to load to transaction quality). These are shown in Figure 5.3. Although these do not show clear linearity, there is no evidence of a curvilinear relationship, suggesting that factor analysis would be inappropriate.

Figure 5-3: Sample scatter plots of variables

5.3.7 Ensure each variable is suitable for inclusion

5.3.7.1 *KMO values*

We examined the KMO value for individual correlations between variables by examining the diagonal elements of the anti-image correlation matrix. This test will determine if there are any individual variables with very low correlations with others that should be considered for removal from the analysis. A minimum value of 0.5 is suggested for variables to be retained for the factor analysis (Field, 2005). The individual correlations between variables in this study all had a KMO value of 0.7 or higher. This suggests that all the variables should be retained for the analysis.

5.3.7.2 *Correlation matrix*

We then examined the correlation matrix. It has been suggested that the correlation matrix should have coefficients of .3 or above for the data to be considered “factorable” (Field, 2005). There should be reasonably high inter-correlations between variables relating to sub-traits of the same factor, and any variables that don’t correlate with any other should be excluded from the factor analysis run. It has also been suggested that any variables that correlated too highly (>0.9) with another should be excluded, because it is impossible to determine the individual contribution of these variables to the total variance (Field, 2005). This data set meets these criteria. All the coefficients in the correlation matrix are >0.3 and less than 0.9, although navigability (9) correlates with easy to use (13) at .817.

5.3.7.3 *Determinant of the correlation matrix*

The determinant of the correlation matrix is not $>.0001$. Some theorists suggest that this means some variables should be eliminated (Field, 2005). However, this needs to be interpreted in the context of other criteria and an overall judgement made (Field, 2005).

5.3.8 Determine the number of factors to extract

Having established the suitability of the data for factor analysis, the next step is to determine the number of factors to extract. There are a number of criteria available for determining the number of factor to extract. It has been suggested that there is no single “best” rule, and that it is valuable to use a range of methods (Habing, 2003).

5.3.8.1 *The Eigenvalue rule (Kaiser's criterion)*

One of the most commonly used techniques for determining the number of factors to extract is Kaiser's criterion, or the eigenvalue rule. The eigenvalue of a factor represents the total amount of variance in the data that is explained by that data. Using Kaiser's criterion, the number of factors with an eigenvalue of 1.0 or more are retained (Field, 2005)(Field, 2005; Pallant, 2005). This is also the default criterion used in SPSS. Although this is a widely used criterion, it is not without its critics. It has been suggested that retaining as many factors as have an eigenvalue >1.0 can result in too many factors (Pallant, 2005), or too few (Field, 2005). Research has suggested that taking the number of factors with an eigenvalue of >1.0 is most reliable when there are 20-50 variables (Hair, Anderson, Tatham, & Black, 2003) and the communalities are greater than 0.7 (Habing, 2003). If the sample size is >250 , average communalities of 0.6 can be considered acceptable (Field, 2005). Our data-set falls just barely outside these suggested guidelines, with a sample size of 147 and average communalities of .617. This suggests that Kaiser's criterion for the maximum number of factors to extract, using eigenvalues above 1.0, is likely to be appropriate in this situation. The eigenvalues and total variance explained are shown in Table 5.7. This suggests that up to five factors should be retained.

Table 5.6: *Total variance explained*

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.328	41.957	41.957	11.328	41.957	41.957
2	2.325	8.610	50.567	2.325	8.610	50.567
3	1.327	4.915	55.482	1.327	4.915	55.482
4	1.202	4.451	59.934	1.202	4.451	59.934
5	1.015	3.760	63.694	1.015	3.760	63.694
6	.876	3.244	66.938			
7	.820	3.036	69.974			
8	.748	2.771	72.745			
9	.672	2.490	75.234			
10	.663	2.455	77.689			
11	.635	2.351	80.040			
12	.564	2.088	82.128			
13	.521	1.928	84.056			
14	.466	1.726	85.782			
15	.448	1.658	87.439			
16	.416	1.542	88.982			
17	.404	1.496	90.478			
18	.382	1.415	91.893			
19	.337	1.250	93.143			
20	.319	1.180	94.323			
21	.284	1.052	95.375			
22	.253	.936	96.311			
23	.248	.917	97.228			
24	.226	.839	98.066			
25	.216	.801	98.867			
26	.187	.691	99.558			
27	.119	.442	100.000			

Extraction Method: Principal Component Analysis.

5.3.8.2 *Catell's scree criterion*

Another popular test for the number of factors is Catell's scree criterion (Pallant, 2005). This criterion involves plotting the eigenvalues of the factors, and looking for the point at which the plot begins to level off from vertical to horizontal. Some research has suggested that Catell's scree criterion tends to keep one or more factors more than the eigenvalue rule (Hair, Anderson, Tatham, & Black, 2003), or that both the eigenvalue rule and the scree plot are accurate if the sample is >250 and the commonalities are greater than 0.6. In this case, a slightly different result is suggested by the scree plot, which would suggest the retention of three significant factors (Hair et al., 2003). The results of Catell's scree plot are shown in Figure 5.4.

Figure 5-4: *Catell's Scree Plot*

5.3.8.3 *Parallel analysis*

A technique that is becoming increasingly popular in the social sciences, particularly psychology and education, is Horn's parallel analysis (Pallant, 2005). This involves comparing the size of the eigenvalues from the research data-set with those obtained from a randomly-generated data-set of the same size. Only those eigenvalues that are larger than the corresponding values (criterion values) from the random data set are retained (Pallant, 2005). This criterion indicates those factors whose eigenvalues are bigger than the reference, are therefore likely to be indicative of factors that would not appear by chance. Freeware for conducting parallel analysis was used ("Monte Carlo PCA for Parallel Analysis" by Marley Watkins, available from http://freewareapp.com/monte-carlo-pca-for-parallel-analysis_download/). The parallel analysis generated for this data-set is included as Table 5.8, and the comparison between the actual eigenvalues from the PCA and the criterion values from the parallel analysis is shown in Table 5.9.

This criterion suggests a smaller number of factors than the previous two criteria. Based on the results of parallel analysis, only two factors would be retained. In this research, the various heuristics available for determining the number of factors to extract did not produce a consensus. Anything between two and five factors could be supported.

Table 5.7: *Parallel Analysis*

Eigenvalue Number	Random Eigenvalue	Std. Deviation
1	1.6772	.0500
2	1.5696	.0407
3	1.4929	.0403
4	1.4260	.0365
5	1.3628	.0302
6	1.3073	.0277
7	1.2587	.0263
8	1.2126	.0236
9	1.1640	.0216
10	1.1215	.0204
11	1.0803	.0223
12	1.0414	.0197
13	1.0036	.0210
14	0.9672	.0195
15	0.9282	.0186
16	0.8948	.0174
17	0.8618	.0196
18	0.8249	.0190
19	0.7903	.0210
20	0.7551	.0215
21	0.7197	.0215
22	0.6863	.0205
23	0.6496	.0217
24	0.6134	.0222
25	0.5744	.0233
26	0.5333	.0224
27	0.4833	.0289
Number of variables		27
Number of subjects		147
Number of replications		100

Table 5.8: *Comparison between actual eigenvalues and parallel analysis*

Component Number	Actual eigenvalue from PCA	Criterion value from parallel analysis	Decision
1	11.328	1.6772	Y
2	2.325	1.5695	Y
3	1.327	1.4929	?
4	1.202	1.4260	?
5	1.015	1.3628	?

5.3.8.4 *Judgement and previous research findings*

Another acceptable method for determining the number of factors is “a priori” (Habing, 2003), who suggests “*if you have a hypothesis about the number of factors that should underlie the data, then that is probably a good (at least minimum number) to use*” (Habing, 2003, p. 4). Previous research in this area has found support for between three and five factors. We opted to extract four factors, which was the number we initially hypothesised.

5.3.9 Determine the most appropriate method of factor analysis and factor rotation

There are a number of different approaches to discovering factors. Due to the fact that approaches with different underlying mathematics often yield similar results, there tends to be a lack of rigour in the use of the terms such a principal component analysis and factor analysis (Field, 2005). In this section, we briefly review the methods available for discovering factors and our approach for selecting the most appropriate method.

5.3.9.1 *Method of extraction*

Exploratory factor analysis is useful for exploring data rather than testing hypotheses. This can be useful for generating future hypotheses. This research aimed to generate hypotheses about the nature of perceived on-line service quality, for inclusion in a future confirmatory study, so EFA is appropriate.

It should be noted that the terms principal component analysis and factor analysis are often used inter-changeably, although they are based on different mathematics. Principal components analysis is only a data reduction method and is heavily data

driven. It is calculated without regard to any underlying structure caused by latent variables; components are calculated using all of the variance in the data, and all of that variance appears in the components constructed in the final solution (Ford, MacCallum, & Tait, 1986). However, researchers rarely collect and analyse data without some idea about how the variables are, or should be related. In factor extraction the co-variance between variables is partitioned from each variables unique variance and error variance to reveal the underlying factor structure; only shared variance appears in the solution. This partitioning of variance is also the approach taken in co-variance based methods such confirmatory factor analysis.

Field, while acknowledging the underlying differences, notes that for most practical purposes, the solutions that are produced are very similar, and principal component analysis (PCA) has been found to be less complex than factor analysis (Field, 2005). Principal components and factors are often conflated by researchers who use the terms interchangeably, and SPSS includes PCA as an option under the “Factor analysis” tab. This phase of the research uses principal component analysis, primarily because that is the technique used by the eQual authors in their iterations of the instrument. This study aims to compare the revised instrument with results from previous eQual studies. We follow the practice of the eQual authors by describing the components extracted as factors.

5.3.9.2 Method of rotation

Once the factors have been extracted, it is possible to calculate the degree to which variables load onto these factors. There are two commonly-used types of rotation. Orthogonal rotation assumes that all the factors identified are independent and do not correlate with one another at all. Oblique rotation assumes that there may be a degree or correlation between the factors. Within these, there are several methods of rotation that are available in SPSS and used in social science research. Varimax rotation is the most popular orthogonal rotation method, because it aims to maximise the dispersion of variable loadings within factors. This tends to produce a smaller number of variables loading more highly onto each factor. This method often produces a clearer and simpler structure which is easier to interpret (Conway & Huffcut, 2003; Field, 2005). Direct oblimin is the most popular oblique rotation. In overview, this method controls the degree to which the factors are allowed to correlate.

Although orthogonal rotation is very widely used in information systems and other social science research (possibly because it is the default setting in SPSS, and simpler to explain and interpret), there is a school of thought that suggests that it is unrealistic to assume that naturalistic data involving human perceptions will produce entirely independent factors, and therefore oblique rotation should be used for data involving human perceptions (Conway & Huffcut, 2003; Field, 2005). Both orthogonal and oblique rotations often produce similar solutions, especially when the pattern of correlations among the items is clear (Pallant, 2005). Pallant suggests that researchers can conduct both orthogonal and oblique rotations, and report the result that is the clearest and easiest to interpret.

In this research, we opted to carry out both orthogonal varimax rotation and oblique direct oblimin rotation. The reasons for this decision include the fact that previous studies, in particular, the eQual research used orthogonal varimax rotation, and the results from this sample can be more easily compared with those studies if the same analysis method is used. However, we also find the argument that factors relating to human perceptions are unlikely to be entirely uncorrelated convincing, suggesting that oblique direct oblimin is appropriate. Further, the exploratory nature of the research suggests that it is appropriate to use multiple methods in order to derive the clearest and simplest factor structure (Pallant, 2005).

5.3.10 Consider Item and Factor loadings

Once a factor structure has been found, the researcher can consider which variables are associated with each factor. It should be relatively easy to determine the significance of a variable loading on a factor, but in fact, as with many other aspects of EFA, there are a range of opinions.

5.3.10.1 Sample size and strength of factor loadings

Interpretation of factor loadings can vary depending on sample size, and whether the amount of variance in the variable that is explained by that factor can be considered practically significant.

A loading of more than 0.3 is often considered to be important (Field, 2005). However, this varies considerably with sample size (see Table 5.10). Smaller loadings can be considered significant with increasing sample sizes (Field, 2005; Pallant, 2005).

Table 5.9: *Sample size and factor loading*

Sample size	Factor loading considered significant
50	.722
100	.512
200	.364
300	.298
600	.21
1000	.162

A factor loading of 0.4 explains only 16% of the variance in the variable. Factor loadings that are below 0.4 are not of much practical interest. Another commonly-cited guideline (Table 5.11) suggests the following guidelines for determining the practical significance of factor loadings (Hair et al., 2003).

Table 5.10: *Factor loading and practical significance*

Factor loading	Interpretation
≥ 0.3	Minimal significance
≥ 0.4	More important
≥ 0.5	Practically significant

5.3.10.2 Initial solution

The initial solution, using both orthogonal and oblique rotations, was inconclusive, with a large number of cross-loading items at above 0.4, which is considered significant (Hair et al., 2003). The rotated component matrix, using Principal component Analysis and Varimax rotation (Table 5.12); and the Pattern Matrix (Table 5.13) and Structure Matrix (Table 5.14) using Principal component analysis and oblimin rotation, are included.

Table 5.11: *Initial solution: PCA with Varimax (orthogonal) rotation (values below 0.4 not reported)*

Rotated Component Matrix	Component			
	1	2	3	4
x14 relevance	.747			
x2 accuracy	.740			
x7 believable	.688			
x10 timeliness	.685			
x 17 easy to understand	.642			
x20 competency	.597			
x27 goods-srvcs will be dlvr'd	.577			
x18 design appropriate	.555			
x23 level of detail	.538			
x26 appropriate format	.472			
x9 navigability		.833		
x1 learnability		.807		
x13 easy to use		.787		
x5 clear and understandable	.400	.730		
x19 in control		.669	.402	
x4 txn that are important		.631		
x3 reputation		.587		
x21 enjoyability			.765	
x22 positive experience		.418	.582	
x15 personalisation		.482	.487	
x24 communicate with org			.483	
x16 attractive			.451	
x25 response time	.429		.440	
x6 txns feel safe				.738
x11 personal info feels secure	.423			.663
x8 txns save time-money				.598
x12 range of txns	.474			.486

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 7 iterations.

Table 5.12: *Initial solution, Pattern Matrix: PCA with Oblimin (oblique) rotation (values below 0.4 not reported)*

Pattern Matrix	Component			
	1	2	3	4
x1 learnability		-.890		
x2 accuracy	.824			
x3 reputation		-.597		
x4 txns that are important		-.664		
x5 clear and understandable		-.738		
x6 txns feel safe			.635	
x7 believable	.765			
x8 txns save time-money			.478	
x9 navigability		-.911		
x10 timeliness	.713			
x11 personal info feels secure	.592		.459	
x12 range of txns	.605			
x13 easy to use		-.845		
x14 relevance	.801			
x15 personalisation		-.499		.422
x16 attractive			-.474	
x17 easy to understand	.627		-.446	
x18 design appropriate	.541			
x19 in control		-.684		
x20 competency	.576			
x21 enjoyability				.724
x22 positive experience				.463
x23 level of detail	.644			
x24 communicate with org	.432			.406
x25 response time	.477			
x26 appropriate format	.575			
x27 goods-srvcs will be dlvr'd	.643			

Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization. Rotation converged in 19 iterations.

Table 5.13: *Initial solution, Structure Matrix: PCA with Oblimin (oblique) rotation 9 values below 0.4 not reported)*

Structure Matrix	Component			
	1	2	3	4
x1 learnability	.412	-.845		
x2 accuracy	.782	-.423		
x3 reputation	.485	-.710		.411
x4 txns that are important	.478	-.688		
x5 clear and understandable	.622	-.853		
x6 txns feel safe	.437		.686	
x7 believable information	.780	-.481		
x8 txns save time-money	.439	-.468	.531	
x9 navigability	.424	-.886		
x10 timeliness	.758	-.523		
x11 personal info feels secure	.654		.553	
x12 range of txns	.701			.494
x13 easy to use	.448	-.851		
x14 relevance	.821	-.519		
x15 personalisation		-.609		.574
x16 attractive	.404			
x17 easy to understand	.601			
x18 design appropriate	.692	-.554		.406
x19 in control	.513	-.799		.482
x20 competency	.763	-.631		.457
x21 enjoyability		-.479		.795
x22 positive experience	.616	-.649		.673
x23 level of detail	.698			
x24 communicate with org	.607	-.408		.570
x25 response time	.614	-.411		.503
x26 appropriate format	.635			
x27 goods-srvcs will be dlvr'd	.741	-.472		.403

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

The components were correlated (Table 5.15). This initial factor structure did not meet the recommended heuristics for the convergent and discriminant validity of the constructs (the degree to which items thought to reflect a construct converge, and the degree to which items posited to reflect the construct differ from those that are not believed to make up the construct). Validation heuristics for survey research suggest that for convergent validity, items should load on the posited construct at least at a 0.4

level, and for discriminant validity, there should be no cross-loading items above the 0.4 level (Straub et al., 2004). Accordingly, we continued to refine the analysis.

Table 5.14: *Initial solution, Component Correlation Matrix*

Component	1	2	3	4
1	1.000	-.537	.158	.334
2	-.537	1.000	-.065	-.335
3	.158	-.065	1.000	.066
4	.334	-.335	.066	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

5.3.11 Refine the analysis

The initial solution, using both orthogonal and oblique rotations, had a number of cross-loading items. We continued to explore the factor structure using exploratory factor analysis in SPSS. The most promising results came from oblique rotations, suggesting a degree of correlation between the factors. We experimented with omitting a number of cross-loading questions to improve the clarity and simplicity of the factor loadings. The most promising solutions were identified and compared in order to arrive at a consensus.

The initial oblique rotation using oblimin with kaiser normalisation showed four factors, with questions 11, 15, 17, and 24 cross-loading (Table 5.13). Removing questions 11 and 24 (alone, or along with other cross-loading questions) did not improve the solution. However, removing questions 15 and 17 provided a clear four-factor structure, with a high degree of consensus for factor-item loadings. The only cross-loading item that remained was question 11. Excluding additional questions (13, 15, and 17, or 15, 17, and 20) yielded a clean four-factor solution. However, there is a reasonably high degree of consensus across several of the best solutions that justify the inclusion of question 11 as part of factor 3, so weak cross-loading in one solution was considered to be acceptable. Further, there was no clear justification for excluding questions 13 or 20, which had each loaded cleanly to a single factor across a range of different solutions, apart from the fact that they seemed to resolve the weak cross-loading of question 11. Therefore, these questions were retained.

Extensive exploratory analysis was carried out, excluding various combinations of variables, to ensure that this solution was robust. This is the only clear factor structure that can be argued for this dataset, and represents a clear “consensus” that confirms, with minor variations, across multiple analyses. The pattern and structure matrices, dropping questions 15 and 17; using principal component analysis and Oblimin rotation, are included as Tables 5.16 and 5.17.

Table 5.15: *Best solution, Pattern Matrix: PCA with Oblimin (oblique) rotation*

Pattern Matrix	Component			
	1	2	3	4
x2 accuracy	.848			
x7 believable	.804			
x14 relevance	.772			
x10 timeliness	.733			
x27 goods-srvcs will be dlvr'd	.514			
x26 appropriate format	.457			
x18 design appropriate	.453			
x20 competency	.447			
x23 level of detail	.430			
x9 navigability		-.887		
x1 learnability		-.871		
x13 easy to use		-.835		
x5 clear and understandable		-.721		
x19 in control		-.685		
x4 txns that are important		-.638		
x3 reputation		-.570		
x6 txns feel safe			.646	
x8 txns save time-money			.618	
x11 personal info feels secure	.417		.536	
x16 attractive			-.465	
x21 enjoyability				.743
x22 positive experience				.554
x24 communicate with org				.510
x25 response time				.508
x12 range of txns				.401

Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization. Rotation converged in 19 iterations.

Table 5.16: *Best solution, Structure Matrix: PCA with Oblimin (oblique) rotation*

Structure Matrix	Component			
	1	2	3	4
x14 relevance	.829	-.491		
x7 believable	.821	-.445		
x2 accuracy	.821			
x10 timeliness	.780	-.495		
x20 competency	.711	-.618		.645
x27 goods-srvcs will be dlvr'd	.704	-.448		.570
x18 design appropriate	.659	-.542		.595
x12 range of txns	.633		.495	.592
x23 level of detail	.629		.405	.546
x26 appropriate format	.604			.456
x9 navigability	.401	-.885		.407
x13 easy to use	.423	-.853		.405
x1 learnability		-.847		
x5 clear and understandable	.601	-.844		.418
x19 in control	.426	-.801		.576
x3 reputation	.442	-.704		.500
x4 txns that are important	.474	-.682		
x6 txns feel safe	.405		.714	
x8 txns save time-money		-.455	.682	
x11 personal info feels secure	.609		.678	
x21 enjoyability		-.467		.762
x22 positive experience	.520	-.637		.755
x25 response time	.543			.657
x24 communicate with org	.491		.459	.649
x16 attractive				.498

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

The comparison table, showing the best solution; and three variations using Oblimin rotation and excluding various combinations of cross-loading variables; is included in Appendix V. The factor structure is clearly visible and consistent across all these analyses, but the best solution is cleaner with few or no weak or cross-loading variables.

It is also interesting that there is a reasonably high level of correlation between the factors, especially between factor 1 and the other factors. Factor three is the most independent (Table 5.18).

Table: 5.17: *Best solution, Component Correlation Matrix*

Component	1	2	3	4
1	1.000	-.474	.323	.435
2	-.474	1.000	-.194	-.426
3	.323	-.194	1.000	.174
4	.435	-.426	.174	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

In factor analysis, we do not necessarily expect that the factors will extract all variance from the items, rather, they will only extract that proportion that is due to the common factors and shared by several items. The proportion of variance of an item that is due to common factors, and therefore shared with other items is called *communality* (Darlington, 2010).

Low communalities are evidence that the variables analyzed have little in common with one another (Darlington, 2010). In our rotated solution, the items have relatively high communalities (Table 5-19). This is expected, as very similar item sets have showed significant loadings on common factors in past studies in the e-Qual stream of research (Barnes, 2001; Barnes & Vidgen, 2001, 2002).

Table 5.18: *Communalities*

Communalities		
	Initial	Extraction
x1	1.000	.728
x2	1.000	.685
x3	1.000	.558
x4	1.000	.571
x5	1.000	.767
x6	1.000	.541
x7	1.000	.718
x8	1.000	.598
x9	1.000	.791
x10	1.000	.632
x11	1.000	.641
x12	1.000	.615
x13	1.000	.749
x14	1.000	.701
x16	1.000	.500
x18	1.000	.609
x19	1.000	.715
x20	1.000	.694
x21	1.000	.640
x22	1.000	.720
x23	1.000	.525
x24	1.000	.565
x25	1.000	.513
x26	1.000	.436

Extraction Method: Principal
Component Analysis.

5.4 Questions of Validity and Reliability

5.4.1 Overall evaluation of the instrument development process

This section summarises our instrument development process, and compares it to best practice (Grover, 1997). A summary of this evaluation is included as Table 5.20.

Table 5.19: *Our instrument development process compared to best practice*

Ideal Survey attributes (Grover's measures)	My score	Comments
Type of research		
Exploratory or Explanatory?	Exploratory	Exploratory research carrying out conceptual refinements to existing theory
Cross-sectional or longitudinal?	C	Cross-sectional
General		
Is the unit of analysis clearly defined?	Y	Individual user perceptions of the on-line service quality of the VUW website
Does the instrumentation clearly reflect that unit of analysis?	Y	Yes, focus on personal response rather than external/objective qualities of the website
Is the respondent(s) chosen appropriate for the research question?	Y	Respondents must have experience of using the website. This is ensured by placing a link to survey on the website.
Is any form of triangulation used to cross validate the results?	Y	Literature, further supplemented by post-hoc focus groups
Measurement Error		
Are multi-item variables used?	Y	Yes, there are a minimum of four items per variable.
Is content validity assessed?	Y	Yes, content validity was assessed by selecting pre-tested items, anchoring new items in literature, and by a field test of question wording.
Is field-based pre-testing of measures performed?	Y	On-line pre-test and feedback
Is reliability assessed?	Cronbach's alpha	This is described in section 5.4.2
Is construct validity assessed?	EFA	This is described in section 5.3
Is pilot data used for purifying measures, or are existing validated measures used?	Partly	Existing validated measures are used. Pre-testing was also carried out for new items, but not a full pilot test.
Sampling error		
Is the sample frame defined and justified?	Y	Users of the VUW web-site, predominantly current and prospective students and staff.
Is random sampling used from the sample frame?	N	Voluntary participation in an on-line survey. Possibility of some degree of response bias.
Is the response rate over 20%?	N	Discussed in Chapter 4
Is the non-response bias identified?	Hard to say	
Internal validity Error		
Are attempts made to establish the internal validity of the findings?	Y	Not applicable – causality not hypothesised.
Statistical conclusion error		
Is there sufficient statistical power to reduce statistical conclusion error?	Y	Sample size in relation to number of questions, and strength of results conform to heuristics for statistical conclusion validity.

Overall, the development and design of the survey followed best practice closely. The type of research is clearly defined, the unit of analysis is clearly identified, the instrumentation reflects the unit of analysis, and appropriate respondents are chosen

Measurement error was controlled as described in Table 5-19. Multi-item variable were used, with between four and nine items per variable. Content validity was assessed using a combination of approaches, including selecting existing, pretested items from eQual, anchoring new items in literature, and pre-testing new items with a field test. Following this, the questionnaire was administered and reliability and construct validity were assessed using Cronbach's alpha and exploratory factor analysis. A limited pilot was conducted, which tested the length and wording of the survey, although it was not sufficiently extensive to be used for purifying the measures.

Questions of internal validity (checking whether differences in the dependant variable are fully attributable to the independent variable) are not applicable to this study. This study does not hypothesise causality, but aims to establish a reliable scale for the dependent variable for the wider study.

This study controlled for statistical conclusion error, by conforming to established heuristics relating to sample size, number of questions, and strength of results.

5.4.2 Reliability

Reliability measures the degree of correlation WITHIN constructs rather than between constructs, in other words, the extent to which measures correlate or move together (Straub et al., 2004). The most commonly used reliability index, and in many published studies, the only reliability test that is used is Cronbach's α . Research validation heuristics suggest that in order for the factor structure to be considered reliable, the α should be > 0.7 for confirmatory research and > 0.6 for exploratory research (Straub et al., 2004). These authors also note that more is not necessarily better with regard to reliability. A very high Cronbach's α . (>0.95) may suggest

common methods bias in the study. Although some of our Cronbach's α 's are high, they are not greater than 0.95.

Reliability statistics are affected by the number of items included for a construct. As we noted earlier, various heuristics exist for the number of items required to test the homogeneity of items within a construct. It has been suggested that as few as three items or four items can provide convergent validity (Harvey et al., 1985). It has been suggested that a smaller number of items provides greater confidence, as scales with a large number of items tend to have higher reliability even if item correlations are quite low (Hinkin, 1998).

The reliabilities for the four factors emerging from this study are included below as Tables 21, 22, 23, and 24. Factor one and two, which have a larger number of items overall, and contain only pre-tested items (with the exception of question 19 in factor 2), and make the largest contribution explaining the overall variance in the data, have very high values for Cronbach's α (0.893 and 0.920).

Factors 3 and 4 contain the majority of the new items (questions 4, 8, 12, 19, 21 and 25), and a smaller number of items (four and five items respectively). The reliability for these factors conforms to reliability heuristics for exploratory research (> 0.6) and is acceptable considering there are a small number of items. Factor three is interesting in that one of the items is negatively correlated with the others (attractiveness).

Table 5.20: *Reliability Statistics for Factor 1*

Reliability Statistics				
No. of items	9			
Cronbach's Alpha	.893			
Factor 1 Item-Total Statistics	Scale mean if item deleted	Scale variance if item deleted	Corrected item total correlation	Cronbach's Alpha if item deleted
x2 accuracy	43.66	50.718	.685	.879
x7 believable	43.37	51.292	.689	.879
x10 timeliness	44.22	48.579	.669	.880
x14 relevance	43.77	49.653	.743	.875
x18 design appropriate	43.81	52.155	.594	.886
x20 competency	44.09	49.767	.699	.878
x23 level of detail	44.21	49.929	.599	.886
x26 appropriate format	43.83	52.144	.571	.888
x27 goods-srvcs will be dlvr'd	44.17	49.118	.650	.882

Table 5.21: *Reliability Statistics for Factor 2*

Reliability Statistics				
No. of items	7			
Cronbach's Alpha	.920			
Factor 2 Item-Total Statistics	Scale mean if item deleted	Scale variance if item deleted	Corrected item total correlation	Cronbach's Alpha if item deleted
x1 learnability	30.06	50.250	.781	.905
x3 reputation	30.04	52.362	.664	.916
x4 txns that are important	29.72	52.878	.623	.920
x5 clear and understandable	29.91	49.365	.789	.904
x9 navigability	30.31	45.577	.847	.898
x13 easy to use	30.09	48.350	.808	.902
x19 in control	30.18	49.759	.754	.908

Table 5.22: *Reliability Statistics for Factor 3*

Reliability Statistics				
No. of items	4			
Cronbach's Alpha	.600			
Factor 3 Item-Total Statistics	Scale mean if item deleted	Scale variance if item deleted	Corrected item total correlation	Cronbach's Alpha if item deleted
x6 txns feel safe	15.90	7.853	.510	.449
x8 txns save time-money	16.29	7.012	.358	.557
x11 personal info feels secure	16.03	6.866	.537	.402
x16 attractive	16.72	9.265	.174	.670

Table 5.23: *Reliability Statistics for factor 3*

Reliability Statistics				
No. of items	5			
Cronbach's Alpha	.798			
Factor 4 Item-Total Statistics	Scale mean if item deleted	Scale variance if item deleted	Corrected item total correlation	Cronbach's Alpha if item deleted
x12 range of txns	18.79	18.747	.516	.779
x21 enjoyability	20.34	17.011	.541	.775
x22 positive experience	19.34	16.918	.700	.723
x24) communicate with org	19.11	16.596	.637	.741
x25 response time	18.75	18.443	.521	.778

5.4.3 Validity

Understanding of validity guidelines for information systems research has progressed considerably in recent years. Straub has been advocating for improved validity in survey research since 1989 (Straub & Carlson, 1989), but a recent “state of the art” study (Boudreau et al., 2001) found plenty of room for improvement. Specifically, they found that many published studies do not validate their instruments (apart from reliability assessment), and that there has been a trend towards using existing instrumentation, possibly because of an increased awareness of the effort involved in developing and validating a new instrument.

In this section, we evaluate the research against the heuristics for mandatory validities suggested in the seminal paper by (Straub et al., 2004). A summary is provided in Table 5.25. Construct validity aims to establish whether the items used to measure a construct, considered together, and compared to other latent constructs, form a reasonable operationalisation of the construct (Straub et al., 2004). Construct validity is normally considered to consist of discriminant and convergent validity. Convergent validity tests whether items that are expected to be related to one another actually are. Convergent validity can be thought to exist when items thought to reflect a construct converge. This can be evaluated by checking whether the items load strongly on the constructs they are associated with. Discriminant validity tests that the constructs posited to be different to one another are observed not to be related. This can be evaluated by checking that there is not undue cross-loading between items.

Straub et al. suggest that both principal component analysis (PCA) and confirmatory factor analysis (CFA) are appropriate for establishing convergent and discriminant validity (Straub et al., 2004). In combination, when using principal components analysis or confirmatory factor analysis, these two are described as factorial validity (Straub et al., 2004). At this stage of the research, we carried out principal component analysis. Confirmatory factor analysis using SEM will be carried out in a later stage. The validation heuristics suggested by Straub et al. align with those we identified from more in-depth studies of exploratory factor analysis. The four factors we extracted had eigenvalues of 11.328, 2.325, 1.327 and 1.202 respectively, which are all above the validation heuristic of 1. As we discussed earlier, in section 5.3 after some refinement, including dropping questions 15 and 17, a clean four-factor structure was established, with all items loading at above 0.4, and only one cross-loading item (question 11), that on the balance of evidence from the exploratory analysis, we decided to retain.

The degree to which the items loaded on the a priori factor structure is open to debate as the results were somewhat mixed. As we have discussed earlier, previous research in this area has been characterised by unstable dimensionality.

Table 5.24: *Mandatory validities applied to our instrument*

Validity component	Technique	Heuristic	This research
Mandatory Validities			
[Construct validity] Discriminant validity	MTMM	Relatively low number of matrix violations; SEM estimates of error attributable to method.	SEM not used at this stage of the research
	PCA	Eigenvalues above 1, or scree test Loadings of at least .4, no cross-loading items above .4 Items that do not load properly dropped	Yes, described in detail in section 5.3
	CFA as used in SEM	GFI > .9, NFI > .9, AGFI > .8, insignificant χ^2 , combined with significant <i>t</i> -values for item loadings.	CFA as used in SEM not used at this stage of the research
[Construct validity] Convergent validity	MTMM	Significant homomethod, homotrait correlations	MTMM not used
	PCA	Eigenvalues above 1 Loadings of at least .4, items load on posited constructs; items that don't load properly are dropped	Yes, described in detail in section 5.3
	CFA as used in SEM	GFI > .9, NFI > .9, AGFI > .8, preferably an insignificant χ^2 ; item loading should be above .707 so that over half the variance is captured by the latent construct; also the residuals (item variance that is not accounted for by the measurement model) should be below 2.56	CFA as used in SEM not used at this stage of the research
Factorial validity	PCA	Pattern of factor loadings should correspond with the a priori structure of latent constructs.	Largely
	CFA as used in SEM	See above	CFA as used in SEM not used at this stage of the research
Mandatory Validities (where appropriate)			
[Reliability] Internal consistency	Cronbach's alpha; SEM reliability coefficients; Randomised question order	CA above .6 for exploratory, .7 for confirmatory; in PLS should be above .7; in LISREL, EQS and AMOS should be above .7	Yes, described in detail in section 5.3
	Inter-rater reliability	Co-efficient Kappa > .7	Not used

5.5 Discussion

5.5.1 The four-factor structure

This exploratory research provides some support for a four-factor structure for perceived on-line service quality in a university website. The emergent factors are content quality, usability, transaction safety and efficiency, and interaction quality. These are somewhat different both to the previous eQual factors, and the hypothesized (a priori) factor structure. In this section, we first introduce the four emergent factors, and then discuss the implications of these results. In particular, we contrast the results with previous eQual results and the a priori factor structure. We then discuss the implications of these variations, and conclude with our revised instrument.

Content quality: encompasses accuracy, believability, timeliness, relevance, level of detail, and appropriateness of format. It also includes the degree to which the design is appropriate for the type of site, the degree to which the site conveys organisational competency, and confidence that goods and services will be delivered as requested.

Usability: includes learnability, site reputation, the ability to complete useful transactions, clear and understandable interaction, navigability, and a sense of control.

Transaction safety and efficiency: includes feeling safe when completing transactions on the website, security of personal information, believing that transactions on the site will be efficient (will save the user time or money). These items are strongly negatively correlated with the belief that the site has an attractive appearance.

Interaction quality: includes the range of transactions offered, a sense of enjoyability or entertainment, the degree to which the site creates a positive experience, the degree to which the site makes it possible to communicate with the organisation, and whether, overall, the user considered the response time acceptable.

These factors confirm some of the latent constructs of the eQual instrument on which this research was based, and disconfirm others.

5.5.2 Factor structure compared to eQual

5.5.2.1 *The evolution of the eQual constructs*

The latent constructs posited by eQual have been evolutionary. The web-qual authors have opted for what Straub et al. describe as “nomological validity”: validity that devolves from the very existence of a well-developed theoretical research stream.

“If theoretically-derived constructs have been measured with validated instruments and tested against a variety of persons, settings, [and] times...the argument that the constructs themselves are valid becomes more compelling. The argument is even stronger when researchers chose different methods for measuring their constructs.” (Straub et al., 2004, p. 27)

In some cases (for example eQual 3.0), the authors have posited latent constructs based on reliability analysis using Cronbach’s alpha, and triangulation of data from multiple methods, rather than factor analysis. They have also argued for adjusting the instrument to reflect domain specific considerations (for example, “buying” and “selling” quality in an auction site. The authors note,

“we see ...a modular approach to instrument design as a very pragmatic; in each application of the tool in a given domain, there may be both general qualities...and those which are context-specific.” (Barnes & Vidgen, 2001, p. 210)

The history of the latent constructs posited by the web-qual/eQual instrument is summarised in Table 5.26. This shows that information quality, and ease of use (sometimes including wider look and feel or “design” considerations) have been the most consistent constructs across multiple studies and business domains. The third latent construct, “service interaction quality” is something of a catch-all, which has at times, split out into more than one factor in exploratory factor analysis, for example, eQual 4.0 (Barnes & Vidgen, 2002). The most recent revision of the eQual instrument, which was used as the starting point for developing the instrument used in this study, posits three latent constructs (Barnes & Vidgen, 2002).

Table 5.25: *The evolution of eQual (formerly WebQual) factors*

WebQual 1.0	WebQual 3.0	WebQual 4.0
<i>Ease of use</i>	Site quality (navigation, look and feel)	Usability: <i>Site Design, Usability</i>
<i>Experience</i>	Interaction quality (trustworthiness, customer relationship)	Service Interaction Quality: <i>Trust, Empathy</i>
<i>Information</i>	Information quality	<i>Information quality</i>
<i>Communication and integration</i>	Auction quality (buying, selling)	
<p><i>Italics:</i> Factor emerging from factor analysis</p> <p>Bold: Researcher-defined construct based on argument and triangulation across multiple methods. Reliability assessment using Cronbach's alpha carried out.</p> <p>Shaded: Constructs that are relatively stable across multiple studies and survey instrument versions</p>		

5.5.2.2 eQual 4.0 latent constructs

Usability: Qualities associated with site design (attractive appearance, appropriate design, competency, positive experience), and usability (learnability, understandability, navigability, ease of use)

Information quality: Quality of the content of the site and suitability for the user's purpose (accuracy, believability, timeliness, relevance, level of detail, appropriateness of format).

Service Interaction: The quality of the service interaction as the user delves deeper into the site, in particular, trust and empathy.

A comparison between the factor structure emergent from our exploratory study, and the posited latent constructs, and associated factor structure of Web-Qual 4.0/eQual is included in Appendix V

Two factors, information quality and usability, have remained relatively stable. For information quality, all existing questions loaded consistently to the same factor in our study, with the exception of question 17 (easy to understand information). Factor

one also picked up three additional variables that had previously loaded to other constructs. Two questions had previously loaded to the “design” factor (included in the usability construct); question 16 (the design is appropriate for the type of site) and question 20 (this site conveys a sense of competency). Question 27 (confidence that the goods and services will be delivered as promised), moved from the “trust” factor that had formed part of the service interaction quality construct.

For usability, all existing questions associated with usability confirmed in our study, although questions associated with the “design” factor previously included by the authors under the umbrella of usability, did not. This factor also picked up a question formerly associated with “trust” (3, reputation of the site) and a new question (19, sense of control).

The remaining two factors were considerably more unstable. This is consistent with a history of greater instability in eQual, and the fact that these factors also included the majority of the new questions, suggesting that adding items to the scale has had an impact on the factor structure.

An entirely new factor emerged, which we have called transaction quality and safety. This picked up (not surprisingly) questions previously associated with “trust” (the safety of transactions and the security of personal information). One of the new questions (8, transactions that will save time or money) also loaded to this factor. A surprising result was the inclusion of question 16 (this site has an attractive appearance) with a strong negative association with the other items. This question had previously loaded to “design”.

The final factor could probably still be described as “interaction quality”, or perhaps “responsiveness”, although the items that loaded to it are almost completely different to the web-qual results. This factor has lost all the variables associated with “trust”, and retained one of the variables associated with “empathy” (24, ease of communication with the organisation). The other variable associated with empathy (15, a sense of personalisation) cross-loaded and was dropped.

The a priori factor structure that we posited for this research was extensively based on eQual, although with some notable variations.

5.5.2.3 Factor structure compared to the a priori factor structure

We posited a factor structure based on the three latent constructs from eQual 4.0 (information quality, usability and design and service interaction quality). In addition to these latent constructs, we posited an entirely new construct, “transaction quality”, reflecting an increasing transaction focus in the design of e-commerce websites and the goals of web-site users. A comparison between our hypothesised factor structure and the results of this study are shown in Appendix V.

Since this structure was extensively based on web-qual 4.0, supplemented by literature, many of the comparisons between the predicted and emergent factor structures have already been discussed in the previous section, particularly with respect to factors 1 and 2. This section will focus on our posited new factor, and the changes to the service interaction quality construct.

We posited, based on an extensive literature review, a new latent construct for “transaction quality” which would reflect the increasing focus in the government and tertiary sector on transaction-capable, feature-rich websites. We added three new questions to measure this construct (4, the ability to complete transactions that are important to me; 8, transactions that will save me time or money; and 12, the range of transactions I expected). In our study, these new questions (with the exception of question 8), did not load as expected. Interestingly, our emerging factor structure does support a latent construct associated with transaction quality and safety, but a revision to our hypothesised scale items is required. There is an element of user pragmatism showing in this factor. This factor acquired two of the questions previously associated with trust (6 and 11) and also the new question 8, and a negative association with the attractiveness of the site.

A similar effect occurred with our final factor. It does appear to capture issues associated with interaction quality; site and organisational responsiveness, and the degree to which the interaction can be considered enjoyable as opposed to a necessary chore, and the overall range of functionality provided. In eQual 4.0, service

interaction quality was always the most unstable latent construct, and it is perhaps not surprising, particularly given the introduction of several new questions, that this construct did not confirm in its previous form.

5.5.3 Implications of the EFA

There are a number of important implications of these results, both for the methodology of the wider study and the insights provided into the tertiary sector business domain.

5.5.3.1 *Impact of the objective qualities of the website*

Exploratory factor analysis using principal component analysis enables the researcher to make generalisations about the characteristics of the data-set collected. These results can form the basis for generating hypotheses, which need to be confirmed with an additional, confirmatory study. Our study was carried out by users of the same website doing the same tasks. This ensures that users have a consistent basis for reporting their experiences, but it also means that all their experiences are equally influenced by the objective qualities of the same website. In our view, the objective qualities of the VUW website explain some aspects of our results. The website (the home-site, as opposed to some of the portals and applications accessed via the home-site) is information-rich, with a strong focus on information provisioning rather than entertainment or self-service transactions. Even where self-service transactions are available, they are generally associated with obtaining documentary information in various forms, or with the relatively early, information gathering stages of a transaction (e.g. prospectuses, application forms, catalogues).

This may explain the loading of question 27 (confidence that goods and services will be delivered as promised) with information quality, because the goods and services available on the site are mainly information goods anyway. This, and the location of the survey within the tertiary sector, may also explain the loading of question 20 (organisational competency) with information quality. Universities are in the information business, and their competency can stand or fall based on the quality, accuracy, timeliness, and so on of the information they purvey.

The surprising loading of the new question 4 (“I believe I will be able to complete transactions that are useful to me”) is likely to be the result of specific question wording interacting with objective characteristics of the site. Anecdotally, one of the draw-backs of the website is that some of the rather arcane processes associated with enrolment and fee payment at a New Zealand tertiary institution can be started on-line, but not carried through to their conclusion. At certain points, the user is handed off to a fulfilment process that requires extensive off-line involvement. This is often described, in informal feedback to the VUW Marketing group, as frustrating. It seems likely that including the word “complete” in the question reminded many users that some transactions they have attempted have not, in fact, completed.

5.5.3.2 Impact of sample

As we noted earlier, in our descriptive statistics, the vast majority of survey respondents were current students (84.1%) and used the site at least once per week (64.1%). This has both advantages and limitations for the study. On the one hand, we can feel confident that the sample is highly appropriate for measuring the phenomenon of interest – the majority of respondents are very experienced with the website and are regular and continuing users. Therefore their perceptions of the service quality are informed. On the other hand, it is likely that their perceptions will reflect, to some degree, their pre-occupations and patterns of use. We believe that the unlikely negative loading of attractiveness with transaction quality and safety reflects this. Our sample was predominantly experienced, pragmatic users, interested in achieving their goals.

5.5.3.3 The implications of the emerging factor structure for research

As we have discussed earlier, the emerging factor structure is not what we hypothesised, nor does it confirm the previous scale on which it was based. An initial, obvious explanation is that the latent constructs have proved somewhat unstable anyway, and this study is simply another example which confirms this trend.

This study provides another illustration of the effect observed by Keller and Dansereau, of the destabilising effect of adding items to scales (Keller & Dansereau, 2001). We note that the two factors where the new questions were concentrated showed the most difference with previous and hypothesised factor structures.

However, since eQual itself was evolutionary, the authors altered and refined their scale and encouraged others to do the same, and no confirmatory studies replicating the same instrument and utilising confirmatory factor analysis have been carried out; it is impossible to say to what extent our results are the result of modifying the scale.

We also confirm the concerns raised by Keller and Dansereau about the difficulties of meta-analysis across studies using different scales (Keller & Dansereau, 2001). It is, in fact, very difficult to reliably position this study in the context of previous research. Our factor-structure compared to web-qual 4.0, and the factor structure of web-qual 4.0 compared to previous versions of the same instrument, are sufficiently different that it would be difficult to perform meta-analysis on these results, despite being part of the same research stream and utilising many of the same variables. In essence, Barnes and Vidgen make an argument for their three latent constructs based on nomological validity based on the balance of evidence across multiple studies. This is quite persuasive, but many of their individual studies lack exhaustive assessment of validity, and no confirmatory studies, using confirmatory factor analysis as used in structural equation modelling, have been conducted.

These difficulties are multiplied many times when different instruments, with different latent constructs and variables are used, that purport to measure the same phenomenon. This is complicated even further because many published studies of perceived on-line service quality also suffer from content validity issues, in that they fail to adequately differentiate the unit of measure (the user's internal perceptions), from the objective design characteristics of websites.

Our study exemplifies the difficulties involved in carrying out conceptual refinements to existing theory (Straub & Carlson, 1989), especially where that theory is not stable. This study is part of a wider study, and aimed to validate the dependant variable independently of the confirmatory study that would test the whole research model. In an ideal world, given that we were making modifications to an existing scale, and applying it in a different business domain; we would have carried out multiple rounds of testing and validation to refine our instrument (Hinkin, 1998; Moore & Benbasat, 1991; Straub et al., 2004). Pragmatism does not permit this, and in general, previously published research in this business area does not demonstrate this degree of rigour

either. Although it is outside the scope of this current study, we feel that large-scale confirmatory studies, studies that replicate previous research, and studies that go back to first principles with regard to our understanding of the latent constructs of perceived on-line service quality would all add significantly to our understanding of this research area.

5.6 Conclusion

With some effort and a certain amount of talking around issues we were able to shoe-horn our data into a somewhat acceptable four-factor structure which met widely-used EFA quality criteria, and supported our addition of a construct for “transaction quality” (Tate et al., 2007). There was considerable debate at this point over what to do next. It would have been possible to continue to develop and test a structural model in PLS. We chose not to do this; concerned about the “piling up” of studies in this research area. This decision was somewhat controversial at the time.

However, with the benefit of hindsight, several things are clear. First, previous EFA-based factor structures were not replicable in a new context. Second, although we were able, with some judicious dropping of indicators, to derive a four-factor structure, the indicators covered by each of our factors cannot be defended as interchangeable indicators of any recognisable latent internal state. Providing a name for a factor derived by an EFA does not make it a really existing psychological entity, and does not mean that items that load on it are a consequence of a real psychological state. Third, we committed (or perhaps continued) the sin of interpretational confounding and meaning variance. By arguing equivalence or near equivalence between our factors and equal factors, despite their different item sets, we uttered the same words with different meanings, “so any logical comparison of their utterances is precluded” (Curd & Cover, 1998, p. 222). Fourth we illustrate the pitfalls in the common approach of “modifying” or “extending” existing instruments, unless this is done with great care. Relatively minor modifications and extensions can change the factor loadings and preclude any meaningful quantitative meta-analysis. In effect, it is an entirely new instrument. Finally, this experience illustrates the high likelihood that a traditional approach to this research area, based on a narrative literature review,

extensions to an existing model, and EFA; will produce yet another study that adds to the general confusion.

6 Factor-analysis-based SEMs

6.1 Overview

This chapter presents the set of Confirmatory Factor Analysis (CFA) models we developed in the process of developing a deeper and more rigorous understanding of measurement theory, modelling, and its applicability to our research area. We noted in our introduction that it can take some time for advances in reference disciplines such as psychometrics to become main-stream in information systems. As Alzola and Robaina noted in their literature review paper, out of 17 studies of online service quality, eleven initially used principal components analysis with varimax rotation (Alzola & Robaina, 2005). Eight of the studies used confirmatory factor analysis.

The papers reviewed by Alzola and Robaina, and the references we cited in Chapter 5 show that EFA still has a considerable body of literature support, despite the “anxiety discourse” about its use in the psychometric research community. Co-variance-based structural equation modelling is increasing in popularity within information systems research, but PLS is also widely used despite its limited popularity in other fields (Rouse & Corbitt, 2008).

Also, PLS does not have an inherent ability to test models using a statistical test, but can only fit given models to data. PLS estimates tend to overemphasize measurement loadings and under-emphasize structural relationships between latent variables (Lohmöller, 1989). This suggests that rather than exploring uncertain theory, researchers should have a strong reason to believe in the basic model structure when using PLS. Since our initial EFA results had been weak, we had no reason to be confident in the correctness of the models tested in this section. We selected co-variance-based structural equation modelling because it explicitly tests the fit of the data to the researcher’s model rather than attempting to fit the model to the data.

When co-variance-based structural equation modelling is used, information systems researchers typically use the “two step” approach we present here, of evaluating the “measurement model” followed by the “structural model”.

In this chapter, we also provide an overview of SEM constructs and metrics. We illustrate the currently mainstream “two-step” approach to co-variance-based SEM modelling, by developing alternate CFA measurement models. We present our models, evaluate the individual parameters and evaluate the measurement models.

6.2 SEM Introduction

In overview, Structural Equation Modelling (SEM) is a statistical methodology that takes a confirmatory (hypothesis testing) approach to the structural analysis of data representing some phenomena. SEM is based on two important assumptions; first, that the theories modelled in the study are represented by a series of structural (regression) equations; and second, that these structural equations can be modelled to create a clearer conceptualisation of the theory under study. A SEM is often conceptualised as consisting of two parts, the measurement model and the structural model.

While it has greatly increased in popularity compared to the “older generation” of multi-variate analysis procedures, best practice in the use of SEM and the interpretation of the metrics it produces, is not settled and continues to evolve.

6.2.1 The SEM Measurement Model

Latent variables represent concepts that cannot be observed directly. Since they cannot be observed directly, they cannot be measured directly, but must be measured in terms of observed variables (indicators) that are believed to represent the latent variable by being in an assumed causal relationship with them. The measurement model is that part of an SEM that deals with latent variables and their indicators.

Latent variables can be exogenous or endogenous. Exogenous latent variables are independent variables, they are assumed to cause, or precede, fluctuations in the values of other latent variables in the model. Endogenous latent variables are dependent variables, and are affected directly or indirectly by the exogenous variables in the model, or by other endogenous variables. Each endogenous variable will

explain a proportion of the overall variance in the model. This is represented by the R^2 , or squared multiple correlation (SMC) of the variable.

A pure measurement model is a confirmatory factor analysis (CFA) model in which there is hypothesised to be unmeasured co-variance between each possible pair of latent variables. The measurement model should include covariances between all pairs of exogenous variables, unless there is a strong theoretical reason not to do so.

CFA assumes that there are direct effects from the latent variables to their indicators (if the indicators are modelled reflectively), and direct effects from the error and disturbance terms to their respective variables, but there are no hypothesized causal effects connecting the latent variables.

6.2.2 The SEM Structural Model

The structural model comprises the set of exogenous and endogenous latent variables in the model, together with the regressions (straight arrows) connecting them and the disturbance terms for those variables. The disturbance terms represent the effect of unmeasured variables not in the model.

The saturated model is the trivial but fully explanatory model where there are as many parameter estimates as there are degrees of freedom – in other words, the exact statistical representation of the structure of the observed data.

The saturated model is that model that assumes freely estimatable covariance between all pairs of observed variables. This model has of course as many parameters to estimate as data points (observed covariances) and so has zero degrees of freedom and perfect fit.

The independence model by comparison assumes all relationships among measured variables are equal to 0. This is equivalent to no correlation at all, or a null model.

The null model is that model which assumes fixed zero covariances between all pairs of observed variables. It has therefore as many degrees of freedom as data points (observed covariances).

The default model is the model specified by the researcher – the hypothesised pattern of correlations in data based on theory. It will always be more parsimonious than the saturated model, and (almost always) better fitting than the null model.

6.2.3 Two step modelling

Kline urges SEM researchers to test the measurement model first, then proceed to testing the structural model by comparing its fit with that of different structural models (for example, models generated by model trimming or building) (Kline, 2005).

6.2.4 SEM modelling approaches.

Testing procedures for SEM models are frequently divided into three main types; strictly confirmatory; alternative; and model generating (Byrne, 2001; Garson, 2008b).

A strictly confirmatory approach is where the researcher postulates a single model and then tests it. This approach is defined as “a model is tested using SEM goodness-of-fit tests to determine if the pattern of variances and co-variances in the data is consistent with a structural (path) model specified by the researcher.” There is an important caveat to a strictly confirmatory model. “As other unexamined models may fit the data as well or better, an accepted confirmatory model is only *a not-disconfirmed model*” (our emphasis) (Garson, 2008b, p. 1, our emphasis). This is widely acknowledged problem with SEM, and provides a strong argument for testing a range of alternative models, especially in a research area plagued with dimensional instability. Given that different models (with different constructs) can potentially fit the data equally well, this tends to cast doubt on the claims of objective existence of the hypothesized constructs.

In the alternative models approach (ATM), the researcher postulates several alternative models all grounded in theory. This approach involves testing two or more causal models to determine which has the best fit, based on appropriate use of goodness of fit indicators. It is noted that “Although desirable in principle, the ATM approach runs into the real-world problem that in most specified research topic areas, the researcher does not find in the literature two well-developed alternative models to

test” (Garson, 2008b, p. 1). Fortunately with perceived on-line service quality, several different models can be argued from literature.

The model generating approach by contrast involves progressively modifying and re-estimating the model in order to improve the model fit. This is an exploratory, rather than a confirmatory process.

In practice, much SEM research combines confirmatory and exploratory purposes: a model is tested using SEM procedures, found to be deficient, and an alternative model is then tested based on changes suggested by the SEM modification indexes (modification indices are discussed in more detail later in this chapter). The problem with this model development approach is that the models confirmed in this manner are post-hoc ones, which may not be stable, or on other words, may not fit new data, having been created based on the uniqueness of an initial dataset.

6.3 SEM Concepts

6.3.1 Model specification

This is the process by which the researcher specifies the model based on theory. The researcher specifies which effects are null, which are fixed to a constant (e.g. 1.0), and which will vary. Hypothesised variable effects are shown by arrows in the model, while null effects are shown by the absence of an arrow. Fixed effects are either those whose parameter has been clearly established in the literature, or are set to 1.0 to establish the metric of a latent variable (usually the strongest loading item).

6.3.2 Model parsimony

A model in which no effect is constrained to 0 will always fit the data. The closer one is to this most-complex model, the better the fit – adding paths will tend to improve fit. Therefore, many fit measures penalise for lack of parsimony (Byrne, 2001). The parsimony ratio can be used to measure model complexity: $PR = df(\text{model})/df(\text{maximum possible } df)$.

6.3.3 Metrics

In SEM, each unobserved latent variable must be assigned a metric (a range). This is usually done by constraining one of the paths from the latent variable to one of its indicator variables by assigning a value of 1.0 to this path. Given this constraint, the remaining paths can then be estimated. The indicator restrained to 1.0 is the reference item. This should be the item that loads most heavily on the latent variable. In this research, the highest loading item for each factor based on the EFA has been constrained to 1.0

6.3.4 Error terms

The error term is the measurement error associated with a given indicator. In SEM, error terms are explicitly modelled. As a result, the path coefficients modelled in SEM are unbiased by error terms on the observed indicators.

6.3.5 Parameter Estimates

The adequacy of individual parameter estimates can be considered before considering the model as a whole. Individual parameter estimates should be feasible – in other words they should exhibit the correct sign and size and be consistent with the underlying theory. Correlations > 1.00 and negative variances may be indicators that the model is wrongly specified (Byrne, 2001). The standard errors associated with each parameter should not be excessively large or small (approaching zero) (Byrne, 2001). Each parameter should be statistically significant at the 0.05 level in order to be retained, and non-significant parameters should be considered for deletion from the model to improve model parsimony.

6.3.6 Structural or path coefficients

These are the regression coefficients calculated by the model estimation program. Maximum likelihood estimation (MLE) is the most common method of SEM estimation, and makes estimates based on maximising the probability that the observed co-variances are drawn from a population that is the same as that reflected in the coefficient estimates. In other words, this method selects those estimates that maximize the likelihood of a model with those coefficients giving rise to the observed covariances. Important characteristics of MLE include: un-correlated error terms are not assumed; a large sample size is assumed; indicator variables are assumed to have

multivariate normal distribution; and indicator variables are assumed to be continuous. If ordinal data is used it should have at least five categories and not be strongly skewed. This means that a likert scale with five or seven categories is also suitable for use with MLE provided the data is normally distributed.

6.4 Our approach to SEM modelling

In this section we describe the application of the principles, techniques and metrics reviewed above. An overview is included as Figure 6.1. Our approach uses the following seven steps: select the approach to SEM modelling (strictly confirmatory, alternative model testing, model generating); model specification (build the model(s) based on theory); evaluate data (evaluate the data for missing values and normal distribution); select fit metrics appropriate for the data-set); evaluate individual model parameters; evaluate the measurement model using confirmatory factor analysis (CFA); evaluate the structural model; and plan the next steps.

uiui

6.4.1 Step 1: Select approach to SEM

We opted to carry out an alternative model approach, due to the inconclusive nature of the literature in this research area, and the fact that slightly different factor structures can be argued based on our literature review, the previous eQual research, and the results of our EFA.

6.4.2 Step 2: Model specification

We developed and specified four alternative models, one based on eQual 4.0 using a three-factor structure, one based on eQual 4.0 using a five-factor structure, one based on our hypothesised factor structure, and one based on the results of our EFA. The coefficients for the highest loading items from eQual 4.0 or from our EFA (as appropriate) were set to 1.0.

Figure 6-1: *Our approach to Confirmatory factor Analysis using SEM*

6.4.2.1 Model 1: Hypothesised 4-factor model from literature

The factor structure in this model (Figure 6.2) is based on our literature review in chapter 2. The highest loading variable for each factor in previous eQual research is constrained to 1. For the hypothesised new factor, we selected one variable to constrain.

Figure 6-2: Model 1: Hypothesised 4-factor model from literature review

6.4.2.2 Model 2: Revised 4-factor model following EFA

The factor structure in this model (Figure 6.3) is based on revisions following the EFA reported in chapter 4. The highest loading variable for each factor from the EFA research is constrained to 1.

Figure 6-3: Revised 4-factor model following EFA

6.4.2.3 Model 3: 3-factor model based on eQual 4.0

The factor structure in this model (Figure 6.4) is based on the eQual 3-factor structure: The highest loading variable for each factor constrained to 1.

Figure 6-4: Model 3: 3-factor model based on eQual 4.0

6.4.2.4 Model 4: 5-factor model based on eQual 4.0

This model (Figure 6.5) is based on the eQual 5-factor structure (Barnes & Vidgen, 2002, p. 122). The highest loading factor constrained to 1.0

Figure 6-5: Model 4: 5-factor structure based on eQual 4.0

6.4.3 Step 3: Evaluate data

Before conducting the analysis, the dataset was checked for missing data, a normal distribution; and the sample size and the model complexity were evaluated.

6.4.3.1 *Missing Data*

We received 247 responses (“cases” in SPSS), of which 147 were complete. Of the remainder, eight, or 3% were empty, twenty, or 8% had three or more items missing, and one hundred, or 40% had at least one item missing. The most commonly omitted items (omitted in more than 10% of all cases) were question 3 (reputation), question 12 (range of transactions), question 25 (response time), and question 27 (goods/services will be delivered as promised).

We carried out a process of scrubbing our dataset in order to deal with missing data. We had the option of eliminating missing data listwise, which means that a case with any missing values is ignored in all calculations. However, this involved discarding 40% of our data and reducing our sample size. The other option was to carry out data imputation for the missing values. Imputation is a process of replacing the unobserved score with an estimated value. AMOS uses the maximum likelihood method of imputation, which involves using an algorithm to obtain the maximum likelihood estimates of the mean vector and covariance matrix for a set of variables. This has been shown to have fewer disadvantages, and be less biased than other forms of imputation (Byrne, 2001, p. 292). However, data imputation in many software programs, including AMOS, requires data that is multi-variate normal. A popular AMOS help site notes that “with skewed and categorical data, multivariate ML solutions are really difficult to compute even with complete data”. The same site notes that future algorithms for non-normal data will probably initially be restricted to “complete data” scenarios <http://www.smallwaters.com/amos/faq/faq-missdat.html> (accessed 8/10/2009). As we discuss in the next section, when we evaluated the data for multi-variate normality we found that the data was not multi-variate normal. We therefore opted for listwise deletion resulting in a usable sample of 147 cases.

6.4.3.2 *Normal distribution*

We then evaluated the data for normal distribution. In SEM analysis, non-normal data tends to inflate the value of the χ^2 statistic of maximum likelihood estimation, especially with smaller sample sizes. As sample size decreases and non-normality increases, analyses may fail to converge. Fit indices, such as the TLI and CLI can produce values that are somewhat underestimated, and non-normality can lead to incorrectly low standard errors (Byrne, 2001, p. 268).

Our dataset had significant skew and kurtosis. The table showing the assessment of data normality is included in Appendix VI (in AMOS the value for multi-variate kurtosis is known as Mardia's coefficient). A value of 1.96 or less means there is no significant kurtosis. Our dataset had a value of 158.468, indicating significant non-normality. The implications of this were twofold. First, we needed to perform bootstrapping, provided by AMOS to assist with the analysis of non-normal data. Second, since these procedures require a complete dataset and cannot operate on missing values, we needed to move forward with the smaller dataset that we had created using listwise deletion of cases with missing values.

Sometimes improvements in normality can be gained by removing outliers (titled in Amos "Observations farthest from the centroid"). There were no cases that clearly stood out as outliers that should be considered for deletion.

6.4.3.3 *Bootstrapping*

Bootstrapping is a procedure that is particularly well supported by the AMOS tool (Byrne, 2001) which can be used when the sample size is relatively small, or when the assumption of multi-variate normality is violated. Bootstrapping assumes that the sample is representative of the underlying population, and that the observations are independent. These assumptions are reasonable in the case of this research. Bootstrapping will also assist with the potential issue of sample size, since listwise deletion required that we delete 40% of our cases in order to obtain complete data. The basic principle of bootstrapping is that it enables the creation of multiple sub-samples from the original database. The bootstrapped estimates then provide the data for calculating of the variability of parameter estimates and indices of fit.

Bootstrapping techniques are recommended for data such as ours (Byrne, 2001, pp. 270-271), although with some caveats. In particular, it is critical that the researcher constrain one factor loading path per factor to a non-zero value.

Where data is not multi-variate normal, an alternative to the χ^2 test known as the Bollen-Stine bootstrap can be used. This is a variation on the χ^2 developed for bootstrapped data. Non-normal data tends to inflate the χ^2 statistic, leading to the impression that the model is worse-fitting than it actually is (Kline, 2005).

6.4.3.4 AMOS terminology

In AMOS, the saturated model is the trivial but fully explanatory model where there are as many parameter estimates as there are degrees of freedom. Overall goodness of fit measures for the saturated model will usually be 1.0, but since the saturated model is the most un-parsimonious model possible, parsimony-based measures will be 0.0

In AMOS, the independence model assumes all relationships among measured variables are 0.0 and the correlations among the latent variables are also 0.0 (effectively the null model).

6.4.4 Step 4: Select metrics

A summary of the range of fit indices available, guidelines for use, and other relevant comments is provided in the following tables. Broadly, SEM fit indices are of several types.

6.4.4.1 Overall fit indices

These measures compare the overall fit of your model to the covariance matrix created based on your data to a model in which no constraints are specified. These are summarised in Table 6.1.

6.4.4.2 Goodness of fit tests comparing the given model with an alternative (null) model

These compare the fit of your model to the independence model, where variables are assumed to be uncorrelated. These are summarised in table 6.2.

6.4.4.3 Goodness of fit tests based on predicted versus observed covariance but penalising for lack of parsimony

More complex models will, all things being equal, generate better fit than less complex ones. These are used when comparing models; the higher parsimony measure represents the better fit. These are summarised in Table 6.3.

6.4.4.4 Goodness of fit measures based on information theory

These measures are appropriate when comparing models that have been estimated using maximum likelihood estimation. As a group, this set of measures is less commonly used in literature, although increasing in popularity. They do not have a cut-off, but are used to compare models, especially non-nested models as in our case – the lower value representing the better fit. Since we are carrying out model comparison, these indicators are appropriate to compare the fit of our four alternative models to the dataset. These are summarised in table 6.4.

Table 6.1: Summary of SEM fit indices and guidelines for use: Overall fit indices

Fit index	Definition	Guidelines for use	Notes and comments	References
Chi-squared χ^2	<p>Chi-squared is a “badness of fit” indicator - a finding of significance means the model’s covariance structure is significantly different to the observed covariance matrix.</p> <p>The chi-sqaure should not be significant if there is a good model fit. Amos calls this indicator CMIN</p> <p>The more complex the model, the more likely a good fit – the closer the model is to being just-identified, the more likely a good fit will be found.</p> <p>Chi-squared is very vulnerable to violations of the assumption of multi-variate normality. Non-normal data will tend to inflate the χ^2 statistic</p>	If the chi-square is significant at $p < .05$ the model is rejected.	AMOS calls this index CMIN. Using CMIN/DF instead of CMIN addresses some of the known issues of complexity and sample size.	
Relative chi-squared	Chi-square fit index divided by degrees of freedom, in an attempt to make it less dependant on sample size.	Heuristics range between 2:1 and 5:1 (see pg 9).	AMOS calls this CMIN/DF	(Wheaton, Muthen, Alwin, & Summers, 1977)
Bollen Stein bootstrap	A variation on the χ^2 test developed for bootstrapped data.	Should be significant at the 0.05 level		(Bollen & Stine, 1993)
Hoetler’s critical N	The size the sample must reach for the researcher to accept the model by chi-square at the .05 or .01 level.	Hoetler’s N should be greater than 200		(Hoetler, 1983)
GFI	GFI varies from 0 to 1, and can theoretically yield meaningless negative values. A large sample size pushes GFI up. GFI is the percent of observed co-variances explained by the covariances implied by the model.	GFI is often high compared to other fit models, so some suggest using .95 as the cut-off. .90 is generally required to accept the model.	For these reasons, use of GFI has declined in popularity compared to other fit measures (Byrne, 2001)	(Hu & Bentler, 1995)
AGFI	<p>A variant of GFI which adjusts GFI for degrees of freedom.</p> <p>AGFI >1.0 is associated with just-identified models and models with almost perfect fit.</p> <p>AGFI <0 is associated with models with extremely poor fit or small sample size</p> <p>AFFI tends to increase as sample size increases and to underestimate fit for small samples.</p>	AGFI should be at least .90.	<p>The popularity of AGFI has also been declining compared to other fit measures. (Byrne, 2001)</p> <p>Since we have a relatively small sample, AGFI is unlikely to be an appropriate fit index.</p>	(Hu & Bentler, 1995)
RMS, RMR, RMSR	RMR is the square root of the mean of the squared residuals, which are the amounts by which the sample variances and covariances differ from the estimated variances and covariances based on the model	The closer the RMR is to 0, the better the model fit.		(Hu & Bentler, 1995)

Table 6.2: Summary of SEM fit indices and guidelines for use: comparing a given model with the null model

Fit index	Definition	Guidelines for use	Notes and comments	References
Comparative fit index (CFI)	Compares the covariance matrix predicted by the model with the observed covariance matrix and compares the null model (covariance matrix of 0's) with the observed covariance matrix, to gauge the percent of lack of fit which is accounted for by going from the null model to the researchers SEM model. CFI penalises for sample size, and along with RMSEA are among the measures least affected by sample size.	CFI should be greater than .90 – indicating that 90% of the covariation can be reproduced in the given model.	Since these measures are least affected by sample size, they are appropriate for our data set which has a relatively small sample size.	(Bentler, 1990)
Incremental fit index (IFI)	Incremental fit index (IFI)	Should be greater than .90 to accept the model		(Bollen, 1989)
Normed fit index (NFI)	An alternative to CFI, which does not require making chi-square assumptions. The proportion by which the researcher's model improves fit compared to random variables (the null model). Varies from 0 to 1. May underestimate fit for small samples	By convention should be >.90, although sometimes a more liberal cut-off of 0.8 is used.	Values greater than 0.8 could be considered acceptable, since we have a small sample.	(Bentler & Bonnett, 1980)
Non-normed fit index (NNFI)	Similar to NFI but penalises for model complexity. Computed as $(\chi^2/df - \chi^2_{null}/df) / (\chi^2/df - 1)$. Less affected by sample size. NNFI close to 1 indicates good fit.	By convention should be >.90, although sometimes a more liberal cut-off of 0.8 is used, although recently NNFI $\geq .95$ has been suggested as the cut-off for good model fit.	Called TLI in AMOS	(Bentler & Bonnett, 1980)

Table 6.3: Summary of SEM fit indices and guidelines for use: comparing predicted versus observed covariance but penalising for lack of parsimony

Fit index	Definition	Guidelines for use	Reference
Parsimony Ratio (PRATIO)	Ratio of the degrees of freedom in your model to degrees of freedom in the null model – used in measures that reward parsimonious models.		(James, Mulaik, & Brett, 1982)
Root mean square error of approximation (RMSEA)	<p>Also called RMS or RMSE or discrepancy per degree of freedom.</p> <p>RMSEA is a popular measure of fit, because it does not require comparison with the null model.</p> <p>It is one of the fit indexes less affected by sample size, although for smallest sample sizes it overestimates goodness of fit.</p> <p>RMSEA computes average lack of fit per degree of freedom based on χ^2 – one could have near-zero lack of fit in both a complex model and a simple model, and most methodologists would judge the simpler model to be better on parsimony grounds.</p>	<p>By convention, there is a good model fit if RMSEA is less than or equal .05, adequate fit if RMSEA $\leq$.08.</p> <p>More recently, Hu and Bentler have suggested RMSEA $\leq$.06 as the cut-off for good model fit.</p>	<p>(Browne & Cudeck, 1993)</p> <p>(Hu & Bentler, 1999)</p>
Parsimony goodness of fit index (PGFI)	<p>This is a variant of GFI which penalises GFI by multiplying it times the ratio formed by the degrees of freedom in your model divided by degrees of freedom in the independence model.</p> <p>Parsimony normed fit index (PNFI) is equal to the PRATIO times NFI. The closer your model is to the all-explaining but trivial saturated model, the more the NFI is penalised.</p>	There is no commonly agreed cut-off for an acceptable model.	(James et al., 1982)
Parsimony comparative fit index (PNFI)	PRATIO times CFI. The closer your model is to the saturated model, the more the CFI is penalised.	There is no commonly agreed cut-off for an acceptable model.	(James et al., 1982)

Table 6.4: Summary of SEM fit indices and guidelines for use: measures based on information theory

Fit index	Definition	Guidelines for use	Reference
Akaike Information Criterion (AIC)	<p><i>AIC adjusts chi-square to penalise for model complexity (over parameterisation). Reflects the discrepancy between model-implied and observed covariance matrices.</i></p> <p><i>AIC can be used to compare non-nested models as well as nested models.</i></p> <p><i>For nested models, you stop modifying when the AIC starts rising. A variation is CAIC, which penalises for sample size as well as model complexity. The penalty is greater than AIC and BCC but less than BIC.</i></p>	<i>An AIC close to zero reflects good fit, and between two AIC measures, the lower one reflects the model with the better fit.</i>	<i>(Akaike, 1987)</i>
Browne-Cudeck criterion (BCC)	<p><i>BCC is very similar to the AIC, except that it imposes greater penalties for model complexity</i></p> <p><i>Needs to be close to .9 to be considered a good fit.</i></p>	<i>Penalises for model complexity.</i>	<i>(Browne & Cudeck, 1993)</i>
Expected cross-validation index (ECVI)	<p><i>Measures the discrepancy between the fitted covariance matrix in the analysed sample, and the expected covariance matrix that would be obtained in another sample of equivalent size. Assumes a comparison of models, the model with the smallest ECVI having the most potential.</i></p> <p><i>Useful for comparing different but non-nested models.</i></p>	<i>Lower ECVI is better when comparing models.</i>	<i>(Browne & Cudeck, 1993)</i>
Bayesian Information Criteria (BIC or ABIC)	<i>Penalises for sample size as well as model complexity. More strongly favours parsimonious models with fewer parameters</i>	<p><i>≤ 2: weak evidence that the model should be ruled out</i></p> <p><i>2-4: positive evidence that the model should be ruled out</i></p> <p><i>6-10: strong evidence that the model should be ruled out</i></p> <p><i>>10 very strong evidence that the model should be ruled out</i></p>	<i>(Rafferty, 1993)</i>

6.4.4.5 Metrics applied to our sample

All of the above fit metrics are included in the output of AMOS software and will be reported. Of the overall fit metrics, CMIN/DF, Hoetler's N, and RMR are the most appropriate for our purposes. The use of GFI and AGFI has been declining in popularity, and the AGFI is sensitive to small sample size, so may be less appropriate for our sample, which is on the small side after discarding cases with missing data.

Goodness of fit measures that compare your hypothesised model with the null model such as the CFI, IFI, NFI and TLI are less affected by sample size, so are likely to offer valuable insights for our study.

Goodness of fit metrics based on predicted versus observed covariance but penalising for lack of parsimony (PRATO, RMSEA, PGFI, PNFI) are reported since they are appropriate when comparing models. Goodness of fit measures based on information theory (AIC, BCC, ECVI and BIC) are metrics which are increasing in popularity for comparing models using maximum likelihood estimation, and are therefore appropriate for this study.

6.4.5 Step 5: Evaluate individual model parameters

6.4.5.1 Overview

Since our data was non-normal, we used the AMOS bootstrapping model for non-normal data, using our smaller data set with missing cases deleted listwise, for all subsequent calculations.

6.4.5.2 Metrics for model parameters

There are a range of metrics available for evaluating individual parameters in a model. These are summarised in Table 6.5. We have opted to present the model parameters first before overall fit metrics, as this is the approach recommended in the literature we consulted. However, we note that since the overall model fit statistics are poor; the individual parameter estimates do not provide useful insights in isolation.

Table 6.5: *Evaluating individual model parameters*

Name	Short description	Recommended cut-off	Comments
<i>R</i> ²	<i>There is one r^2 or multiple correlations for each endogenous variable in the model. This is the percent variance explained in that variable.</i>	<i>By convention should be greater than 0.4</i>	<i>Called “Squared Multiple Correlations” in AMOS output reports</i>
Factor loadings	<i>Standardised value for the strength of loading each observed variable on its exogenous factor</i>	<i>By convention should be greater than 0.7 for CFA (Garson, 2008b)</i>	
Standardised residuals	<i>There is one residual for each pair of observed variables Estimates of the number of standard deviations between the observed residuals and the zero residuals that would exist if the model were a perfect fit</i>	<i>Values >2.58 are considered to be large (Byrne, 2001, p. 89)</i>	
Modification indices	<i>For each fixed parameter in the model, AMOS provides an MI, which is the expected drop in chi-squared if this path were added (these parameters were allowed to co-vary). For co-variances, the MI shows the decrease in chi-square if the two error terms are allowed to correlate. The parameter change (Par Change) is the regression co-efficient for the added arrow.</i>	<i>A rule of thumb is to consider adding paths associated with parameters whose modification index exceeds 100 (Garson, 2008b)</i>	<i>High MI’s in error terms can indicate redundant content, or a common cause not shown in the model.</i>

6.4.5.3 Discussion of results for evaluation of individual model parameters

A summary of the key points from the evaluation of the individual parameters in the four models is provided in Table 6.6.

The regression weights for all observed variables are significant at the 1.96 level for all models. This metric does not support an argument for dropping any items.

The standardised regression weights (effectively factor loadings) provide a much weaker result. All of the models have a significant number of variables (10, 14, 7, 9, between a third and a half of the total number of items) that do not load strongly onto their hypothesized exogenous variables at the 0.7 level or above.

The standardised residuals are also unacceptably high. Models 3 and 4 in particular include a large number of residuals above the 2.58 level, indicating that a significant amount of co-variation is not explained by these models.

Modification indices can be conceptualised as an χ^2 statistic with one degree of freedom. For each fixed parameter, the MI is the expected drop in overall χ^2 value of the parameter were freed (Joreskog & Sorbom, 1988). Although the model fit was poor, the MIs did not point to any specific problem areas. The modification indices for the fixed parameters and the error term co-variances for all the models were nowhere near “rule of thumb” value of 100; the point at which it is suggested that the model should be respecified.

Overall, the regression weights and modification indices do not give any clear guidelines for modifying any of the parameters of any of the models, and appear to be within acceptable guidelines. Nevertheless, the low factor loading and high standardised residuals suggest that there are issues with the model parameters in all four models. Models 1 and 2 are slightly superior to models 3 and 4.

Although the evaluation of individual model parameters indicated some cause for concern, there were no clear candidates for omission or inclusion. We therefore continued to the next step of examining overall model fit.

Table 6.6: *Summary of important results from analysis of individual model parameters*

<i>Metric</i>	<i>Model 1: Lit review</i>	<i>Model 2: EFA</i>	<i>Model 3: eQual 3-factor</i>	<i>Model 4: eQual 5-factor</i>	<i>Comments</i>
<i>Standardised regression weights</i>	<i>All significant at the 1.96 level</i>	<i>All significant at the 1.96 level</i>	<i>All significant at the 1.96 level</i>	<i>All significant at the 1.96 level</i>	<i>Important validity criteria</i>
<i>Factor loadings</i>	<i>See Table 6.7 below</i>	<i>See Table 6.8 below</i>	<i>See Table 6.9 below</i>	<i>See Table 6.10 below</i>	<i>Weak loadings across the board. Cut-off of 0.707 recommended.</i>
<i>R² (squared multiple correlations in AMOS)</i>	<i>See Table 6.11 below</i>	<i>See Table 6.11 below</i>	<i>See Table 6.11 below</i>	<i>See Table 6.11 below</i>	<i>Model 1 and 4: Matrix not positive definite – solution not admissible. This indicates likely high multi-co linearity.</i>
<i>Residuals</i>	<i>none</i>	<i>x16 x21 x5</i>	<i>Multiple variables with values above 2.58, see report in Appendices</i>	<i>Multiple variables with values above 2.58, see report in Appendices</i>	<i>Values above 2.58 are considered to indicate problems with model fit.</i>
<i>Modification indices:</i>					<i>There is no clear evidence that dropping one or more paths would improve model fit.</i>
<i>Added paths</i>	<i>No serious issues</i>	<i>No serious issues</i>	<i>No serious issues</i>	<i>No serious issues</i>	
<i>Error co-variance</i>	<i>No serious issues</i>	<i>No serious issues</i>	<i>No serious issues</i>	<i>No serious issues</i>	

The four following tables (6.7 – 6.10) show the factor loadings of the items on the hypothesised factors. All of the hypothesized factors for all of the models contain weak loading items. Item loadings less than 0.7 are highlighted in grey.

Table 6.7: *Factor Loadings (Model 1)*

Model based on Lit review			
x1 0	<-- -	Content quality	.761
x1 4	<-- -	Content quality	.828
x2 6	<-- -	Content quality	.568
x2 3	<-- -	Content quality	.662
x1 7	<-- -	Content quality	.549
x5	<-- -	Usability	.831
x1	<-- -	Usability	.720
x2 1	<-- -	Service interaction quality	.569
x2 4	<-- -	Service interaction quality	.633
x1 9	<-- -	Service interaction quality	.767
x7	<-- -	Content quality	.786
x8	<-- -	Transaction quality	.573
x4	<-- -	Transaction quality	.600
x1 2	<-- -	Transaction quality	.664
x2	<-- -	Content quality	.750
x1 3	<-- -	Usability	.766
x9	<-- -	Usability	.775
x1 6	<-- -	Usability	.435
x2 5	<-- -	Usability	.599
x1 8	<-- -	Usability	.711
x2 2	<-- -	Usability	.768
x2 0	<-- -	Usability	.787
x1 5	<-- -	Service interaction quality	.591
x3	<-- -	Service interaction quality	.702
x1 1	<-- -	Service interaction quality	.566
x6	<-- -	Service interaction quality	.454
x2 7	<-- -	Service interaction quality	.686

Table 6.8: *Factor Loadings (Model 2)*

Model based on EFA			
x4	<-- -	Usability	.635
x5	<-- -	Usability	.850
x9	<-- -	Usability	.858
x13	<-- -	Usability	.831
x8	<-- -	Transaction quality	.548
x25	<-- -	Responsiveness	.661
x22	<-- -	Responsiveness	.803
x21	<-- -	Responsiveness	.569
x12	<-- -	Responsiveness	.708
x1	<-- -	Usability	.785
x11	<-- -	Transaction quality	.821
x3	<-- -	Usability	.686
x19	<-- -	Usability	.800
x6	<-- -	Transaction quality	.668
x2	<-- -	Content quality	.723
x7	<-- -	Content quality	.735
x16	<-- -	Transaction quality	.295
x10	<-- -	Content quality	.718
x14	<-- -	Content quality	.786
x23	<-- -	Content quality	.673
x26	<-- -	Content quality	.611
x18	<-- -	Content quality	.717
x20	<-- -	Content quality	.801
x27	<-- -	Content quality	.740
x24	<-- -	Responsiveness	.686

Table 6.9: *Factor Loadings (Model 3)*

Model based on eQual (3-factor)			
x2	<---	information	.731
x7	<---	information	.773
x1	<---	usability	.741
x5	<---	usability	.845
x9	<---	usability	.802
x26	<---	information	.581
x23	<---	information	.651
x10	<---	information	.790
x17	<---	information	.534
x22	<---	usability	.732
x20	<---	usability	.768
x18	<---	usability	.696
x13	<---	usability	.801
x16	<---	usability	.438
x24	<---	service interaction quality	.651
x15	<---	service interaction quality	.552
x11	<---	service interaction quality	.631
x6	<---	service interaction quality	.503
x3	<---	service interaction quality	.691
x27	<---	service interaction quality	.724

(The item loading fixed to 1 for each factor is highlighted).

Table 6.10: *Factor Loadings (Model 4)*

Model based on eQual (5-factor)			
x26	<---	information	.597
x24	<---	empathy	.622
x11	<---	trust	.648
x6	<---	trust	.500
x3	<---	trust	.665
x27	<---	trust	.739
x15	<---	empathy	.558
x17	<---	information	.535
x7	<---	information	.766
x2	<---	information	.738
x23	<---	information	.663
x10	<---	information	.769
x1	<---	usability	.774
x9	<---	usability	.888
x13	<---	usability	.878
x5	<---	usability	.818
x16	<---	design	.457
x18	<---	design	.752
x20	<---	design	.848
x22	<---	design	.773

In the following table (6.11) we report the squared multiple correlations for each endogenous variable (in this case the items). This is the proportion of variance in these variables that is explained by our models. If these items were truly reflective of the underlying factors that we hypothesised, we would expect that a high proportion of the variance in these items would be explained by our models. However, in many cases, a very low proportion of the variance is explained by the models, suggesting that the hypothesised factors are not having a significant effect.

Table 6.11: *Squared multiple correlations*

<i>Lit review</i>		<i>EFA</i>		<i>3-factor</i>		<i>5-factor</i>	
<i>x2</i>	.562	<i>x20</i>	.642	<i>x20</i>	.589	<i>x18</i>	.565
<i>x15</i>	.350	<i>x23</i>	.453	<i>x18</i>	.484	<i>x1</i>	.599
<i>x24</i>	.401	<i>x14</i>	.617	<i>x17</i>	.285	<i>x9</i>	.788
<i>x22</i>	.589	<i>x10</i>	.516	<i>x14</i>	.000	<i>x10</i>	.592
<i>x20</i>	.620	<i>x16</i>	.087	<i>x24</i>	.424	<i>x14</i>	.000
<i>x18</i>	.506	<i>x18</i>	.514	<i>x15</i>	.305	<i>x2</i>	.545
<i>x16</i>	.189	<i>x27</i>	.548	<i>x5</i>	.714	<i>x5</i>	.670
<i>x13</i>	.587	<i>x26</i>	.374	<i>x2</i>	.534	<i>x24</i>	.387
<i>x9</i>	.601	<i>x7</i>	.541	<i>x22</i>	.536	<i>x22</i>	.597
<i>x1</i>	.519	<i>x2</i>	.523	<i>x16</i>	.192	<i>x20</i>	.720
<i>x5</i>	.690	<i>x3</i>	.471	<i>x13</i>	.641	<i>x16</i>	.209
<i>x26</i>	.323	<i>x11</i>	.675	<i>x9</i>	.643	<i>x13</i>	.770
<i>x17</i>	.302	<i>x12</i>	.502	<i>x1</i>	.549	<i>x26</i>	.357
<i>x3</i>	.493	<i>x21</i>	.324	<i>x26</i>	.337	<i>x17</i>	.287
<i>x4</i>	.360	<i>x22</i>	.645	<i>x3</i>	.477	<i>x3</i>	.442
<i>x12</i>	.441	<i>x24</i>	.471	<i>x6</i>	.253	<i>x15</i>	.311
<i>x21</i>	.324	<i>x25</i>	.437	<i>x11</i>	.398	<i>x6</i>	.250
<i>x25</i>	.359	<i>x6</i>	.447	<i>x27</i>	.524	<i>x11</i>	.419
<i>x6</i>	.206	<i>x8</i>	.300	<i>x23</i>	.424	<i>x27</i>	.546
<i>x8</i>	.328	<i>x19</i>	.639	<i>x10</i>	.625	<i>x23</i>	.439
<i>x11</i>	.320	<i>x13</i>	.690	<i>x7</i>	.597	<i>x7</i>	.587
<i>x19</i>	.588	<i>x9</i>	.736				
<i>x27</i>	.470	<i>x5</i>	.722				
<i>x23</i>	.439	<i>x4</i>	.403				
<i>x14</i>	.685	<i>x1</i>	.616				
<i>x10</i>	.579						
<i>x7</i>	.618						

6.4.6 Step 6: Evaluate measurement models using CFA

Our four models are compared using the fit indices we discussed in step 3. The full Amos output reports are available in Appendix VII. Important results are included below in table 6.10.

6.4.6.1 Discussion of results for evaluation of fit indices for measurement model

Overall, these fit indices suggest a very poor fit for all four models, and continue the themes identified in the discussion of the individual model parameters. The Chi-squared for overall model fit are all poor. Model 2 the best overall fit to the data (although still not a good fit), which is not surprising since it is the result of modifications following EFA.

The relative chi-squared values (CMIN/DF) for all the models exceed the heuristic of 2.0, which suggest that too many paths may have been dropped. In conjunction with the weak factor loadings, this may suggest cross-loading and poor specification of the measurement model suggests incorrect paths.

Hoetler's critical N for all the models is unacceptably low, suggesting that these models could only be accepted on a sample of between 50-70. By convention, Hoetler's N should be at least 200.

The GFI and AGFI values are all well below recommended cut-off for good fit, with not much to chose between the four models. We note also that these are likely to be overestimated because of the relatively small sample size.

The RMR values for the models range between .141 and .255, when values close to 0.0 would indicate a good fit. Models 1 and 2 are slightly better fitting than models 3 and 4, but none of the values indicate an acceptable fit to the data.

By convention the CFI and IFI values should be $>.9$ for good fit. Only Model 1 approaches (but does not reach) the .9 level. Once again, this model was respecified following EFA, and should be expected to be the best fitting.

The NFI and TLI values also do not reach currently accepted heuristics for good fit. Generally 0.9 is used as the cut-off, although sometimes the more liberal value of 0.8 is used especially in explanatory research. None of the four models achieve values higher than 0.9.

The RMSEA values for all four models exceed the current heuristic of 0.08, although Model 2 is the closest at .09.

Interestingly, the set of metrics that compare model fit for alternative models using maximum likelihood estimation (for example AIC, ECVI) indicate a better fit for models 3 and 4 (based on eQual) than models 1 and 2 (based on this research). This reverses the findings of many of the other indices which suggested that model 2 had the superior fit. However, none of the models approach a good fit to the data.

6.4.7 Step 7: Specify and Evaluate structural model

Since the fit of the measurement model was so poor, we could not see any justification for continuing with this step.

6.4.8 Step 8: Plan next steps

The next steps for SEM modelling depend on the outcomes of step 6. A good fitting measurement model can be used to move on to the next step of evaluating the structural model.

If the measurement model is a poor fit, at this point the researcher can respecify the model by model building or trimming. Model trimming is deleting one path at a time until a significant chi-square difference indicates trimming has gone too far. The goal is to find the most parsimonious model that is well fitting with a selection of goodness of fit tests. As paths are trimmed, chi-square tends to increase, so other goodness of fit measures need to be used as well.

One focus for model trimming is to delete arrows that are not significant, by looking at the CR's for regression weights. Those below 1.96 are non-significant at the .05 level. More weight should be given to parsimony with regard to structural arrows connecting the latent variables than to measurement arrows from the latent variables to their respective indicators.

Model building is the opposite strategy of starting with a null model or simple model and adding paths, retaining those that yield a significant chi-square difference. A

significant chi-square difference indicates the fit of the complex model is significantly better than the simple one.

When considering Model building versus Model trimming, the usual procedure is to overfit the model and change only one parameter at a time. A common process is to add one arrow at a time, then once this process has gone as far as possible, to erase one arrow at a time based on non-significant structural paths.

The outcome of the analysis of the individual parameter estimates did not provide a good case for model building or trimming, and the fit indices of the measurement model were sufficiently weak that we did not feel we were justified in continuing to build a structural model based on using this scale for measuring perceived on-line service quality.

6.4.8.1 Summary of overall fit indices of interest

Table 6.12 provides a summary of AMOS fit indices for the four CFA models presented in this chapter.

Table 6.12: Summary of fit indices for the four CFA models presented in this chapter

Metric	Model 1: Lit review	Model 2: EFA	Model 3: eQual 3-factor	Model 4: eQual 5-factor	Comments
Chi squared	834.062	604.412	611.396	518.333	Reject if > 5
Degrees of freedom	318	269	187	180	
Relative Chi-squared	2.623	2.247	3.269	2.880	You may have dropped too many paths on the default model (from the saturated model) if this exceeds 2-3
Bollen-Stine bootstrap	$p = .007$	$p = .014$	$p = .007$	$P = 0.007$	Should be significant at the 0.05 for a fitting model
Hoetlers critical N	64	75	53	60	Should be greater than 200
GFI	.683	.765	.719	.777	Should be >.95 or at least .9
AGFI	.623	.716	.653	.714	Should be > .9
RMR	.137	.138	.245	.237	The closer to 0 the better
CFI	.782	.849	.758	.807	Should be >.9
IFI	.785	.851	.761	.810	Should be >.9
NFI	.694	.760	.688	.736	Should be >.9, sometimes .8 used
NNFI (TLI)	.760	.832	.728	.775	Should be >.9, sometimes .8 used
PRATO	.906	.897	.890	.857	Ratio of how many paths you have dropped, compared to how many you could have dropped
RMSEA	.105	.092	.125	.113	Good fit if $\leq .05$, adequate fit if $\leq .08$
PGFI	.571	.633	.582	.605	GFI Index adjusted to reward simple models
PNFI	.628	.682	.613	.631	NFI Index adjusted to reward simple models
AIC	954.062	716.412	699.396	620.333	Close to zero is a good fit
BCC	982.537	740.679	715.009	638.430	Needs to be close to .9 to be considered a good fit
ECVI	6.535	4.907	4.790	4.249	Lower is better
ABIC (BIC)	954.062	883.876	830.975	772.	≤ 2 : weak evidence that the model should be ruled out 2-4: positive evidence that the model should be ruled out 6-10: strong evidence that the model should be ruled out >10 very strong evidence that the model should be ruled out

6.5 Discussion

As can be seen from the previous section, the model fit for all of the model versions we tested was very poor. This was sufficiently wide-spread that trimming a small number of offending endogenous variables did not appear to be an option. Our revised model, following our exploratory factor analysis represented a slight improvement on previous versions but could not be described as a good fitting model. We note also that many of the fit metrics were likely to be slightly overestimated due to characteristics of the model and the sample. A more detailed discussion of the implications of these findings, and a comparison of this research phase with the other analyses conducted, is included in Chapter 9.

6.6 Conclusion

Using a more rigorous co-variance-based approach we were unable to do what we had (marginally) been able to do using EFA: make our data fit the initial theory. This chapter provides empirical support for our assertions in Chapter 3; that our initial conceptualisation of this construct was badly mis-specified, and the majority of the models on which our research was based (to the extent they share the same characteristics) were mis-specified also. This goes back over an extended period of time, and affects the original work of PZB in defining service quality, and subsequent instruments such as eQual.

The fit of all the CFA-based models based on eQual, the Literature Review we presented in Chapter 2, and the EFA we conducted in Chapter 4 was extremely poor. This supports assertions about the issues and limitations of this research area and the traditional approaches used to investigate it. However, it is a “negative” finding, which does not offer a way forward. We therefore considered how the models might be respecified in a way that was more theoretically sound. Some of these alternative specifications are presented in Chapters 6 and 7. The various models are then compared in Chapter 9

7 Formative Models

7.1 Overview

In this chapter, we reconceptualise the constructs from our revised eQual instrument in two formative models. This is based on our earlier discussion, in Chapter 3, that previous authors likely mis-specified the constructs as reflective. The purpose of this chapter is to: review the modelling issues and metrics for formative models; test the degree to which the indicators cover the scope of the e-service quality domain; test a model based on the eQual indicators that is specified formatively, and produce results that can be compared with the outcome of our factor-based models and multi-indicator structural models.

7.2 Formative Model Concepts

In Chapter 3, we defined formative and reflective indicators and the differences between them. This section provides a brief recap of some of the salient points of this discussion, in order to enable this chapter to be read without excessive recourse to earlier sections.

In contrast to reflective indicators, which are caused by the underlying latent variable, formative indicators cause the latent variable they are measuring (Diamontopoulos & Winklhofer, 2001). A formatively specified latent construct is basically a multiple regression equation, as follows.

$$\eta = \gamma_1 x_1 + \gamma_2 x_2 + \dots \gamma_n x_n + \zeta$$

Where η is the latent variable, x is the item score; γ is the regression co-efficient, and ζ = the disturbance term.

7.2.1 Co-variance and partial-least squares regression “SEMs”

In this chapter we use both co-variance-based and partial-least squares regression-based “SEM” models. In this section we provide a brief explanation of the differences between the two types of models and the reasons for our selection of methods. Within the information systems field, the term “structural equation modelling” is used

broadly to incorporate both covariance-based and regression based structural models. However these approaches are very different.

Partial least squares regression was developed in the 1960's and 70's as a tool to help overcome the issues of multi-collinearity within multiple regression, which tends to lead to instability. To address this, PLS calculates a set of uncorrelated constructs from correlated individual measures, similar to principal components analysis (Rouse & Corbitt, 2008). The PLS algorithms maximise and account for both the variance of the component, and the variance of the dependent variable at the same time (Rouse & Corbitt, 2008). The components that result are constructed from the sample data (and therefore have weak claim to separate ontological existence) and are constructed in such a way as to maximise the prediction of the dependant variable and item loadings in that particular sample (Rouse & Corbitt, 2008). A detailed description of the PLS method is offered in a text by (Lohmöller, 1989).

Of the two methods, co-variance-based SEM is by far the most widely used in reference disciplines, and is considered the most current and mainstream (Rouse & Corbitt, 2008). This is the main focus of our study. However, the somewhat less exacting PLS approach has some potential to offer insights, and some of these are explored in this section.

A potential advantage of PLS is that it handles formative measures (especially error-free formative measures) without the identification issues of co-variance-based models (Gefen et al., 2000; Petter et al., 2007). In this section, based on our theoretical discussion in Chapter 3, we respecify our constructs of interest as formative variables. Some of the models we wished to investigate were not identified using co-variance-based models, so these were specified in PLS.

7.2.2 What are Formative Indicators?

We compared the characteristics of formative and reflective indicators in chapter 3. The characteristics of formative indicators are reprised below.

- The indicators are not interchangeable – omitting an item is omitting part of the construct. Indicators do not necessarily share a common theme

- The indicators are the set of causes of the construct
- A complete census of indicators is required. Items must cover all possible causes of the latent variable (DeVellis, 1991, p. 55)
- Eliminating or changing an indicator may change the conceptual domain of the construct (Jarvis et al., 2003)
- Causality is from measure to construct
- Correlations are not explained by the measurement model. Indicators are exogenously determined
- The formative construct may possess “surplus” meaning in the form of an error term that represents the portion not explained by the indicators (Jarvis et al., 2003). This is a subject of debate and is discussed in more detail later in the chapter.
- Indicators may vary independently (Jarvis et al., 2003).
- There is no reason to expect a specific pattern of signs and magnitude of correlation – internal consistency is of minimal importance – even negatively related variables can be meaningful indicators of a construct
- Indicators do not have error terms; error variance is represented by the disturbance, which is uncorrelated with the indicators. The error (disturbance) is at the construct level, not the item (indicator) level
- Variance in true scores is higher than variance in observed scores.
- In isolation, the formative model is under identified. The model can only be estimated if it is placed within a larger model that incorporates consequences (effects) of the latent variable
- “Traditional” concepts of validity and reliability – e.g. factor analysis and assessment of internal consistency) are not appropriate. A measure cannot be unreliable, when it has no measurement error associated with it.
- Items do not necessarily have the same antecedents and consequences (Jarvis et al., 2003).
- Change in the indicators precedes changes in the latent construct (Edwards & Bagozzi, 2000)

- Logic or common sense (rather than statistical testing) suggests that a change in the latent variable will not necessarily effect changes in the indicators (Edwards & Bagozzi, 2000)

7.2.3 Achieving Model Identification for a Formative Model (co-variance-based)

As we have discussed earlier, formative models alone are frequently not identified – it is impossible to calculate the model parameters from the observed values. A number of different approaches have been advocated for achieving identification of formative models. This section expands on the overview of this issue provided in Chapter 3.

7.2.3.1 *Model identification*

There are two necessary but not sufficient conditions for formative model identification. 1) The *t*-rule – that the number of non-redundant elements in the covariance matrix of the observed variables needs to be greater than or equal to the number of unknown parameters in the model. 2) The latent construct needs to be scaled, by either fixing a path from a formative indicator to the latent construct, or fixing a path from the formatively measured construct to a reflectively measured endogenous variable, or standardising the formatively measured construct by fixing its variance to unity (Diamontopoulos, Riefler, & Roth, 2008).

In addition to these, the formative measurement model needs to be placed within a larger model that incorporates consequences. The formative latent variable needs to have at least two paths to other reflective constructs or indicators. This is referred to as the “2+ emitted paths rule”. There are generally three possible approaches for achieving this (Diamontopoulos et al., 2008; Jarvis et al., 2003; Petter et al., 2007). The basic methods for achieving identification are described here. A discussion of the issues arising from these methods is included in the section on issues with formative models, below.

7.2.3.2 *Adding two reflective indicators to the formative construct*

This allows for different conceptual interpretations. 1) A MIMIC model. 2) An endogenous construct with two reflective indicators that is caused by exogenous observed variables 3) a formatively measured construct that influences indicators of another construct.

7.2.3.3 Adding two reflective constructs

Add structural relations between the formative latent variable to two reflectively measured constructs. They need to be justifiable based on the nomological net of the formative construct. Note that the parameters for the observed variables to their purported formative latent construct are affected by the endogenous latent (reflectively measured) variables and their indicators – in other words the reflective indicators added for identification can change the meaning of the formative latent.

7.2.3.4 Adding one reflective indicator and one reflective construct

This is a combination of the previous two and allows for one reflective indicator (perhaps a summative or global measure) and a structural relationship to another latent.

7.2.4 The Disturbance term of a Formative Construct (co-variance-based)

As we discussed in Chapter 3, the disturbance term in formative models occurs at the construct level rather than the indicator level. There is considerable debate over the meaning of the disturbance term. It is generally agreed that it is “surplus” meaning not captured by the formative indicators, but there is less agreement about what that surplus meaning is (Diamontopoulos et al., 2008; Wilcox, Howell, & Breivik, 2008).

A number of explanations have been suggested, for example, that the disturbance term is: 1) The collective (overall) random error of all the formative indicators taken as a group, 2) all the invalidity of a set of measures, including measurement errors, interactions among the measures and aspects of the construct domain not represented by the constructs, or 3) The impact of all remaining causes other than those represented by the indicators in the model (Diamontopoulos, 2006). In a discussion of the issues, Diamontopoulos et al. refute the suggestion of measurement error (Diamontopoulos et al., 2008). They suggest that the error term simply represents the aspects not covered by the indicators. The more comprehensive the indicators, the lower the error term should be. A high error term indicates ambiguity in the formative construct. The issue of ambiguity arising from interactions among the measures is discussed in more detail below.

In a MIMIC model, the disturbance term can be conceptualised as the shared variance between the outcomes not accounted for by the formative indicators. The surplus meaning has to do with the endogenous variables, not the formative construct. In other words, it is the gap between the index as created by the formative variables and the equivalent reflective construct as measured by its indicators.

This explanation is compatible with our business domain. We suggested in our literature review that overall affect towards an online website might be the result of a number of descriptive beliefs (for example, about the response time, the accuracy of the information, and so on) and a number of inferential beliefs (for example, about past experiences with technology). So it is possible that a portfolio of descriptive beliefs, such as those described in eQual, might largely, but not entirely predict a response on a question about an affect towards the overall service quality. The disturbance term in that case would represent the potential unmeasured causes (experiential residue) (Fishbein & Ajzen, 1975).

7.3 Issues with Formative Models

Awareness of formative variables has increased in social science and behavioural research. As we discussed earlier in Chapter 3, increasing numbers of researchers are challenging the routine use of reflectively specified SEM models and their associated fit indices. The consequences of mis-specifying formative constructs as reflective have been examined in detail see for example (Diamontopoulos & Siguaw, 2006; Podsakoff, Shen, & Podsakoff, 2006). Despite this, the appropriate use of formative models in social science and behavioural research has been the subject of heated debate in modelling research literature. Issues include ontology, the assumption of error free measurement, the constraints required to achieve model identification, interpretational confounding, correlation and multi-collinearity, the degree to which formative constructs function as a point variable, and whether formative constructs have a place in behavioural models at all.

7.3.1 The Ontology of Formative Models

In our view, the most significant of these issues is ontological, and arises from the difficulties of achieving identification of formative models. Although we noted above that a MIMIC model could be considered as an index of descriptive beliefs that effectively predicts an affect towards the online service quality of a website, it is difficult to explain away the ontological confusion of this conceptualisation. A latent construct in a MIMIC model is doing double duty as both an index, formed from its formative measures, and compatible with instrumentalist and operationalist explanations (as we explained in chapter 3) and as an unobserved latent with a separate realist ontology, which is capable of causing responses on its reflective indicators. These explanations are incompatible, and cannot be describing exactly the same construct, regardless of what it is called or whether the statistics do or do not work.

Bagozzi demonstrates that when the formative construct predicts a single endogenous latent variable (Figure 7.1), even if it is measured with reflective indicators, then the model containing the formative construct can be shown to be equivalent to the model in which the so-called formative indicators directly predict the endogenous latent (Bagozzi, 2007) (Figure 7.2), which now resembles a MIMIC model. In our view, a MIMIC model is in fact “short-hand” for a model such as that in Figure 7.1. Does the introduction of the formative indicator add any additional meaning? It may help resolve the ontological confusion, as it makes more ontological sense to suggest that a certain measurement instrument, which measures subjects responses on a range of factors predicts a large proportion of their response to a question about their affective beliefs about an online service than it does to ask a single latent construct to do double duty as both a formative and reflective variable. In the psychological domain the problem is quite specific: a construct cannot claim to be as a sum-score of external numbers, as well as an internal cognitive concept at the same time.

Figure 7-1: *A formative latent variable predicts a single reflective latent variable*

Figure 7-2: *Formative indicators directly predict the latent variable*

7.3.2 Issues with error free measures

As we have discussed earlier, a formative model assumes that there is no individual error term for the each item that makes up a formative construct. If this is theoretically unconvincing, a possible solution is to introduce a latent variable with a single reflective indicator (and an error term) which corresponds to every formative indicator (Diamontopoulos et al., 2008).

This may be an appropriate form for modelling beliefs and perceptions. If we want to claim a realist ontology, then reintroducing unobserved latent variables to represent user's descriptive beliefs (perceptions) is necessary. Descriptive beliefs are quite specific and we have also previously established that users are relatively unlikely to

be wrong about such beliefs, so a single “gold-standard” indicator could be considered. This approach tends to lead us towards multi-indicator structural models, which are described in more detail in Chapter 8.

7.3.3 Issues with constraining the model to achieve identification

In a formative model, it is common to assume that while the formative indicators of a single construct may correlate, the indicators of one formative construct are perfectly uncorrelated with indicators of all other formative constructs. This may not be a theoretically justifiable assumption. Ideally, all exogenous items in a formative model should be allowed to freely correlate with all others. However, this is not very parsimonious, it may be difficult to identify the model, and goodness of fit indices will suffer (Petter et al., 2007).

This issue is related to the issue of whether the formative construct acts effectively as a point variable. In other words, do we achieve anything in our model by introducing (for example) trust or usefulness as formative constructs, or should we simply combine our measurement items to predict POLSQ directly? There are two possible reasons for including a point variable. One is to create a useful index – a basket of related characteristics that have proved useful and salient for some business purpose, perhaps for benchmarking organisational performance from year to year. The other is to make some claim about reality – for example, that secure information and a good reputation are antecedent to a feeling of trust. The theoretical explanation for the model needs to make this clear.

Many modellers find that in order to enable the model to be identified, structural paths may need to be removed, or construct error terms constrained to zero. This puts constraints on the theory. To address this issue in our research we examined correlation and collinearity between the exogenous items in the model in SPSS and Smart PLS, which are not covariance based and do not require the model to be identified.

7.3.4 Interpretational Confounding

We noted earlier in Chapter 3 that service quality research and online service quality research has widespread issues with interpretational confounding; that is, using the

same factor label to connote slightly different constructs. The differences need not be as extreme as measuring the factor with different indicators.

Howell et al. make a distinction between the nominal meaning and the empirical meaning of constructs (Howell, Breivik, & Wilcox, 2007). Nominal meaning is the meaning assigned without reference to empirical information; the inherent definitional nature of the construct that forms the basis for hypothesising linkages with other constructs, developing indicators, and so on. The empirical meaning derives from its relationship to one or more observed variables in a model. These may be measures of the construct itself, or relationships to other constructs in the model. In our view, the empirical meaning should be as close as possible to the nominal meaning. The problem is that the nominal meaning is usually significantly under-defined (for example, it is unlikely to allow the researcher to hypothesize the relative strengths of indicators of a construct), so that the empirical meaning is the only one we have. Problems arise when that is inconsistent across time or across studies and contexts.

If the nominal and empirical meanings diverge, there is interpretational confounding – the construct does not mean what it purports to mean. This is particularly a risk with formative constructs, where model changes can change the meaning of the construct substantially (Howell et al., 2007). This is because the formative construct has relatively little meaning in its own right; any combination of values can be combined into a composite. The meaning of a formative construct is derived from its contribution to a theory net.

“A construct measured formatively is not just a composite of measures, but the composite that best predicts the value of the dependent variable. The meaning of the latent construct is as much a function of the dependent variable as it is a function of its indicators.” (Edwards & Bagozzi, 2000, p. 159)

This is a potential issue for our formative models, as we had a fixed set of indicators and observations to work with. For this reason, we selected the most conceptually inclusive indicators, with the strongest literature support, to form our latent constructs, for example x13 “I find the site easy to use” for the “Ease of Use” latent, and x27 “I feel confident that the goods and services I have ordered will be delivered as promised” for the “Trust” latent.

Aspects of this discussion will be revisited in Chapter 8 in Multi-indicator structural modelling, as the issue of unclear meaning is a major motivating factor for this more granular and structural modelling style.

7.3.5 Co linearity and correlation

Collinearity and correlation pose issues for formative models. Our approach to evaluating these issues is described in this section.

7.3.5.1 *Co linearity*

Co linearity is expected in reflective models. If two indicators have a common cause, they may be expected to be collinear to some extent. By contrast, formative measurement is based on multiple regression, therefore the stability of the indicator coefficients (the items) is affected by the sample size and the strength of indicator intercorrelations (Diamontopoulos & Siguaw, 2006). Excessive co linearity means it is difficult to assess the impact of an individual item x on the latent variable. Since each y is a direct structural relation between x and η , the magnitude of the y 's can be interpreted as validity coefficients, so high multi-collinearity makes this problematic. Further, since each item (x) is supposed to contribute directly to the formation of the latent dependent variable η then an item that is a perfect linear combination of others contains redundant information (Diamontopoulos & Siguaw, 2006).

In formative measures each latent variable is constructed from a set of parts, so it can be interpreted as a multiple regression model. For example, in a simple model it means that $\eta = b_1 x_1 + b_2 x_2 + \text{error}$ where b_1 and b_2 are regression coefficients and the x_1 and x_2 are two formative "indicators". In a path diagram this would look like: $x_1 \rightarrow \eta \leftarrow x_2$. If x_1 and x_2 are collinear, i.e. have non-zero covariance, multiple explanations are possible (because x_1 might cause x_2 or vice versa), so there might be indirect effects as well as direct effects, as follows.

$x_1 \rightarrow x_2 \rightarrow \eta$

$x_2 \rightarrow x_1 \rightarrow \eta$

$x_2 \rightarrow x_1 \rightarrow \eta \leftarrow x_2$

$x_1 \rightarrow x_2 \rightarrow \eta \leftarrow x_1$

7.3.5.2 *Exogenous variable intercorrelations*

As we discussed earlier, it is frequently necessary to constrain correlations between exogenous variables in formative models.

In reflective models, a common approach is to free all covariances among exogenous latent variables to allow for intercorrelations. In formative models, this approach leads to a number of additional parameters, including correlation estimates of covariances between 1) formative indicators within a construct, 2) formative indicators between constructs and 3) exogenous latent constructs. Two approaches are possible:

- 1) Allowing formative indicators of the same construct to be correlated with each other but uncorrelated with indicators of other constructs. This retains model parsimony as no non-hypothesised paths are added. The GOF indicators are based on relationships of interest. However, this assumes that the paths not specified are perfectly uncorrelated. This may not be correct or appropriate. If any common cause of the variables concerned exists (but is excluded by not specifying the paths) it will contribute to poor fit.
- 2) Specifying formative indicators to be correlated with each other as well as with indicators of other constructs or exogenous variables. The advantage of this is that it does not assume complete independence, which may not be theoretically justifiable. Also, the additional parameters decrease parsimony and but may as a consequence lead to better fit.

For this research, we used SPSS and Smart PLS to evaluate correlations and multi-collinearity, as we were not able to achieve model identification in AMOS without imposing additional constraints.

7.3.6 Summary of issues and our approach

A summary of modelling issues that need to be addressed when developing formative models, and the approach taken in this research, is provided as Table 7.1

Table 7.1: *Issues with formative models and our approach*

Issues	Research approach
Ontological confusion (Borsboom, 2005; Borsboom et al., 2003)	Noted. We believe this issue is not resolvable with MIMIC models, unless the composite variable modelled separately from the reflectively measured latent.
Error-free measures (Diamontopoulos et al., 2008)	We introduce additional latents to address this issue, as proposed by Diamontopoulos, in the next chapter on Multiple Indicator Structural Models
Model identification (Bagozzi, 2007)	Both linear regression tools and co-variance-based tools were used to evaluate the models, since some tests could not be performed using covariance-based software due to problems with model identification. MIMIC models were used.
Formative constructs as point variables	Alternative models, with formative constructs acting as point variables, and without point variables, were tested in PLS and AMOS.
Interpretational confounding (Bagozzi, 2007)	Noted as a potential limitation of formative models. This is addressed in the more granular modelling approach taken in the next chapter.
Correlation	Tested in SPSS and Smart PLS
Multi-collinearity	Tested in SPSS and Smart PLS
Generalisability. The measurement of formative latents depends on their consequences (emitted paths) as well as their measures. This makes statements of generalisability problematic. (Bagozzi, 2007)	Noted. Where possible, reflective measures with good literature support have been selected from our item set. We note, however, that this is also a potential issue with reflectively measured constructs if the item-factor loadings vary considerably from study to study.

7.4 Our approach to formative model development and evaluation

Although formative models have been attracting increased attention in measurement literature, there is no widely agreed methodology for specifying and evaluating formative models. Further, formative models can be specified and analysed in different ways using different software analysis tools. In this research, we use as our primary source a recent MISQ paper by (Petter et al., 2007), supplemented by other references.

7.4.1 Evaluating Formative Models

At this stage, there is no widely agreed process for evaluating formative models. A number of different processes have been suggested. In many cases, the approach to evaluation is significantly different to evaluating reflective models. In this section, we consider these approaches, and synthesise them in terms of their suggestions with regard to content validity (do the formative indicators cover the domain?), construct validity (is the correct indicator set associated with the correct construct?), reliability (do the indicators reliably measure the construct?), evaluation of alternatives, and model fit.

7.4.1.1 *Content Validity*

Content validity tests that the domain of the focal construct has been adequately covered by the indicators. This can include the use of expert panels, literature review, and card sorting (Petter et al., 2007).

7.4.1.2 *Construct validity*

Formative models are not expected to demonstrate convergent validity; in fact, excessive collinearity suggests that redundant items may be included in the model.

Collinearity

We are aiming for complete coverage of the domain, without redundancy. In formative models indicators that are highly collinear are likely to contain redundant information which reduces the meaning of the unique contribution that they make to the formative construct (Diamontopoulos et al., 2008).

SEM problems arising from collinearity

- The beta-weights and r^2 values cannot be interpreted reliably. As a rule of thumb, intercorrelation among independents of 0.8 signals a possible problem (Bagozzi, 2007; Garson, 2008b).
- High multi-collinearity leads to large variances and covariances, large confidence intervals, and insignificant coefficients (Bagozzi, 2007).
- All standardised regression weights should be in the range of -1.0 to 1.0. Weights close to 1.0 show that two variables are close to being identical. When these two nearly identical variables are used as causes of a third latent variable, SEM may have difficulty computing separate regression weights from the two paths from two nearly-equal variables and the third variable. It may come up with one regression weight greater than 1 and one less than -1 for these two paths.
- If there are two nearly identical latent variables, and these two are used as causes of a third variable, the difficulty in computing separate regression rates may be reflected in much larger standard errors for the paths than for other paths in the model.
- The same difficulty may be reflected in high covariances of the parameter estimates for these paths.
- *Variance estimates.* A further effect may be negative variance estimates – if two nearly identical latent variables cause a third, the variance of the third one may be estimated to be negative.

Measuring and managing multicollinearity

We opted to measure the degree of multicollinearity using the variance inflation factor (VIF) and use it to consider eliminating indicators. A $VIF > 10$ or its tolerance equivalent is often used as a cut-off. The risk is that this eliminates indicators on purely statistical grounds (Garson, 2008a). The Variance Inflation Factor (VIF), and Tolerance are both useful for identifying collinearity. VIF is the reciprocal of tolerance. If $VIF > 4.0$ (Garson, 2008a), or 3.3 (Petter et al., 2007), multicollinearity may be a problem. Some researchers use a more lenient cut-off of 5.0 (Garson,

2008a), or 10.0 (Petter et al., 2007). Tolerance is $1 - r^2$, where r^2 is the multiple r of a given independent, regressed on all other independent variables. If the tolerance value is less than some cut-off value, usually 0.2, the independent should be dropped from the analysis. Tolerance considers the independent variable in relation to all other independents and takes interaction effects into account as well as correlations (Garson, 2008a).

$$VIF(\hat{\beta}_i) = \frac{1}{1 - R_i^2}$$

$$Tolerance(\hat{\beta}_i) = 1/VIF = 1 - R_i^2$$

Where β_i is a typical regression coefficient and $\hat{\beta}$ is the estimate.

In the collinearity diagnostics available in SPSS, eigenvalues close to zero (< 0.01) indicate items that explain little variance in the focal construct. Multiple eigenvalues close to zero may suggest a possible problem with collinearity (Garson, 2008a).

Condition indices (CIs) can also indicate collinearity problems. Generally, CI's over 15 may indicate collinearity, and CI's over 30 indicate serious collinearity. This is potentially ambiguous, because it is possible for CI's to be unacceptably high even when VIFs and tolerances do not indicate problems with collinearity. If this is the case, the researcher should check if two or more variables have a variance proportion above .5 or higher on a factor with a high condition index (Garson, 2008a). We regressed each eQual item against the dependant variable to test for collinearity.

Resolving collineraity

If the highly correlated items are measuring the same facet of the same construct, and content validity would not be affected, then all of the correlated items except one should be removed. If correlated items are measuring related but slightly different aspects of the DV, then it is possible that the DV could be split into more than one construct.

7.4.1.3 Reliability (internal consistency within a construct)

Measuring the reliability of formative constructs is also somewhat problematic. Normally reliability assesses consistency and measurement error among the individual measures of a construct. However, this is not relevant for formative constructs, as we have established, the disturbance can only be measured at the level of the formative construct. It has been suggested that formative constructs can be considered reliable if each indicator makes a significant contribution to the formative construct, and the error term at the construct level is low. Note that this only applies to covariance-based models that explicitly separate the error term.

Error term

As we discussed previously, the error term of a formative latent variable indicates the amount of variance that is surplus to that explained by the indicators. A high error term indicates that the domain is not fully captured by the existing indicators (Diamontopoulos, 2006). Note: as we have established earlier, this means the domain of the formative indicator **as defined by** the reflective indicators.

Indicator coefficients

The indicator coefficients explain the amount of variance in the formative construct that is explained by the indicator. A high co-efficient suggests that the indicator is making a significant contribution to the formative latent variable (Diamontopoulos, 2006). On the other hand, a high disturbance (error) term on the formative construct means that constructs' domain of content may not be fully captured by the set of items. Table 7.2 summarises the impact of error terms of different magnitudes on the model.

Table 7.2: *Impact of error terms of different magnitudes on the model*

Size of the effect (error term)		
Small	$f^2 = 0.02$	corresponds to an R^2 of 0.0196
Medium	$f^2 = 0.15$	corresponds to an R^2 of 0.15
Large	$f^2 = 0.35$	corresponds to an R^2 of 0.26

7.4.1.4 Evaluation of alternatives

It has been noted that assuming independence between all exogenous constructs can impose unreasonable restrictions on the model. Petter et al. suggest that alternative models are evaluated, one where all exogenous variables are allowed to co-vary with one another, and one where they are not, followed by chi-square difference test to evaluate the best fitting model (Petter et al., 2007). The major issue with this approach is that the first model (all exogenous variables allowed to freely covary) may not be identified, it is often necessary to impose constraints in order to enable the model to be identified. Instead, we have tested for the presence of collinearity using linear regression.

7.4.1.5 Evaluate estimated model

Two approaches have been suggested for evaluating formative models. One approach uses a linear regression-based method such as PLS to evaluate the model weights for the formative constructs and the r^2 values for the dependent variable in the model. This is based on guidelines from (Chin, 1998) and (Diamontopoulos, 2006).

A co-variance-based approach requires the model to be identified and uses standard SEM fit statistics. It has been suggested that the linear-regression approach can be used as a pre-cursor to a covariance-based approach (Gefen et al., 2000).

7.4.2 Summary: Evaluation criteria for formative models

7.4.2.1 Evaluation Heuristics

The first set of criteria, summarised in Table 7.3 can be considered as simple guidelines for evaluating the coverage, reliability and predictive strength of the set of formative indicators on the dependent variable, based on the work of (Chin, 1998) and (Diamontopoulos, 2006). These heuristics work equally well in PLS and in AMOS.

Table 7.3: *Simple evaluation criteria for formative models (revisited)*

Error term	Coefficients	Comments
Small	High	Good formative model
Small	Some Low	Test collinearity, some indicators may be redundant. Delete some indicators?
Large	High	Check if the domain is fully covered – may need additional indicators?
Large	Low	Re-conceptualisation required

Since PLS is a principal components model, the same loading cut-off heuristics can be used as for other factor loading evaluations – e.g. above 0.7 is considered a high loading, and a loading of more than 0.4 is often considered to be worth reporting (Field, 2005).

If the magnitude of the error term is small, and the indicator coefficients are high, then you have a good-fitting formative model. If the magnitude of the error term is large, but all the indicator coefficients are all significant, then you may not have fully covered the domain of the construct of interest. If the magnitude of the error term is large, but all the indicator coefficients are non-significant, then you have made a mess and should start again (Diamontopoulos, 2006).

7.4.2.2 *Validity, Reliability and Model Quality*

The next set of guidelines are a more comprehensive set of evaluation criteria covering the whole development cycle of a formative model, based primarily on the work of Petter, supplemented by other research where relevant. These are summarised in Table 7.4.

Table 7.4: *More comprehensive evaluation criteria for formative models*

Evaluation Criterion	Guidelines	Reference
Content validity Check that the indicators cover the domain of the construct based on theory	Card sorting	(Petter et al., 2007; Straub et al., 2004)
	Expert validation to see if the measures can be categorised as per theoretical predictions.	(Petter et al., 2007; Straub et al., 2004)
	Literature review	(Petter et al., 2007; Straub et al., 2004)
Construct validity: Discriminant How clearly differentiated is the construct from others?	Indicator coefficients Strong loading on target items.	(Diamontopoulos, 2006; Diamontopoulos & Winklhofer, 2001; Petter et al., 2007)
Construct validity: Convergent	Not appropriate	
Reliability (internal consistency within a construct)	Non significant disturbance term on the formative construct VIF > 10 (traditional) VIF > 3.3 (more restrictive) May indicate collinearity problems	(Diamontopoulos, 2006; Garson, 2008a; Petter et al., 2007)
	Tolerance less than 0.2 may indicate collinearity problems	(Diamontopoulos, 2006; Garson, 2008a; Petter et al., 2007)
	Multiple eigenvalues close to zero (<0.01) may indicate collinearity problems	(Garson, 2008a)
	Condition Index above >15 may indicate collinearity problems	(Garson, 2008a)
Ensure model is identified	See guidelines for model identification	
Assess model quality	Use standard fit statistics, OR	(Gefen et al., 2000)
	Examine model weights for formative constructs, AND	(Chin, 1998)
	Evaluate R2 for dependent variables in model.	(Chin, 1998)

7.4.2.3 Selection of Software and Metrics

The linear regression function within SPSS 16.0 was used to evaluate the VIF and tolerance of the indicators; and eigenvalues and condition indices. These tests were used to assess the collinearity of the formative indicators.

Smart PLS 2.0 was used to provide an initial evaluation of the path coefficients, R-square, and statistical significance of the paths, since the PLS algorithm allows evaluation of models that would not be identified within AMOS.

Finally, co-variance-based SEMs were developed using AMOS 6.0 and AMOS 16.0. Both overall fit indices, and the model weights and R-square values for dependent variables were calculated.

7.5 Specifying the Models

We have argued earlier that although online service quality models have been specified reflectively in the past, these models are probably mis-specified. In this section, we respecify them formatively. In this section we discuss the process of developing and specifying the models. This includes the following steps (summarised in Figure 7.3): developing a census of the indicators, identifying directionality and causality, identifying reflective indicators to allow the model to be identified, and specifying the model(s).

Figure 7-3: *The process of developing and specifying formative models*

7.5.1 Census of indicators

7.5.1.1 *Literature Review*

Our literature review suggested that the eQual instrument that we used for our study provided a fairly comprehensive census of indicators in its initial form. One additional construct, and some additional indicators of existing constructs, were added as a result of our literature review.

It is likely that a different item set that might have been developed for formative and reflective models if the item set was developed from scratch (Diamontopoulos & Siguaw, 2006). This is a potential limiting factor for the model fit that can be achieved.

7.5.1.2 *Expert Validation*

Four faculty members from marketing (1), information systems (2), and organisational behaviour (1) were invited to consider the items and to hypothesise latent formative variables. These people were selected because they were familiar with the process of developing research models and hypothesised relationships, but were previously unfamiliar with the eQual instrument and online service quality literature, and therefore likely to be free from pre-conceptions. Each expert was given the set of twenty-seven indicators and invited to group them into either formative or reflective variables. A template was supplied to each expert that included each question, and two further columns (Table 7.5). In the second column, the expert noted whether the question was a component of their hypothesised latent (a formative indicator) or reflective of it. In the final column, the experts included their hypothesised latent variable.

Table 7.5: *Expert validation of indicators*

Q18. The design is appropriate to the type of site	<Is reflective of> OR <Is a component of>	<latent variable assigned by the expert>
--	--	--

The independent results from each expert were then compared by the researcher. There was a high degree of consistency between the results, with the main difference being that some of the participants proposed a more detailed set of constructs than others. For example, one respondent included formative latent variables for navigation quality and information quality, while another related the same questions directly to a broader formative construct for ease of use.

7.5.1.3 Empirical Evidence

The original eQual instrument was extensively based on empirical data gathering in the target domain (Barnes, 2001; Barnes & Vidgen, 2001, 2002). Subsequently, the scope of the online service domain in a university context was revisited in a series of focus groups conducted with major stakeholders (current and prospective students, graduate students, post-experience students, and staff) by Tate et al. (Tate et al., 2008b; Tate et al., 2007). In combination, these studies give a high degree of confidence that the domain is adequately covered.

7.5.2 Identify causality

Analysis of causality and “directionality” of indicators was conducted using multiple methods. These included: application of the heuristics for formative and reflective indicators, as outlined in Chapter 3, and application of the principle of moving from more specific, descriptive beliefs (perceptions) to more generalised, inductive attitudes and affects), also outlined in Chapter 3. A summary of formative indicators and the composite latent variables they predict is provided in Table 7.6.

Table 7.6: *Formative indicators*

Question / survey item: formative indicator	Composite variable
x1. I find the site easy to learn to operate	Perceived Ease of Use
x17. The site provides easy to understand information	Perceived Ease of Use
x18. The design is appropriate to the type of site	Perceived Ease of Use
x24. This site makes it easy to communicate with the organization	Perceived Ease of Use
x26. The site presents the information in an appropriate format	Perceived Ease of Use
x9. I find the site easy to navigate	Perceived Ease of Use
x10. The site provides timely information	Perceived Usefulness
x12. This site offered the range of on-line transactions I expected	Perceived Usefulness
x14. The site provides relevant information	Perceived Usefulness
x15. This site creates a sense of personalization	Perceived Usefulness
x2. The site provides accurate information	Perceived Usefulness
x23. The site provides information at the right level of detail	Perceived Usefulness
x25. Overall, the response time was acceptable to me	Perceived Usefulness
x4. I believe I would be able to complete transactions that are useful to me	Perceived Usefulness
x7. The site provides believable information	Perceived Usefulness
x11. My personal information feels secure	Trust
x19. I feel in control on this site	Trust
x20. The site conveys a sense of competency	Trust
x3. This site has a good reputation	Trust
x6. It feels safe to complete transactions on this site	Trust

7.5.3 Identify reflective Indicators to enable model identification

To achieve identification of the SEM models, we needed to create a MIMIC model, and to specify reflective indicators for the formative latent constructs. To select these, we used a combination of the most direct indicators (e.g. “I find the site easy to use” for Ease of Use), indicators that had been validated in other large-scale studies, and in two cases, indicators we selected after evaluating the available indicators. The indicators we chose to add as reflective indicators to allow our models to be identified and summarised in Table 7.7.

Table 7.7: *Summary of indicators specified reflectively to identify the models*

Formative variable	Reflective measure	Comment / Literature source
Perceived Usefulness	x8. Completing transactions on this site will save me time or money	Davis et al.'s items related were worded for an organisational context but related to personal productivity. This is the nearest equivalent (Davis et al., 1989).
Perceived Ease of Use	x5. My interaction with the site is clear and understandable x13. I find the site easy to use	From the TAM model Direct measure, also used in the TAM model (Davis et al., 1989).
Perceived Trust	x27. I feel confident that goods/services I have requested will be delivered as promised	Very similar to items for "Trusting intentions towards web provider" (McKnight et al., 2004)
POLSQ	x22. The site creates a positive experience for me x28. What is your overall view of the quality of this website	Selected by the authors as representing a generalised affect towards the site. Direct measure

7.5.4 Overview of the Models

Our initial model to test for collinearity was not identified in AMOS. We therefore opted to conduct our modelling in two stages, first using multiple regression in SPSS and PLS, followed by MIMIC models in AMOS.

7.5.4.1 *Formative Model 1 (PLS): Evaluate measures*

We first evaluated multiple regression models using PLS to evaluate the set of measures. To evaluate convergent and discriminant validity, we first regressed all independent variables on the overall summative item (how do you rate the overall quality of this website). Reliability was tested using tolerance, the variance inflation factor (VIF), eigenvalues, and condition indices. The R-square of each of the dependent variable was also evaluated to obtain an initial indication of model fit.

7.5.4.2 Formative Model 2 (AMOS): MIMIC model

A MIMIC model was developed in AMOS. Overall fit was evaluated using the magnitude of the error term of the endogenous variables, and the amount of variance in each endogenous construct that was predicted by the model. Model fit statistics were also considered. Some purification was carried out. The r^2 statistic for perceived overall service quality was improved by dropping the fun construct and some indicators. The purified formative model with the highest r^2 and the lowest error terms we were able to achieve is presented.

7.5.5 Evaluate Models

Each of the models was evaluated using the appropriate subset of the evaluation criteria. The results for each model are reported below.

7.6 Formative Model 1 (PLS and SPSS): Evaluate Measures

In this model (Figure 7.4), all the indicators in the dataset load on perceived online service quality, apart from two summative indicators which are as reflective indicators of the dependent variable POSQ.

Figure 7-4: Model with all the indicators loading on perceived online service quality

7.6.1 Results: Construct Validity

7.6.1.1 Correlation

Examining the correlation between the items in SPSS showed that all items were correlated with the others to some degree, but there no patterns of significant correlation.

7.6.1.2 Collinearity

The overall set of indicators used on this research to measure POLSQ does not seem to have significant issues with collinearity. In Formative Model 1 (done in PLS), all the indicators are regressed on the overall measures of perceived quality (x22 and x28). Of these, x28 is the most direct measure, and was chosen as the dependent variable for the regression conducted to test collinearity in SPSS (Table 7.8).

Table 7.8: *Evaluating collinearity using regression*

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Constant)	-.924	.457		-2.022	.045		
x1	.141	.070	.156	1.999	.048	.346	2.893
x2	.172	.079	.164	2.183	.031	.373	2.683
x3	.255	.069	.271	3.704	.000	.394	2.539
x4	.027	.057	.030	.462	.645	.487	2.052
x5	.187	.079	.216	2.363	.020	.254	3.940
x6	.211	.075	.198	2.802	.006	.423	2.363
x7	.093	.088	.085	1.055	.294	.326	3.070
x8	-.021	.052	-.027	-.399	.691	.472	2.120
x9	-.114	.078	-.151	-1.462	.146	.198	5.062
x10	-.111	.065	-.131	-1.714	.089	.361	2.769
x11	-.155	.069	-.181	-2.259	.026	.331	3.026
x12	.002	.069	.002	.032	.975	.365	2.742
x13	.083	.082	.101	1.012	.314	.214	4.677
x14	.047	.084	.046	.555	.580	.312	3.204
x15	-.040	.047	-.058	-.869	.387	.470	2.129
x16	.134	.052	.151	2.592	.011	.620	1.613
x17	-.072	.065	-.069	-1.104	.272	.541	1.849
x18	.034	.076	.034	.454	.651	.370	2.706
x19	-.039	.074	-.045	-.524	.601	.282	3.540
x20	.106	.086	.106	1.223	.224	.281	3.563
x21	.018	.053	.024	.345	.731	.432	2.316
x23	.071	.059	.084	1.217	.226	.445	2.249
x24	.063	.056	.076	1.132	.260	.463	2.158
x25	.086	.060	.099	1.446	.151	.451	2.216
x26	.032	.067	.033	.479	.633	.451	2.219
x27	-.030	.065	-.034	-.462	.645	.386	2.589

a. Dependent Variable: x28

Various metrics are recommended by various authors to assess collinearity. None of these indicators exceed the more generous cut-off figure of 10.0. If the most conservative figure of 0.33 is used then x5, x9, x13, x19, and x20 could have potential issues with collinearity.

Eigenvalues and condition indices can also be used to evaluate collinearity (Table 7.9).

Table 7.9: *Evaluating collinearity using eigenvalues and condition indices*

Model	Dimension	Eigenvalue	Condition Index
1	1	26.010	1.000
	2	.176	12.171
	3	.110	15.385
	4	.087	17.320
	5	.071	19.169
	6	.058	21.159
	7	.048	23.176
	8	.047	23.424
	9	.043	24.481
	10	.040	25.605
	11	.036	26.877
	12	.032	28.558
	13	.031	29.009
	14	.028	30.567
	15	.024	32.864
	16	.021	34.801
	17	.020	35.728
	18	.018	37.978
	19	.016	39.861
	20	.016	40.668
	21	.015	42.087
	22	.013	45.351
	23	.012	47.266
	24	.008	56.102
	25	.008	58.153
	26	.007	61.803
	27	.006	66.977
a. Dependent Variable: x28			

The majority of the eigenvalues are significantly different from 0 (above 0.01). The condition indices however exceed the recommended guidelines. However condition indices can be high while other indicators are acceptable. The evidence of issues with collinearity was borderline based on the most restrictive metrics. We opted to retain all the indicators, as they have strong empirical and theoretical support.

7.6.1.3 Item loadings on the dependent variable

All the items load on the dependent variable at the 0.4 level or above (Table 7.10). There do not seem to be any redundant items that do not make a significant contribution to explaining the dependant variable.

Table 7.10: *Item loadings on dependant variable (online service quality): PLS*

	Original Sample (O)	Sample Mean (M)	Std. Deviation (STDEV)	Standard Error (STERR)	T Statistics (O/STERR)
x1 -> Composite formative latent: online service quality	0.673831	0.629844	0.086769	0.086769	7.765834
x2 -> Composite formative latent: online service quality	0.609773	0.558263	0.104078	0.104078	5.858790
x3 -> Composite formative latent: online service quality	0.755701	0.716363	0.061978	0.061978	12.193134
x4 -> Composite formative latent: online service quality	0.609155	0.585531	0.080663	0.080663	7.551887
x5 -> Composite formative latent: online service quality	0.790254	0.753347	0.062697	0.062697	12.604365
x6 -> Composite formative latent: online service quality	0.487312	0.458601	0.123500	0.123500	3.945837
x7 -> Composite formative latent: online service quality	0.615472	0.562775	0.115850	0.115850	5.312655
x8 -> Composite formative latent: online service quality	0.540964	0.509821	0.095070	0.095070	5.690193
x9 -> Composite formative latent: online service quality	0.653248	0.618998	0.077100	0.077100	8.472738
x10 -> Composite formative latent: online service quality	0.591817	0.555268	0.093062	0.093062	6.359382
x11 -> Composite formative latent: online service quality	0.462082	0.421039	0.107223	0.107223	4.309540
x12 -> Composite formative latent: online service quality	0.602174	0.559083	0.110530	0.110530	5.448066
x13 -> Composite formative latent: online service quality	0.647543	0.612511	0.088240	0.088240	7.338389
x14 -> Composite formative latent: online service quality	0.609401	0.553954	0.106639	0.106639	5.714604
x15 -> Composite formative latent: online service quality	0.547153	0.521605	0.084520	0.084520	6.473677
x16 -> Composite formative latent: online service quality	0.477437	0.467478	0.140473	0.140473	3.398790
x17 -> Composite formative latent: online service quality	0.427112	0.411023	0.127513	0.127513	3.349552
x18 -> Composite formative latent: online service quality	0.685519	0.639252	0.091679	0.091679	7.477353
x19 -> Composite formative latent: online service quality	0.732371	0.696176	0.074367	0.074367	9.848052
x20 -> Composite formative latent: online service quality	0.806441	0.760920	0.067340	0.067340	11.975643
x21 -> Composite formative latent: online service quality	0.636531	0.602917	0.068533	0.068533	9.288006
x22 <- POLSQ	0.905485	0.904128	0.019691	0.019691	45.985490
x23 -> Composite formative latent: online service quality	0.625005	0.589210	0.086664	0.086664	7.211862
x24 -> Composite formative latent: online service quality	0.619111	0.588086	0.084597	0.084597	7.318380
x25 -> Composite formative latent: online service quality	0.646434	0.607328	0.100983	0.100983	6.401434
x26 -> Composite formative latent: online service quality	0.561473	0.541005	0.110055	0.110055	5.101756
x27 -> Composite formative latent: online service quality	0.660486	0.624848	0.092339	0.092339	7.152811
x28 <- POLSQ	0.913254	0.908837	0.019920	0.019920	45.846705

7.6.1.4 r^2 and Error Term on Overall Service Quality Measure (PLS)

The r^2 of the dependent variable (perceived online service quality) is provided in Table 7.11 (from PLS). This shows that the items explain approximately 80% of the variance in the dependant variable. The error term is “medium to large” according to our heuristics. Since the error term is medium-large, but all the item loadings are significant, we can conclude that we need to retain all the indicators, but it is possible they do not fully cover the scope of the domain. This is not surprising given that it is a highly generalised construct. The error term is calculated as $1-r^2$ in PLS, which gives an error term of 0.199. According to our heuristics in Table 7.2, this is considered medium-large.

Table 7.11: *The r^2 of perceived online service quality from Formative Model 1 (PLS).*

	AVE	Composite Reliability	R Square	Cranach's Alpha	Communality	Redundancy
Composite formative latent: Online service quality					0.390773	
POLSQ	0.826969	0.905289	0.801183	0.790848	0.826968	0.662207

7.7 Formative Model 2 (Covariance-based): Purified MIMIC model

The second model we present here (Figure 7.5) represents the best example of a covariance-based SEM model using formative latent variables to predict perceived online service quality. We chose ease of use, usefulness and trust as formative latent variables, since these had good literature support, and we had reflective items available, that had been validated in previous studies, that we could use to identify the model. We dropped the formative latent variable for “fun” and its associated indicators, as the fit of the model increased once these were deleted.

Figure 7-5: *Formative model 2, with formative latent variables for ease of use, usefulness and trust.*

7.7.1 Results: Model Quality

We can see that the composite formative variables of trust, usefulness and ease of use explain more than 90% of the variance of POLSQ. The formative indicators for ease of use, trust and usefulness all explain more than 80% of the variance in the formative constructs (Table 7.12). The error terms on the formative latents are low and statistically insignificant (Table 7.13). The regression weights of the formative constructs on POLSQ are mixed (Table 7.14), with Trust being fairly high at 0.728, while ease of use is low. These results suggest that we can cautiously accept this formative model, according to evaluation heuristics for formative models. A full report of the metrics produced by AMOS for this model is included as Appendix VIII. A caveat is that the regression weights for the items forming the composite measures are extremely variable (see Appendix VIII). This was fore-shadowed by relatively weak loadings for some items in PLS.

Table 7.12: *Squared Multiple Correlations:*
(Group number 1 - Default model)

	Estimate
Trust	.853
Usefulness	.839
Ease of Use	.991
POLSQ	.926
x27	.280
x22	.360
x28	.409
x8	.261
x13	.598
x5	.479

Table 7.13: *Variances (error terms): (Group number 1 - Default model)*

	Estimate	S.E.	C.R.	P	Label
error ease of use	.006	.042	.133	.894	par_133
error Trust	.059	.066	.901	.368	par_134
error usefulness	.078	.161	.483	.629	par_135
error POLSQ	.022	.057	.384	.701	par_136
error x5	.655	.088	7.483	***	par_137
error x13	.515	.081	6.378	***	par_138
error x8	1.361	.226	6.018	***	par_139
error x28	.430	.067	6.379	***	par_140
error x22	.580	.084	6.926	***	par_141
error x27	1.034	.138	7.519	***	par_142

Table 7.14: *Regression Weights: (Group number 1 - Default model)*

			Estimate	S.E.	C.R.	P	Label
POLS	<-	Ease of Use	.110	.044	2.531	.011	par_16
Q	--						
POLS	<-	Usefulness	.339	.067	5.063	***	par_17
Q	--						
POLS	<-	Trust	.728	.102	7.159	***	par_18
Q	--						

7.7.2 Model Fit Summary

Although scholars are divided on the usefulness of conventional fit metrics for evaluating formative models, the AMOS fit metrics for formative model 2 are included here for comparative purposes (Tables 7.15 to 7.23).

Table 7.15: *CMIN*

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	81	1987.854	236	.000	6.716
Saturated model	377	.000	0		
Independence model	52	2731.755	325	.000	8.405

Table 7.16: *Baseline comparisons*

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.272	.201	.305	.228	.297
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Table 7.17: *Parsimony adjusted measures*

Model	PRATIO	PNFI	PCFI
Default model	.911	.248	.271
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

Table 7.18: *NCP*

Model	NCP	LO 90	HI 90
Default model	1691.854	1554.418	1836.727
Saturated model	.000	.000	.000
Independence model	2406.755	2243.647	2577.254

Table 7.19: *FMIN*

Model	FMIN	F0	LO 90	HI 90
Default model	13.615	11.588	10.647	12.580
Saturated model	.000	.000	.000	.000
Independence model	18.711	16.485	15.367	17.652

Table 7.20: *RMSEA*

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.198	.190	.206	.000
Independence model	.225	.217	.233	.000

Table 7.21: *AIC*

Model	AIC	BCC	BIC	CAIC
Default model	2149.854	2186.611		
Saturated model	754.000	925.076		
Independence model	2835.755	2859.352		

Table 7.22: *ECVI*

Model	ECVI	LO 90	HI 90	MECVI
Default model	14.725	13.784	15.717	14.977
Saturated model	5.164	5.164	5.164	6.336
Independence model	19.423	18.306	20.591	19.585

Table 7.23: *HOELTER*

Model	HOELTER .05	HOELTER .01
Default model	25	27
Independence model	20	21

The traditional model fit statistics for this model are poor. However, this is not unexpected with formative models (Petter et al., 2007). The full report for this model as generated by AMOS is included in Appendix VIII.

7.8 Discussion

7.8.1 Model 1

A comparison of these models enables us to consider whether the item set identified gives appropriate overall coverage of the domain of POLSQ, and to consider some alternative conceptualisations of the nature and structure of the relationships between the items. We first consider the item set.

In model 1, the complete set of indicators used to measure online service quality were all regressed on an overall measure of POLSQ. The r^2 for overall service quality was

0.8012. This means that the item set included in the study explained more than 80% of the variance in the dependent variable. However, since the r^2 tends to improve with an increased number of indicators, this does not necessarily tell us that we have the most salient items.

All of the items load reasonably strongly (>0.4) on the dependent variable, and all loadings are statistically significant at the 0.0005 level, suggesting that there are no redundant items.

Using the more restrictive guidelines suggested by Diamontopoulus and Petter for VIF in formative models, items x5, x9, x13, x19 and x29 have VIF values > 3.3 , although all of them are well below the more generous guideline of VIF not > 10 .²

There do not appear to be significant issues with collinearity suggesting redundant indicators. The tolerance, VIF, and eigenvalue figures are all within the outer bounds of recommended guidelines, although collinearity issues with some items might be identified if more restrictive guidelines are used. Of these in subsequent models, x5 and x13 are used as reflective indicators for the latent construct that x9 contributes to, so some degree of collinearity is not surprising. The indicators x19 and x20, with a VIF of 3.3 barely exceeded the most restrictive guideline of VIF not > 3.3 .

The condition indices (CI's) are the only metric that is higher than the recommended guidelines, but it has been noted that these may be high when other metrics are acceptable (Garson, 2008a).

In summary, the item set explains a large proportion of the variance in the dependent variable, and there do not appear to be redundant items or issues with collinearity that will affect the explanatory power of the model. It is possible additional items could be added.

² x5, my interaction with the site is clear and understandable; x9, I find the site easy to navigate; x13, I find the site easy to use; x19, I feel in control on this site; x20, this site conveys a sense of competency.

7.8.2 Model 2

Model 2 offers a more exacting co-variance-based MIMIC model using composite formative variables for Trust, Ease of Use and usefulness to predict overall service quality. The composite indicators appear to have strong predictive power for perceptions of overall online service quality, and the error terms are low.

This model is a better fit according to heuristics for formative model fit than any of our reflectively specified models, and it fits a number of the criteria for a good-fitting formative model in that it has relatively high predictive power and low error variance. The overall model fit, based on the Bollen Stein bootstrap statistic, is acceptable.

There are two caveats to be noted, relating to the contribution of the formative constructs to the overall perception of POLSQ, and the contribution of the indicators to the formative constructs. Of the three formative constructs, Trust contributed the largest proportion of perceptions of overall service quality with a loading of 0.7. Usefulness was “worth reporting”, with a loading of 0.3, but the contribution of ease of use was only 0.1.

The indicator set predicted a large proportion of the variance of each of the formative constructs (91% for Trust, 90% for Usefulness, and 99% for Ease of Use), with very low error terms. However, the individual indicators had very low loadings on the formative constructs.

7.8.3 Revisiting the issues with formative models

Using the models in this chapter as illustrations, we revisit and discuss some of the known issues with formative models.

7.8.3.1 *Ontological issues*

The potential for ontological confusion remains a major concern for formative models of attitudes and perceptions (the issues may be quite different for economic models). It is extremely difficult to make a credible case that a response on a survey item (this is what the indicators are) *causes* a latent internal belief or perception. If you carefully separate the composite formative variable from the reflective variable (as per our model 2), then the ontological claims are that a series of observations we have taken

(we make no claims about the causes of the observations), formed into a composite measure, will be effective in predicting a high proportion of the variance on unobserved latent variables.

This is an instrumentalist approach, which may have salience for purposes such as bench-marking performance between organisations, or bench-marking performance from year to year, but does not make any sensible realist truth claims from a research perspective.

Far more likely, in measures of attitudes and perceptions, is that the indicators themselves are reflective of further, un-modelled latent variables. For example, a response on x2. “The site provides accurate information”, is almost certainly caused by some unobserved latent variable which is the respondent’s perception of the site’s information accuracy at the time they answered the survey (a model similar to that shown in Figure 7.6). Where the latent has only one indicator, the model in Figure 7.7 can be used as a “short-hand” form, because it is mathematically equivalent. Models with a larger number of single indicator latents are explored in the next chapter.

Figure 7-6: *Example of formative latent variables with one indicator*

Figure 7-7: *Formative indicators are used as “shorthand” for single indicator formative latent variables*

7.8.3.2 Modelling issues

Constraints required to achieve parsimony and model identification using SEM do place restrictions on our ability to fully explore interactions between the indicators. We needed to specify that our formative indicators were independent of all other formative latents. In an environment of complex interaction of multiple attitudes and perceptions leading to an overall affect towards a site, these assumptions may not be correct. Consideration of collinearity mitigates this risk to some degree.

We also note that many of the item loadings for individual questions on the formative indicators were weak, although collectively they explained a large proportion of the variance. This may be due to unexamined structural issues between the indicators. Other possible reasons are scaling and error variance on the endogenous latent variables. Each endogenous latent is scaled to the same scale as a single indicator. With 5 indicators contributing, their contributions will each be 20% of that of a single

indicator. This means that to get to a score of 0.7 on the endogenous latent, you need each indicator to only contribute 20% of 0.7. Another issue might be errors. Large errors on the endogenous indicators make for low variance of the endogenous latent, i.e. there is not much to explain as most of the endogenous variance is error variance. The fact that the indicator set was not developed specifically for formative modelling may also have influenced these results.

7.8.3.3 *Generalisability*

The use of established reflective latents such as perceived usefulness, as consequences, measured by existing, tested reflective indicators, improves the potential for generalisability of the model. By generalisability, we mean not so much the ability to generalise POLSQ in a university context to other contexts, but the ability to perform meaningful meta-analysis to compare the findings of this study with those of other studies.

We recollect that the measurement of formative latents depends on their consequences (emitted paths) as well as their measures, which makes statements of generalisability problematic (Bagozzi, 2007). We further recollect our model showing the increasing generalisability of measures. If we use established, generalised measures as consequences, then we are able to establish that *in this context* beliefs and perceptions about information accuracy, information relevance and so on (the formative indicators in the models) are likely to lead to a set of more generalised beliefs and perceptions *about this site*. This enables this model to be compared with the process of formation of generalised perceptions about other technologies in other contexts.

7.8.3.4 *Correlation*

Although formative indicators are not expected to be correlated, because they have no common cause, of course, they may still be correlated. Our analysis suggested that our indicators all had some degree of correlation with others, but did not have significant patterns of correlation.

7.9 Conclusion

We can conclude some things with a reasonable degree of confidence from this analysis. Most importantly, the item set used in the survey appears to provide reasonable domain coverage and reasonable predictive power for overall perceptions of online service quality, predicting more than 80% of the variance. However, the structure and inter-relationships between the constructs and item are not clearly established. Our formative constructs provided a reasonably good-fitting model, according to heuristics for evaluating formative models, but did not provide a great deal of confidence that the ideal structure had been reached for the individual formative constructs themselves.

We can see, by using a real-world example, the unconvincing ontology implied by formative models of attitudes, beliefs and perceptions. A formative model requires us to assert that a response on a survey causes a psychological state, unless it is interpreted as short-hand for a more granular model with single indicator latents. Even that explanation is not very satisfying, because it merely admits a pragmatist/instrumentalist interpretation – i.e. that the model is useful for prediction, rather than that the model makes truth claims about reality.

Since our formative models had reasonable fit metrics, this supports our assertions that previous reflective models of this domain are likely to be mis-specified, or at least, have weaker explanatory power than our formative models. A formative index may be useful, as it can measure the relative quality of an IS artefact according to a range of criteria. It could be argued that unless these measures are also included in further theory, or a more extended nomological net, their structure, and the existence, or otherwise or latent variables with a separate ontology is not actually very important. As we suggested in Chapter 3, the perceived salience of the measures to respondents, the value to survey sponsors, and the usefulness of formative indexes for bench-marking may be sufficient justification. However, we still have many unanswered questions about the structure of online service quality, which will be explored further in Chapter 8. A more detailed discussion of the implications of these findings, and a comparison of this research phase with the other analyses conducted, is included in Chapter 9.

8 Multi-indicator structural models

8.1 Overview

This section applies an alternative approach, proposing a more granular, non-factor analytic model with fewer indicators per construct, to explain our data. This approach is based on debates within the SEM methods community about factor-analytic versus non-factor analytic models, and their strengths and weaknesses.

We note the following caveats about the models in this chapter. As with the previous chapters, we were constrained by the indicator-set that was available to us. Also, we were unable to develop a comprehensive and well-fitting model that utilised all our indicators. There are literally hundreds of alternative models that could be considered, and we did indeed develop many that are not reported here. The models presented in this chapter should more properly be considered as model fragments that illustrate aspects of our argument than as “finished” models of POLSQ.

8.1.1 Non-factor analytic models

As we discussed in our literature review, the general approach to developing this model is based on work by Leslie Hayduk, primarily his 1996 book, his contributions to the debate between leading SEM methodologists from *Structural Equation Modelling* (2000) vol 7 issue 1, and discussion on the SEMNET forum.

To recap briefly, there is considerable similarity in how factor models and multiple indicator structural models account for the observed covariances among items. The key difference is that the factor model’s primary explanatory mode is item correlations (covariances) arising from an (assumed) dependence on the common-factor cause, while SEM permits explanations based on hypothesised causal effects, and may or may not include multiple indicators. The key point is that the factor model requires 3 or more indicators per factor, and requires a distinction in thought, if not mathematically, between measurement part and structural part. Neither of these is necessary for the SEM approach.

In extremely simple form Hayduk's argument is that there is no fundamental difference between a measurement arrow (between latent and indicator) and a structural arrow (between latent and latent), they both represent theoretical assertions by the researchers, and affect, and are affected in turn, by other relationships in the model (Hayduk, 1996).

Hayduk demonstrates that a common cause does not necessarily need to be a common factor, and suggests that it is important to provide some justification as to why the posited identity of a variable should correspond to the most proximal (in the researcher's model) of the potentially many common causes of the items. Each of the causal forces emerging from the (hypothesised) common factor and taking separate routes to the indicators needs to be justified – effectively, a separate hypothesis needs to be formed for the strength and likely error variance on every factor-item loading (Hayduk, 1996).

Hayduk suggests the following four steps for modelling. First, for each latent, chose the indicator you believe is the best single available indicator of the concept in question. This selects which part of the shared external world (the indicators) is most similar to your hypothesised latent variable η .

1. Select the best or most direct indicator
2. Fix the λ for that indicator at 1.0
3. Fix the variance (θ) for that indicator at a specific value justified from theory
4. Enter free λ and θ variances for the second and third best indicators of the concept, unless you are confident that you can fix these also.

Step 1 and step 2 set the scale for that concept to correspond to the scale for the observed indicator, by making each real unit of change in the concept correspond exactly to one unit of change in the indicator. Step 3 quantifies your assessment of how similar or dissimilar your concept is to the best indicator. Step 4 is only possible if there are two or more indicators of a concept, and it provides a test for the conceptualisation you have asserted in steps 1-3 (Hayduk, 1996).

Our approach to the development of our non-factor model is based on these four steps. Hayduk suggests that if a multiple indicator model fails (has a large χ^2), start with a model containing only the single best indicator of each construct, and add additional indicators in separate runs. Hayduk notes that it can be hard to get even 2 or 3 indicators to co-operate, and he discourages (initially) using many sets of multiple indicators (Hayduk, 1996).

8.2 Our Approach to model development

8.2.1 Revised eQual instrument and items

Our starting point for the non-factor analytic model was the same as that for our other models – the 28 items and our original dataset.

8.2.1.1 Step 1: Abstraction of latent variables from items

In this step, the researcher evaluated every item in order and identified the most likely proximal cause for a variation in the indicator. This proximal cause was abstracted into a latent variable. The latent variable was assumed to be an internal (to the user), unobservable, perception, attitude, or belief held by the survey respondent about the website. We assumed that the reflective latent variables we had previously selected to identify our models (for example usefulness, ease of use and overall perceptions of service quality) were the proximal cause of their indicators.

8.2.1.2 Step 2: Initial definitions of candidate latent variables

In this step, the researchers created initial working definitions of latent variables based on vernacular usage, including dictionary definitions, and their own beliefs and experience. We used an emergent approach, without reference to existing theory (at this stage). Many of our constructs were sufficiently granular that they did not have a dedicated literature in the IS area. In other cases, the existing theoretical nets for the constructs were ambiguous or even contradictory.

In 1982, Blalock noted that

“It will be to our advantage to define concepts as clearly and concisely as we can, to attempt to assess the degree to which a common vocabulary has been attained, and to press for additional clarification as soon as ambiguities arise or further distinctions need to be made”, (Blalock, 1982, p. 12).

Unfortunately, a factor-analytic approach sometimes appears to act as a substitute for this precision.

By identifying the most immediate cause of each indicator, we were able to achieve greater precision in the definition of each latent, and provide a richer and more descriptive theory net. In some cases, these were subsequently collapsed again if the nomological nets (in the context of our model) were the same.

Our final set of definitions (Table 8.1) was developed using vernacular usage, wordnet (<http://wordnetweb.princeton.edu/perl/webwn>) and dictionary definitions (<http://www.oed.com/>). More specific constructs and more general constructs that incorporate them were defined separately.

Table 8.1: *Candidate latent construct definitions*

Indicator	Definition
Cognitive fit (x5)	The degree to which the user's mental model (beliefs about) of the domain matches the designer's mental model (beliefs about) of the domain.
Self efficacy belief (x4)	The degree to which the user believes they possess capacity to achieve a desired result using this website.
Perceived appropriateness of information format (x26)	The degree to which the user perceives the visual presentation of the information to be appropriate to the purpose of the site/page.
Perceived ease of learning (x1)	The (inverse of the) effort that an individual perceives is required to acquire knowledge about operating the web site.
Perceived appropriateness of information detail (x23)	The degree to which the user perceives that the information (<u>content</u>) is sufficiently detailed to be useful.
Perceived ease of navigation (x9) (Navigability = Ease of Navigation)	The (inverse of the) effort that the user perceives is required to reach desired information or functionality within the site.
Perceived information understandability (x17)	The degree to which the user perceives he/she can comprehend the meaning of the presented information.
Perceived information believability (x7)	The degree to which the user perceives he/she can believe the presented information.
Perceived ease of use (x13)	The (inverse of the) effort that the user perceives is required to accomplish their intended goals on the web site.
Perceived ease of communication (x24)	The (inverse of the) effort that the user perceives is required to establish an alternative communication channel with the organization (the web site being the first channel).
Perceived technical response time (x25)	The perceived waiting time for visual feedback from the web-site.
Perceived information accuracy (x2)	The degree to which the user perceives the information presented on the web site to be consistent with his/her beliefs about the state of the domain.
Perceived information relevance (x14)	The degree to which the user perceives the information presented on the web site to be within the scope of the business domain of the website (belief about the scope).
Perceived information timeliness (x10)	The inverse of the perceived time lag between changes in the domain and corresponding changes in the presented information.
Perceived control (x19)	The degree to which the user perceives the web site to behave in the way the user expects (includes absence of unexpected side-effects).
Perceived attractiveness (x16)	The degree to which the user perceives the site to have visual beauty.
Perceived security (x11)	The degree to which the user perceives the web site protects his/her personal information from <i>external</i> threats.

Perceived safety (x6)	The degree to which the user perceives the web site to protect him/her from internal, accidental, self-inflicted damage.
Perceived usefulness (x8)	The degree to which the user perceives the web site to enable him to accomplish intended goals, (including but not limited to (ultimate) goals of saving time and money).
Perceived enjoyability (x21)	The degree to which the user perceives that interaction with the web-site is fun or pleasurable.
Believed competency (x20)	The degree to which the user believes the web site is able to efficiently and effectively carry out its perceived purpose(s).
Positive experience (x22)	User's affect towards the website. (attitude)
Belief about site reputation (x3)	(=Attitude Towards Site) Affective evaluative sum of beliefs (Fishbein) The beliefs or opinions that are generally held about something or someone (Oxford American Dictionary) noun: The general estimation that the public has for a person; "he acquired a reputation as an actor before he started writing"; "he was a person of bad report" (WordNet) noun: The state of being held in high esteem and honour (WordNet)
Belief about organizational reputation	See above.
Trust (x27) (but not clear, as may also incorporate items such as x11 and x6)	The degree to which the user believes the organization operating the web site to be able and willing to deliver promised services.

8.2.1.3 Step 3: Comparison with Literature Review and Previous Models

Following the procedures recommended by Hayduk (Hayduk & Glaser, 2000a, 2000b), we fixed the error term for the most important (in some cases only) indicator for each latent. This was based on general theory of attitude formation (Fishbein & Ajzen, 1975), and specific studies which included equivalent latent constructs and indicators.

From general theories of attitude formation, we determined that many of our single-item latents, while not quite fully descriptive (Fishbein & Ajzen, 1975), do not require a high degree of inference. *“Strictly speaking...a belief is formed as soon as an object is linked to an attitude...irrespective of the subjective probability associated with the link”* (Fishbein & Ajzen, 1975, p. 134). In setting a fixed error term, we are making an assessment about the subjective probability that the user’s perception is accurate. Our

latent variables were specifically designed as the likely proximal cause of the observed variables, and this means that the link between the user's descriptive beliefs about the actual information contained in the website (the user's direct perceptions of the object), and the user's perceptions of accuracy and timeliness (for example) is direct. Since there was not an extensive chain of inference, these perceptions are not likely to be extensively modified by other beliefs and therefore our measures have a high validity.

Another argument for assuming a high degree of accuracy is that there are no other, better measures for someone's perceptions of a range of service quality attributes than their own report. People's perceptions are their perceptions and they and only they can judge what they are. Likely sources of error will only occur if the user accidentally selects a different response from the one they intended.

We also returned to existing literature for assistance in setting the error term for our indicators. The literature was examined to identify previous item loadings on latents (where the same items had been used in previous studies) in order to provide an indicative magnitude for the error term of the indicators.

Several of our latent variables and indicators had been included in previous models, in particular, the UTAUT (x13, x1, x5, x3 and x4) (Venkatesh et al., 2003). Further, some of our indicators had been used explicitly in a similar business domain by Gefen (Gefen, 2003; Gefen et al., 2003) in studies of experienced (continuing) online shoppers (x13, x1, x5, x3). In both studies these indicators had error terms less of less than 0.2. Accordingly, we fixed the error term for our observed indicators to 0.2.

8.2.1.4 Step 4: Model Development

With twenty-five candidate latent variables and unclear or conflicting evidence, or no previous studies at all, to explain their relationships, developing candidate models was a challenging prospect. There are literally thousands of candidate models that could be justified based on previous studies or vernacular usage (nomological nets) of the candidate constructs. The authors developed and tested many variants. A large number of these models were not identified, or produced extremely poor fit indices.

These models were loosely based on the formative models presented in Chapter 5, but avoid the implied ontological confusion that requires an indicator to “cause” a psychological state. We abandoned the eQual-based latent constructs hypothesised in Chapter 4, because our improved understanding of psychological measurement suggests these were probably mis-specified, and the fit indices for the measurement models were extremely poor anyway. We initially concentrated on Ease of Use, Usefulness, and Trust and their nomological nets, but we were unable to develop any good-fitting models incorporating Trust and its antecedents.

This approach allows us to evaluate our theory of user attitudes and perceptions towards technology. In broad terms, we posit that the user brings all their previous experience into interactions with a particular technology (which influence the degree of “fit” between the user and the technology). Based on the characteristics and cognitive affordances of the technology, a range of descriptive beliefs will be possible and relevant for the interaction. These in turn will influence affective beliefs and attitudes towards the technology. A selection of the best-fitting or most interesting models are presented in this chapter.

8.3 Multi-indicator structural models

8.3.1 Overview

In this section we present seven models that attempt to fit an increasing number of the observed variables into multi-indicator structural models. These models are presented as illustrative examples, and are not intended to be the “last word” on the subject. In many cases we have been required to impose constraints or make assumptions that could be open to disagreement in order to achieve model identification. In reality, it is likely that the majority of the latent variables co-vary to a greater or lesser extent, and that a complicated network of interactions is occurring. It is also possible, due to the somewhat imprecise nature of the initial indicators, that some latent variables may load on more than one indicator, which further complicates the picture.

In this series of models (MISM1-MISM7), we begin with a relatively simple model that is similar to one of our formative models, showing direct perceptions of

Accuracy, Believability, Timeliness and Relevance leading to a more inclusive latent Information Quality, which in turn predicts an overall attitude towards the Perceived Usefulness of the site. We use this model as the core for a number of variants where we attempt to fit an increasing number of additional indicators. In the next sections, we first describe the development of the models, and then present and discuss the fit indices and comparative fit indices. A summary of the indicators included in each of the models is included in Table 8.2.

Table 8.2: *The indicators included in the models in this chapter.*

	MISM-1	MISM-2	MISM-3	MISM-4	MISM-5	MISM-6	MISM-7
1. I find this site easy to learn to operate							
2. The site provides accurate information	X	X	X	X	X	X	X
3. This site has a good reputation							
4. I believe I would be able to complete transactions that are useful to me	X	X	X	X	X	X	X
5. My interaction with the site is clear and understandable							
6. It feels safe to complete transactions							
7. The site provides believable information	X	X	X	X	X	X	X
8. Completing transactions on this site will save me time or money	X	X	X	X	X	X	X
9. I find the site easy to navigate							
10. The site provides timely information	X	X	X	X	X	X	X
11. My personal information feels secure							
12. This site offered the range on on-line transactions I expected							
13. I find the site easy to use				X	X	X	X

14. The site provides relevant information	X	X	X	X	X	X	X
15. This site creates a sense of personalisation							
16. This site has an attractive appearance					X		X
17. The site provides easy to understand information				X	X	X	X
18. The design is appropriate to the type of site					X		X
19. I feel in control on this site							
20. The site conveys and sense of competency			X		X		X
21. This site conveys a sense of enjoyability or entertainment							
22. The site creates a positive experience for me			X		X		X
23. The site provides information at the right level of detail		X	X	X	X	X	X
24. This site makes it easy to communicate with the organisation							
25. Overall, the response time was acceptable to me							
26. The site presents the information in an appropriate format		X	X	X	X	X	X
27. I feel confident that the goods/services I have requested will be delivered as promised							
28. What is your overall view of the quality of this website			X		X		X

8.3.2 MISM-1

In this model (Figure 8.1) we concentrate on descriptive beliefs about the information presented, and the influence of those beliefs on overall perceptions of usefulness.

Figure 8-1: *MISM-1*

We hypothesise single indicator latent variables for Perceived Information Accuracy (x2), Perceived Information Relevance (x14), Perceived information Timeliness (x10), and Perceived Information Believability (x7). These latents are allowed to co-vary. We measure each of these with direct indicators, and a low error variance fixed at 0.2. We suggest that collectively, these contribute to a formative variable we have called Perceived Information Quality, which in turn contributes to Perceived Usefulness. Perceived Usefulness is measured by two direct, reflective indicators, x4 and x8. Both of these are closely related to existing indicators of perceived usefulness from previous literature, for example (Gefen et al., 2003). This produces a well-fitting model.

In this model (Figure 8.2) we introduce a latent variable to represent the degree of contextual “fit” between the user and the technology based on their past experience.

Figure 8-2: *MISM-2*

Rather than allowing Perceived Information Accuracy (x2), Perceived Information Relevance (x14), Perceived Information Timeliness (x10), Perceived Information Believability (x7) to co-vary, we hypothesise a common cause, Perceived Presentation Fit, which we measure using Perceived Appropriateness of Information Detail (x23) and Perceived Appropriateness of Information Format (x26). Although we have previously hypothesised candidate latent variables for both Perceived Appropriateness of Information Detail (x23) and Perceived Appropriateness of Information Format (x26), we collapse these together, since they do not have any independent relationships in our model. Perceived presentation fit may arise from the shared context and domain understanding of the users and site designers, and also from prior experience with the site, as the user has a chance to adjust their mental models to avoid cognitive dissonance. This also yields a good-fitting model, although we needed to free the error variance on x26 and x23.

8.3.4 MISM-3

This model (Figure 8.3) extends MISM-2. We add a latent variable for overall affect towards the service quality of the site (Perceived Overall Service Quality).

Figure 8-3 : *MISM-3*

We measure this using a direct indicator “What is your overall view of the quality of this website” (x28), and two other indicators that capture overall affect and belief about the site, Believed Competence (x20) and Positive Experience (x22). We need to add an indirect effect from Perceived Presentation Fit to Perceived Usefulness in order for the model to converge. We attempt to fix the error variances on our indicators, but we need to free most of them, apart from the original four, which remain fixed at 0.02. This also provides a good-fitting model.

8.3.5 MISM-4

In this model (Figure 8.4) we add a new latent for ease of use and attempt to fit it. This model is an example of the insights that can be gained from a more granular model.

Figure 8-4: *MISM-4*

In this case, we illustrate possible issues with our question wording in the context of our particular site and user sample. Based on the outputs of other model variants, it seems likely that we can improve the fit of our model by allowing one of our indicators to load on more than one latent. Conventional wisdom in modelling suggests that this is undesirable, and shows poor discriminant validity. In this model, we allow x4 (I believe I would be able to complete transactions that are useful to me) to load on Perceived ease of Use as well as Perceived Usefulness, and this produces an improvement in our fit indices.

This might appear to be “capitalising on chance”, but there is an explanation. At the time this sample was taken, there was considerable dissatisfaction associated with some of the functionality implemented through the website. Enrolling for courses, could be initiated on the website, but at a certain point on the process, the user would be redirected to a face-to-face process to complete their enrolment. The enrolment process was also not effective at storing partially completed details for the user to return to at a later point. The online process was not easy to use, and was notorious for failing part-way through. Our particular wording “I believe I will be *able to complete* transactions that are important to me” was open to a different interpretation in an environment where completing online transactions was a known problem.

8.3.6 MISM-5

In this model (Figure 8.5) we need to impose additional constraints to enable the model to be identified. The cross-loading on x4 is removed again. This variant is a TAM-type model, with ease of use contributing to usefulness, and both of these contributing to POLSQ. Metrics are not reported for this model as it did not converge and the solution is not admissible. However, this could act as a starting point for future research.

Figure 8-5: *MISM-5*

8.3.7 MISM-6

In this model (Figure 8.6) we attempt to fit a new latent variable, “Perceived Design Fit”, which is reflected by .x16 (The site has an attractive appearance) and x18 (The design is appropriate to the type of site). The cross-loading on x4 is removed. As with perceived presentation fit, we posit that design fit is positively associated with ease of use, because it reduces the cognitive load on the user. We move x17 to perceived presentation fit. Previous models had shown some cross-loading on this item. This also improves the fit statistics. Unlike x4, in this case, we acknowledge that there may be issues with discriminant validity. The fit indices begin to drop at this point as we try to fit more items and latent variables.

Figure 8-6: MISM-6

8.3.8 MISM-7

This model (Figure 8.7) is the most complete model we have chosen to present, including fifteen of our original twenty-eight indicators, and ten latent variables. In broad terms, it hypothesizes that the degree of cognitive fit between the user and the site characteristics, (which may be inherent, or the result of the user's previous experience with the site) will influence the user's descriptive beliefs about the site, which in turn will influence their overall affect towards the site.

This model has potential for strong explanatory power for the processes involved in forming perceptions of online service quality, but unfortunately, the fit indices have begun to drop. This may be useful as a starting point for future research.

Figure 8-7: MISM-7

8.4 Results

These models all use bootstrapping, as recommended for non-normal data. We are mostly looking for overall and comparative model fit, with a particular emphasis on the structural paths. Factor loadings are not important for this type of “non-factor” model. Squared multiple correlations are included to show the explanatory power of the models for latent variables, and to enable these models to be compared with other variants in our other chapters. The fit indices previously identified as being the most appropriate for our sample are highlighted.

Overall fit metrics, for each of the models are shown in Tables 8.3. and 8.4. Squared multiple correlations are shown in Table 8.5. The full reports for these models as generated by AMOS are included in Appendix IX.

Table 8.3: *Overall fit metrics-1*

Metric	MISM-1	MISM-2	MISM-3	MISM-4	MISM-5	MISM-6	MISM-7	Comments
Chi squared	4.917	34.477	74.698	44.931		152.700	260.082	Reject if >.5
Degrees of freedom	4	18	40	31		52	83	
Bollen Stein bootstrap	p = .473	p = .234	p = .085	p = .423		p = .005	p = .005	p > .05 indicates a good fit
Relative Chi-squared (CMIN/DF)	1.229	1.915	1.867	1.417		2.937	3.134	You may have dropped too many paths on the default model (from the saturated model) if this exceeds 2-3
Hoetlers critical N	282	123	109	150		67	60	Should be greater than 200
GFI	.989	.944	.921	.947		.873	.823	Should be >.95 or at least .9
AGFI	.941	.889	.870	.905		.809	.744	Should be > .9
RMR	.037	.089	.067	.065		.273	.286	The closer to 0 the better
CFI	.997	.964	.957	.977		.849	.838	Should be >.9
IFI	.997	.965	.958	.977		.852	.841	Should be >.9
NFI	.986	.929	.914	.927		.792	.783	Should be >.9, sometimes .8 used
NNFI (TLI)	.990	.944	.941	.966		.809	.795	Should be >.9, sometimes .8 used
PRATO	.267	.643	.727	.689		.799	.790	Ratio of how many paths you have dropped, compared to how many you could have dropped
RMSEA	.040	.079	.077	.053		.115	.121	Good fit if <= .05, adequate fit if <=.08
PGFI	.188	.472	.558	.534		.582	.659	GFI Index adjusted to reward simple models
PNFI	.263	.597	.664	.639		.624	.619	NFI Index adjusted to reward simple models
AIC	38.917	70.477	126.698	91.931		204.700	334.082	Closer to zero is a good fit
BCC	40.630	72.842	131.355	95.842		209.783	343.190	Needs to be close to .9 to be considered a good fit
ECVI	.267	.483	.868	.630		1.402	2.288	Lower is better

Table 8.4: *Overall fit metrics-2*

Metric	MISM-1	MISM-2	MISM-3	MISM-4	MISM-5	MISM-6	MISM-7	Comments
RMSEA	.040	.079	.077	.053		.115	.121	Good fit if $\leq .05$, adequate fit if $\leq .08$
PGFI	.188	.472	.558	.534		.582	.659	GFI Index adjusted to reward simple models
PNFI	.263	.597	.664	.639		.624	.619	NFI Index adjusted to reward simple models
AIC	38.917	70.842	126.698	91.931		204.700	334.082	Close to zero is a good fit
BCC	40.630	72.842	131.355	95.842		209.783	343.190	Needs to be close to .9 to be considered a good fit
ECVI	.267	.483	.868	.630		1.402	2.288	Lower is better

Table 8.5: Squared multiple correlations

MISM-1		MISM-2		MISM-3		MISM-4		MISM-5		MISM-6		MISM-7	
infQual	.976	accuracy	.592	accuracy	.582	Accuracy	.571			Perceived presentation fit	.000	Perceived presentation fit	.000
Usefulness	.599	timeliness	.582	timeliness	.579	Timeliness	.594			accuracy	.573	accuracy	.693
x8	.321	relevance	.682	relevance	.677	Relevance	.700			timeliness	.592	timeliness	.653
x4	.412	believability	.657	believability	.651	believability	.629			relevance	.699	relevance	.827
x14	.843	infQual	.958	infQual	.988	infQual	.969			believability	.631	believability	.778
x10	.893	usefulness	.557	Overall SQ	.995	ease of use	.498			Perceived design fit	.000	Perceived design fit	.000
x7	.822	x23	.395	x20	.673	x8	.421			Ease of use	.357	Ease of use	.743
x2	.835	x26	.284	x22	.604	x4	.499			usefulness	.595	usefulness	.657
		x4	.320	x28	.686	x13	.302			x16	.205	Overall SQ	.999
		x14	.411	x23	.412	x17	.425			x18	.624	x28	.622
		x10	1.00	x26	.311	x23	.420			x17	.301	x22	.507
		x7	1.00	x8	.283	x26	.287			x8	.539	x20	.632
		x2	1.00	x4	.395	x14	1.00			x4	.344	x16	.225
				x14	1.00	x10	1.00			x13	.447	x18	.538
				x10	1.00	x7	1.00			x23	.402	x17	.300
				x7	1.00	x2	1.00			x26	.291	x8	.323
				x2	1.00					x14	1.00	x4	.346
										x10	1.00	x13	.696
										x7	1.00	x23	.419
										x2	1.00	x26	.307
												x14	.842
												x10	.893
												x7	.821
												x2	.835

8.4.1 Model Comparison

This section shows the overall fit indices for the models in this section ranked from best to worst (table 8.6 and 8.7).

Table 8.6: *Overall fit indices compared-1*

Metric	MISM-1	MISM-2	MISM-3	MISM-4	MISM-5	MISM-6	MISM-7	Comments
Chi-squared	4.917	34.477	74.698	44.931		159.703	260.082	
Degrees of freedom	4	18	40	31		52	83	
Bollen Stein bootstrap	p = .473	p = .234	p = .085	p = .423		p = .005	p = .005	p > .05 indicates a good fit
Relative Chi-squared	1.229	1.915	1.867	1.417		2.937	3.134	You may have dropped too many paths on the default model (from the saturated model) if this exceeds 2-3
NNFI (TLI)	.990	.944	.941	.966		.809	.795	Should be >.9, sometimes .8 used
RMSEA	.040	.079	.077	.053		.115	.121	Good fit if <= .05, adequate fit if <=.08
RANK	1	2	4	3		5	6	

Table 8.7: *Overall fit indices compared-2*

Metric	MISM-1	MISM-2	MISM-3	MISM-4	MISM-5	MISM-6	MISM-7	Comments
AIC	38.917	70.842	126.698	91.931		204.700	334.082	Close to zero is a good fit
ECVI	.349	.483	.868	.630		1.402	2.288	Lower is better
RANK	1	2	4	3		5	6	

8.5 Discussion

8.5.1 Overview

These examples show that it is possible to generate good-fitting models of our business domain using a sub-set of our existing indicators and a multi-indicator-structural model approach. All of these models are better-fitting than those in the previous chapters, although this is partly explained by having fewer indicators to fit. These models also provide some support for aspects of our arguments about modelling perceptions and attitudes towards technology, and the structure and antecedents of POLSQ. Our models also suggest that the explanatory power of information systems models of online shopping and information systems continuance is stronger than that of eQual-based models, which look back to ServQual.

8.5.2 Model fit

Unsurprisingly, the best-fitting models are the simplest. The only model that can be considered extremely good-fitting based on a broad range of fit criteria is MISM-1, which is a simple, almost trivial model of the structure of perceived information quality and its effect on perceived usefulness.

Fit indices for MISM-2 through MISM-5 are consistently good across a range of indicators, apart from those that penalise for lack of parsimony, such as BCC. This is not surprising, as this modelling approach deliberately sacrifices parsimony for accuracy.

As we established earlier in Chapter 5, of the overall fit metrics, CMIN/DF, Hoetler's N, and RMR are the most appropriate for our purposes. Goodness of fit measures that compare our hypothesised model with the null model such as the CFI, IFI, NFI and TLI are less affected by sample size, so are likely to offer valuable insights for our study.

Of these, only MISM-1 includes a Hoetler's N which is above the recommended cut-off of 200, but MISM2-MISM-5 are well within recommended criteria for the other key indicators and can be considered well-fitting models.

The fit indices for MISM-6 and MISM-7 have dropped as we try to fit more indicators, and fall just outside recommended cut-off criteria for good-fitting models, although not by a very wide margin.

8.5.2.1 Modelling perceptions and attitudes towards technology

These models provide some support for the salience of MISM's (following Hayduk's "four-step" process as closely as possible) in modelling perceptions and attitudes towards technology.

We have demonstrated previously that many of the items modelled reflectively could vary independently (and are therefore mis-specified), and that formative models require dubious ontological claims. By decomposing the latent variables that had previously been hypothesised (for example, content quality, or usability) into a larger number of latent variables, and following the "four-step" approach, we were able to address these issues.

Adding additional latent variables allowed us to replace many measurement arrows with structural arrows, to make our theory more explicit. This approach makes the correspondence between theory and the world extremely close. If we model a large number of clearly defined and granular latent constructs, we are not required to make great leaps of faith about our unobserved latent variables. When we follow steps 1 and 2 of the "four-step" we end up with a single, direct indicator of our hypothesised latent, with a loading of 1.0 (for example x2 "The site provides accurate information" as an indicator of the hypothesised latent variable Perceived Information Accuracy). By hypothesising Perceived Information Accuracy as its own latent, we are positing that a user is able to distinguish perceptions of the accuracy of the information on the site from their perception of other quality features, which seems intuitively correct.

We had some success in following step three of the "four-step", which advocates fixing the error term on the most direct indicator based on theory. We were able to fix the error terms for x2, x7, x10, and x14 to 0.2, based on previous studies, and still achieve good-fitting models. However, we were required to free the other error terms.

The “four-step” method proved extremely rigorous, and if we had followed it rigidly, we would have had nothing to present.

We believe these models are superior to the formative models presented in Chapter 5. They avoid the ontological confusion that requires an indicator to “cause” a psychological state by positing a psychological state (latent variable) that is the direct cause of every conceptually distinct indicator. They are more explicitly specified, better fitting, and address many of the known issues with formative models, such as the ambiguity over the meaning of the error term, and the difficulty in selecting meaningful fit indices.

This approach allows us to test our general theory of user attitudes and perceptions towards technology. We posited that the user brings all their previous experience into interactions with a particular technology (which influence the degree of “fit” between the user and the technology). Based on the characteristics and cognitive affordances of the technology, a range of descriptive beliefs will be possible and relevant for the interaction. These in turn will influence affective beliefs and attitudes towards the technology.

MISM-1 identifies a range of descriptive beliefs associated with information, including accuracy, believability, timeliness and relevance, and models these as antecedents of perceived information quality, which in turn precedes perceived usefulness.

Models MISM2 through MISM5 add a latent for contextual fit (perceived presentation fit). MISM3 completes the model by including a measure of overall affect towards the site. These models are better fitting than any of our previous models, suggesting that this general theory has merit, and can be applied effectively in specific contexts to explain what characteristics of a technology will likely contribute to a user’s overall effect towards it. Although the fit indices for MISM6 and MISM7 are marginal, they continue the same theme.

In summary, we feel that our MISMs address the issues of some of the previous approaches, provide some support for our general theory of user attitudes and

perceptions towards technology, and provide a direction for future models of user attitudes and beliefs towards technology. However, there are a number of disadvantages. These models become very complex very quickly, and privilege accuracy over parsimony. In their purest form, they are extremely restrictive and very difficult to fit.

8.5.3 Comparison with our other models

We are able to make some observations about the similarities between the results of these models and the models presented in our previous chapters, but since we were only able to fit fifteen of our twenty-eight indicators, we cannot compare every construct. We will focus on information quality, usefulness, and ease of use as these were the only constructs we were able to fit in our MISMs.

Difficulties in achieving model identification and model fit with the MISM's mean that direct comparisons are difficult. We have some degree of "interpretational confounding", in that some indicators have moved from being formative to reflective, although they are in the same nomological net. For example, in MISM4 and 5, we use x13 and x17 to reflect ease of use, while in our formative model; we use x5 and x13, while x17 is specified formatively. This issue arises because we are using the same indicators and data set for all our models, rather than indicators developed specifically for each model, so in some case, we have been required to make use of the best available indicator. These comparisons should therefore be taken as indicative, rather than as a true meta-analysis of the models.

We first consider information quality. In our reflective models, indicators for accuracy (x2), believability (x7), timeliness (x10) and relevance (x14) were high-loading on content quality. Level of detail (x23) and appropriate information format (x26) are also included as reflective indicators of content quality. In our model, x2, x7, x10 and x14 are represented as separate latent variables, with a causal effect on Perceived Information Quality. We include the indicators x23 and x26 are in the same nomological net as Perceived Information Quality but we posit them as indicators of cognitive fit rather than direct indicators of CQ.

The indicators x2, x7, x10, x14, x23 and x4 are included as formative indicators of usefulness in our formative model; our MISM also places them in the same nomological net.

The similarity of the nomological nets across the three models allows us to make a comparison, and the better fit of the MISMs suggests that these provide a more accurate representation of reality with regard to information quality and usefulness, and that the structural relationships are in the correct direction.

For Ease of Use, the picture is less clear. In our CFA, x17 is not associated with ease of use, but with information quality. Some of our MISMs (not presented) showed x17 cross-loading with ease of use and perceived presentation fit. In MISM6 and MISM7 we move x17 to perceived presentation, which places it in the nomological net for information quality and usefulness.

Our reflective items for perceived ease of use x13, and x17 are both modelled as reflective of usability in our CFA, while in our formative model, x13 is modelled reflectively to identify the model, while x17 is formative. When we add perceived design fit to the nomological net for perceived ease of use, we bring x16 and x18. These were modelled as reflective of usability in our CFA. In the formative model x18 was modelled as formative of ease of use, while x16 was dropped.

The inconsistency in placement of indicators, and the marginal fit indices for MISM6 and 7 mean that it is difficult to draw any clear observations about this construct. The general direction of the models is similar to that for ease of use, but the evidence is much more confused and weaker.

A further discussion of the implications of these findings, and a comparison of this research phase with the other analyses conducted, is included in Chapter 9.

8.5.4 Comparison with models from literature

Our models support previous research that finds well-established constructs from information systems have strong explanatory power in predicting overall user effect towards a websites (Gefen et al., 2003; Tate et al., 2008a).

Our MISMs support previous research in the online shopping context (Gefen et al., 2003) that hypothesises the TAM-based constructs of Perceived Usefulness and Perceived Ease of Use as antecedents of affect towards a website.

The clearer effect of perceived usefulness on POLSQ supports the findings of IS continuance literature, which finds that the importance of ease of use diminishes over time compared to perceived usefulness (Bhattacharjee, 2001). Since we set out to sample, experienced, continuing users, this finding is expected.

Unfortunately we were not able to fit indicators for trust and fun in our MISMs so we are unable to compare these with previous studies.

8.6 Conclusion

We present these model fragments with some caution; we invite you to consider them as illustrations of a modelling process, rather than as finished pieces of research in the on line service quality area. For readers who would like to argue about the specifics of the models and suggest alternatives – we agree with you – there are many. Many of them we have tried and chosen not to present. With the subset of models we have selected, we can make some tentative observations about modelling processes and methods, and about the structure of POLSQ.

With regard to modelling methods, these models illustrate both the precision and the frustration inherent in this modelling method. We were able, in a series of models, to create multiple alternatives to explore very detailed and precise aspects of the phenomenon. By modelling the cross-loading on x4 (I believe I would be able to complete transactions that are useful to me) in MISM-4, to capture perceptions that are unique to this sample in this context in a way that is useful and interesting,

because it provides a very specific diagnosis of a user satisfaction issue. We could show the gradual development by building from simpler to more complex models.

We can also see why this form of modelling is not popular in the information systems community compared to methods such as PLS. It is extremely exacting. We were unable to fit all our indicators. Many of our candidate models did not converge or had very poor fit indices. The number of models published in information systems would be likely to reduce dramatically if this modelling approach were to become more widespread.

Despite the fact that we were only able to fit model fragments, we are able to make some observations about the structure of POLSQ. Our models seem to support development from more descriptive to more influential beliefs. There is also evidence that a common cause, such as the degree of fit between the software and the user's requirements influence their beliefs.

Descriptive beliefs cause an inferential belief about the information quality of the website which causes an overall affect towards the perceived usefulness of the site. We free the error term on the perceived usefulness construct, as it is possible that there are unknown other beliefs that contribute to overall beliefs about usefulness. It is likely that our model will explain only a portion of the variance of perceived usefulness of the information website, as there are likely to be other, un-modelled constructs that also contribute.

We were able to add constructs for perceived presentation fit and perceived design fit to our models. This is an area with potential for further research. It is possible that ongoing users of a service tend to adjust their expectations as far as possible to match what they believe they will receive, to reduce cognitive dissonance (Boulding et al., 1993), creating a degree of "fit". More experienced users will likely have a higher "perceived fit" and therefore more positive perceptions. We have suggested that the cognitive affordances of the website define the scope of the relevant descriptive beliefs, but it is possible that aspects of the design and presentation, and their fit to user's mental models, also influence the strength of those beliefs. Established theory in social psychology (Fishbein & Ajzen, 1975) suggests that previous experience

influences current belief formation. There appear to be opportunities for further research on the “left hand side” of the model, in exploring how design features, user’s mental models, and experience with the site combine to influence belief formation.

The meaning, scope and likely generalisability of the theory are made clear in these examples. We have eschewed a generic label such as “perceived usefulness” and made explicit exactly what we mean by perceived usefulness in this theory net. We have modelled what we believe are the antecedents to perceived usefulness that have salience in this context. We have allowed different characteristics of information to vary independently, which provides the model with high diagnostic power. It is possible that the contribution of the various qualities of information might vary from one information provisioning site to another. For example, timeliness might make greater contribution to the perceived information quality of a traffic report website than to an advanced mathematics homework revision site, where accuracy and relevance to the syllabus might be more highly valued.

Although our findings are tentative, these models find support for the argument that POLSQ is an overall affect towards a website, which is influenced by generalised beliefs about the usefulness and ease of use of the site. These beliefs in turn are influenced by more descriptive beliefs. These models are both better fitting, and have a stronger theoretical foundation, than the EFA-based models used for Equal and ServQual.

9 Discussion

9.1 Overview

In this chapter we return to our research questions and discuss our findings. We finish with implications for research and practice and the limitations of the research.

We recollect that our research questions are:

1. What is the structure of POLSQ?
2. What are the antecedents of POLSQ?
3. What is the ontology of POLSQ?
4. What are the most appropriate theoretical, paradigmatic and measurement methods for measuring online service quality quantitatively, and what insights can be gained from psychometrics?
5. What insights does this offer IS researchers for the measurement of user attitudes and perceptions towards technologies?

9.2 The Structure of POLSQ

At the beginning of the study, we noted that despite decades of research, there was no consensus about the structure of POLSQ. We noted that information systems research in the area relied heavily on techniques such as narrative literature reviews, in many cases followed by exploratory and confirmatory factor analysis.

Our first approach to this question considered the possibility that this approach was sound, and the instability in the structure of POLSQ was the result of changes in the affordances of online technologies. This appeared to be supported by our narrative literature review, which suggested that the range and usefulness of online transactions might emerge as additional service quality criteria following the wide availability in recent years of online self-services.

To evaluate this possibility, we selected a leading instrument and added additional indicators, to enable us to compare the factor structure emerging from our study with

the factors identified in previous research. In chapter 4 we carried out exploratory factor analysis on the dataset gathered using our revised instrument, using principal component analysis and oblique and orthogonal rotation. This approach was selected in order to provide a direct comparison with previous eQual studies (Barnes, 2001; Barnes et al., 2001; Barnes & Vidgen, 2001, 2002, 2004, 2007). Our factor analysis identified four orthogonal components, and provided some support for our suggestion that unstable dimensionality was a result of changing affordances, as our new items associated with transaction range and usefulness loaded on a new factor at the level of 0.4 or above which was considered “worth reporting” (Straub et al., 2004), while the Cronbach’s α ’s for each of the four factors was 0.6 or above which is considered appropriate for exploratory analysis (Straub et al., 2004). However, the item factor loadings for existing survey items were not consistent with the findings of previous eQual studies.

The exploratory factor analysis provided the basis for a further evaluation of the structure of POLSQ using confirmatory factor analysis, as we now had several different factor structures. In Chapter 5, we conceptualised POLSQ as three, four, or five factors with reflective indicators. This followed a widely adopted practice in online service quality research of conducting an EFA followed by a CFA (Alzola & Robaina, 2005). The CFA models we developed were based on the eQual stream of research, our narrative literature review, and our exploratory factor analysis

The analysis in chapter 5 allowed us to establish with a reasonably degree of certainty what the structure of POLSQ is NOT. POLSQ should NOT be modelled reflectively as a set of 3-5 factors composed from the set of indicators we selected. It does not appear to be justifiable to model it as consisting of content quality, usability, transaction quality, interaction quality, or any of the other variants included in the models we present in chapters 4 and 5. These models did not produce stable results when compared with previous eQual studies, and the fit indices for all our model variants based on this approach were extremely poor.

Our second line of enquiry, that our existing theoretical and measurement approaches are flawed, was also supported. The application of recent insights from psychometric literature to the domain of POLSQ suggests that the reflective specification used for

CFA models (where variation in the hypothesised factor precedes variation on the indicators) is conceptually incorrect. Many of the supposedly reflective indicators are not interchangeable (for example information timeliness and information relevance), and can vary independently. Therefore they cannot be reflective of the same underlying latent construct.

So changing affordances (the wider availability of online self-service transactions) are likely affecting user's perceptions of what constitutes online service quality but this change alone does not account for the unstable dimensionality in previous studies, it seems our theoretical and measurement approaches were also flawed. However, this is a negative finding. We turn next to more positive results.

Rather than being a “multi-dimensional construct”, or a set of several factors, we suggest that POLSQ is best conceptualised as an overall affect (attitude) that the user holds towards the service quality provided by the website. Concepts in a theory net do not or should not have “components”, or “facets” or “dimensions”: they have antecedents and consequences. The antecedents to this attitude are the perceptions (descriptive beliefs) and inferential beliefs the user holds about specific characteristics of the service. Our justification for this conceptualisation is partly theoretical, based on social psychology, and partly empirical, based on our formative models and multi-indicator structural models.

9.3 The Antecedents of POLSQ

If we define perceived service quality as overall attitude towards the service quality of the site, rather than as a multi-dimensional construct, we can next examine the antecedents to that attitude.

We first separated out two items that appeared to be reflective indicators of overall attitude (x22, “This site creates a positive experience for me” and x28 “What is your overall view of the quality of this website?”). We then examined the remaining item and developed a series of alternative models based on theory and previous studies.

9.3.1 An unstructured set of antecedents

In Formative Model 1 we posited that the twenty-six items included in our revised instrument (once the items reflective of overall attitude were taken out) formed an unstructured set of antecedents of POLSQ.

We had reasons to believe that the items would form a convincing set of antecedents. The item set was based on extensive qualitative research. The eQual item-set was originally developed based on focus group research (Barnes, 2001) which was triangulated extensively with research literature (Barnes & Vidgen, 2002). In addition we conducted independent qualitative research to confirm the scope of the phenomenon and the relevance of the items to our research context (Tate et al., 2008b; Tate et al., 2007). The items also appeared to have real-world salience. The eQual instrument was selected by major government department within the United Kingdom to evaluate the quality of their web services (Barnes & Vidgen, 2007).

We note that our proposed new items were all supported by our study. Our EFA suggested that all the items (including the new items) should be retained, and our formative model 1 found that all the items loaded on POLSQ at a level of 0.4 or above.

If we triangulate the item-set in our study with the cognitive affordances of websites, we can see that these are fully covered. Experienced web users are very unlikely to have perceptions of attributes that a website cannot possess (affordances it does not have). As we discussed in our literature review in Chapter 2, the cognitive affordances of websites have evolved from information provisioning to transaction processing. Service delivery online typically affords interaction using a hyperlinked interface, information provisioning, communication, and increasingly, the ability to conduct online transactions. All of these affordances are included in the item set.

Modelled as an unstructured set of antecedents using PLS, the items explained approximately 80% of the variance in the dependant variable. There is little evidence of collinearity. All the items load on the dependent variable at the 0.4 level or above. There do not seem to be any redundant items that do not make a significant

contribution to explaining the dependant variable. However, the error term was “medium to large” according to our heuristics. This suggested that all the indicators should be retained, but there were likely to be further unmeasured causes of POLSQ.

This is not surprising given that it is a highly generalised construct. Our discussion based on social psychology suggests that attitudes are more generalised than perceptions, tend to have a larger error term, and have more unmeasured causes. Given that we are measuring individuals in a social context we are very unlikely to capture every potential antecedent to an attitude they hold. Overall, our 26 items seem to form a reasonable set of antecedents to the formation of an overall attitude towards the service quality of a website. However, this model does not provide any useful insights about the potential relationships between these 26 items.

9.3.2 Formative model of trust, usefulness and ease of use

We have established that previously posited factors such as “content quality” and “transaction quality” did not appear to be appropriate. Further, many of the indicators included in eQual had been used in earlier studies as reflective indicators of latent variables widely used in information systems research such as usefulness, ease of use, and trust. Since the items appeared to form part of the (unstructured) set of antecedents for POLSQ, then it is possible that the constructs they reflected were also antecedent to POLSQ.

We modelled latent variables for trust, usefulness and ease of use as a formative model in AMOS (Formative Model 2). Modelled formatively, the selected indicators predicted more than 80% of the variance for Trust, Usefulness and Ease of Use. Collectively, these predicted more than 90% of the variance in perceived overall service quality.

We dropped the latent variable for perceived fun or enjoyability from our formative model. Our descriptive statistics found that these items had the highest standard deviations in terms of their importance to our respondents, so it is not surprising that it was difficult to fit these items, as there did not seem to be a wide consensus within the sample.

Although Formative Model 2 does not have good SEM fit statistics, experts are divided as to the extent to which conventional fit statistics can be used as quality indicators for a formative model. If we use the model weights, r^2 and error terms to evaluate the formative model, as proposed by Chin, and Petter et al. (Chin, 1998; Petter et al., 2007), then Formative Model 2 can be considered a good-fitting formative model with good predictive power.

Allowing for the ongoing debates about formative models, this result provides some support for the assertion that perceptions of Trust, Ease of Use, and Usefulness are antecedents of Online Service Quality. This finding supports previous qualitative research by the author which developed cognitive maps of user's perceptions of online services (Tate et al., 2008a). This study found that information systems theory, particularly continuance theory, had superior explanatory power for user's perceptions of online services than consumer behaviour theories.

This model is also consistent with our assertion that attitudes towards technologies (such as POLSQ) build up from descriptive beliefs (such as beliefs about the response time, or the accuracy of the information), to inferential beliefs (for example, whether the service is useful), and ultimately towards an overall affect.

However, we have previously discussed the ontological confusion associated with formative models. As it stands, if we strictly adopt a realist ontology for Trust, Ease of Use, and Usefulness (i.e. if we claim that they are separately existing internal psychological states) then our Formative model 2 makes the claim that our respondents' responses on our instrument "caused" these states, which is clearly ridiculous. If we want to adopt a realist ontology, then we need to hypothesise additional unobserved latent variables as the causes of our observed indicators. This can be achieved using multi-indicator structural models.

9.3.3 More nuanced models of the antecedents

Our final modelling approach was to develop a number of variants of multi-indicator structural models, based on a greatly expanded number of latent variables which we posited as the direct cause of many of our indicators. Our ontological assumption was that these represented descriptive beliefs about characteristics of the website. We

continued to develop the approach we took with our formative models based on Usefulness, Ease of Use, and Trust.

We were unable to fit all twenty-six indicators using this approach. In particular, we were unable to fit models involving the items associated with trust, but several model fragments based on a sub-set of the items for Ease of Use and Usefulness provided good-fitting models. These models also offer cautious support for the assertion that perceived usefulness and perceived ease of use are antecedents to POLSQ, and that descriptive beliefs about (for example) the accuracy and timeliness of the information may be antecedents to beliefs about the perceived usefulness of the online service.

9.3.4 Summary

Overall, the structure of online service quality was much less well understood than previous largely factor-based research had led us to believe. As we discussed above, perceptions towards the Trust, Usefulness and Ease of Use appear to be important antecedents of the Perceived Service Quality of a website. Our multi-indicator structural models provide some support for extending this nomological net further, based on our generalised model from social psychology. In summary, we suggest that the user brings all their previous experience (attitudes and mental models) into interactions with a particular technology (which influence the degree of “fit” between the user and the technology). The technology has independently existing characteristics and affordances. Attitudes and perceptions towards a technology are formed a result of an interaction between the user and the technology. Based on the characteristics and cognitive affordances of the technology, a range of descriptive beliefs will be possible and relevant for the interaction. These will be held to varying degrees of strength depending on the user’s prior context (mental models and prior experiences). These descriptive beliefs in turn will influence affective beliefs and attitudes towards the technology. This model was originally presented as Figure 3.2. We revisit it now, and have included it as Figure 9.1.

Figure 9-1: *The formation of attitudes towards websites (revisited)*

This general model was operationalised in our MIS5 and 6. These models suggested that “fit” between the user’s mental models and the site (e.g. perceptions of appropriateness) may be antecedent to beliefs about specific site characteristics (e.g. timeliness, relevance). These in turn may be antecedent towards inferential beliefs such as Perceived Usefulness, which may be antecedent to overall attitude towards the Service Quality of the site. We acknowledge that these models had marginal fit indices and therefore cannot be regarded as conclusive evidence, but they offer some insights and areas for future research with regard to the antecedents and nomological net for POLSQ.

9.4 The ontology of POLSQ

9.4.1 Ontological assumptions

In this study, we explored the ontological assumptions we make about user attitudes and perceptions, and the relationship between those assumptions and various modelling and measurement approaches. This enables us to distinguish the truth-claims of various models from the issues of statistical fit. It is possible for models that make no sense ontologically to have excellent fit indices, but this does not make them good models.

Overall, there are two common ontological positions that are adopted with regard to the quantitative measurement of POLSQ. The first is instrumentalism, and the second is realism. Each of these has different implications from a modelling and measurement perspective.

Instrumentalism assumes that theories are merely instruments, tools for predicting and summarising data. This neatly avoids making any truth claims about POLSQ – all we are claiming is that a certain set of observations (such as responses on a survey) seem to be effective in predicting the values of other observations (for example, an increase in web traffic). This assumption can be useful as a bench-marking approach to compile an index or sum-score of indicators that are salient in the real-world domain. Rather than hypothesising an independent existence for the latent constructs, they are simply sum-scores used for data reduction purposes. eQual, which is used to calculate an index that can be used for comparison or bench-marking, sometimes appears to be viewed in this way by its authors.

An instrumentalist ontology is consistent with any modelling method; all that is required for a quality model is that the observations taken by the researcher have good predictive power for the dependent variable when the model is applied.

The problem is that it is very difficult to carry out a quantitative study of POLSQ without introducing a realist discourse at some stage in the process. Definitions of constructs, and qualitative studies aimed at eliciting items tend to refer to beliefs, attitudes, perceptions and experiences as if they existed, as we discussed in our introduction, see for example, (Gefen et al., 2003).

Studies of perceived online service quality usually explicitly or implicitly assume a realist ontology: that there are real, but unobserved psychological states within the user of online services, that can be measured meaningfully by observed indicators, and that changes on these states cause responses on the indicators. We also assume that there is no external standard by which the responses can be evaluated, no independent bench-mark, such as exists, for example, with IQ tests. We assume that people are qualified to make their own judgements about their own internal attitudes towards technologies and will do so with a high degree of accuracy. This means that

measurement error, to the extent it occurs, should mainly arise from carelessness, lack of thoughtful attention when completing the instrument, and other uncontrolled nuisance factors.

A realist ontology requires a reflective relationship between latent variable and indicators. We previously established that reflective models of user attitudes and perceptions towards technology require that the indicators of the latent constructs are fully interchangeable. It makes no sense to hypothesise that there is a single psychological state that can cause responses on two or more indicators when it can be demonstrated that these indicators can vary independently and have independent causality. We showed that “factors” which are initially derived from factor-item loading patterns in exploratory factor analysis are particularly vulnerable to this sort of confusion. We suggest that both ServQual and eQual, which were based on factor-analytic approaches, reified “factors” (for example, interaction quality) derived from patterns in their data, and mis-specified them as reflective latent variables.

Ontologically, formative models are incompatible with realist assumptions. A realist approach to formative latent variable models hypothesises that specific responses on a series of indicators cause or predict an internal psychological state, which makes no sense at all.

MIMIC models share the same ontological issues as formative models, except that the option of treating the hypothesised latent simply as an index no longer exists, because the latent needs to cause a response on its reflective indicators. The confused ontology of MIMIC models, which requires responses on indicators to cause a psychological state, which in turn causes a response on other indicators, is a major argument against their use in this context, regardless of the fact that they solve statistical issues of identification of formative models.

We suggest that the approach most consistent with a realist ontology is offered by multi-indicator structural models, which can be used to model a range of unobserved (but assumed to exist) internal states (latent variables) associated with every construct with independent meaning or causation. These latents cause responses on indicators. Ideally, there is a very direct relationship between the hypothesized latent and the

indicator(s) used to measure it. For example we hypothesise a latent variable Perceived Information Timeliness, which is measured by the indicator “The site provides timely information” (x10).

Previous researchers in this area have frequently not stated their ontological assumptions explicitly, nor matched them to the most appropriate modelling method. This has resulted in a considerable amount of confusion and widespread reification of factors that have been constructed from data using exploratory factor analysis as “dimensions” of online service quality.

9.4.2 These assumptions applied to POLSQ

What ontological positions are defensible? An anti-realist instrumentalist position is always defensible because it does not make any sweeping truth claims. According to this position, there does not need to be any such thing as POLSQ, or indeed, any of the other latent variables in any of the models. The danger is that the development and operationalisation of measures is often couched in realist terms.

From a realist perspective, any concept or internal state that can be argued to be conceptually distinct from others can be argued to “exist”. In modelling and measurement terms, this assumption requires the researcher to hypothesise a separately existing latent variable for every distinct concept in their model, measured with at least one or more reflective variables. A change in the latent variable is hypothesised to precede a change on the item.

This means that it is possible to adopt a realist ontology for POLSQ, but not if it is conceptualised as a “composite” of multiple “dimensions”. To the extent that POLSQ can be claimed to exist, it is best conceptualised as a generalised attitude (affect) towards the service quality of a website at a point in time. It is also possible to adopt a realist ontology for a set of antecedents of POLSQ, which are the internal psychological states formed by the user, which we have modelled as descriptive and inferential beliefs about aspects of the service. Each conceptually distinct indicator is posited to have a separate existence.

9.5 Theoretical, paradigmatic and measurement approaches

9.5.1 What is the most appropriate theoretical approach?

We noted in our narrative literature review that there was a family tree of research in service quality that typically formed the provenance of online service quality research. We offered some alternative conceptualisations based on information systems theories of technology adoption and continuance, task technology fit, and affordance theory from human computer interaction. We discuss the insights offered by this research for the various streams of literature that are salient to the study of POLSQ.

9.5.1.1 *Service Quality, Service Encounters and Satisfaction*

In our narrative literature review we identified a certain amount of confusion around the distinction between service quality and satisfaction and service encounters and overall service quality. Our research had not advanced the satisfaction versus service quality dialogue; we agree that they are both evaluative responses with unclear and overlapping definitions. However, we are able to make some comments on the debate between service encounters and overall service quality. We noted that longitudinal research on face-to-face services by Boulding et al. found that customers' perceptions of each service encounter were affected by both idealised ("should") expectations and previous experience ("will" expectations), which jointly contributed to updating the overall affect towards the service after each encounter (Boulding et al., 1993). Further, "should" and "will" expectations tended to converge over time to reduce cognitive dissonance. We concur that POLSQ is an overall affect towards an online service that is updated by individual encounters, which are in turn influenced by experiential residue from previous encounters. This is consistent with the generalised model of attitude formation from (Fishbein & Ajzen, 1975). We also pointed out that in a technology-mediated context; individual encounters are likely to be consistent with each other anyway. This, in common with the desire to reduce cognitive dissonance, suggests that the gap between perceptions of individual encounters and overall affect towards a service will tend to be quite small. The same process will reduce the gap between expectations (prior to a service encounter) and perceptions after the event.

9.5.1.2 Face-to-face Service Quality Literature

We noted that research in POLSQ tended to look back to consumer behaviour research, particularly the ServQual stream of literature for its provenance. This occurred despite acknowledged difficulties in establishing equivalences between face-to-face and online services, and the lack of consensus or replicability in the findings across multiple studies. This appears to arise from the desire to establish a cumulative research tradition. E-commerce research that looked back to ServQual for its provenance typically conceptualised online service quality in a similar way, as a multi-dimensional construct measured with several reflective latent variables. We found this approach to be theoretically and empirically unsound.

In our empirical studies conducted using EFA and CFA and using the eQual instrument as a starting point, we found no support for this conceptualisation of online service quality. The EFA produced results that were not consistent with previous eQual studies, and CFA models did not fit the data. The CFA models were probably mis-specified, with confusion between formative and reflective indicators, which was carried over from previous studies.

From a theoretical point of view, we found the service quality research “family tree” and ServQual in particular, was unsound. ServQual was originally developed as the dependant variable for the gaps model, specifically to measure gap five. The “self-service” nature of online services means this model is inappropriate. The organisation does not “deliver” a series of unique service encounters, it provides a website with a series of predefined, and (in principle) predictable and consistent behaviours and affordances, from which the users serve themselves. This means the user has a much stronger role in constructing their own service delivery experience, and the gaps model, which is the original theoretical foundation of ServQual, does not accurately describe the delivery process of online self-services, and it does not provide a solid theoretical foundation for this research. Subsequent to the present study, the author has presented a detailed refutation of ServQual “descendents” in online service quality research (Tate & Evermann, 2009a), which contains a further elaboration of this argument.

Although online service quality research frequently looks back to ServQual, we noted the “Nordic” model of service quality in our literature review as an alternative to ServQual. This model conceptualises service quality as comprised of process and outcome quality. We argued that perceived usefulness represents “outcome quality” and perceived ease of use represents “process quality” in an online context. This means that our formative models and MISMs (chapters 7 and 8) are closer to the Nordic model of the structure of POLSQ than to the ServQual tradition. However, both traditions have their origins in face-to-face service encounters, and should be used with caution when informing an online study.

9.5.1.3 The “family tree” of online service quality literature: service quality in IS departments, end-user computing, website quality and online service quality

As we discussed in our literature review, there was a reasonable consensus at the item level about the determinants of user perceptions of POLSQ. Our research suggests that the items identified from the extensive family tree are all salient, and collectively predict a large proportion of the variance in overall affect towards the service quality of a website. Considerations of technology features such as presentation format and response time were introduced with studies of end-user computing in the 1990’s, before the widespread availability of web-based services, and have remained on the “menu” of POLSQ measures ever since. They have been joined by additional technology characteristics such as personalisation and transaction quality as these affordances became available. The failure of previous research appears to have been in understanding and modelling the relationships between the determinants of POLSQ, and in grouping them into generalised “factors” rather than analysing causality and antecedents.

9.5.1.4 Information Systems theory

Our study provides considerable support for the explanatory power of information systems theory when studying POLSQ. This acknowledges that consumers of online systems are users of information systems. In particular, our formative and MISMs (Chapters 7 and 8) confirm previous research on the importance of perceptions of Usefulness, Ease of Use (Gefen et al., 2003; Tate et al., 2008a) and Trust (Gefen et al., 2003) in forming attitudes towards websites.

We were unable to fit “playfulness”, which is another TAM extension that had been found to be significant in an online context (Ahn et al., 2007; Lin et al., 2005) into our models. This is likely due to the fact that online services are less hedonic than online shopping or entertainment.

We note that information systems continuance is a special case of information systems adoption. Almost by definition, anyone with any significant experience of an online service is a continuing user. Our study strongly supports the conclusions of IS continuance studies (Bhattacharjee, 2001), that perceptions of useful outcomes are critical determinants of continuing use of information systems (which include online services). Our study also supports Bhattacharjee’s conclusions that ease of use was less important than usefulness for continuing users. Our formative model found that perceived ease of use had an insignificant effect on overall POLSQ.

We are not in a position to carry out a comprehensive comparison between our research and the Task Technology Fit Model (TTF) (Goodhue & Thompson, 1995), since many of the constructs included in TTF are significantly different to those that we used. However, we note that constructs such as information quality, ease of use, and learnability are included by Goodhue and Thompson as indicators of fit, as well as likely being antecedents of POLSQ.

The likely explanation for this commonality lies in affordance theory. We suggest that Goodhue and Thompson have “black-boxed” the process of negotiation and adaptation between users and software, and so to some extent have we. At the left-hand side of our generalised model we show that both the user’s mental models and the objective characteristics of the technology contribute to the formation of beliefs. In reality, there is likely to be a complex interplay between aspects of the interface design, the actual functional capabilities of the system, the user’s task requirements, the user’s previous experience with the software, and the user’s perceptions of the perceptual, cognitive, and functional affordances of the system, that all contribute to the formation of attitudes and beliefs about the technology.

We were able to include latent variables for perceived design and presentation fit in our multi-indicator structural models. There is further potential for research in this

area. However, overall in this study we have de-emphasised these aspects and concentrated on the process that occurs once the user has formed perceptions about the functional affordances of the software.

From affordance theory, we found that an understanding of the functional affordances of self-service technologies provided useful insights into the likely scope and domain of the phenomenon, which could be triangulated with research literature and qualitative data. New affordances such as transaction quality (as we identified in Chapter 2) need to be incorporated into the scope of the phenomenon, and will improve the explanatory power.

From social psychology, we found that generic theories of attitude and belief formation provided many valuable insights into the formation of attitudes and beliefs towards technology. Although Fishbein and Ajzen's theories are widely cited by information systems researchers, we found that their 1975 book informed the development of our general model, and is worth revisiting in the original (Fishbein & Ajzen, 1975).

9.5.2 What is the most appropriate paradigmatic approach?

Considered in the context of wider considerations from philosophy of science and social science theory, we found that research into online service quality often contains implicit "normal science" assumptions that we are extending our existing understanding of service quality. Anecdotally, researchers have remarked "you can't get anything published in the service quality area unless you mention ServQual". We suggest we have in fact made relatively little progress in our understanding of this area despite a very large number of published papers. Reasons for this include a very high degree of semantic ambiguity in the constructs, which makes meta-analysis very subjective and structural relationships are hypothesised (and supported) by models that are contradictory. Additionally, there is a lack of theory testing and an emphasis on data-driven techniques such as EFA and PLS.

Further, to the extent that theory does exist, it is frequently descriptive (Type 1 theory), not predictive. In previous research, we have suggested that service quality and online service quality research may be more accurately represented as a post-

modern discourse than as a normal science research tradition, and that considerations such as strong and ongoing advocacy by the ServQual authors have been equally as influential on the current situation as the strength of the results (Sylvester & Tate, 2008; Sylvester et al., 2007).

On balance, we advocate a post-positivist, critical realist paradigm for researchers of online service quality. We assert that there is a reality of IT artefacts and user perceptions thereof, and more specifically, online services and user evaluations thereof; that is independent of our thinking about it. However, following (Longino, 1990) we believe that our explanations of these phenomena are fallible “best efforts” constructions that are negotiated between researchers, rather than facts accumulating in a cumulative research tradition.

9.5.3 Measurement approaches for POLSQ.

In this section, we assume, adopting the critical realist approach we advocated in the previous section, that POLSQ is best conceptualised as a theory net of increasingly generalised perceptions and attitudes towards the service obtained via a website. This means that measurement approaches to POLSQ need to be done in the context of general theories of user attitudes and perceptions towards technology.

9.5.3.1 *A Comparison of Modelling Approaches*

The following table (Table 9.1) compares the various modelling methods used in this study according to a range of criteria, including the nature of their truth claims, their potential for parsimony, generalisability and extensibility, a summary of issues, and some recommendations for use. Overall, despite the popularity of this approach, we do not recommend approaches based on exploratory factor analysis, followed by a confirmatory factor model and a structural model for this business domain. In our view, this has led historically to mis-specified models, a tendency to reify factors, and ontological assumptions and truth claims that cannot be supported.

We believe that researchers are better to hypothesise a larger number of latent constructs, and structural relationships (as in multi-indicator structural models), as these have a much more direct and accurate correspondence with the real-world situation they purport to represent. These models also allow us to represent most

closely the implementation of our generalised model of the formation of beliefs and attitudes towards technology, in our specific context.

Since the number of potential variants of multi-indicator structural models is very large, formative models may be useful as short-hand, to provide some initial, high-level understanding of the amount of variance of a specific latent that is predicted by a set of indicators, so long as the limitations of formative models are borne in mind.

Formative models may also be useful for developing indices and sum-scores for comparison and bench-marking purposes.

Table 9.1: *A comparison of modelling methods*

	Reflective Factor analytic models	Formative model	MIMIC model	Multi-indicator structural models	Comments
Truth claims	A latent variable with separate ontological existence causes the effect on the indicators. Requires assumptions about the number of latents.	Responses on observed variables (indicators) cause and can be used to predict the variance on another latent construct OR can be used to form the sum-score (index) of another construct	Somewhat confused. There is a latent variable which is caused by its indicators (or is the sum of its indicators), but also has separate existence and is the cause of reflective indicators.	Explicit and direct. Close correspondence between indicators and the hypothesised latents that they represent. Structural relationships explicitly specified.	Reflective FA model is clearly untrue in a POLSQ context. Hypothesised latent cannot be the common cause, because they can vary independently. Reflective FA model has causality from more generalised construct to less generalised, and from attitude to perception, which cannot be correct. Formative models and MIMIC models are likely ontologically muddled in this business domain.
Parsimony	Can be used for data reduction. Parsimony is illusory unless the latents are meaningful.	Useful for data reduction	Can be mathematically equivalent to a non-factor model and more parsimonious. However, it is less extensible.	Sacrifices parsimony for precision and flexibility	Different modelling approaches reflect a different balance between accuracy and parsimony, with multi-indicator structural models the least parsimonious and the most accurate.
Generalisability	Only if the latents are meaningful and the item loadings are consistent across studies	Yes, if the formative variable is treated as a sumscore, then it can be generalised to all instances of its indicators.	Only to the extent that the hypothesised formative indicators fully describe and predict the same latent across multiple contexts.	No. Individual causal nets are unique, but have maximum precision, explanatory power and predictive power for the theory net that they implement	It has been suggested that excessive claims for generalisability are often made in information systems research (Seddon & Lyytinen, 2008).

Extensability of constructs and the theory net	Constructs cannot be extended. The theory net can potentially be extended by adding new constructs so long as they are clearly distinct from existing ones.	No. By definition the indicators of a formative latent are supposed to fully describe it.	Similar to the issues for formative models.	Constructs cannot be extended; however, the theory net can easily be extended.	
Recommendations for use	When there is a good theoretical reason to believe in the existence of the latent, and the indicators are truly reflective (measure the same things)	To clarify the scope of a domain and ensure adequate coverage before hypothesising structural relationships. May be useful for benchmarking and comparisons over time.	Possibly to simplify a non-factor model and aid with model identification.	Preferred for causal modelling. Suggest for IS that models build from detailed perceptual beliefs to overall affect (attitude)	Different modelling approaches may be appropriate at various points during theory development. As we have demonstrated, using different approaches for the same dataset may yield valuable insights. Formative models may be useful to produce sum scores for bench-marking purposes, and can be used as a “short-hand” for mathematically equivalent multi-indicator structural models.
Issues	Reifying factors	Does not explain the causes of co-variance between indicators.	Muddled ontological assumptions.	Very challenging to achieve model identification. Some constraints may not be theoretically justified.	
Explains	There are items that co-vary	There are a collection of items, that may co-vary (we don’t know why), that collectively have some predictive power for another item. OR there are a collection of items that have some real-world salience that can be expressed as a sum-score.	That formative items predict responses on a latent, as it is defined in terms of its reflective indicators.	A specific theory with very high precision	Explanatory power of reflective models is often over-stated if the indicators are not truly reflective. There is a danger of reifying factors based on patterns in the data. MISMs have the highest explanatory power.

9.5.3.2 Issues with operationalising measures

To return briefly to our starting point in eQual, we note that the eQual authors observed that “for each application of the tool in a given context there may be both general qualities...and those that are context specific” (Barnes & Vidgen, 2001, p. 10). Well, we agree in principle! However, the implementation of this principle using exploratory factor analysis, and extensions to existing instruments seems unsound. Our approach for modelling the process by which specific contexts of use lead to generalised attitudes and affects is better grounded in social psychology and psychometric theory. We suggest that context-specific qualities are much more likely to be antecedents to overall POLSQ – descriptive beliefs that are salient to the context of each study – than “extensions” at the same level of generalisation.

We note also that there appear to be unresolved issues with determining the reliability of measures of beliefs and attitudes towards technology. The widely followed practice of using multiple items (three to five reflective items per latent variable) to increase the reliability of measures appears to us to be problematic.

It is extremely difficult to make multiple measures truly interchangeable. As soon as individual items that purport to measure the same construct become conceptually distinct (i.e. it is possible to make an argument that they could vary independently or have different antecedents) then arguably we have different underlying latent variables. The likelihood of this occurring increases with the number of items.

We note that two of the major sources referred to in this research have argued for the use of fewer indicators. Fishbein and Ajzen noted that since people rarely doubt the validity of their own senses, descriptive beliefs are usually held with maximum certainty, and that attitude is frequently measured by a single indicator of affect for or against an object (Fishbein & Ajzen, 1975). Hayduk makes a compelling case for single-item “gold standard” indicators and a rethink of measures of validity and reliability (Hayduk, 1996; Hayduk & Glaser, 2000a, 2000b).

9.5.4 What insights can be gained from psychometrics?

This study required an extensive reading program in contemporary psychometric research. This is a very large and complex area of study, and we can only claim to have scratched the surface. However, based on our limited experience, we would like to share some observations with others within the information systems community, since this literature does not seem to be very widely cited in our field except by a few SEM specialists.

The use of exploratory factor analysis, followed by two-step process that involves first assessing a “measurement model” using confirmatory factor analysis, then assessing a “structural” model using co-variance-based SEMs is the received wisdom of many popular SEM guides and textbooks. Within information systems, when SEMs are used at all, they tend to be specified reflectively, and we have only recently begun to see influential papers published, for example, the MISQ paper by Petter et al. examining formative measures (Petter et al., 2007).

However, within the SEM community, there are heated debates which have largely bypassed information systems researchers: for example, the appropriate use of formative measures, and the effectiveness of the “two-step” approach (measurement model followed by structural model) versus the multi-indicator structural model approach) advocated in particular by Leslie Hayduk (Hayduk, 1996; Hayduk & Glaser, 2000a, 2000b). Many of these debates are ongoing and inconclusive, so it is not reasonable to expect non-specialist information systems researchers to have the last word on these issues. We do believe, however, that it behoves information systems researchers to regularly refresh their knowledge of modelling and measurement approaches by referring back to reference disciplines such as psychometrics. The wider information systems research community needs to understand the implications of recent advances in psychometrics for our research, or run the risk of conducting poor-quality research using discredited techniques.

Our two important learning points are, first that measurement theory has continued to advance steadily since the initial publication of widely-cited and seminal studies such as ServQual, yet these studies are not critically reassessed, and, second, that

exploratory factor analysis is widely mis-used in information systems research. Our reassessment of the “family tree” of online service quality research is presented in Chapter 3. Issues with the inappropriate use of EFA, which is used without a sound understanding of the underlying measurement theory, continue to be legion in information systems research.

The discussion of factor analysis dates back at least seventy years, and has serious implications for latent variable modelling approaches (Schonemann, 1996a). Since there can potentially be a large number of different factor structures that fit the same data, researchers need to be very careful about the truth claims made on the basis of factor analysis. Two major explanations are widely used – the “alternative solution” position (ASP) which considers that the latent common factor is determined by constraints inherent in the model (in other words, it is a mathematical construction) and the “posterior moment” theory, which considers the latent factor to be a single random and separately existing entity (Conway & Huffcut, 2003). The issue is that factor models frequently do double duty by acting as a criterion (test) for the existence of the factor, AND measuring its effects. The meaning and measurement of constructs tend to be conflated (Bickhard, 2001). This problem is widespread and deserving of serious discussion.

At a more detailed level, a better understanding of the nature of latent variables, the differences between formative and reflective indicators, the ontological implications of various modelling methods, and the possibilities of multi-indicator structural models all have potential to improve the quality of information systems research into attitudes and perceptions towards technology.

9.6 Implications for IS Research

9.6.1 Use of Formative Latent Variables in Theories of Attitudes and Behaviour towards Technology.

Recently, there has been a great deal of attention to formative models in leading information systems journals (Cenfetelli & Bassellier, 2009; Petter et al., 2007). These high profile papers will hopefully reduce the incidence of latent variables being specified reflectively when the indicators are clearly not reflective. Our contribution to this discourse relates to the ontology of formative constructs when used in psychometric research.

In psychometric research, there is a distinction between observed measures (such as questions on surveys) and the internal unobserved psychological states (latent constructs) they supposedly represent. This poses multiple problems for the ontology of formative models. First, from an ontological perspective, the formative indicators themselves cannot exist in isolation – some internal state must be antecedent to each subject's survey responses. Therefore, there is an implied unobserved latent variable behind every formative item. Effectively, each formative indicator could be seen as shorthand for a latent variable with a single reflective indicator.

Second, there are issues with the notion of formative latent variables. There are several positions which can be taken. An anti-realist or instrumentalist position can be adopted, which does not assume any separate existence for the formative latent, but treats it as a point variable or index formed from its basket of antecedents. This index may have predictive power for other variables in a model. We note that it makes sense to talk about “covering the scope” of a construct, as a formative index which covers the scope of a business domain thoroughly might be expected to have higher predictive power for other dependent variables.

The second position that can be adopted for formative models is that expressed in MIMIC models. In these, reflective indicators are added to the formative latent variable. As they stand, MIMIC models are cannot be resolved ontologically. They require a single construct to simultaneously act as a mathematical index with no

independent ontology, and an internal psychological state. As we discussed, in Chapter 7 (Figure 7.1 and 7.2), a model with a formative point variable which has only one path out to a reflectively specified latent variable is mathematically equivalent to a MIMIC model, but makes the ontological assumptions more explicit. In this formulation, a basket of indicators form an index which predicts the subject's response on a more generalised psychological latent variable. A MIMIC model can be used as a form of short-hand for a more ontologically consistent specification

Considering formative constructs as an index or point variable rather than as an internal psychological latent construct also helps to resolve the debate over which formative items to retain under what circumstances, and whether changing the formative items changes the “meaning” of the variable. In our interpretation, the “meaning” does change, but since we are not assuming an independent ontology anyway, it does not matter. The only thing that matters is whether retaining or deleting a formative item improves the predictive power of the point variable.

9.6.2 The Scope and Nature of Information Systems Theories of Attitudes and Perceptions Towards technology

As we have previously discussed, the diversity of social phenomena creates issues with theory generalisability and parsimony (simplicity). The quest for parsimoniousness is often the cause of a gap between models and the complexities of the real world.

“[Although social scientists] strive for theories that are simultaneously parsimonious, highly general, and therefore applicable to a wide range of phenomena, yet precise enough to imply rejectable hypotheses, it does not appear possible...to achieve simultaneously all three of these ideal characteristics...my own position is that of the three, parsimony is the most expendable.” (Blalock, 1982, p. 28)

We suggest that with regard to the measurement of attitudes and perceptions towards technology, there are a number of potential issues for information systems research, which include normal science assumptions and a pre-occupation with generalisability, parsimony and originality. To this extent our research supports the assertions of Seddon and Lyytinen in their recent panel debate (Seddon & Lyytinen, 2008).

A pre-occupation with parsimony and generalisability tends to give rise to core IS constructs, such as “usefulness” and “ease of use” which are parsimonious and highly generalisable but not very useful; nor are they a particularly good representation of the real world, as these constructs can apply to all kinds of technologies from electric hedge-clippers to ERP systems. Ideally, IS theory should explain the processes by which the interaction between users and specific technology features leads to the formation of attitudes, rather than simply describing and measuring those generalised attitudes at a considerable remove from the interactions that give rise to them. A pre-occupation with originality tends to lead to excessive assumptions of generalisability for models that may only have been tested in a very limited range of contexts (Seddon & Lyytinen, 2008). It is often perceived that it is difficult to get replicatory studies published, but these would provide a much greater degree of confidence in the truth claims of our theories.

9.6.3 Approaches for information systems researchers

A possible solution to these issues is to carry out more specific and precise research. Our theoretical approach, implemented with multi-indicator structural models, provides a framework and methodology for this sort of research (Evermann & Tate, 2009b; Tate & Evermann, 2009b). The affordances of a specific technology can be used to identify candidate descriptive beliefs and the contribution of these to generalised attitude and affect can be measured.

From a practical point of view, we believe it would be valuable to report more “failed” models, and to routinely make the covariance matrices for SEM models available, so that future researchers could do further analysis on the data. Perhaps research datasets could be shared more freely (assuming ethical considerations permit this).

We also suggest that more alternative models are considered and compared. This has become more popular with the increasing popularity of co-variance-based SEMs.

9.6.4 The provenance and further development of theories of attitudes and perceptions towards technology

We discussed earlier a range of theories that had salience for this research. As a result of this study, we were able to propose some generalised heuristics for developing theories of user attitudes and perceptions towards technologies. These heuristics are summarised below. A more complete discussion is provided in our paper (Tate & Evermann, 2009b).

In our research, we combined theories of cognitive affordances and cognitive fit of the IT artefact (from human computer interaction), to explain the *antecedents* of inferential beliefs such as perceived ease of use and perceived usefulness *in a specific context*. This provides accuracy and real-world salience for IS theories.

Figure 9-2: *Affordance theory allows us to identify the salient features of technologies for inclusion in theory*

Social psychology informs our understanding of how affordances form descriptive beliefs, which lead to inferential beliefs and ultimately to attitudes towards technologies, as shown in Figure 9.2. In combination, we can propose some general heuristics for theory development. Perceived functional affordances of a technology lead to descriptive beliefs about a technology. These in turn lead to inferential beliefs, which lead to generalised attitudes and affects.

Multi-indicator structural models offer a practical way of modelling and testing these theories (se Figure 9.3). The causal progression from perceived affordances to attitudes is shown in solid lines; while the possible approaches to modelling this process using MISM’s is shown with dotted lines.

Figure 9-3: *Theory development and modelling*

9.7 Implications for Practice

The measurement of online service quality is extremely, perhaps increasingly salient for practice as online services become more ubiquitous. ServQual variants, and to a lesser extent, versions of eQual have both been applied in a wide range of business domains. Business users would have every justification for believing that these were rigorous, reliable, and useful and proven measures.

Our empirical research challenges these assumptions, and suggests that the instruments are probably mis-specified, while our literature review suggests that four decades of research has not created a consensus.

We suggest that practitioners should not uncritically adopt online service quality measurement instruments, but should engage with the research community to ensure that they develop measurements appropriate to their needs and business context.

9.8 Limitations

We acknowledge a number of limitations to this study, which fall into three major areas, the scope of the study, the research design, and the data analysis.

9.8.1 Scope

There are several significant restrictions to the scope. The contribution is mainly theoretical, as the empirical analysis was based on a single sample. The theoretical contribution based on psychometrics concentrates mainly on the ontological implications of different modelling methods.

Conducting our analysis on a single dataset ensures a direct comparison between modelling methods, which is essential to our study. However, it raises the possibility that all our results were affected by chance characteristics of the dataset. This means that the empirical analysis potentially has more value as an illustration of the various modelling methods than as a source of generalisable insights about POLSQ. To this extent, our contribution is more theoretical than empirical.

While we believe the theoretical argument is well supported, we do not feel we can claim it is definitively “proved” by our illustrations. We offer a number of alternative conceptualisations of POLSQ, illustrated by models and data analysis, but we are not able to test them all rigorously. All of our models use exactly the same (relatively limited) sample and data-set. The sample size is relatively small and drawn from a highly specific context.

With regard to the theoretical arguments we have presented, we acknowledge that psychometrics and measurement theory have extremely large bodies of knowledge. In this thesis we have concentrated mainly on the ontological issues - the issues associated with the truth claims that can be supported by different modelling methods.

9.8.2 Research Design

There were two limitations in our research design, the fact that we conducted a single cross-sectional study, and the fact that we were limited to our set of twenty-eight items.

In our data gathering we conducted a single cross-sectional study rather than a longitudinal study. Previous studies of perceptions of face-to-face service quality (Boulding et al., 1993) and technology acceptance (Venkatesh et al., 2003) have found that user's perceptions of technologies and of services change over time. Social psychology suggests that the formation of generalised attitudes typically includes unmeasured causes from "experiential residue" from the respondent's past life experience. We were unable to evaluate these effects in our study.

We were limited to the set of twenty-eight items we collected using our modified eQual instrument. A limitation of this approach is that the item set was not purposely developed for each type of model. Many of our indicators have been used as indicators for more than one hypothesised constructs in our model. These decisions have theoretical support, but we did not start with a clean slate to develop an optimised set of indicators. Previous studies have suggested that when developed from scratch, the item sets for formative and reflective models may potentially be quite different (Diamontopoulos & Siguaw, 2006). We were limited in our more generalised constructs (for example "Ease of Use") to those we had meaningful reflective indicators for. Even our "better fitting" models had some weaknesses. Our formative MIMIC model provided good explanatory power for the dependent variable, but the item loadings on the formative constructs were variable. Some of our MISMs were reasonably well-fitting but were only model fragments, and were subject to a number of constraints.

9.8.3 Data analysis

Some of the decisions we made during our data analysis have potential limitations. We would like to highlight issues with our handling of missing data, choice of factor extraction method, the validity and reliability of indicators for multi-indicator structural models, and the process of respecifying previously ill-fitting models.

Approximately 40% of the cases in our original sample had missing data. We used list-wise deletion for the missing data. This required us to discard 40% of the cases. Many statisticians recommend data imputation (using an algorithm to estimate missing values) as preferable to list-wise deletion to minimise the loss of statistical power (Howell, 2009). However, AMOS does not support data imputation for non-normal data, so we were unable to adopt this approach.

In our exploratory factor analysis we continued the practice of the eQual researchers by confusing principal components analysis (PCA) and factor analysis. We used PCA as our extraction technique when carrying out “factor analysis” in SPSS. We acknowledge that PCA is a data-driven technique that “constructs” components from the data set, and that other extraction techniques are frequently preferred for exploratory factor analysis. However, it has been suggested that the results of the two methods are often similar (Field, 2005). Also, we intended to compare our results with those previously obtained by the eQual authors, so we used the same extraction technique. We note that this confusion is widespread in online service quality research using factor analysis (Alzola & Robaina, 2005).

We were unable to provide a thorough discussion of issues of validity and reliability of the indicators in our MISMs, apart from arguing for the conceptual distinctiveness of the constructs (for example, they can be shown to vary independently) and the relatively small distance between the hypothesised latent and the indicator. We note that theorists of MISMs such as Leslie Hayduk have suggested that non-factor models require a different conception of validity and reliability.

“The implicit model underlying all the classical discussions of reliability and validity is the factor model. In fact, factor model jargon is so pervasive we have difficulty imagining any standard alternative” (Hayduk, 1996, p. 104).

At this stage we were not able to identify any widely agreed reliability and validity heuristics that we could apply to this type of model.

We acknowledge that we developed multiple models based (partly) on the outcomes of previous (poorly fitting) models, although all our models could be supported by

previous research literature. Fitting your model to your data is often called “capitalizing on chance” and widely considered to be bad practice.

While there may be situations where this is undoubtedly true, we would like to suggest an alternative view. So long as you have clearly specified the context of your research, clearly understand the characteristics of your sample, and develop a detailed, granular multi-indicator structural models with defensible latents in defensible relationships to one another; any model that achieves a good fit to your data has the potential to provide original insights about the underlying reality (Evermann & Tate, 2009a).

Of course this approach is context-specific, but so is a case study. We could consider generating alternative models and seeking explanations as being somewhat like carrying out a quantitative case study. After all, it is likely that different samples in different contexts will have slightly different concerns and priorities.

There are some cautions that require further investigation of this line of reasoning is to be pursued further. In a model adjusted post-hoc, the p-value (probability of a hypothesis) becomes a conditional hypothesis, that is, conditional on the knowledge that some other model did or did not fit. The conditional probability is not the same as the unconditional one. The researcher should correct the p-values (as for example is widely done with post-hoc tests for ANOVAs). At this stage, we are not aware of any heuristic that suggests how to adjust the p-values.

9.8.4 Alternatives

As we have noted in the introduction, and at other points throughout the discussion, this study diverged somewhat from its original intent. If we had planned a comparative study of modelling methods from the outset, we would have included some additional, well established items in our questionnaire to allow improved measurement of constructs such as ease of use, trust, fun, and usefulness, and overall perceived service quality as reflective variables. This would have improved the accuracy of measurement of those constructs, and also made it easier to compare our results those from with previous studies.

10 Conclusion

In this thesis we have presented a series of linked reflections on POLSQ: structure, antecedents, ontology, theory and measurement. These reflections include a mixture of theoretical argument and empirical evidence. As we mentioned in the introduction, this was the result of an iterative process of reading and data analysis. In addition, some lines of argument that originated as part of this process have been developed further for publication elsewhere. In this section, we present a summary of our conclusions organised by various themes associated with our research questions, and a summary of the evidence that supports them. A summary of these is presented in Tables 10.1 to 10.6.

I would like to note that although I have used the pronoun “we” throughout the thesis, to acknowledge the contribution of my supervision team, the work I have presented here represents my own contribution to this stream of research. However, some of the themes which were first raised as part of the research for this dissertation have been developed further in other contexts, in collaboration with co-authors. Much of this work has been cited at points throughout this report. In many cases, this further work provides additional legitimacy for our conclusions, so I have included co-authored papers in the supporting evidence.

Returning to our starting point, both our initial lines of enquiry were justified, it seems that changing affordances in the underlying technology, such as the increasing range of online self-service transactions do indeed affect user’s perceptions of POLSQ. It also seems that our existing theoretical and measurement approaches are flawed, with mis-use of exploratory factor analysis, reification of factors derived from data, and little or no understanding of the differences between formative and reflective indicators. Further, promising streams of research from psychometrics, such as the use of multi-indicator structural models have not been adopted by researchers in this area.

Overall, the eQual research that we used as our starting point was in the right track. The eQual indicators appear to provide a reasonably comprehensive set of antecedents

to POLSQ, and the eQual authors correctly recognised that there is an important difference between generalised and context-specific constructs. However, the choice of EFA, and principal components analysis, was not theoretically sound.

If we adopt a realist ontology for POLSQ, that is, if we assume that we are researching real, but unobserved psychological states, then we suggest that rather than being a “multi-dimensional construct”, POLSQ is an overall attitude to the service quality of a website. This attitude likely has a complicated net of antecedents. These antecedents are also likely to be conceptually distinct psychological states. Various causal models could be argued. We offer some examples and suggest that there is theoretical and empirical evidence for a progression from the perceived functional affordances of the technology, through descriptive beliefs about the technology, to inferential beliefs about the technology, and ultimately towards an overall attitude towards the service quality.

Our strongest conclusions from this research are negative and provide a refutation of previous approaches, especially the ServQual “tradition” and the widespread use of EFA. We should not keep doing “more of the same” (that is, narrative literature reviews and factor-based models) in this research area. It is not getting us anywhere, and it is based on normal science assumptions that are largely illusory; we are not building a cumulative body of knowledge. Nor are we concentrating on the “core business” of information systems, the interface between people and technology, or generating accurate, salient theory in this area.

We offer some suggestions and illustrations of possible alternative approaches, which have a strong theoretical basis and some initially promising empirical results. Our hope is that this will stimulate debate and lead future researchers to consider fresh approaches to this important research area. It seems that online services, using a range of technologies, will become increasingly ubiquitous in our lives. Improving their quality will surely benefit the organisations that offer them and the people who use them.

Table 10.1: *Theme 1: Online service quality research*

Conclusion	Weight	Supporting evidence
Extensive previous research has not resulted in a consensus on the dimensions and structure of POLSQ. New studies continue to propose new dimensions and new re-combinations of existing indicators	Strong	<u>Chapter 2</u> : Evidence of multiple studies with little consensus
User perceptions of POLSQ change with the affordances of the underlying Internet technologies. For example, as technologies have evolved from information and communication to transactional capability, user perceptions of online service quality have evolved to include these affordances	Strong	<p><u>Chapter 2</u>: Previous studies suggest transaction range and usefulness emerging as a quality factor.</p> <p><u>Chapter 5</u>: Our EFA suggests that new indicators associated with transaction quality should be retained.</p> <p><u>Chapter 7</u>: Formative model 1 suggests that new indicators associated with transaction quality contribute to the variance on POLSQ.</p>

Table 10.2: *Theme 2: The structure and antecedents of POLSQ*

Conclusion	Weight	Supporting evidence
POLSQ is best conceptualised as an overall affect (attitude) towards the website rather than as a multi-dimensional construct.	Tentative	<u>Chapter 3</u> : Discussion of the evolution of attitudes from descriptive and inferential beliefs. Chapter 8: Multi-indicator structural model fragments that illustrate this approach.
The posited eQual structure of composite “dimensions” such as content quality measured reflectively with the existing indicator set cannot be justified using a realist ontology: those dimensions do not “exist”.	Strong	<u>Chapter 3</u> : These dimensions are reified from exploratory factor analysis and mis-specified as latent constructs with reflective indicators
Twenty-six items of our revised instrument appear to form a reasonably comprehensive set of antecedents for POLSQ. Two items can be used to measure overall attitude towards the service quality of the website.	Strong	<u>Chapter 2</u> : A literature review did not significantly extend the item set <u>Chapter 7</u> : Formative model 1 suggests that all the indicators contribute to the variance of POLSQ and that collectively they explain more than 80% of the variance. Our papers reporting a qualitative study of stakeholder’s expectation of quality in a university web portal did not significantly extend the item set: Tate, M., Evermann, J., & Hope, B. (2008). Stakeholder expectations of service quality in a university web portal. In D. Oliver, C. Romm & F. Sudweeks (Eds.), <i>Self-Service and the Internet</i> : Springer. Tate, M., Evermann, J., Hope, B., & Barnes, S. (2007). <i>Perceived Service Quality in a University Web Portal: Revising the eQual instrument</i> . Paper presented at the Hawaii International Conference on System Sciences (HICSS), Waikoloa, Hawaii.
The relationships between these antecedents are not clear	Strong	<u>Chapter 6</u> : A confirmatory factor analysis based on factor structures from eQual, our literature review, and our EFA produced poor-fitting models <u>Chapter 7</u> : A formative model with trust, ease of use and usefulness provided a model with reasonably strong explanatory power. <u>Chapter 8</u> : A variety of multi-indicator structural model fragments had reasonable fit indices.
POLSQ is an overall attitude towards the service quality of a website that is built up from descriptive beliefs based on the affordances of the website, through inferential beliefs about the usefulness, trustworthiness, and ease of use of the website.	Tentative	There is strong theoretical support but the empirical evidence is fairly limited at this time. <u>Chapter 3</u> : We discuss the foundations of the formation of attitude and beliefs from social psychology. <u>Chapter 8</u> : We offer a number of multi-indicator structural models that illustrate this progression. This theme is developed further in our paper: Tate, M., & Evermann, J. (2009). <i>Perceptive Users with Attitudes: some Heuristics of Theorising</i> . Paper presented at the ICIS, San Francisco.

Table 10.3: *Theme 3: The ontology of POLSQ*

Conclusion	Weight	Supporting evidence
There is a realist discourse in research in POLSQ that is not always correctly matched by the underlying ontology of the models. Realist assumptions need to be implemented using reflective models with correctly specified reflective indicators (including single indicators as in MISMs).	Tentative	<u>Chapter 3</u> : We discuss the ontological assumptions of modelling methods, and then apply and evaluate those assumptions in subsequent chapters.
There is a low awareness in this research area of the ontological assumptions compatible with different modelling methods, and a tendency to reify factors constructed from data as if they had a separate ontological existence.	Strong	<u>Chapter 3</u> : We discuss the issues

Table 10.4: *Theme 4: Theoretical and paradigmatic approaches*

Conclusion	Weight	Supporting evidence
Implicit normal science assumptions and a desire for a cumulative tradition lead researchers to try to demonstrate the provenance of online service quality research in research on face-to-face service quality, particularly ServQual. This is often unsound.	Strong	<p><u>Chapter 2</u>: We discuss the “family tree” of online service quality research</p> <p><u>Chapter 3</u>: We discuss the limitations of ServQual and the Gaps model as a basis for ongoing research in this area.</p> <p>We explore this further in our paper:</p> <p>Tate, M., & Evermann, J. (2009). <i>Descendents of ServQual in Online Services Research: The End of the Line?</i> Paper presented at the Americas Conference in Information Systems (AMCIS), San Francisco.</p>
Online services are information systems as well as services and IS theory has strong explanatory power	Strong	<p><u>Chapter 2</u>: We present a review of information systems theory with salience to this research</p> <p><u>Chapters 7 and 8</u>: Information systems constructs such as ease of Use, Usefulness and Trust appear to predict variance in POLSQ in our formative and MISM models</p> <p>We present empirical evidence for this assertion in a qualitative study:</p> <p>Tate, M., Evermann, J., & Hope, B. (2008). <i>Old Theory and New Service Quality: An Exploratory Study of the Nature and Nomological Net of Online Service Quality and Continuing Use using Information Systems Theory</i>. Paper presented at the Australasian Conference in Information Systems (ACIS).</p> <p>We discuss this from a theoretical perspective in our paper:</p> <p>Tate, M., & Evermann, J. (2009). <i>Descendents of ServQual in Online Services Research: The End of the Line?</i> Paper presented at the Americas Conference in Information Systems (AMCIS), San Francisco.</p>
There are issues with narrative literature reviews and there is widespread semantic ambiguity surrounding key constructs	Strong	<p><u>Chapter 2</u>: We show how difficult it is to broker a consensus with a large, diverse and contradictory literature.</p> <p><u>Chapter 7</u>: We show how some of our items have been used as reflective measures for other constructs</p> <p>We discuss alternative ways of presenting literature reviews in our papers:</p> <p>Sylvester, A., & Tate, M. (2008). <i>Beyond the “Mythical Centre”: An Affirmative Post-Modern View of ServQual Research in Information Systems</i>. Paper presented at the European Conference in Information Systems (ECIS), Galway, Ireland.</p> <p>Sylvester, A., Tate, M., & Johnstone, D. (2007). <i>Re-presenting the literature review: A Rich Picture of Service Quality Research in Information Systems</i>. Paper presented at the Pacific Asia Conference in Information Systems (PACIS), Auckland, New Zealand.</p>
Use of formative models: formative constructs are not consistent with a realist ontology and therefore do not have a separate existence as internal psychological latent variables. They can, however, be used as “short-hand” for other types of models.	Strong	<p><u>Chapter 3</u>: We discuss the ontology of formative models</p> <p><u>Chapter 7</u>: We illustrate the use of formative constructs as point variables and in MIMIC models</p> <p><u>Chapter 9</u>: We provide some recommendations for the use of modelling methods.</p>
Extending theories of POLSQ to incorporate the affordances of technologies improves contextualisation, salience and accuracy for this theory, but this may be achieved at the expense of generalisability.	Tentative	<p><u>Chapter 8</u>: We illustrate this process</p> <p><u>Chapter 9</u>: We discuss the implications for theory building</p> <p>This theme is developed further in our paper:</p> <p>Tate, M., & Evermann, J. (2009). <i>Perceptive Users with Attitudes: some Heuristics of Theorising</i>. Paper presented at the ICIS, San Francisco.</p>

Table 10.5: Theme 5: The measurement of POLSQ

Conclusion	Weight	Supporting evidence
There is an over-reliance on factor analysis, particularly EFA in this research area.	Strong	<u>Chapter 3</u> : We discuss the theoretical issues <u>Chapter 4</u> : Our EFA does not produce strong results
There is widespread confusion in this research area about the differences between formative and reflective indicators	Strong	<u>Chapter 3</u> : We discuss the differences and illustrate our points with examples from the online service quality research “family tree”.
There is a low awareness of debates in contemporary psychometric research in this research area.	Strong	<u>Chapter 3</u> : We discuss the issues
Quantitative research and co-variance-based SEMs can potentially be used for theory building as well as theory testing	Tentative	<u>Chapter 8</u> : We offer a range of alternative models to iteratively test theory We explore this further in our paper: Evermann, J., & Tate, M. (2009). <i>Building Theory from Quantitative Studies: or how to fit SEM models</i> . Paper presented at the ICIS.
Multi-indicator structural models provide an opportunity to develop nuanced models with high accuracy and salience Issues of reliability of measures may need to be revisited for multi-indicator structural models.	Tentative	<u>Chapter 8</u> : We present and discuss a range of models. We include a number of latent variables with single-item measures.

Table 10.6: Theme 6: IS research into user attitudes and perceptions towards technology

Conclusion	Weight	Supporting evidence
Many of the issues identified with regard to POLSQ are common to IS research	Tentative	<u>Chapter 9</u> : We discuss the implications for theory building in information systems theories of user attitudes and perceptions towards technology. This theme is developed further in our papers: Tate, M., & Evermann, J. (2009). <i>Perceptive Users with Attitudes: some Heuristics of Theorising</i> . Paper presented at the ICIS, San Francisco. Evermann, J., & Tate, M. (2009). <i>Constructs in the Most: The Lost World of the IT Artefact</i> . Paper presented at the ICIS.

The title of this thesis is “Reflections on perceived online service quality: Structure, antecedents, ontology, theory and measurement”. We have presented a wide-ranging discussion leading to conclusions grouped around the major themes of the research. Is it possible to make any concluding comments that link these themes together? If there are two things I believe we have demonstrated beyond reasonable doubt, it is that we should never cease to re-evaluate existing “knowledge”, and we should never uncritically use modelling and measurement techniques in our research without a detailed deconstruction of the underlying assumptions.

References

- Abels, E. G., White, M. D., & Hahn, K. (1997). Identifying user-based criteria for Web pages. *Internet research*, 7(4), 252 - 262.
- Accenture_Report. (2003). *eGovernment leadership: Engaging the Customer*: Accenture Consulting.
- Ahn, T., Ryu, S., & Han, I. (2007). The impact of Web quality and playfulness on user acceptance of online retailing. *Information and Management*, 44(3), 263-275.
- Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- Akaike, H. (1987). Factor Analysis and AIC. *Psychometrika*, 52, 317-332.
- Aladwani, A. M., & Palvia, P. C. (2002). Developing and validating an instrument for measuring user perceived web quality. *Information & Management*, 39(6), 178-190.
- Allen, M., & Yen, W. (1979). *Introduction to Measurement Theory*. Monterey, California: Brooks/Cole Publishing Company.
- Alzola, L., & Robaina, V. (2005). SERVQUAL: Its Applicability in Electronic Commerce B2C. *The Quality Management Journal*, 12(4), 46-57.
- Asubonteng, P., MacCleary, K., J., & Swan, J. E. (1996). ServQual revisited: a critical review of service quality. *Journal of Services Marketing*, 10(6), 62-81.
- Bagozzi, R. (2007). On the Meaning of Formative Measurement and How it Differs from Reflective Measurement: Comment on Howell, Breivik and Wilcox (2007). *Psychological Methods*, 12(2), 229-237.
- Baierova, P., Tate, M., & Hope, B. (2003). *The Impact of Purpose for Web Use on User Preferences for Web Design Features*. Paper presented at the PACIS, 2003, Adelaide, Australia.
- Barki, H. (2008). Thar's Gold in them Thar Constructs. *The Database for Advances in Information Systems*, 39(3), 9-20.
- Barnes, S. (2001). An Evaluation of Cyber-Bookshops: The WebQual Method. *International Journal of Electronic Commerce*, 6(1), 6-25.
- Barnes, S., Liu, K., & Vidgen, R. (2001). *Evaluating WAP News Sites: the WebQual/M approach*. Paper presented at the European Conference on Information Systems (ECIS).
- Barnes, S., & Vidgen, R. (2001). *Assessing the quality of auction web sites*. Paper presented at the Hawaii Conference on System Sciences, Hawaii.
- Barnes, S., & Vidgen, R. (2002). An Intergrative Approach to the Assessment of E-Commerce Quality. *Journal of Electronic Consumer Research*, 3(3).
- Barnes, S., & Vidgen, R. (2004). Interactive e-government: evaluating the Web site of the UK inland revenue *Journal of Electronic Commerce in Organizations* 2(1), 42-63.
- Barnes, S., & Vidgen, R. (2007). Interactive e-government: evaluating the Web site of the UK Inland Revenue *International Journal of Electronic Government Research Jan-March*, 19-37.
- Barret, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, 42(5), 815-824.

- Barret, P. (2008). The Consequence of Sustaining a Pathology: Scientific Stagnation—a Commentary on the Target Article 'Is Psychometrics a Pathological Science?' by Joel Michell. *Measurement*, 6, 78-123.
- Baskerville, R. L., & Myers, M. D. (2002). Information Systems as a Reference Discipline. *MIS Quarterly*, 26(1), 1-14.
- Bell, H., & Tang, N. K. (1998). The effectiveness of commercial Internet Web sites: a user's perspective. *Internet Research*, 8(3), 219 - 228.
- Bentler, P., M. (1988). Causal modelling via structural equation systems. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology*, 2nd edition (pp. 317-355).
- Bentler, P., M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin*, 107(2), 238-246.
- Bentler, P., M. (2007). On tests and indices for evaluating structural models I. *Personality and Individual Differences*, 42(5), 825-829.
- Bentler, P., M., & Bonnett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-600.
- Berry, L. (2000). Cultivating Service Brand Equity. *Journal of the Academy of Marketing Science*, 28(1), 128-137
- Bhattacharjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly*, 25(3), 351-370.
- Bickhard, M. H. (2001). The Tragedy of Operationalism. *Theory & Psychology*, 11(1), 35-44.
- Bitner, M. J. (1990). Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses. *Journal of Marketing*, 54(2), 69-82.
- Bitner, M. J., Brown, S., & Meuter, M. L. (2000). Technology infusion in service encounters. *Academy of Marketing Science Journal*, 28(1), 138-149
- Bitner, M. J., Ostrom, A. L., & Meuter, M. L. (2002). Implementing successful self-service technologies. *Academy of Management Executive*, 16(4), 96-108.
- Blalock, H. (1982). *Conceptualisation and Measurement in the Social Sciences*. Beverley Hills: Sage.
- Blinkhorn, S. (1997). Past imperfect, future conditional: Fifty years of test theory. *British Journal of Mathematical and Statistical Psychology*, 50(2), 175-186.
- Bollen, K. (1984). Multiple indicators: internal consistency of no necessary relationship?' *Quality and Quantity*, 18, 377-385.
- Bollen, K. (1989). New incremental fit index for general structural equation models. *Sociological Methods and Research*, 17, 303-316.
- Bollen, K. (1990). Overall fit in covariance structure models: two types of sample size effects. *Psychological Bulletin*, 107(2), 403-421.
- Bollen, K., & Lennox, R. (1991). Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin*, 110(2), 305-314.
- Bollen, K., & Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 111-135). Newbury Park, CA: Sage.
- Borsboom, D. (2005). *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*. Cambridge, England: Cambridge University Press.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The Theoretical Status of Latent Variables. *Psychological Review*, 110(2), 203-219.
- Boudon, R. (1991). *Theories of social change*. Cambridge: Press.
- Boudreau, M.-C., Gefen, D., & Straub, D. (2001). Validation in Information Systems Research: A State-of-the-Art Assessment. *MIS Quarterly*, 25(1), 1-16.

- Boulding, W., Ajay, K., Staelin, R., & Zeithaml, V. (1993). A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions. *Journal of Marketing Research*, 30(1), 7-27.
- Brady, M. K., & Cronin, J. J. (2001). Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach. *Journal of Marketing*, 65(3), 34-49.
- Brannick, M. T., Chan, D., Conway, J., Lance, C. E., & Spector, P., E. (2010). What is Method variance and How Can we Cope with It? A Panel Discussion. *Organisational Research Methods*, 13(3), 407-420.
- Bressolles, G. (2008). Measurement of Electronic Service Quality: Improvements and Application. *International Journal of E-Business Research*, 4(3), 1-19.
- Bridgman, P. (1927). *The Logic of Modern Physics*. New York, New York: MacMillan.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 445-455). Newbury Park, CA: Sage.
- Byrne, B. M. (2001). *Structural Equation Modelling with AMOS*. Mahawh, New Jersey: Lawrence Erlbaum Associates.
- Carman, J. (1990). Consumer Perceptions of service quality: An assessment of the SERVQUAL dimensions. *Journal of Retailing*, 66(1), 33-55.
- Cenfetelli, R., & Bassellier, G. (2009). Interpretation of Formative Measurement in Information Systems Research. *MIS Quarterly*, 33(4), 689-708.
- Cheung, C., M.K., & Lee, M., K.O. (2005). *Consumer satisfaction with Internet Shopping: A research framework and propositions for Future Research*. Paper presented at the International Conference in Electronic Commerce (ICEC).
- Chin, W. W. (1998). Issues and Opinions on Structural Equation Modeling. *MIS Quarterly*, 22(1), vii-xvi
- Cohen, J. (1988). *Statistical power Analysis for the Behavioral Sciences (2nd edition)*. Mahwah, New Jersey,: Lawrence Erlbaum Associates.
- Conway, J., & Huffcut, A. (2003). A Review and Evaluation of Exploratory Factor Analysis Practices in Organisational Research. *Organisational research Methods*, 6(2), 147-168.
- Cote, J. A., & Buckley, R. (1987). Estimating Trait, method and error variance: generalizing across 70 construct validation studies *Journal of Marketing Research*, 24 315-318.
- Cronin, J. J., & Taylor, S. (1994). SERVPERF versus SERVQUAL: reconciling Performance-Based and Performance-Minus-Expectations Measurement of Service Quality. *Journal of Marketing*, 58(1), 125-131
- Curd, M., & Cover, J. A. (1998). *Philosophy Of Science: The Central Issues*. New York, New York: W.W. Norton & Company.
- Dabholkar, P. (1995). A Contingency Framework for Predicting Causality Between Customer Satisfaction and Service Quality. *Advances in Consumer Research*, 22(1), 101-108.
- Dabholkar, P. (1996). Consumer evaluations of new technology-based self-service options: an investigation of alternative models of service quality. *International Journal of research in Marketing*, 13(1), 29-51
- Dabholkar, P., & Bagozzi, R. (2002). An Attitudinal Model of Technology-Based Self-Service: Moderating Effects of Consumer traits and Situational Factors. *Journal of the Academy of Marketing Science*, 30(3), 184-201
- Darlington, R. (2010). *Factor Analysis*: University of Cornell.

- Davenport, T. H., Harris, J. K., & Kohli, A. (2001). How do they know their customers so well? *MIT Sloan Management Review*, 42(2), 63-73.
- Davis, F., Bagozzi, R., & Warshaw, P. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982-1003
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60-95.
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9-30.
- Department_of_Broadband_Communications_and_the_Digital_Economy. (2008). Online statistics, Document ID: 57471 [Electronic Version].
http://www.archive.dbcde.gov.au/2008/01/statistical_benchmarking/online_statistics from
http://www.archive.dbcde.gov.au/2008/01/statistical_benchmarking/online_statistics.
- DeVellis, R. F. (1991). *Scale Development: Theory and Applications*. Newbury Park, California: Sage Publications.
- Diamontopoulos, A. (1996). Viewpoint. Export performance measurement: Reflective versus formative indicators. *International Marketing Review*, 16(6), 444-457.
- Diamontopoulos, A. (2006). The error term in formative measurement models: interpretation and modelling implications. *Journal of Modelling in Management*, 1(1), 7-17.
- Diamontopoulos, A., Riefler, P., & Roth, K. (2008). Advancing formative measurement models. *Journal of Business Research*, 61(12), 1203 -1218.
- Diamontopoulos, A., & Siguaw, J. (2006). Formative Versus reflective Indicators in Organisational Measure Development. *British Journal of Management*, 17, 263-282.
- Diamontopoulos, A., & Winklhofer, H. (2001). Index Construction with Formative Indicators: an alternative to Scale Development. *Journal of Marketing Research*, 38(2), 269-277.
- Doll, W., Xia, W., & Torkzadeh, G. (1994). A Confirmatory Factor Analysis of the End-User Computing Satisfaction Instrument. *MISQ*, 18(4), 453-461
- Dubin. (1978). *Theory Building*. New York: Free Press.
- Durova, V., & Amin, N. (2009). *Using Webqual 4.0 in the evaluation of the Russian B2C Cosmetic websites*. Paper presented at the International Conference on Web Information Systems and Technologies, Lisbon, Portugal.
- e-marketer.com. (2006). US Retail E-commerce Sales: US Department of Commerce Statistics with calculations by e-marketer [Electronic Version]. *obtained from* <http://www.emarketer.com/> from <http://www.emarketer.com/>.
- Edwards, K., R., & Bagozzi, R., P. (2000). On the Nature and Direction of Relationships between Constructs and Measures. *Psychological Methods*, 5(2), 155-174.
- Eighmey, J. (1997). Profiling user responses to commercial websites. *Journal of Advertising Research*, 37(3), 59-66.
- Evermann, J., & Tate, M. (2009a). *Building Theory from Quantitative Studies: or how to fit SEM models*. Paper presented at the International Conference in Information Systems (ICIS).

- Evermann, J., & Tate, M. (2009b). *Constructs in the Mist: The Lost World of the IT Artifact*. Paper presented at the International Conference in Information Systems (ICIS).
- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modelling fit indexes. *Structural equation modelling*, 6(1), 56-83.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press. .
- Field, A. (2005). *Discovering Statistics Using SPSS* (2nd ed.). London: Sage.
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The Application of Exploratory Factor Analysis in Applied Psychology: A Critical Review. *Personnel Psychology*, 39(2), 291-314.
- Garson, G. D. (2008a). Regression: StatNotes from North Carolina State University [Electronic Version].
- Garson, G. D. (2008b). Structural Equation Modelling: StatNotes from North Carolina State University: North Carolina State University.
- Garver, W. W. (1991). *Technology Affordances*. Paper presented at the CHI Conference on Human Factors in Computing Systems.
- Gefen, D. (2003). TAM or just plain habit: A look at experienced online shoppers. *Journal of End User Computing. Hershey: Jul-Sep 2003.*, 15(3), 1-13.
- Gefen, D., Karahanna, E., & Straub, D. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, 27(1), 51-90
- Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the AIS* 4(7), 1-79.
- Gehrke, D., & Turban, E. (1999). *Determinants of Successful Web Site Design: Relative Importance and Recommendations for Effectiveness*. Paper presented at the HICSS, Hawaii.
- Gibson, J. J. (1977). The Theory of Affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology* (pp. 67-82). Hillsdale, NJ: Lawrence Erlbaum.
- Gilbert, D., Balestrini, P., & Littleboy, D. (2004). Barriers and Benefits in the Adoption of e-Government *The International Journal of Public Sector Management*, 17(4), 286-301.
- Goodhue, D. L., & Thompson, R., L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 19(2), 213-236.
- Gounaris, S., & Dimitiadis, S. (2003). Assessing Service Quality on the Web: Evidence from Business to Consumer Portals. *The Journal of Services Marketing* 17(5), 529-547.
- Grapentine, T. (1998). The History and Future of Service Quality Assessment. *Marketing Research*, 10(4), 4-21.
- Gregor, S. (2004). *Information Systems Theory, Constructing and Criticizing*. Retrieved 23 September, 2008, from http://epress.anu.edu.au/info_systems/mobile_devices/index.html.
- Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly*, 30(3), 611-642.
- Gronroos, C. (1988). Marketing Services: the case of a missing product. *Journal of Business and Industrial Marketing*, 13(3-4), 322-338.

- Gronroos, C. (2001). The perceived service quality concept - a mistake? *Managing Service Quality*, 11(3), 150-152.
- Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, 18(4), 37-44.
- Grover, V. (1997). *A Tutorial on Survey research: from Constructs to Theory*: Association of Information Systems.
- Guzzo, R. A., Jackson, S. E., & Katzell, R. A. (1987). Meta-Analysis analysis. *Research in Organisational Behavior*, 9, 407-422.
- Habing, B. (2003). *Exploratory Factor Analysis* (Working paper): University of South Carolina.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2003). *Multivariate Data Analysis*. Singapore: Pearson Education.
- Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour and Information Technology*, 22(5), 315-338.
- Harvey, R. J., Billings, R. S., & Nilan, K. J. (1985). Confirmatory Factor Analysis of the Job Diagnostic Survey: Good News and Bad News. *Journal of Applied Psychology*, 70(3), 461-469.
- Hayduk, L. (1996). *LISREL: Issues, Debates and Strategies*. Baltimore, Maryland: The John Hopkins Press Ltd.
- Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! one, two, three – Testing the theory in structural equation models! *Personality and Individual Differences*, 42(5), 841-850.
- Hayduk, L., & Glaser, D. (2000a). Doing the Four-Step, Right-2-3, Wrong-2-3: A Brief reply to Mulaik and Millsap; Bollen; Bentler; and Herting and Costner. *Structural Equation Modelling*, 7(1), 111-123.
- Hayduk, L., & Glaser, D. (2000b). Jiving the Four-Step, waltzing Around Factor Analysis, and Other Serious Fun. *Structural Equation Modelling*, 7(1), 1-35.
- Hinkin, T. R. (1998). A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organisational Research Methods*, 1(1), 101-121.
- Hoetler, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods and Research*, 11, 325-344.
- Howell, D. (2009). *Treatment of Missing Data*. Burlington: University of Vermont.
- Howell, R., Breivik, E., & Wilcox, J. (2007). Is Formative Measurement Really Measurement? reply to Bollen (2007) and Bagozzi (2007). *Psychological Methods*, 12(2), 238-245.
- Hsu, M. H., Chiu, M. C., & Ju, T. L. (2004). Determinants of continued use of the WWW: an integration of two theoretical models. *Industrial Management and Data Systems*, 104(9), 766-775.
- Hu, L.-t., & Bentler, P., M. (1995). Evaluating Model Fit. In R. H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues and applications* (pp. 76-99). Thousand Oaks, CA: Sage.
- Hu, L.-t., & Bentler, P., M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural equation modelling*, 6(1), 1-55.
- Iacobucci, D., & Ostrom, A. L. (1995). Distinguishing Service Quality and Customer Satisfaction: The Voice of the Consumer. *Journal of Consumer Psychology*, 4(3).
- James, L. R., Mulaik, S. A., & Brett, J. M. (1982). *Causal analysis: Assumptions, models and data*. Beverley Hills, CA: Sage.

- Jarvis, C. B., MacKenzie, S., & 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.
- Jon, A. (2000). Customer Connections. *Electric Perspectives*, 15(4), 14-22.
- Joreskog, K. G., & Sorbom, D. (1988). *LISREL 7: A guide to the program and applications*. Chicago: SPSS Inc. .
- Keen, P. (1980). *MIS research: Reference disciplines and a cumulative tradition*. Paper presented at the International Conference on Information Systems (ICIS), Philadelphia, United States.
- Keller, T., & Dansereau, F. (2001). The Effect of Adding Items to Scales: An Illustrative Case of LMX. *Organisational Research Methods*, 4(2), 131-143.
- Kettinger, W. J., & Lee, C. C. (1995). Exploring a "gap" model of information services quality. *Information Resources Management Journal*, 8(3), 5-16
- Kettinger, W. J., & Lee, C. C. (1997). Pragmatic Perspectives on the measurement of information systems service quality. *MIS Quarterly*, 21(2), 195-208.
- Kim, E. (1999). *Model of an Effective Web*. Paper presented at the Americas Conference on Information Systems.
- King, W. R., & He, J. (2005). Understanding the Role and Methods of Meta-Analysis in IS Research. *Communications of the Association for Information Systems*, 16(1), 665-686.
- Kline, R. B. (2005). *Principles and practice of Structural Equation Modelling*. New York, New York: The Guildford Press.
- Kuhn, T. S. (1983). Rationality and Theory Choice. *Journal of Philosophy*, 80, 563-570.
- Kuhn, T. S. (1996). *The Structure of Scientific Revolutions, 3rd edition*. Chicago: University of Chicago Press.
- Kuo, T., Lu, I.-y., Huang, C.-h., & Wu, G.-c. (2005). Measuring User's Perceived Portal service Quality: An Em-pirical Study. *Total Quality Management and Business Excellence*, 16(3), 309-321.
- Lagrosen, S., Seyyed-Hashemi, R., & Leitner, M. (2004). Examination of the Dimensions of Quality in Higher Education. *Quality Assurance in Education*, 12(2), 61-69.
- Lance, C. E., Dawson, B., Birkelbach, D., & Hoffman, B. (2010). Method Effects, Measurement Error, and Substantive Conclusions. *Organisational Research Methods*, 13(3), 435-455.
- Landauer, T. K. (1995). *The trouble with computers: Usefulness, usability, and productivity*. Cambridge, MA: The MIT Press.
- Lawrence, E., Corbitt, B., Fisher, J., Lawrence, J., & Tidwell, A. (1998). *Internet Commerce*. Milton, Queensland, Australia: John Wiley and Sons.
- Leavitt, N. (2009). Is Cloud Computing Really Ready for Prime Time? *Computer*, 42(1), 15-20.
- Lee, G.-G., & Lin, H.-F. (2005). Customer perceptions of e-service quality in online shopping. *International Journal of Retail and Distribution Management*, 33(2/3), 161-175.
- Lemon, K., Rust, R., & Zeithaml, V. (2001). What Drives Customer Equity. *Marketing Management*, 10(1), 20-25.
- Li, Y. N., Tan, K. C., & Xie, M. (2002). Measuring web-based service quality. *Total Quality Management*, 13(5), 685-700.

- Lin, C., Wu, S., & Tsai, R. (2005). Integrating playfulness into expectation-disconfirmation model for web portal context. *Information and Management*, 42(5), 683-693.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1198-1202.
- Liu, C., Arnett, K. P., Capella, L. M., & Taylor, R. D. (2001). Key dimensions of Web design quality as related to consumer response. *Journal of Computer Information Systems*, 42(1), 70-78.
- Lohmöller, J.-B. (1989). *Latent Variables Path Modeling with Partial Least Squares*. Heidelberg Physica-Verlag.
- Loiacono, E. T., & Taylor, N. J. (1999). *Factors Effecting Perceptions of Web Site Quality*. Paper presented at the Americas Conference on Information Systems.
- Longino, H. E. (1990). *Science as Social Knowledge: values and Objectivity in Scientific Enquiry*. Princeton, New Jersey: Princeton University Press.
- Lyytinen, K., & King, J. L. (2004). Nothing at the centre? Academic legitimacy in the Information Systems field. *Journal of the Association for Information Systems*, 5(6), 220-264.
- MacCallum, R., Browne, M. W., & Sugawara, H. (1996). Power Analysis and determination of Sample Size for Covariance Structure Modelling. *Psychological Methods*, 1(2), 130-149.
- Maraun, M. (1996). Metaphor Taken as Math: Indeterminacy in the Factor Analysis Model. *Multivariate Behavioral Research*, 31(4), 517-538.
- Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences*, 42(5), 851-858.
- Marsh, H., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers on overgeneralising Hu and Bentler's (1999) findings. *Structural equation modelling*, 11(3), 320-341.
- McGrenere, J., & Ho, W. (2000). *Affordances: Clarifying and evolving a concept*. Paper presented at the Graphcis Interface2000, Canadian Human-Computer Communications Society, Toronto.
- McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typography. *Information Systems Research*, 13(3), 334-359.
- McKnight, D. H., Kacmar, C., & Choudhury, V. (2004). Dispositional Trust and Distrust Distinctions in Predicting High- and Low-Risk internet Expert Advice Site Predictions. *E-Service Journal*, 3(2), 35-58.
- Moon, J. M., & Kim, Y. G. (2001). Extending the TAM for a World-Wide-Web context. *Information and Management*, 38(4), 217-230.
- Moore, G. C., & Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information technology Innovation. *Information Systems Research*, 2(3), 192-222.
- Nielsen, J. (1999). *Designing Web Usability: The Practice of Simplicity*. Indianapolis: New Riders Publishing.
- Nielson, J. (1993). *Usability Engineering*. San Diego: Academic Press.
- Norman, D. A. (1999). Affordances, conventions, and design. *Interactions*(May/June), 38-42.

- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd Edition)* New York, New York: McGraw-Hill.
- Oliver, R. (1981). Measurement and evaluation of satisfaction processes in retail settings. *Journal of Retailing*, 57(25-48).
- Oliver, R., & Swan, R. (1989). Equity and Disconfirmation Perceptions as Influences on Merchant and Product Satisfaction. *Journal of Consumer Research*, 16(3), 372-383.
- Pace, V., L. (2010). method Variance from the Perspectives of Reviewers: Poorly Understood Problem or Overemphasized Complaint? *Organisational Research Methods*, 13(3), 421-434.
- Pallant, J. (2005). *SPSS survival manual* (2nd ed.). Crows Nest NSW: Allen & Unwin.
- Parasuraman, A., & Grewal, D. (2000). Serving customers and onsumers effectively in the twenty-first century: A conceptual framework and overview. *Academy of Marketing Science Journal*, 28(1), 9-18.
- Parasuraman, A., & Zeithaml, V. (2001). Understanding and improvng service quality: a literature review and research agenda. In R. Wensley & B. Weitz (Eds.), *Handbook of Marketing* (pp. 339-367). London, England: Sage.
- Parasuraman, A., Zeithaml, V., & Berry, L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49(4), 41-50.
- Parasuraman, A., Zeithaml, V., & Berry, L. (1988). SERVQUAL: A Multi-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 64(Spring).
- Parasuraman, A., Zeithaml, V., & Berry, L. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, 67(4), 420-451.
- Parasuraman, A., Zeithaml, V., & Berry, L. (1994). Reassessment of Expectations as a Comparison Standard in measuring Service Quality: Implications for Further Research. *Journal of Marketing*, 58(1), 111-124.
- Parasuraman, A., Zeithaml, V., & Malhotra, A. (2005). E-S-Qual: A multiple-item scale for assessing electronic servive delivery. *Journal of service Science*, 7(3), 213-233.
- Petter, S., Straub, D., & Rai, A. (2007). Specifying Formative Constructs in Information Systems Research. *MIS Quarterly*, 31(4), 623-656.
- Pitt, L., & Watson, R. (1994). *Service Quality in Information Systems: A longitudinal Case Study*: Henly Management College.
- Pitt, L., Watson, R., & Kavan, B. (1995). Service Quality: A Measure of Information Systems Effectiveness. *MIS Quarterly*, 19(2), 173-187.
- Pitt, L., Watson, R., & Kavan, B. (1997). Measuring Information Systems Service Quality: Concerns for a Complete Canvas. *MISQ*, 21(2), 209-221.
- Podsakoff, N. P., Shen, W., & Podsakoff, P. M. (2006). The Role of Formative Models in Strategic Management Research: review, Critique, and Implications for Future Research. *Research Methodology in Strategy and Management*, 3, 197-252.
- Podsakoff, P. M., Podsakoff, N. P., MacKenzie, S., & Lee, J.-Y. (2003). Common Methods Biases in Behavioural Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), 879-903.
- Popper, K. R. (1959). *The Logic of Scientific Discovery*. New York, New York: Basic Books.

- Rafferty, A. E. (1993). Bayesian model selection in structural equation models. In K. Bollen (Ed.), *Testing Structural Equation Models*. Newbury Park, CA: Sage.
- Rice, M. (1997). What Makes Users Revisit a Web Site? . *Marketing News*, 31(6), 12-12.
- Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A Tale of Three Perspectives: Examining Post Hoc Statistical Techniques for detection and Correction of Common Method Variance. *Organisational Research Methods*, 12(4), 762-800.
- Rouse, A. C., & Corbitt, B. (2008). *There's SEM and "SEM": A Critique of the Use of PLS Regression in Information Systems Research*. Paper presented at the Australasian Conference in Information Systems (ACIS).
- Rust, R., & Oliver, R. (1994). Service Quality: Insights and Managerial Implications from the Frontier. In R. Rust & R. Oliver (Eds.), *Service Quality: New Directions in Theory and Practice*. Thousand Oaks, CA: Sage Publications.
- Sandhu, K., & Corbitt, B. J. (2003). *Web-based Electronic Service Adoption Model (E-SAM)*. Paper presented at the PACIS, Adelaide, South Australia.
- Scharl, A., & Bauer, C. (1999). *Explorative analysis and evaluation of commercial web information systems*. Paper presented at the ICIS, Charlotte, North Carolina.
- Schonemann. (1996a). The Psychopathy of Factor Indeterminacy. *Multivariate Behavioral Research*, 31(4), 571-577.
- Schonemann. (1996b). Syllogisms of Factor indeterminacy. *Multivariate Behavioral Research*, 31(4), 651-654.
- Seddon, P. (2006). *MISQ New Author's Workshop*. Paper presented at the Pacific-Asia Conference in Information Systems (PACIS), Auckland, New Zealand.
- Seddon, P., & Lyytinen, K. (2008). *Panel on Generalisability in Information Systems Theory*. Paper presented at the Australasian Conference in Information Systems (ACIS).
- Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology*, 15(201-293).
- Spector, P., E. (2006). Method variance in organisational research: Truth or urban legend? . *Organisational Research Methods*, 9, 221-232.
- Spreng, R. A., & Mackoy, R. D. (1996). An Empirical Examination of a Model of Perceived Service Quality and Satisfaction. *Journal of Retailing*, 72(2), 201-214.
- SPSS_Inc. (2007). *SPSS Missing Values 17.0*: SPSS_Inc.
- Srivihok, A., Ho, R., & Burstein, F. (2000). *An Instrument for Web Measurement: End User Evaluation of Web Application Effectiveness*. Paper presented at the ACIS, Brisbane, Australia.
- Stafford, M. R., Prybutok, V., Wells, B., & Kappelman, L. (1999). Assessing the Fit and Stability of Alternative Measures of Service Quality. *Journal of Applied Business Research*, 15(2), 13-31.
- Stafford, T. F., & Stafford, M. R. (2001). Identifying Motivations for the Use of Commercial Web Sites. *Information Resources Management Journal*, 14(1), 22-31.
- Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the ACM*, 13, article 24, 1-79.
- Straub, D., & Carlson, C. L. (1989). Validating Instruments in MIS research. *MIS Quarterly*, June, 147-169.

- Sylvester, A., & Tate, M. (2008). *Beyond the "Mythical Centre": An Affirmative Post-Modern View of SERVQUAL Research in Information Systems*. Paper presented at the European Conference in Information Systems (ECIS), Galway, Ireland.
- Sylvester, A., Tate, M., & Johnstone, D. (2007). *Re-presenting the literature review: A Rich Picture of Service Quality Research in Information Systems*. Paper presented at the Pacific Asia Conference in Information Systems (PACIS), Auckland, New Zealand.
- Szymanski, D. M., & Hise, R. T. (2000). E-satisfaction: An initial examination. *Journal of Retailing*, 76(3), 309-322.
- Takeuchi, H., & Quelch, J. (1988). Quality is more than making a good product. *Harvard Business Review*, 61(4), 139-145.
- Tan, C.-W., Benbasat, I., & Cenfetelli, R. (2008). *Building Citizen Trust Towards eGovernment Services: Do High Quality Websites Matter?* . Paper presented at the Hawaii International Conference in System Sciences (HICSS).
- Tate, M., & Evermann, J. (2009a). *Descendents of ServQual in Online Services Research: The End of the Line?* Paper presented at the Americas Conference in Information Systems (AMCIS), San Francisco.
- Tate, M., & Evermann, J. (2009b). *Perceptive Users with Attitudes: some Heuristics of Theorising*. Paper presented at the ICIS, San Francisco.
- Tate, M., Evermann, J., & Hope, B. (2008a). *Old Theory and New Service Quality: An Exploratory Study of the Nature and Nomological Net of Online Service Quality and Continuing Use using Information Systems Theory*. Paper presented at the Australasian Conference in Information Systems (ACIS).
- Tate, M., Evermann, J., & Hope, B. (2008b). Stakeholder expectations of service quality in a university web portal. In D. Oliver, C. Romm & F. Sudweeks (Eds.), *Self-Service and the Internet*. Springer.
- Tate, M., Evermann, J., Hope, B., & Barnes, S. (2007). *Perceived Service Quality in a University Web Portal: Revising the E-Qual instrument*. Paper presented at the Hawaii International Conference on System Sciences (HICSS), Waikoloa, Hawaii.
- Taylor, S., & Barker, T. (1994). An Assessment of the Relationship between Service Quality and Customer Satisfaction in the Formation of Consumer's Purchase Intentions. *Journal of Retailing*, 70(2), 163-179.
- Teas, R. (1994). Expectations as a comparison standard in measuring service quality: An assessment of a reassessment" *Journal of Marketing*, 58(1), 132-139.
- Tiecher, J., Hughes, O., & Dow, N. (2002). E-government: A New Route to Public Sector Quality. *Managing Service Quality* 12(6), 384-393.
- Tilson, R., Dong, J., Martin, S., & Kieke, E. (1998). *Factors and principles affecting the usability of four e-commerce sites*. Paper presented at the Conference on Human Factors and the Web,, Basking Ridge, NJ.
- Turban, E., Lee, J., Warkentin, M., & Chung, H. (2002). *Electronic Commerce: A Managerial Perspective*. Upper Saddle River, New Jersey: Prentice Hall.
- Van Dyke, T. P., Kappelman, L., & Prybutok, V. (1997). Measuring Information Systems Service Quality: Concerns on the SERVQUAL Questionnaire. *MIS Quarterly*, 21(2), 195-208.
- Venkatesh, V., Morris, M., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Towards a Unified View. *MIS Quarterly*, 27(3), 425-478.

- Wang, Y., Tang, T., & Tang, J. E. (2001). An Instrument for Measuring Customer Satisfaction Toward Web Sites that Market Digital Products and Services. *Journal of Electronic Commerce Research*, 2(3), 89-102.
- Watson, R., Pitt, L., & Kavan, B. (1998). Measuring Information Systems Quality: Lessons from Two Longitudinal Case Studies. *MISQ*, 22(1), 61-79.
- Weber, R. (1997). *Ontological Foundations of Information Systems* Melbourne, Australia: Coopers & Lybrand and the Accounting Association of Australia and New Zealand.
- Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological Methodology* (pp. 84-136). San Francisco, CA: Jossey-Bass.
- Wilcox, J., Howell, R., & Breivik, E. (2008). Questions about formative measurement. *Journal of Business Research*, 61(12), 1219-1228
- Wolfenbarger, M., & Gilly, M. (2003). eTailQ: dimensionalizing, measuring and predictingetail quality. *Journal of Retailing*, 79, 183-198.
- Yang, Z., Jun, M., & Peterson, R. T. (2004). Measuring customer perceived online service quality. *International Journal of Operations & Production Management*, 24(11), 1149-1174.
- Zeithaml, V. (2000). Service Quality, profitability, and the economic work of customers: What we know and what we need to learn. *Academy of Marketing Science Journal*, 28(1), 67-86
- Zeithaml, V., Berry, L., & Parasuraman, A. (1996). The Behavioural Consequences of Service Quality. *Journal of Marketing*, 60.
- Zeithaml, V., Parasuraman, A., & Berry, L. (1988). Communication and Control Processes in the Delivery of service Quality. *Journal of Marketing*, 52(2), 35-48.
- Zeithaml, V., Parasuraman, A., & Malhorta, Y. (2002). Service quality delivery through websites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science*, 30(4), 262-375.

Appendix I: Cross-reference

This table provides a cross-reference between

1. the question (item) numbers on the survey instrument, which are also used as the variable numbers in the dataset,
2. the full wording (long text) of the question, and
3. the short text of the question

This table can be used for reference for models, AMOS reports, and other diagrams and tables included in this report.

Question/Variable number/Long text/Short text cross reference

Question Variable number	Long text	Short Text
x1	I find the site easy to learn to operate	Learnability
x2	The site provides accurate information	Accuracy
x3	This site has a good reputation	Reputation
x4	I believe I would be able to complete transactions that are useful to me	Txns that are important
x5	My interaction with the site is clear and understandable	Clear and understandable
x6	It feels safe to complete transactions on this site	Txns feel safe
x7	The site provides believable information	Believable
x8	Completing transactions on this site will save me time or money	Txns save time-money
x9	I find the site easy to navigate	Navigability
x10	The site provides timely information	Timeliness
x11	My personal information feels secure	Personal info feels secure
x12	This site offered the range of on-line transactions I expected	Range of txns
x13	I find the site easy to use	Easy to use
x14	The site provides relevant information	Relevance
x15	This site creates a sense of personalization	Personalisation
x16	The site has an attractive appearance	Attractive
x17	The site provides easy to understand information	Easy to understand
x18	The design is appropriate to the type of site	Design appropriate
x19	I feel in control on this site	In control
x20	The site conveys a sense of competency	Competency
x21	This site conveys a sense of enjoyability or entertainment	Enjoyability
x22	The site creates a positive experience for me	Positive experience
x23	The site provides information at the right level of detail	Level of detail
x24	This site makes it easy to communicate with the organization	Communicate with org
x25	Overall, the response time was acceptable to me	Reponse time
x26	The site presents the information in an appropriate format	Appropriate format
x27	I feel confident that goods/services I have requested will be delivered as promised	Goods-srvcs will be dlvr'd
x28	What is your overall view of the quality of this website	Overall POLSQ

Appendix II: Survey Questions Pre-test

This document provides a report on the pre-test carried out for the new questions which were added to the eQual 4.0 survey instrument

Survey Questions Pre-test

Six participants assisted with the survey question pre-test. All were tertiary educated with considerable e-commerce experience. The questions were pre-tested using the following protocol:

“The scenario is an e-services site, for example banking, insurance, local body or education. These are questions where you would be asked to provide a response on a scale from strongly agree to strongly disagree. At this stage I am pre-testing the questions. I just want you to tell me what you think I am asking when I ask this question.” (results 1-4)

“For the first four questions I was considering the same set of questions, but using the word transaction rather than task. Can you tell me what the words “transaction” and “task” mean to you?” (results 5)

Now we will continue with the rest of the questions (results 6-8).

Results of pre-test

1. Question: I was able to complete tasks that are useful to me

1	Consider separating what I was able to do from what was useful. Some tasks I could complete might not be useful.
2	Whether I could do what I had come to the site to do
3	Something about usability – does it require me to do something as defined by the interface
4	I was able to do the things that I wanted to do
5	I was able to complete what I was willing to do – and things that I would appreciate, like checking my bank balance
6	People have a perception of the service and they have their customer needs. This is to what level their needs are fulfilled and how pleased they are

2. Question: Completing tasks on this site will save me time or money

1	I would want the save time and the save money parts to be split – they are two questions
2	This is independent of whether it was easy to do the task – but if it was it would have saved time or money – you need to make sure that is clear
3	I think this is asking about the efficiency of the task
4	Consider adding “compared to alternatives”. Doing what I want to do on the Internet is preferable to doing it any other way. There are two reasons I might do something on-line – because it is there, or because it is easier or cheaper than doing it any other way – I need to do this somehow and its best to do it on-line
5	The site is a better way of doing something that I wanted to do. I would accomplish the task on a more efficient way than doing it any other way.
6	Mostly if I do something on-line I just want it solved immediately. I you have time you usually prefer f2f. This is how simple, how responsive each step is. If it takes 10 minutes to pay a bill on-line I’d rather go into the bank. If you want a problem solved, you care about how quick, how responsive, you don’t care about the design.

3. Question: This site offered the [range of] on-line tasks I expected

1	Question is clear
2	What I want, when I want it, didn't frustrate my expectations
3	Completeness – the overall thing will determine the range. There is a possibility that uses expectations are not accurate
4	Having decided what I want to do, this is what the site allows, or makes it easier for me to do – this is the difference between being able to do something [at all] and how easy it is to do.
5	The site is offering the mix of features that I expect for this kind of service – for example, a bank
6	I would expect the site to bring in the range of tasks from a brick and mortar shop. I would have a perception of the business objectives from that.

4. Question: I was able to accomplish the tasks I wanted to

1	OK
2	Consider “I was able to accomplish the tasks that I had come to the site in order to accomplish”
3	OK
4	Consider “I was able to accomplish the tasks I set out to. Looks like you have two questions 1) Did it have the services I want 2) Was I able to use them Ideally the questions should be proximate so the difference is clear
5	I had a goal and I made it
6	I think you are asking how reliable the website is – how many errors

5. Using the same question-set using the word “transaction” instead of “task”

- ☐ *I was able to complete transactions that are useful to me*
- ☐ *Completing transactions on this site will save me time or money*
- ☐ *This site offered the on-line transactions I expected*
- ☐ *I was able to complete the transactions I wanted to on-line*

1	To me transaction means spending money or receiving money or transferring money A task may include transactions
2	You may need to explain your terms and give examples To me transaction means actually doing something on the site that changes something – e.g. banking, registering, using the interactivity. A task might include getting information
3	A transaction means something like writing a cheque, exchange of money or commitment. A transaction is a sub-category of task, defined by the nature of the task
4	Task is more generic, with broader appeal. The word transaction is quite well understood in the financial and information processing areas, well understood in a more formal business sense. It may not have the same meaning for people in other fields
5	If you do a web survey, you could add a hyperlink to explain the keywords in a textbox. To me a task is an activity the site allows me to do. A transaction is information exchange, or financial exchange.
6	Transaction – I put in input and get output Task – I may not get output directly

6. Question: *I feel in control*
I feel in control on this site

The wording was changed to “*I feel in control on this site*” after participants made jokes about “*in control of what, my life?*”

1	OK, clear
2	OK. Maybe you could consider the negative “I felt this site was controlling my experience”
3	Control is a slippery term. I would not be too sure what I was being asked. The user can only control what the interface and the design allows – for example I might be required to put in a middle initial when I don’t want to, or I might want to put one in and I can’t. Control could relate to many things. Privacy is also a factor.
4	I like this question
5	I feel I’m behind the steering wheel. I can go wherever I want. I dominate the technology
6	This procedure is closely aligned with the task description. I don’t get unexpected things. It keeps “on task” and doesn’t go outside. People get frustrated, for example with pop-up windows.

7. Question: *This site conveys a sense of fun or enjoyability*
This site conveys a sense of entertainment or enjoyability

The wording was changed to “*This site conveys a sense of entertainment or enjoyability*” after feedback

1	OK, but I don’t go to a service website to be entertained. I do what I have to do.
2	“Fun” is too lively, try “entertainment”
3	The question is clear, but this is not very important to me
4	Clear
5	Its pleasant to complete my tasks, fun, not boring
6	Some sites don’t aim for enjoyment – its all about speed and performance

8. Question: *The waiting time was acceptable to me*
Overall the waiting time was acceptable to me

The wording was changed to “*Overall, the waiting time was acceptable to me*” after the notion of response time attracted discussion.

1	The time from giving an instruction to having it executed
2	You might have quite variable experiences that this question doesn’t capture
3	There are lots of types of wait – I’d want to qualify it, its hard to make an overall assessment
4	I’d consider the whole session if asked for overall waiting time. Am I happy th6at it was an acceptable amount of time for me to do what I wanted to
5	I think that I was able to do what I wanted. The time of sending and receiving my response was in a range that I considered acceptable
6	I would judge against a similar service from a different company. Its also how anxious you are to do the task. For example if I have to pay a bill now or my phone will be cut off.

Appendix III: Online Survey Pre-test

This document provides a report on pre-test carried out for the online version of the survey

Survey Online Pre-test results

Overview

Number of respondents: 8

Sample of respondents: SIM staff and post-graduates

Time taken (minutes)

35, 20, 35, 40, 15, 20, 15, 20

Shorter times appear not to have been very thorough. Approx 35 minutes is probably realistic.

Question Feedback

Q1-4. Some questions were quite hard to respond to. [Probably referring to the exercise – Mary's note]

Q4a. Not clear if this differs for international students or domestic students

Q4a. I had to look up two separate pages and use a calculator.

Q4a-7. Only refers to a specific group of people

Q5a. The first link using the search function came up to an outdated SIM page (1999)

Q 8-8b, 9-9B are the same

Q8. Difficult to respond to the “learn to operate” question as I already know

Q9b. Reputation – nothing to judge this by in the experience part of the exercise

Q9b. Consider replacing “I was able to” with “I believe I would be able to”.

Q9. What scale is considered to evaluate the accuracy of the site

Q12. Not sure of meaning.

Q.12. My interaction with the site is clear and understandable

I'm not completely sure about this question. Am I being asked if I understood the interaction I had with the website or if the interaction the website provided me is understandable? Also, when talking about interaction with a website, I'm not sure about what a clear interaction means.

Q13. Completing the transactions has two different aspects 1) the nature of the transaction – internal(?) or financial, and 2) personal preference. Here it is not clear which aspect it is referring to.

Q16, Q20. These questions look similar

Q.17. The site provides timely information and **Q.32.** Overall the response times were acceptable to me. What is the difference between those two questions?

Q17. Timely information – possibly encompassed in accurate information? But maybe not? Felt to some extent like a doubled-up question.

Q.27. and Q.27b. have a typo error conveyys instead of conveys. My spelling is shocking, I noticed this error because I wasn't sure about the meaning of the word "convey" and I looked at it in a dictionary; this brings me to comment on the use of the "convey" word. I don't know if it is just my lack of vocabulary or maybe the word could be replaced for communicate or any other synonym .

Q27. Competency of what?

Q29. It is difficult to follow up.

Q32. Somewhat vague – response times to what?

Q34. “The goods and services I have requested”. I didn’t request any.

General comments

R1:

- ☐ The task to perform weren't that simple for me =] Well, the truth is that I haven't used the VUW site much and I didn't know where to look at to start with. There are a few places where the information could've been located and getting to the right place took me a few minutes. I completed all four task because I knew it was for your; if the survey was from someone I didn't know, I don't honestly know if I would've gone through all the tasks. To be precise, the task that was tricky for me was 6a.
- ☐ In general the survey is well layout and clear to follow through. One question about the survey; in a survey do you have to state what the scale means? E.g. I know you used a Likert test, and I know the values from 1 to 7; but, it is necessary to state all the values in the survey? I don't really know about this.

R2:

- ☐ Add more information and a disclaimer about the research being anonymous
- ☐ The tasks are hard and time-consuming – many people would give up. Possibly you should make it clearer that it is an experiment as well as a survey so people are prepared.
- ☐ Consider having the tasks in a more logical order – admission, enrollment, course information.
- ☐ Consider adding a hyperlink to define key terms
- ☐ Consider allowing pages to be loaded and submitted one at a time – this would assist people on slow connections and allow you to see where people gave up (if they did)
- ☐ Consider a push-type survey e.g. to all SCS or all SIM students. Could then do a follow-up and a non-respondent analysis.
- ☐ Consider linking from Bb rather than from the home page.

R3:

- ☐ Get rid of text “a PhD student”. Improve impact of opening section – need a click-thru section to accept the terms and conditions (HEC).
- ☐ Q 1-4 quite easy and quick. Fee question time-consuming – lost the survey – may need to open a separate window.
- ☐ I think the difficulty of the experience questions would cause people to give up. It would have put me off if it had been anyone else.
- ☐ I did not see the distinction rating vs importance till I was at Q12. Consider:
 - More white space separate questions
 - Less space between part a and part b of each question – make a clearer pairing.
 - Slightly different wording, e.g.
 - Agree.....disagree
 - Very important....totally unimportant
 - Important....unimportant

R4:

- ☐ I kept having to return to the top to check what each of the two questions was actually asking
- ☐ Lengthy and wondering what purpose each question was.

R5:

- ☐ Some questions quite similar
- ☐ Questions 4a-7 are only relevant to a specific group of people

Appendix IV: Survey instrument

This document provides the information sheet, instructions to participants, and survey instrument used in the data gathering.

Validating the dimensions of e-service quality

This questionnaire was prepared by Mary Tate, from the School of Information Management, Victoria University of Wellington. The aim of this study is to validate the dimensions of e-service quality that are important to users.

The questionnaire is to be completed by people who have used the VUW home site. Completing the exercise and questionnaire will take approximately 20 minutes.

If you are interested in seeing the findings of this survey, summarised results will be posted when they are available for a finite period on the VUW website (www.vuw.ac.nz)

Information and Consent

1. Your participation is voluntary, and you may decline to participate or stop completion at any time without consequence.
 2. The information collected in this survey will be made available, in summarised form, to the VUW Marketing Department, and may be published in academic conferences or journals. Detailed responses will only be seen by the researcher and will be anonymous. (For the on-line survey only) Your e-mail address is collected by the system just to ensure that people do not complete the survey more than once, and for the administration of the prize draw, but e-mail addresses are kept separate from the survey responses and will not be made available to the researcher.
 3. By filling out this survey, you are indicating that you have read this statement and have agreed to voluntarily participate. You agree that the data collected may be used, in summarised form, for the purposes explained above.
 4. Electronic data collected will be stored in a secure, password-protected folder. Paper documents will be stored in a locked cabinet. Data will be destroyed on completion of the research.
 5. All information provided by you for this research is confidential. Information will be used only in summarised form to create a general profile of the people who responded to this questionnaire.
 6. Be spontaneous in your responses. There are no right or wrong answers.
-

Personal information

All information provided by you for this research is confidential. Information will be used only to create a general profile of the people who responded to this questionnaire.

1. How often do you usually access the VUW Home Site?

At least once a week ☐

At least once a month ☐

Less than once a month ☐

This is my first visit ☐

2. What is your age?

Under 20 ☐ 26-30 ☐ 41-50 ☐

20-25 ☐ 31-40 ☐ Over 50 ☐

3. Are you

Male ☐ Female ☐

4. Are you a

Prospective student ☐ Current student ☐

Prospective staff member ☐ Current staff member ☐

Former student ☐ Other ☐

A: Instructions to participants

This survey is evaluating the VUW home site. This includes all pages that contain this in their URL.

<http://www.vuw.ac.nz/home>

The sections of the site that are relevant for the survey are shown on the overhead projector.

B. Exercise

To ensure you have a reasonable opportunity to judge the service quality offered by the VUW home site you are requested to complete two tasks.

1. Obtain a form to assess a how much credit you will get for qualification you have from another institution. What is the name of the form? What format is it available in?

Name of form:.....

Format:.....

2. What time is the lecture for ELCM 201 in trimester 1,2004. What room is it in?

Time:.....

Room:.....

What sort of experience did you have? Please complete the survey.

C: E-quality survey of the VUW home site

Quality	Description
1	I find the site easy to learn to operate
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
2	The site provides accurate information
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
3	This site has a good reputation
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
4	I believe I would be able to complete transactions that are useful to me
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
5	My interaction with the site is clear and understandable
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
6	It feels safe to complete transactions on this site
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
7	The site provides believable information
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
8	Completing transactions on this site will save me time or money
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
9	I find the site easy to navigate
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
10	The site provides timely information
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
11	My personal information feels secure
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
12	This site offered the range of on-line transactions I expected
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
13	I find the site easy to use
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
14	The site provides relevant information
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
15	This site creates a sense of personalization
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
16	The site has an attractive appearance
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
17	The site provides easy to understand information
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
18	The design is appropriate to the type of site
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
19	I feel in control on this site
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
20	The site conveys a sense of competency
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
21	This site conveys a sense of enjoyability or entertainment
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree
22	The site creates a positive experience for me
Score	Strongly agree 1 2 3 4 5 6 7 N/A Strongly disagree

23	The site provides information at the right level of detail										
Score	Strongly agree	1	2	3	4	5	6	7	N/A	Strongly disagree	
24	This site makes it easy to communicate with the organization										
Score	Strongly agree	1	2	3	4	5	6	7	N/A	Strongly disagree	
25	Overall, the response time was acceptable to me										
Score	Strongly agree	1	2	3	4	5	6	7	N/A	Strongly disagree	
26	The site presents the information in an appropriate format										
Score	Strongly agree	1	2	3	4	5	6	7	N/A	Strongly disagree	
27	I feel confident that goods/services I have requested will be delivered as promised										
Score	Strongly agree	1	2	3	4	5	6	7	N/A	Strongly disagree	
28	What is your overall view of the quality of this website										
Score	Highest quality	1	2	3	4	5	6	7	N/A	Lowest quality	
29	Comments: Please add any comments you would like to make about any aspect of this site										

Thank you for your participation

Appendix V: Factor structure comparisons

These diagrams provide summaries of various comparisons that were carried out during the exploratory factor analysis.

1. The diagram titled “Factor structure” compares the “best” 4-factor structure from the EFA compared with several other results. This shows a fairly consistent factor structure but with weak and cross-loading items resolved in the “best” factor structure.
2. The diagram entitled “Factor structure compared with hypothesised factor structure” compares the factor-item loadings from the “best” solution, with the hypothesised a-priori structure (e.g. a-P_CQ represents the items hypothesised to load on Content Quality, and so on). From this it can be observed that there was not a very close fit to the hypothesised factor structure, with the exception of some of the questions associated with content quality and usability
3. The diagram titled “Factor structure compared with eQual 4.0 compares the factor-item loadings from the “best” solution, with the eQual 4.0 factor-item loadings. From this it can be observed that there was not a very close fit to the eQual factor structure, with the exception of some of the questions associated with content quality and usability.

Factor structure

[illegible]

Factor structure compared with hypothesised factor structure

1	a-P_CQ	2	a-P_U	3	a-P_TQ	4	a-P_SIQ
2	2						
7	7						
10	10						
14	14						
18	17						
20							
23	23						
26	26						
27							
		1	1				
		3					
		4					
		5	5				
		9	9				
		13	13				
		19	16		4		
			18	6			3
			20	8	8		6
			22	11	12		11
			25	16			15
						12	19
						21	21
						22	
						24	24
						25	
							27
Legend							
	content quality						
	usability (design and usability)						
	transaction quality						
	service interaction quality (trust and empathy)						
	new questions: 4,8,12,19,21,25						
15,17	dropped						

Factor structure compared with eQual 4.0

1	EQ-CQ	2	EQ_U	3	a priori-TQ	4	EQ-SIQ	
2	2							
7	7							
10	10							
14	14							
18	17							
20								
23	23							
26	26							
27								
		1	1					
		3						
		4						
		5	5					
		9	9					
		13	13					
		19	16		4			
			18	6			3	
			20	8			6	
			22	11	8		11	
				16	12		15	
						12		
						21		
						22		
						24	24	
						25		
							27	
Legend								
	content quality							
	usability							
	design							
	trust							
	empathy							
	new questions: 4,8,12,19,21,25							
	dropped							

Appendix VI: Assessment of normality

This document provides output generated by the AMOS package evaluating the normality of the data in the dataset.

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
x15	1.000	7.000	-.282	-1.394	-.759	-1.877
x24	1.000	7.000	-.710	-3.515	.228	.564
x22	1.000	7.000	-.756	-3.742	.904	2.238
x20	1.000	7.000	-1.022	-5.058	1.608	3.979
x18	1.000	7.000	-1.226	-6.070	2.271	5.621
x16	1.000	7.000	-.604	-2.987	.037	.092
x13	1.000	7.000	-.871	-4.309	.278	.688
x9	1.000	7.000	-.780	-3.860	.064	.159
x1	1.000	7.000	-.915	-4.531	.839	2.077
x5	1.000	7.000	-.988	-4.889	.864	2.138
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x17	2.000	7.000	-.640	-3.168	.092	.227
x3	1.000	7.000	-.670	-3.315	.664	1.643
x4	1.000	7.000	-.820	-4.059	.331	.818
x12	1.000	7.000	-1.181	-5.847	1.673	4.141
x21	1.000	7.000	.021	.106	-.765	-1.894
x25	1.000	7.000	-.942	-4.662	.822	2.034
x6	1.000	7.000	-1.329	-6.579	2.839	7.027
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x11	1.000	7.000	-1.051	-5.201	1.040	2.575
x19	1.000	7.000	-.638	-3.160	.219	.541
x27	1.000	7.000	-.944	-4.674	.866	2.143
x23	1.000	7.000	-.789	-3.904	.144	.357
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
Multivariate					173.540	26.585

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
30	68.960	.000	.002
86	67.700	.000	.000
27	67.481	.000	.000
116	64.808	.000	.000
77	61.747	.000	.000
50	61.156	.000	.000
17	56.662	.001	.000
74	56.487	.001	.000
48	56.156	.001	.000
40	56.065	.001	.000
97	54.833	.001	.000
65	53.639	.002	.000
46	52.515	.002	.000
55	51.548	.003	.000
64	50.082	.004	.000
41	48.621	.007	.000
103	47.736	.008	.000
111	46.081	.012	.000
144	45.456	.015	.000
121	44.768	.017	.000
145	44.434	.019	.000
38	42.750	.028	.000
47	41.966	.033	.000
11	41.670	.035	.000
10	41.577	.036	.000
101	40.307	.048	.000
118	39.221	.060	.000
69	38.601	.069	.000
2	38.383	.072	.000
58	38.209	.075	.000
5	38.180	.075	.000
80	37.633	.084	.000
12	37.605	.084	.000
60	36.286	.109	.000
68	35.161	.135	.001
22	35.055	.137	.000
119	34.985	.139	.000
35	34.969	.140	.000
109	34.574	.150	.000
110	34.270	.158	.000
6	33.650	.176	.001
136	33.596	.178	.001
135	33.596	.178	.000
8	33.371	.185	.001
52	33.286	.188	.000
81	32.240	.223	.008
91	31.931	.235	.012
72	31.864	.237	.009
98	31.811	.239	.006
87	31.464	.253	.011
107	30.304	.301	.129
43	29.776	.324	.248
113	29.286	.347	.397
147	29.208	.351	.367
16	28.588	.381	.600

Observation number	Mahalanobis d-squared	p1	p2
125	28.254	.398	.692
131	27.894	.416	.785
49	27.538	.435	.859
137	27.404	.442	.860
127	27.367	.444	.832
90	26.483	.492	.975
23	26.139	.511	.988
45	26.015	.518	.988
123	25.699	.535	.994
141	25.657	.538	.992
92	25.605	.541	.989
78	25.380	.553	.993
66	24.926	.579	.998
29	24.872	.582	.998
128	24.812	.585	.997
28	24.360	.610	.999
21	24.043	.628	1.000
18	23.751	.644	1.000
122	23.398	.663	1.000
124	23.398	.663	1.000
56	22.681	.702	1.000
70	22.468	.713	1.000
67	22.421	.716	1.000
88	22.288	.723	1.000
126	22.011	.737	1.000
95	21.918	.742	1.000
130	21.872	.744	1.000
1	21.801	.747	1.000
20	21.616	.757	1.000
105	21.202	.777	1.000
61	20.954	.788	1.000
108	19.867	.836	1.000
84	19.866	.836	1.000
34	19.632	.846	1.000
120	19.034	.869	1.000
57	19.004	.870	1.000
51	18.395	.891	1.000
112	18.318	.894	1.000
14	17.965	.905	1.000
117	17.825	.909	1.000
146	17.694	.913	1.000
115	17.680	.913	1.000
54	17.635	.914	1.000
93	17.562	.916	1.000
102	17.527	.917	1.000

Appendix VII: AMOS reports, Confirmatory factor analysis

These documents provide output generated by the AMOS package after calculating the estimates for the four confirmatory factor analysis models included in Chapter 6. Selected subsets of these reports are included in the text.

\\staff\home\fca\tatema\My Documents\PhD\2008\SEM\webqual4.0 measurement model 3-factor.amw

Analysis Summary

Date and Time

Date: Thursday, 12 November 2009

Time: 4:01:19 p.m.

Title

webqual4.0 measurement model 3-factor: Thursday, 12 November 2009 04:01 p.m.

Notes

Bootstrap confidence intervals are not available when the Bollen-Stine bootstrap is performed.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x7
x10
x23
x27
x11
x6
x3
x26
x1
x9
x13
x16
x22
x2
x5
x15
x24
x14
x17
x18
x20

Unobserved, exogenous variables

err-b
err-t
err-lod
err-swbd
err-pifs
err-tfs
err-rep
err-af
err-l
err-n
err-etus
err-attr
err-pe
information
usability
service interaction quality
err-pers
err-cwo
err-a
err-r
err-etun
err-cau
err-da
err-c

Variable counts (Group number 1)

Number of variables in your model: 45
Number of observed variables: 21
Number of unobserved variables: 24
Number of exogenous variables: 24
Number of endogenous variables: 21

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
--	---------	-------------	-----------	-------	------------	-------

Fixed	24	0	0	0	0	24
Labeled	0	0	0	0	0	0
Unlabeled	17	3	24	0	0	44
Total	41	3	24	0	0	68

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 231
Number of distinct parameters to be estimated: 44
Degrees of freedom (231 - 44): 187

Result (Default model)

Minimum was achieved
Chi-square = 611.396
Degrees of freedom = 187
Probability level = .000

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

	Estimate	S.E.	C.R.	P	Label
x2 <--- information	.984	.110	8.926	***	
x7 <--- information	1.000				
x1 <--- usability	.774	.079	9.814	***	
x5 <--- usability	.918	.078	11.699	***	
x9 <--- usability	1.000				
x26 <--- information	.841	.122	6.914	***	
x23 <--- information	1.084	.138	7.845	***	
x10 <--- information	1.318	.135	9.747	***	
x17 <--- information	.720	.114	6.313	***	
x22 <--- usability	.745	.077	9.659	***	
x20 <--- usability	.725	.070	10.280	***	
x18 <--- usability	.654	.072	9.066	***	
x13 <--- usability	.910	.084	10.875	***	
x16 <--- usability	.465	.088	5.311	***	
x24 <--- service interaction quality	1.672	.307	5.447	***	
x15 <--- service interaction quality	1.682	.340	4.949	***	
x11 <--- service interaction quality	1.555	.290	5.354	***	
x6 <--- service interaction quality	1.000				
x3 <--- service interaction quality	1.555	.277	5.616	***	
x27 <--- service interaction quality	1.745	.304	5.747	***	

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
x2 <--- information	.731
x7 <--- information	.773
x1 <--- usability	.741
x5 <--- usability	.845
x9 <--- usability	.802
x26 <--- information	.581
x23 <--- information	.651
x10 <--- information	.790
x17 <--- information	.534
x22 <--- usability	.732
x20 <--- usability	.768
x18 <--- usability	.696
x13 <--- usability	.801
x16 <--- usability	.438
x24 <--- service interaction quality	.651
x15 <--- service interaction quality	.552
x11 <--- service interaction quality	.631
x6 <--- service interaction quality	.503
x3 <--- service interaction quality	.691
x27 <--- service interaction quality	.724

Covariances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
information <--> service interaction quality	.393	.083	4.716	***	
information <--> usability	.769	.132	5.824	***	

usability	<-->	service interaction quality	.572	.120	4.751	***
-----------	------	-----------------------------	------	------	-------	-----

Correlations: (Group number 1 - Default model)

	Estimate
information <--> service interaction quality	.879
information <--> usability	.763
usability <--> service interaction quality	.851

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
information	.670	.126	5.314	***	
usability	1.514	.264	5.735	***	
service interaction quality	.298	.097	3.078	.002	
err-b	.451	.065	6.923	***	
err-t	.699	.104	6.722	***	
err-lod	1.068	.138	7.748	***	
err-swbd	.824	.113	7.284	***	
err-pifs	1.090	.140	7.808	***	
err-tfs	.880	.108	8.174	***	
err-rep	.790	.105	7.513	***	
err-af	.931	.116	7.997	***	
err-l	.745	.096	7.724	***	
err-n	.839	.115	7.325	***	
err-etus	.701	.096	7.335	***	
err-attr	1.382	.165	8.385	***	
err-pe	.729	.094	7.768	***	
err-pers	1.925	.239	8.063	***	
err-cwo	1.131	.147	7.720	***	
err-a	.566	.078	7.298	***	
err-r	1.270	.149	8.544	***	
err-etun	.871	.107	8.116	***	
err-cau	.511	.075	6.854	***	
err-da	.689	.087	7.912	***	
err-c	.554	.073	7.576	***	

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
x20	.589
x18	.484
x17	.285
x14	.000
x24	.424
x15	.305
x5	.714
x2	.534
x22	.536
x16	.192
x13	.641
x9	.643
x1	.549
x26	.337
x3	.477
x6	.253
x11	.398
x27	.524
x23	.424
x10	.625
x7	.597

Matrices (Group number 1 - Default model)**Residual Covariances (Group number 1 - Default model)**

	x20	x18	x17	x14	x24	x15	x5	x2	x22	x16	x13	x9	x1	x26	x3	x6	x11	x27	x23	x10
x20	.000																			
x18	.188	.000																		
x17	.229	.095	.000																	
x14	.784	.695	.613	.000																
x24	.046	.060	.138	.776	.000															
x15	.013	-.033	-.092	.650	-.029	.000														
x5	-.060	-.045	.096	.843	-.107	.262	.000													
x2	.198	.124	-.017	.790	-.084	-.196	-.039	.000												
x22	.115	.041	.081	.622	.248	.323	-.038	.010	.000											
x16	.126	.066	.313	.417	-.001	.150	-.006	.076	.020	.000										
x13	-.097	-.049	.117	.725	-.260	.215	-.001	-.237	-.133	-.096	.000									
x9	-.175	-.174	-.018	.675	-.106	.351	.063	-.209	-.124	-.074	.372	.000								
x1	-.102	-.109	-.023	.595	-.176	.182	.070	-.057	-.082	-.063	.156	.144	.000							
x26	.291	.098	.020	.604	.113	.149	.050	-.054	.177	.085	-.057	-.169	-.083	.000						
x3	.107	.072	-.105	.539	-.025	.091	.096	-.119	.235	.047	.086	.234	.136	-.099	.000					

x6	-.013	-.024	-.197	.379	.002	.014	.010	-.017	.038	-.118	-.218	-.167	-.097	.235	-.022	.000			
x11	.042	.136	-.060	.773	.087	-.150	-.197	-.002	-.049	-.053	-.389	-.432	-.246	.128	-.029	.438	.000		
x27	.201	.061	.034	.848	.055	-.126	-.041	.063	.180	.049	-.131	-.247	-.104	.251	-.075	-.149	.029	.000	
x23	.170	.157	.032	.833	.157	-.191	.039	-.042	.337	-.065	-.227	-.212	-.092	.144	-.149	-.018	.138	.208	.000
x10	.002	.085	.023	.921	.012	-.148	.050	.022	.116	-.018	-.099	.017	-.186	-.127	-.052	-.047	-.022	.065	.006 .000
x7	.085	.118	-.045	.809	-.077	-.041	.061	.091	.018	-.033	-.201	-.197	-.152	-.065	-.008	.032	.084	-.012	-.092 .042 .0

Standardized Residual Covariances (Group number 1 - Default model)

	x20	x18	x17	x14	x24	x15	x5	x2	x22	x16	x13	x9	x1	x26	x3	x6	x11	x27
x20	.000																	
x18	1.495	.000																
x17	2.065	.864	.000															
x14	7.240	6.440	5.952	.000														
x24	.315	.415	1.028	5.937	.000													
x15	.076	-.199	-.587	4.192	-.143	.000												
x5	-.390	-.301	.742	6.759	-.622	1.324	.000											
x2	1.719	1.095	-.157	7.685	-.606	-1.218	-.290	.000										
x22	.836	.302	.678	5.326	1.580	1.772	-.233	.082	.000									
x16	.947	.501	2.579	3.416	-.007	.817	-.042	.618	.140	.000								
x13	-.616	-.320	.871	5.563	-1.464	1.043	-.007	-1.699	-.789	-.598	.000							
x9	-1.008	-1.032	-.122	4.718	-.545	1.556	.306	-1.365	-.671	-.420	1.766	.000						
x1	-.716	-.790	-.190	4.968	-1.093	.973	.418	-.452	-.542	-.428	.901	.757	.000					
x26	2.416	.823	.173	5.470	.780	.881	.357	-.462	1.370	.650	-.394	-1.060	-.629	.000				
x3	.823	.567	-.886	4.699	-.158	.504	.632	-.968	1.690	.342	.544	1.357	.954	-.773	.000			
x6	-.114	-.218	-1.934	3.748	.018	.088	.079	-.159	.322	-.988	-1.641	-1.148	-.799	2.137	-.185	.000		
x11	.301	.986	-.469	6.159	.518	-.765	-1.205	-.018	-.329	-.355	-2.296	-2.324	-1.600	.924	-.192	3.457	.000	
x27	1.437	.448	.266	6.909	.324	-.643	-.251	.475	1.203	.330	-.771	-1.324	-.674	1.824	-.503	-1.183	.180	.000
x23	1.217	1.138	.243	6.554	.933	-.971	.239	-.308	2.240	-.430	-1.341	-1.139	-.595	1.009	-1.003	-.138	.856	1.297
x10	.012	.598	.170	7.237	.069	-.736	.294	.151	.753	-.116	-.566	.087	-1.170	-.860	-.336	-.363	-.134	.388
x7	.760	1.081	-.428	8.191	-.576	-.265	.466	.821	.151	-.275	-1.481	-1.323	-1.233	-.569	-.067	.313	.658	-.090

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

			M.I.	Par Change
err-c	<-->	usability	8.606	-.157
err-c	<-->	information	8.323	.108
err-da	<-->	information	5.177	.093
err-da	<-->	err-c	16.132	.224
err-etun	<-->	usability	5.192	.150
err-r	<-->	information	24.213	.264
err-r	<-->	err-c	6.273	.184
err-pers	<-->	usability	10.390	.314
err-pers	<-->	err-c	5.613	-.219
err-pers	<-->	err-da	5.037	-.227
err-a	<-->	err-c	9.630	.163
err-pe	<-->	service interaction quality	11.387	.091
err-pe	<-->	usability	11.381	-.205
err-pe	<-->	err-c	5.812	.139
err-pe	<-->	err-cwo	7.116	.219
err-attr	<-->	err-etun	8.361	.271
err-etus	<-->	usability	11.019	.202
err-etus	<-->	information	5.581	-.101
err-etus	<-->	err-c	4.482	-.123
err-etus	<-->	err-a	4.681	-.129
err-etus	<-->	err-pe	6.216	-.164
err-n	<-->	usability	5.438	.155
err-n	<-->	information	5.766	-.112
err-n	<-->	err-c	12.070	-.220
err-n	<-->	err-da	9.279	-.212
err-n	<-->	err-r	5.110	-.207
err-n	<-->	err-pers	6.554	.296
err-n	<-->	err-pe	4.524	-.153
err-n	<-->	err-etus	44.529	.482
err-l	<-->	information	4.939	-.096
err-l	<-->	err-c	4.440	-.123
err-l	<-->	err-etus	8.439	.194
err-l	<-->	err-n	5.996	.179
err-af	<-->	service interaction quality	7.151	.080
err-af	<-->	information	5.282	-.107
err-af	<-->	err-c	9.240	.196
err-af	<-->	err-n	4.881	-.178
err-rep	<-->	usability	12.183	.223
err-rep	<-->	information	5.701	-.105
err-rep	<-->	err-n	9.086	.230
err-tfs	<-->	err-etun	8.050	-.214
err-tfs	<-->	err-af	7.852	.220
err-pifs	<-->	service interaction quality	4.460	.066
err-pifs	<-->	usability	14.102	-.278
err-pifs	<-->	err-da	7.982	.218

err-pifs <--> err-etus	6.240	-.201
err-pifs <--> err-n	8.944	-.264
err-pifs <--> err-tfs	32.746	.490
err-swbd <--> information	5.347	.105
err-swbd <--> err-c	7.368	.171
err-swbd <--> err-n	7.331	-.213
err-swbd <--> err-af	5.215	.182
err-swbd <--> err-tfs	5.253	-.175
err-lod <--> err-pe	11.733	.273
err-t <--> err-c	8.000	-.170
err-t <--> err-n	7.675	.209
err-t <--> err-af	4.621	-.163
err-b <--> err-r	5.302	.158
err-b <--> err-cau	4.548	.102
err-b <--> err-a	6.325	.122

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
x20 <--- information	4.741	.184
x20 <--- x18	7.816	.157
x20 <--- x17	6.583	.151
x20 <--- x14	6.273	.145
x20 <--- x2	12.106	.205
x20 <--- x26	12.932	.198
x20 <--- x27	7.460	.135
x18 <--- x20	5.985	.150
x18 <--- x2	4.034	.130
x18 <--- x11	7.063	.141
x18 <--- x7	4.022	.135
x17 <--- x16	8.977	.181
x17 <--- x6	5.690	-.174
x14 <--- service interaction quality	76.789	1.588
x14 <--- usability	59.870	.609
x14 <--- information	88.500	1.137
x14 <--- x20	52.415	.581
x14 <--- x18	41.480	.519
x14 <--- x17	35.432	.503
x14 <--- x24	35.252	.395
x14 <--- x15	17.571	.235
x14 <--- x5	45.682	.472
x14 <--- x2	59.056	.650
x14 <--- x22	28.364	.396
x14 <--- x16	11.667	.244
x14 <--- x13	30.949	.371
x14 <--- x9	22.264	.287
x14 <--- x1	24.682	.361
x14 <--- x26	29.925	.431
x14 <--- x3	22.077	.356
x14 <--- x6	14.048	.322
x14 <--- x11	37.939	.427
x14 <--- x27	47.733	.490
x14 <--- x23	42.955	.449
x14 <--- x10	52.378	.494
x14 <--- x7	67.089	.721
x15 <--- x9	6.005	.188
x22 <--- x24	8.951	.158
x22 <--- x27	4.497	.119
x22 <--- x23	12.300	.190
x16 <--- x17	6.504	.227
x13 <--- information	4.757	-.210
x13 <--- x24	5.686	-.126
x13 <--- x2	8.310	-.194
x13 <--- x9	14.111	.182
x13 <--- x6	5.282	-.157
x13 <--- x11	8.110	-.157
x13 <--- x23	7.675	-.151
x13 <--- x7	6.127	-.174
x9 <--- x20	4.518	-.149
x9 <--- x18	4.507	-.149
x9 <--- x14	5.110	-.163
x9 <--- x2	4.992	-.165
x9 <--- x13	14.208	.219
x9 <--- x26	7.378	-.186
x9 <--- x11	8.477	-.176
x9 <--- x27	6.392	-.156
x9 <--- x23	5.718	-.143
x9 <--- x7	4.713	-.167

x26 <--- x20	5.347	.164
x26 <--- x6	7.901	.213
x26 <--- x27	4.420	.131
x3 <--- x9	7.801	.141
x3 <--- x1	4.311	.125
x6 <--- x17	5.585	-.169
x6 <--- x26	4.828	.147
x6 <--- x11	18.310	.252
x11 <--- x13	6.826	-.168
x11 <--- x9	8.318	-.169
x11 <--- x6	23.604	.402
x27 <--- x26	5.033	.152
x23 <--- x22	5.600	.169
x7 <--- x14	5.302	.124

Minimization History (Default model)

Iteration		Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	8		-1.023	9999.000	1907.263	0	9999.000
1	e	5		-.262	4.036	1094.097	20	.211
2	e*	2		-.190	1.055	885.532	5	.657
3	e	1		-.055	.614	722.499	4	.888
4	e	0	3066.259		.885	636.439	7	.810
5	e	0	552.941		.666	621.553	4	.000
6	e	0	726.409		.399	612.973	1	1.111
7	e	0	1441.761		.250	611.571	1	1.144
8	e	0	2012.815		.132	611.404	1	1.111
9	e	0	2350.520		.033	611.396	1	1.042
10	e	0	2350.065		.003	611.396	1	1.003
11	e	0	2361.678		.000	611.396	1	1.000

Bootstrap (Default model)**Summary of Bootstrap Iterations (Default model)**

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	7	0
10	0	12	0
11	0	11	0
12	0	18	0
13	0	10	0
14	0	13	0
15	0	17	0
16	0	14	0
17	0	8	0
18	0	5	0
19	0	25	0
Total	0	140	0

0 bootstrap samples were unused because of a singular covariance matrix.

0 bootstrap samples were unused because a solution was not found.

140 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 140 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 0 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .007

Bootstrap Distributions (Default model)**ML discrepancy (implied vs sample) (Default model)**

	119.889	-----
	142.094	*
	164.300	**
	186.505	*****
	208.711	*****
	230.916	*****
	253.122	*****
N = 140	275.327	*****
Mean = 248.579	297.533	*****

S. e. = 4.098	319.738	***
	341.943	**
	364.149	
	386.354	*
	408.560	*
	430.765	*

Model Fit Summary**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	44	611.396	187	.000	3.269
Saturated model	231	.000	0		
Independence model	21	1961.618	210	.000	9.341

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.245	.719	.653	.582
Saturated model	.000	1.000		
Independence model	.686	.207	.128	.188

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.688	.650	.761	.728	.758
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.890	.613	.675
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	424.396	353.296	503.100
Saturated model	.000	.000	.000
Independence model	1751.618	1613.529	1897.118

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	4.188	2.907	2.420	3.446
Saturated model	.000	.000	.000	.000
Independence model	13.436	11.997	11.052	12.994

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.125	.114	.136	.000
Independence model	.239	.229	.249	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	699.396	715.009	830.975	874.975
Saturated model	462.000	543.968	1152.790	1383.790
Independence model	2003.618	2011.069	2066.417	2087.417

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	4.790	4.303	5.329	4.897
Saturated model	3.164	3.164	3.164	3.726
Independence model	13.723	12.778	14.720	13.774

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	53	57
Independence model	19	20

Execution time summary

Minimization: .016

Miscellaneous:	.234
Bootstrap:	.640
Total:	.890

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\CFA model based on lit review.amw

Analysis Summary

Date and Time

Date: Thursday, 12 November 2009

Time: 3:08:23 p.m.

Title

Cfa model based on lit review: Thursday, 12 November 2009 03:08 p.m.

Notes

Bootstrap confidence intervals are not available when the Bollen-Stine bootstrap is performed.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x7
x10
x14
x23
x27
x19
x11
x8
x6
x25
x21
x12
x4
x3
x17
x26
x5
x1
x9
x13
x16
x18
x20
x22
x24
x15
x2

Unobserved, exogenous variables

err-b
err-t
Content quality
err-r
err-lod
err-swbd
err-ic
err-pifs
err-stm
err-tfs
err-rt
err-enj
err-rot
err-ttai
err-rep
err-etun
Usability
err-af
err-l
err-n
err-etu
err-attr
err-da
err-com
err-pe
Service interaction quality
Transaction quality
err-p
err-acc
err-cau
err-cwo

Variable counts (Group number 1)

Number of variables in your model: 58
 Number of observed variables: 27
 Number of unobserved variables: 31
 Number of exogenous variables: 31
 Number of endogenous variables: 27

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	31	0	0	0	0	31
Labeled	0	0	0	0	0	0
Unlabeled	23	6	31	0	0	60
Total	54	6	31	0	0	91

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 378
 Number of distinct parameters to be estimated: 60
 Degrees of freedom (378 - 60): 318

Result (Default model)

Minimum was achieved
 Chi-square = 834.062
 Degrees of freedom = 318
 Probability level = .000

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

	Estimate	S.E.	C.R.	P	Label
x10 <--- Content quality	1.000				
x14 <--- Content quality	.898	.087	10.325	***	
x26 <--- Content quality	.648	.095	6.799	***	
x23 <--- Content quality	.869	.108	8.039	***	
x17 <--- Content quality	.584	.089	6.558	***	
x5 <--- Usability	.934	.085	10.992	***	
x1 <--- Usability	.778	.084	9.227	***	
x21 <--- Service interaction quality	1.771	.370	4.790	***	
x24 <--- Service interaction quality	1.801	.357	5.049	***	
x19 <--- Service interaction quality	2.092	.383	5.468	***	
x7 <--- Content quality	.802	.082	9.747	***	
x8 <--- Transaction quality	1.062	.179	5.936	***	
x4 <--- Transaction quality	1.000				
x12 <--- Transaction quality	1.059	.159	6.648	***	
x2 <--- Content quality	.796	.086	9.235	***	
x13 <--- Usability	.901	.091	9.945	***	
x9 <--- Usability	1.000				
x16 <--- Usability	.478	.091	5.251	***	
x25 <--- Usability	.666	.089	7.448	***	
x18 <--- Usability	.692	.076	9.089	***	
x22 <--- Usability	.809	.081	9.962	***	
x20 <--- Usability	.769	.075	10.276	***	
x15 <--- Service interaction quality	1.996	.409	4.884	***	
x3 <--- Service interaction quality	1.751	.331	5.282	***	
x11 <--- Service interaction quality	1.546	.324	4.777	***	
x6 <--- Service interaction quality	1.000				
x27 <--- Service interaction quality	1.832	.350	5.231	***	

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
x10 <--- Content quality	.761
x14 <--- Content quality	.828
x26 <--- Content quality	.568
x23 <--- Content quality	.662
x17 <--- Content quality	.549
x5 <--- Usability	.831
x1 <--- Usability	.720
x21 <--- Service interaction quality	.569
x24 <--- Service interaction quality	.633

x19 <--- Service interaction quality	.767
x7 <--- Content quality	.786
x8 <--- Transaction quality	.573
x4 <--- Transaction quality	.600
x12 <--- Transaction quality	.664
x2 <--- Content quality	.750
x13 <--- Usability	.766
x9 <--- Usability	.775
x16 <--- Usability	.435
x25 <--- Usability	.599
x18 <--- Usability	.711
x22 <--- Usability	.768
x20 <--- Usability	.787
x15 <--- Service interaction quality	.591
x3 <--- Service interaction quality	.702
x11 <--- Service interaction quality	.566
x6 <--- Service interaction quality	.454
x27 <--- Service interaction quality	.686

Covariances: (Group number 1 - Default model)

		Estimate	S.E.	C.R.	P	Label
Usability	<--> Transaction quality	.864	.159	5.420	***	
Content quality	<--> Service interaction quality	.421	.096	4.389	***	
Content quality	<--> Transaction quality	.728	.138	5.271	***	
Content quality	<--> Usability	.965	.166	5.819	***	
Usability	<--> Service interaction quality	.561	.122	4.582	***	
Service interaction quality	<--> Transaction quality	.389	.092	4.242	***	

Correlations: (Group number 1 - Default model)

		Estimate
Usability	<--> Transaction quality	.914
Content quality	<--> Service interaction quality	.822
Content quality	<--> Transaction quality	.881
Content quality	<--> Usability	.782
Usability	<--> Service interaction quality	.957
Service interaction quality	<--> Transaction quality	.992

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Content quality	1.078	.205	5.257	***	
Usability	1.415	.258	5.481	***	
Service interaction quality	.243	.086	2.809	.005	
Transaction quality	.633	.169	3.747	***	
err-b	.428	.060	7.136	***	
err-t	.785	.107	7.352	***	
err-r	.400	.060	6.651	***	
err-lod	1.041	.132	7.870	***	
err-swbd	.918	.115	7.983	***	
err-ic	.746	.098	7.599	***	
err-pifs	1.230	.149	8.257	***	
err-stm	1.463	.185	7.923	***	
err-tfs	.935	.111	8.389	***	
err-rt	1.121	.136	8.232	***	
err-enj	1.589	.193	8.252	***	
err-rot	.899	.123	7.288	***	
err-ttai	1.123	.144	7.778	***	
err-rep	.767	.097	7.927	***	
err-etun	.850	.104	8.173	***	
err-af	.951	.117	8.134	***	
err-l	.794	.100	7.939	***	
err-n	.939	.122	7.696	***	
err-etu	.806	.104	7.743	***	
err-attr	1.387	.165	8.414	***	
err-da	.660	.083	7.970	***	
err-com	.513	.067	7.627	***	
err-pe	.645	.083	7.738	***	
err-p	1.801	.219	8.215	***	
err-acc	.532	.072	7.431	***	
err-cau	.554	.076	7.286	***	
err-cwo	1.177	.145	8.130	***	

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
x2	.562
x15	.350
x24	.401
x22	.589
x20	.620

x18	.506
x16	.189
x13	.587
x9	.601
x1	.519
x5	.690
x26	.323
x17	.302
x3	.493
x4	.360
x12	.441
x21	.324
x25	.359
x6	.206
x8	.328
x11	.320
x19	.588
x27	.470
x23	.439
x14	.685
x10	.579
x7	.618

Matrices (Group number 1 - Default model)

Residual Covariances (Group number 1 - Default model)

	x2	x15	x24	x22	x20	x18	x16	x13	x9	x1	x5	x26	x17	x3	x4	x12	x21	x25	x6	x8
x2	.000																			
x15	-.214	.000																		
x24	-.041	-.064	.000																	
x22	-.047	.135	.143	.000																
x20	.156	-.151	-.037	.054	.000															
x18	.087	-.179	-.014	-.012	.154	.000														
x16	.060	.062	-.039	-.002	.116	.059	.000													
x13	-.241	.081	-.300	-.136	-.078	-.029	-.064	.000												
x9	-.221	.193	-.160	-.139	-.165	-.161	-.046	.476	.000											
x1	-.070	.055	-.223	-.099	-.099	-.105	-.044	.230	.214	.000										
x5	-.062	.100	-.172	-.070	-.068	-.049	.009	.074	.132	.118	.000									
x26	-.057	.160	.174	.152	.278	.087	.086	-.033	-.149	-.071	.059	.000								
x17	-.044	-.107	.168	.038	.197	.067	.301	.113	-.028	-.034	.078	.017	.000							
x3	-.104	.022	-.015	.103	-.004	-.025	-.009	.010	.142	.060	-.004	-.063	-.095	.000						
x4	-.017	-.148	-.128	.054	.039	-.121	-.013	.041	.141	.254	.256	.115	-.088	.117	.000					
x12	.083	-.004	.112	.092	.133	.128	-.127	-.164	-.347	-.218	-.082	.210	.008	-.253	-.142	.000				
x21	-.143	.507	.093	.320	.008	-.013	.264	-.033	-.017	-.134	-.111	-.071	-.027	.159	-.236	-.015	.000			
x25	.164	-.192	.230	.092	.083	.173	.025	-.208	-.087	-.192	-.153	.239	.033	.103	-.109	.189	-.036	.000		
x6	.035	.030	.063	.011	-.029	-.037	-.120	-.203	-.156	-.091	.011	.292	-.160	.017	.088	.069	-.016	.025	.000	
x8	-.087	.110	.223	.153	-.002	-.165	-.338	-.226	-.198	.144	.015	-.084	-.121	-.103	-.012	.183	-.150	.052	.125	.000
x11	.081	-.120	.187	-.088	.020	.118	-.054	-.361	-.410	-.233	-.190	.220	.000	.035	-.040	.307	-.249	.088	.527	.167
x19	-.170	.047	-.045	.128	.019	.047	-.103	.130	.201	.187	.143	-.032	.101	.031	.075	-.228	.140	.038	-.196	.018
x27	.124	-.139	.123	.093	.135	.004	.022	-.149	-.276	-.131	-.084	.328	.077	-.045	.036	.163	-.088	.224	-.073	.019
x23	-.073	-.205	.211	.279	.129	.122	-.078	-.225	-.218	-.100	.021	.147	.008	-.127	.052	.293	.075	.221	.043	.094
x14	.019	-.103	.096	-.079	.118	.095	.002	-.056	-.192	-.079	.033	-.023	.047	-.123	-.075	.025	-.142	-.046	.002	-.075
x10	.032	-.117	.120	.091	-.006	.080	-.008	-.047	.065	-.154	.079	-.084	.029	.017	-.026	-.035	-.178	.070	.050	-.262
x7	.062	-.054	-.028	-.035	.047	.086	-.045	-.199	-.202	-.159	.044	-.062	-.067	.012	-.020	.005	-.234	.072	.087	.039

Standardized Residual Covariances (Group number 1 - Default model)

	x2	x15	x24	x22	x20	x18	x16	x13	x9	x1	x5	x26	x17	x3	x4	x12	x21	x25
x2	.000																	
x15	-1.327	.000																
x24	-.296	-.311	.000															
x22	-.379	.715	.895	.000														
x20	1.334	-.861	-.251	.383	.000													
x18	.763	-1.041	-.093	-.091	1.210	.000												
x16	.491	.336	-.251	-.014	.875	.447	.000											
x13	-1.726	.385	-1.677	-.811	-.499	-.192	-.402	.000										
x9	-1.438	.838	-.815	-.752	-.952	-.963	-.261	2.305	.000									
x1	-.551	.288	-1.369	-.653	-.698	-.757	-.302	1.354	1.146	.000								
x5	-.457	.490	-.991	-.427	-.441	-.330	.057	.402	.654	.709	.000							
x26	-.481	.946	1.213	1.173	2.302	.735	.661	-.230	-.936	-.534	.420	.000						
x17	-.400	-.679	1.262	.312	1.761	.606	2.475	.838	-.192	-.276	.598	.148	.000					
x3	-.849	.122	-.097	.720	-.028	-.195	-.063	.063	.805	.411	-.029	-.495	-.805	.000				
x4	-.129	-.763	-.780	.364	.282	-.887	-.088	.244	.774	1.678	1.591	.849	-.696	.800	.000			
x12	.658	-.020	.701	.634	.984	.970	-.891	-1.011	-1.948	-1.481	-.520	1.602	.069	-1.781	-.949	.000		
x21	-.964	2.276	.494	1.856	.051	-.083	1.548	-.174	-.081	-.763	-.598	-.453	-.186	.947	-1.328	-.087	.000	
x25	1.280	-.998	1.411	.612	.594	1.256	.167	-1.233	-.468	-1.255	-.936	1.785	.264	.712	-.714	1.281	-.201	.000
x6	.342	.196	.485	.089	-.265	-.345	-1.006	-1.533	-1.075	-.751	.089	2.687	-1.579	.146	.716	.582	-.113	.207
x8	-.604	.513	1.224	.931	-.015	-1.094	-2.064	-1.228	-.981	.860	.086	-.561	-.866	-.634	-.071	1.107	-.760	.310
x11	.626	-.612	1.125	-.580	.143	.852	-.360	-2.142	-2.213	-1.518	-1.165	1.611	.000	.239	-.259	2.038	-1.386	.569
x19	-1.255	.229	-.262	.802	.124	.322	-.672	.731	1.021	1.154	.822	-.226	.777	.198	.466	-1.441	.754	.235
x27	.954	-.708	.742	.609	.945	.026	.146	-.872	-1.474	-.848	-.508	2.417	.615	-.302	.230	1.073	-.493	1.449

x23	-.529	-1.039	1.261	1.837	.916	.876	-.518	-1.327	-1.169	-.648	.127	1.030	.057	-.857	.327	1.913	.415	1.414
x14	.159	-.619	.673	-.603	.969	.800	.016	-.383	-1.199	-.601	.235	-.191	.418	-.965	-.559	.188	-.928	-.350
x10	.226	-.584	.705	.584	-.042	.562	-.052	-.272	.338	-.972	.467	-.573	.215	.112	-.163	-.220	-.966	.443
x7	.555	-.346	-.212	-.289	.414	.775	-.380	-1.467	-1.358	-1.294	.333	-.548	-.639	.105	-.159	.038	-1.634	.581

Notes for Model (Group number 1 - Default model)

The following covariance matrix is not positive definite (Group number 1 - Default model)

	Transaction quality	Service interaction quality	Usability	Content quality
Transaction quality	.633			
Service interaction quality	.389	.243		
Usability	.864	.561	1.415	
Content quality	.728	.421	.965	1.078

Notes for Group/Model (Group number 1 - Default model)

This solution is not admissible.

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

	M.I.	Par Change
err-cwo <--> Usability	7.368	-.155
err-pe <--> Service interaction quality	12.843	.064
err-pe <--> Usability	14.671	-.164
err-pe <--> err-cwo	4.585	.165
err-com <--> Content quality	8.509	.130
err-com <--> err-acc	7.331	.131
err-com <--> err-p	5.870	-.207
err-da <--> Content quality	6.419	.126
err-da <--> err-p	4.892	-.210
err-da <--> err-com	12.036	.181
err-attr <--> Transaction quality	5.489	-.145
err-etu <--> Usability	11.577	.163
err-etu <--> err-acc	5.460	-.141
err-etu <--> err-cwo	7.992	-.244
err-etu <--> err-pe	6.216	-.162
err-n <--> Transaction quality	4.306	-.110
err-n <--> Usability	6.120	.128
err-n <--> Content quality	6.319	-.151
err-n <--> err-p	4.837	.253
err-n <--> err-pe	5.585	-.166
err-n <--> err-com	9.918	-.199
err-n <--> err-da	7.161	-.188
err-n <--> err-etu	52.177	.569
err-l <--> Content quality	6.093	-.135
err-l <--> err-cwo	5.066	-.190
err-l <--> err-com	4.154	-.117
err-l <--> err-etu	14.033	.268
err-l <--> err-n	10.516	.251
err-cau <--> err-cwo	4.385	-.154
err-cau <--> err-n	6.146	.166
err-cau <--> err-l	5.616	.144
err-af <--> Content quality	4.122	-.119
err-af <--> err-com	6.847	.163
err-af <--> err-n	4.054	-.169
err-etun <--> err-attr	9.203	.280
err-rep <--> Transaction quality	4.030	-.094
err-rep <--> err-n	5.348	.176
err-ttai <--> Usability	4.830	.124
err-ttai <--> err-da	7.154	-.201
err-ttai <--> err-l	9.920	.260
err-ttai <--> err-cau	10.181	.229
err-rot <--> err-com	5.621	.145
err-rot <--> err-da	5.110	.153
err-rot <--> err-n	12.494	-.290
err-rot <--> err-l	5.962	-.182
err-rot <--> err-rep	11.808	-.251
err-enj <--> Service interaction quality	5.506	.063
err-enj <--> Content quality	4.296	-.157
err-enj <--> err-p	13.992	.541
err-enj <--> err-pe	11.825	.306
err-enj <--> err-attr	5.287	.289
err-enj <--> err-ttai	4.802	-.251
err-rt <--> Usability	6.089	-.138
err-rt <--> err-acc	4.602	.148
err-rt <--> err-p	4.333	-.253
err-rt <--> err-cwo	6.384	.250
err-rt <--> err-da	6.484	.190
err-rt <--> err-etu	7.859	-.235
err-rt <--> err-l	6.694	-.213

err-rt	<-->	err-cau	6.553	-.182
err-rt	<-->	err-ttai	5.003	-.216
err-rt	<-->	err-rot	4.132	.177
err-tfs	<-->	Usability	4.163	-.103
err-tfs	<-->	err-etu	4.563	-.162
err-tfs	<-->	err-af	8.677	.237
err-tfs	<-->	err-etun	7.572	-.209
err-stm	<-->	err-da	4.552	-.182
err-stm	<-->	err-attr	7.279	-.326
err-stm	<-->	err-l	5.206	.214
err-stm	<-->	err-rot	4.094	.201
err-pifs	<-->	Transaction quality	5.046	.131
err-pifs	<-->	Usability	18.581	-.251
err-pifs	<-->	Content quality	7.091	.177
err-pifs	<-->	err-da	6.106	.193
err-pifs	<-->	err-etu	10.119	-.279
err-pifs	<-->	err-n	12.161	-.330
err-pifs	<-->	err-rot	10.908	.300
err-pifs	<-->	err-enj	4.904	-.264
err-pifs	<-->	err-tfs	36.890	.552
err-ic	<-->	Service interaction quality	5.006	-.040
err-ic	<-->	Usability	20.192	.206
err-ic	<-->	Content quality	4.506	-.113
err-ic	<-->	err-etu	4.205	.144
err-ic	<-->	err-n	8.008	.215
err-ic	<-->	err-l	6.635	.178
err-ic	<-->	err-rot	8.138	-.208
err-ic	<-->	err-tfs	8.970	-.219
err-ic	<-->	err-pifs	5.284	-.194
err-swbd	<-->	Service interaction quality	4.295	-.042
err-swbd	<-->	Content quality	11.497	.197
err-swbd	<-->	err-com	6.364	.155
err-swbd	<-->	err-n	9.454	-.255
err-swbd	<-->	err-af	4.871	.180
err-swbd	<-->	err-rot	4.552	.170
err-swbd	<-->	err-rt	6.497	.224
err-swbd	<-->	err-ic	11.050	-.245
err-lod	<-->	Transaction quality	6.140	.136
err-lod	<-->	err-pe	11.948	.255
err-lod	<-->	err-etu	5.789	-.198
err-lod	<-->	err-rot	7.193	.229
err-r	<-->	err-pe	6.736	-.128
err-r	<-->	err-rt	4.978	-.141
err-t	<-->	Transaction quality	5.016	-.110
err-t	<-->	err-com	6.966	-.156
err-t	<-->	err-n	9.136	.241
err-t	<-->	err-stm	6.195	-.240
err-b	<-->	err-rep	4.037	.107
err-b	<-->	err-ic	5.387	-.123
err-b	<-->	err-lod	5.582	-.146

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
x15 <--- x21	9.180	.223
x24 <--- x13	4.441	-.138
x22 <--- x24	4.294	.103
x22 <--- x21	10.173	.145
x22 <--- x23	7.766	.142
x20 <--- x2	9.059	.171
x20 <--- x18	5.614	.128
x20 <--- x26	9.679	.164
x20 <--- x17	5.726	.136
x20 <--- x12	4.990	.110
x20 <--- x27	4.120	.096
x20 <--- x14	4.113	.112
x18 <--- x20	4.154	.122
x18 <--- x25	4.026	.105
x18 <--- x11	4.132	.105
x16 <--- x17	6.418	.225
x16 <--- x8	5.242	-.152
x13 <--- x2	6.858	-.185
x13 <--- x24	6.529	-.142
x13 <--- x9	19.083	.221
x13 <--- x1	6.347	.152
x13 <--- x25	4.885	-.130
x13 <--- x6	4.771	-.157
x13 <--- x11	8.721	-.171

x13 <--- x23	7.657	-.158
x13 <--- x7	4.646	-.158
x9 <--- x2	4.608	-.164
x9 <--- x13	19.843	.268
x9 <--- x1	4.758	.143
x9 <--- x26	7.056	-.189
x9 <--- x12	10.378	-.214
x9 <--- x11	9.839	-.196
x9 <--- x27	6.471	-.163
x9 <--- x23	6.484	-.157
x9 <--- x14	4.267	-.154
x1 <--- x13	5.321	.126
x1 <--- x4	4.666	.124
x1 <--- x12	4.469	-.127
x1 <--- x25	4.157	-.118
x5 <--- x4	6.672	.129
x5 <--- x25	4.082	-.101
x26 <--- x20	5.285	.163
x26 <--- x6	9.543	.235
x26 <--- x27	5.267	.144
x17 <--- x16	9.129	.180
x17 <--- x6	5.177	-.163
x3 <--- x12	7.173	-.158
x4 <--- x1	6.588	.179
x4 <--- x5	4.740	.146
x12 <--- x9	5.869	-.128
x12 <--- x3	6.415	-.166
x12 <--- x11	6.537	.154
x12 <--- x23	5.469	.139
x21 <--- x15	8.797	.189
x21 <--- x22	4.749	.184
x21 <--- x16	4.398	.170
x25 <--- x2	4.024	.162
x25 <--- x24	4.781	.139
x25 <--- x27	4.400	.142
x6 <--- x26	7.104	.182
x6 <--- x11	24.326	.296
x8 <--- x16	6.245	-.195
x11 <--- x13	6.641	-.172
x11 <--- x9	7.457	-.166
x11 <--- x12	7.803	.205
x11 <--- x6	28.811	.461
x11 <--- x7	4.141	.179
x19 <--- x9	5.400	.114
x19 <--- x1	5.302	.134
x19 <--- x12	5.917	-.144
x19 <--- x6	7.014	-.183
x19 <--- x27	5.553	-.134
x19 <--- x7	4.877	-.156
x27 <--- x9	4.102	-.108
x27 <--- x26	7.587	.190
x27 <--- x19	4.165	-.124
x27 <--- x23	4.526	.127
x23 <--- x22	6.206	.174
x23 <--- x12	5.946	.168
x10 <--- x8	4.166	-.108

Minimization History (Default model)

Iteration		Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	12		-1.141	9999.000	2652.249	0	9999.000
1	e*	9		-.591	4.944	1548.225	20	.179
2	e*	4		-.205	.781	1287.842	5	.839
3	e	2		-.228	.733	1083.112	5	.851
4	e	1		-.055	.658	944.882	5	.896
5	e	0	47479.864		.651	868.781	5	.884
6	e	0	3182.652		.825	853.435	6	.000
7	e	0	2493.955		.487	842.896	2	.000
8	e	0	4045.586		.425	836.420	1	1.146
9	e	0	5525.867		.387	834.758	1	1.017
10	e	0	12246.494		.174	834.139	1	1.103
11	e	0	14990.247		.146	834.068	1	1.029
12	e	0	17752.752		.022	834.062	1	1.017
13	e	0	17623.246		.002	834.062	1	1.002
14	e	0	17795.778		.000	834.062	1	.999

Bootstrap (Default model)

Summary of Bootstrap Iterations (Default model)

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	3	0
11	0	0	0
12	0	3	0
13	0	9	0
14	0	16	0
15	0	9	0
16	0	11	0
17	0	8	0
18	0	10	0
19	0	71	0
Total	0	140	0

0 bootstrap samples were unused because of a singular covariance matrix.

0 bootstrap samples were unused because a solution was not found.

140 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 140 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 0 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .007

Bootstrap Distributions (Default model)

ML discrepancy (implied vs sample) (Default model)

	222.150	-----
	257.374	
	292.599	***
	327.823	****
	363.048	*****
	398.273	*****
	433.497	*****
N = 140	468.722	*****
Mean = 437.087	503.947	*****
S. e. = 6.807	539.171	*****
	574.396	*
	609.620	***
	644.845	*
	680.070	*
	715.294	*

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	60	834.062	318	.000	2.623
Saturated model	378	.000	0		
Independence model	27	2722.078	351	.000	7.755

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.137	.683	.623	.574
Saturated model	.000	1.000		
Independence model	.699	.173	.110	.161

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.694	.662	.785	.760	.782
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.906	.628	.709
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	516.062	434.386	605.401
Saturated model	.000	.000	.000
Independence model	2371.078	2208.710	2540.842

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	5.713	3.535	2.975	4.147
Saturated model	.000	.000	.000	.000
Independence model	18.644	16.240	15.128	17.403

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.105	.097	.114	.000
Independence model	.215	.208	.223	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	954.062	982.537	1133.488	1193.488
Saturated model	756.000	935.390	1886.384	2264.384
Independence model	2776.078	2788.891	2856.819	2883.819

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	6.535	5.975	7.147	6.730
Saturated model	5.178	5.178	5.178	6.407
Independence model	19.014	17.902	20.177	19.102

HOELTER

Model	HOELTER	HOELTER
	.05	.01
Default model	64	67
Independence model	22	23

Execution time summary

Minimization: .031
 Miscellaneous: .328
 Bootstrap: 1.077
 Total: 1.436

\\staff\home\fca\tatema\My Documents\PhD\2008\SEM\measurement model based on EFA.amw

Analysis Summary

Date and Time

Date: Thursday, 12 November 2009

Time: 3:49:41 p.m.

Title

measurement model based on efa: Thursday, 12 November 2009 03:49 p.m.

Notes

Bootstrap confidence intervals are not available when the Bollen-Stine bootstrap is performed.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x1
x4
x5
x9
x13
x19
x8
x6
x25
x24
x22
x21
x12
x11
x3
x2
x7
x26
x27
x18
x16
x10
x14
x23
x20

Unobserved, exogenous variables

err-etl
Usability
err-fic
Transaction quality
er-stm
err-fs
Responsiveness
err-rt
err-etc
err-pe
err-e
err-rot
err-atcut
err-cau
err-etn
err-pifs
Content quality
err-b
err-af
err-cid
err-datos
err-aa
err-a
err-r
err-rlod
err-c
err-t
err-rep
err-etu

Variable counts (Group number 1)

Number of variables in your model: 54

Number of observed variables: 25
 Number of unobserved variables: 29
 Number of exogenous variables: 29
 Number of endogenous variables: 25

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	29	0	0	0	0	29
Labeled	0	0	0	0	0	0
Unlabeled	21	6	29	0	0	56
Total	50	6	29	0	0	85

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 325
 Number of distinct parameters to be estimated: 56
 Degrees of freedom (325 - 56): 269

Result (Default model)

Minimum was achieved
 Chi-square = 604.412
 Degrees of freedom = 269
 Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
x4 <--- Usability	.639	.075	8.533	***	
x5 <--- Usability	.863	.065	13.234	***	
x9 <--- Usability	1.000				
x13 <--- Usability	.883	.069	12.731	***	
x8 <--- Transaction quality	1.115	.197	5.648	***	
x25 <--- Responsiveness	1.002	.162	6.182	***	
x22 <--- Responsiveness	1.153	.166	6.966	***	
x21 <--- Responsiveness	1.000				
x12 <--- Responsiveness	1.030	.159	6.466	***	
x1 <--- Usability	.767	.066	11.603	***	
x11 <--- Transaction quality	1.524	.204	7.478	***	
x3 <--- Usability	.641	.068	9.495	***	
x19 <--- Usability	.817	.068	11.948	***	
x6 <--- Transaction quality	1.000				
x2 <--- Content quality	.901	.096	9.341	***	
x7 <--- Content quality	.879	.092	9.532	***	
x16 <--- Transaction quality	.532	.167	3.177	.001	
x10 <--- Content quality	1.107	.120	9.261	***	
x14 <--- Content quality	1.000				
x23 <--- Content quality	1.035	.121	8.562	***	
x26 <--- Content quality	.818	.107	7.653	***	
x18 <--- Content quality	.936	.101	9.240	***	
x20 <--- Content quality	1.051	.099	10.620	***	
x27 <--- Content quality	1.101	.114	9.612	***	
x24 <--- Responsiveness	1.102	.174	6.335	***	

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
x4 <--- Usability	.635
x5 <--- Usability	.850
x9 <--- Usability	.858
x13 <--- Usability	.831
x8 <--- Transaction quality	.548
x25 <--- Responsiveness	.661
x22 <--- Responsiveness	.803
x21 <--- Responsiveness	.569
x12 <--- Responsiveness	.708
x1 <--- Usability	.785
x11 <--- Transaction quality	.821
x3 <--- Usability	.686
x19 <--- Usability	.800

x6 <--- Transaction quality	.668
x2 <--- Content quality	.723
x7 <--- Content quality	.735
x16 <--- Transaction quality	.295
x10 <--- Content quality	.718
x14 <--- Content quality	.786
x23 <--- Content quality	.673
x26 <--- Content quality	.611
x18 <--- Content quality	.717
x20 <--- Content quality	.801
x27 <--- Content quality	.740
x24 <--- Responsiveness	.686

Covariances: (Group number 1 - Default model)

		Estimate	S.E.	C.R.	P	Label
Usability	<--> Content quality	.855	.143	5.981	***	
Usability	<--> Transaction quality	.505	.116	4.364	***	
Transaction quality	<--> Content quality	.487	.094	5.191	***	
Usability	<--> Responsiveness	.871	.172	5.062	***	
Responsiveness	<--> Content quality	.689	.132	5.218	***	
Transaction quality	<--> Responsiveness	.478	.106	4.492	***	

Correlations: (Group number 1 - Default model)

	Estimate
Usability <--> Content quality	.734
Usability <--> Transaction quality	.530
Transaction quality <--> Content quality	.758
Usability <--> Responsiveness	.759
Responsiveness <--> Content quality	.892
Transaction quality <--> Responsiveness	.755

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Usability	1.731	.271	6.378	***	
Transaction quality	.526	.127	4.147	***	
Responsiveness	.761	.213	3.578	***	
Content quality	.784	.141	5.570	***	
err-etl	.633	.084	7.535	***	
err-fic	.653	.088	7.430	***	
er-stm	1.523	.197	7.739	***	
err-fs	.652	.093	6.980	***	
err-rt	.984	.127	7.769	***	
err-etc	1.040	.136	7.651	***	
err-pe	.557	.084	6.639	***	
err-e	1.589	.197	8.071	***	
err-rot	.801	.106	7.527	***	
err-atcut	1.048	.129	8.123	***	
err-cau	.497	.072	6.906	***	
err-etn	.622	.092	6.787	***	
err-pifs	.589	.131	4.498	***	
err-b	.515	.066	7.792	***	
err-af	.879	.108	8.164	***	
err-cid	.783	.101	7.769	***	
err-datos	.650	.083	7.869	***	
err-aa	1.562	.187	8.370	***	
err-a	.580	.074	7.844	***	
err-r	.486	.065	7.511	***	
err-rlod	1.015	.127	8.017	***	
err-c	.483	.065	7.394	***	
err-t	.902	.115	7.864	***	
err-rep	.799	.100	7.988	***	
err-etu	.606	.085	7.141	***	

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
x20	.642
x23	.453
x14	.617
x10	.516
x16	.087
x18	.514
x27	.548
x26	.374
x7	.541
x2	.523
x3	.471
x11	.675
x12	.502

x21	.324
x22	.645
x24	.471
x25	.437
x6	.447
x8	.300
x19	.639
x13	.690
x9	.736
x5	.722
x4	.403
x1	.616

Matrices (Group number 1 - Default model)

Residual Covariances (Group number 1 - Default model)

	x20	x23	x14	x10	x16	x18	x27	x26	x7	x2	x3	x11	x12	x21	x22	x24	x25	x6	x8	x19
x20	.000																			
x23	-.079	.000																		
x14	-.040	.021	.000																	
x10	-.176	.064	.053	.000																
x16	.364	.055	.158	.167	.000															
x18	.135	-.058	-.039	-.065	.284	.000														
x27	.017	.058	-.015	.013	.228	-.094	.000													
x26	.085	.090	-.037	-.095	.174	-.080	.121	.000												
x7	-.083	-.079	.120	.162	.098	-.024	-.085	-.066	.000											
x2	.004	-.058	.084	.109	.195	-.042	-.039	-.078	.130	.000										
x3	.174	-.055	-.010	.146	.288	.140	.130	-.034	.121	-.012	.000									
x11	-.093	.032	.031	-.038	-.066	.023	.022	.035	.043	-.069	.199	.000								
x12	.090	.228	.007	-.050	.049	.097	.136	.129	-.001	.057	-.107	.194	.000							
x21	.047	.009	-.163	-.196	.485	.028	-.059	-.152	-.243	-.171	.353	-.312	-.070	.000						
x22	.097	.135	-.173	-.009	.252	.034	.049	.008	-.108	-.142	.253	-.227	-.072	.244	.000					
x24	-.060	.083	.016	.037	.164	-.026	.089	.043	-.089	-.122	.135	.060	-.011	.029	-.008	.000				
x25	.082	.064	-.159	-.051	.221	.178	.148	.091	-.020	.054	.197	-.064	.013	-.137	-.026	.062	.000			
x6	-.110	-.096	-.107	-.068	-.132	-.105	-.164	.167	-.004	-.068	.118	.101	-.011	-.064	-.087	-.026	-.080	.000		
x8	.133	.204	.077	-.089	-.211	-.037	.179	-.027	.182	.040	.260	-.088	.347	.050	.282	.380	.130	-.048	.000	
x19	.186	.138	.025	-.034	.239	.204	-.050	-.033	-.097	-.099	.013	-.015	-.099	.328	.256	.085	.106	-.101	.422	.000
x13	.108	-.251	-.030	-.013	.308	.145	-.054	-.087	-.165	-.229	-.085	-.260	-.131	.092	.007	-.237	-.129	-.144	.104	-.061
x9	.024	-.264	-.180	.083	.362	.016	-.190	-.223	-.180	-.223	.013	-.313	-.328	.105	.000	-.110	-.018	-.101	.157	-.041
x5	.171	.040	.104	.162	.408	.173	.062	.039	.117	-.009	-.046	-.046	-.001	.064	.130	-.058	-.027	.099	.386	.017
x4	.129	.118	.032	.097	.229	-.034	.147	.140	.083	.070	.089	.069	-.045	-.104	.111	-.041	-.091	.154	.300	-.015
x1	.058	-.126	-.060	-.128	.276	.043	-.054	-.120	-.134	-.063	-.027	-.149	-.193	-.029	.020	-.173	-.128	-.042	.427	.015

Standardized Residual Covariances (Group number 1 - Default model)

	x20	x23	x14	x10	x16	x18	x27	x26	x7	x2	x3	x11	x12	x21	x22	x24	x25	x6
x20	.000																	
x23	-.529	.000																
x14	-.310	.147	.000															
x10	-1.164	.376	.366	.000														
x16	2.851	.369	1.275	1.118	.000													
x18	1.051	-.399	-.317	-.442	2.241	.000												
x27	.118	.351	-.102	.076	1.580	-.656	.000											
x26	.668	.626	-.300	-.647	1.343	-.647	.856	.000										
x7	-.699	-.594	1.049	1.197	.840	-.210	-.645	-.577	.000									
x2	.036	-.422	.712	.779	1.612	-.354	-.288	-.657	1.186	.000								
x3	1.369	-.377	-.080	.989	2.155	1.120	.912	-.271	1.049	-.097	.000							
x11	-.644	.197	.225	-.229	-.438	.162	.137	.246	.332	-.512	1.390	.000						
x12	.661	1.470	.054	-.315	.356	.726	.892	.967	-.011	.452	-.780	1.258	.000					
x21	.294	.047	-1.058	-1.063	2.897	.182	-.331	-.964	-1.693	-1.147	2.172	-1.723	-.402	.000				
x22	.701	.858	-1.288	-.055	1.828	.254	.315	.059	-.872	-1.103	1.829	-1.454	-.477	1.399	.000			
x24	-.397	.484	.114	.212	1.067	-.179	.530	.296	-.658	-.870	.892	.356	-.065	.151	-.046	.000		
x25	.580	.402	-1.173	-.317	1.525	1.294	.943	.662	-.158	.408	1.387	-.403	.086	-.765	-.166	.369	.000	
x6	-.974	-.739	-.984	-.523	-1.100	-.952	-1.298	1.498	-.036	-.646	1.039	.730	-.090	-.445	-.717	-.192	-.638	.000
x8	.892	1.183	.529	-.514	-1.302	-.254	1.062	-1.180	1.345	.286	1.697	-.487	2.148	.257	1.747	2.135	.778	-.339
x19	1.299	.844	.183	-.205	1.627	1.459	-.312	-.236	-.754	-.745	.084	-.097	-.646	1.818	1.650	.504	.667	-.802
x13	.719	-1.472	-.205	-.077	2.018	.995	-.323	-.596	-1.230	-1.641	-.521	-1.570	-.814	.490	.044	-1.343	-.780	-1.098
x9	.143	-1.407	-1.129	.436	2.160	.100	-1.033	-1.383	-1.217	-1.449	.073	-1.719	-1.852	.504	-.001	-.565	-.096	-.700
x5	1.192	.242	.751	.982	2.797	1.235	.387	.277	.910	-.071	-.293	-.289	-.010	.352	.835	-.341	-.170	.787
x4	.949	.757	.241	.612	1.587	-.258	.962	1.035	.677	.552	.602	.449	-.309	-.595	.755	-.254	-.598	1.265
x1	.424	-.812	-.460	-.816	1.972	.322	-.352	-.899	-1.092	-.493	-.184	-.986	-1.321	-.166	.138	-1.072	-.849	-.350

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

			M.I.	Par Change
err-c	<-->	Content quality	4.081	-.063
err-c	<-->	Usability	5.203	.134
err-rlod	<-->	Responsiveness	8.558	.140
err-r	<-->	Content quality	4.374	.065
err-r	<-->	Responsiveness	9.966	-.107

err-t	<-->	err-c	12.679	-.215
err-aa	<-->	Transaction quality	9.968	-.203
err-aa	<-->	err-c	7.089	.206
err-datos	<-->	err-c	10.285	.164
err-cid	<-->	Responsiveness	4.613	.091
err-b	<-->	Responsiveness	6.219	-.086
err-b	<-->	err-c	4.918	-.101
err-b	<-->	err-r	10.106	.145
err-b	<-->	err-t	9.604	.189
err-a	<-->	Content quality	6.031	.082
err-a	<-->	err-r	4.427	.101
err-a	<-->	err-b	9.628	.152
err-pifs	<-->	err-rep	4.672	.159
err-rot	<-->	Content quality	4.653	.086
err-rot	<-->	Transaction quality	4.574	.103
err-rot	<-->	Usability	7.390	-.203
err-rot	<-->	err-aa	4.018	-.198
err-rot	<-->	err-rep	4.536	-.154
err-rot	<-->	err-pifs	7.695	.208
err-e	<-->	Content quality	4.745	-.120
err-e	<-->	err-aa	7.669	.373
err-e	<-->	err-b	4.416	-.168
err-e	<-->	err-pifs	5.783	-.247
err-pe	<-->	Usability	5.341	.150
err-pe	<-->	err-c	4.159	.103
err-pe	<-->	err-r	8.574	-.148
err-pe	<-->	err-a	4.175	-.111
err-pe	<-->	err-pifs	10.064	-.208
err-pe	<-->	err-e	12.580	.312
err-rt	<-->	err-r	6.916	-.166
err-rt	<-->	err-datos	4.079	.145
err-fs	<-->	err-cid	4.147	-.136
err-fs	<-->	err-af	9.596	.214
err-fs	<-->	err-pifs	6.432	.168
er-stm	<-->	Responsiveness	7.795	.165
er-stm	<-->	Transaction quality	5.216	-.146
er-stm	<-->	err-pe	4.204	.180
err-fic	<-->	Responsiveness	8.757	.117
err-fic	<-->	err-rlod	6.369	.187
err-fic	<-->	err-e	4.434	.194
err-fic	<-->	err-pe	4.688	.128
err-fic	<-->	err-fs	5.447	-.145
err-etu	<-->	err-rlod	5.147	-.165
err-etu	<-->	err-datos	4.365	.122
err-etu	<-->	err-a	4.053	-.111
err-etn	<-->	Usability	4.638	.147
err-etn	<-->	err-t	8.316	.206
err-etn	<-->	err-rot	4.957	-.152
err-etn	<-->	err-etu	26.655	.310
err-cau	<-->	Content quality	6.908	.088
err-cau	<-->	err-b	6.767	.125
err-atcut	<-->	err-datos	7.035	-.193
err-atcut	<-->	err-e	4.098	-.227
err-atcut	<-->	err-etu	6.938	-.192
err-atcut	<-->	err-cau	4.143	.136
err-etl	<-->	er-stm	7.634	.247

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
x20 <--- x10	5.752	-.108
x20 <--- x16	5.789	.113
x20 <--- x18	4.687	.115
x20 <--- x13	4.759	.096
x14 <--- x7	4.313	.120
x14 <--- x22	5.370	-.113
x14 <--- x25	6.274	-.116
x10 <--- x20	4.021	-.141
x10 <--- x7	4.082	.156
x16 <--- Content quality	4.473	.258
x16 <--- Responsiveness	4.329	.262
x16 <--- Usability	7.504	.225
x16 <--- x20	9.773	.281
x16 <--- x18	5.769	.217
x16 <--- x21	11.175	.227
x16 <--- x22	4.584	.178
x16 <--- x13	5.999	.183
x16 <--- x9	6.846	.178

x16 <--- x5	7.498	.214
x16 <--- x1	4.764	.177
x18 <--- x13	4.750	.108
x26 <--- x6	6.789	.190
x7 <--- x10	4.339	.095
x7 <--- x2	4.274	.116
x7 <--- x21	4.419	-.085
x2 <--- x7	4.093	.125
x2 <--- x13	4.534	-.100
x3 <--- Transaction quality	6.223	.289
x3 <--- x7	4.361	.151
x3 <--- x11	8.091	.161
x3 <--- x21	6.494	.127
x3 <--- x22	4.440	.128
x3 <--- x25	4.879	.128
x11 <--- x21	6.269	-.130
x11 <--- x22	6.348	-.160
x11 <--- x9	4.815	-.114
x11 <--- x5	4.265	-.123
x12 <--- x3	6.352	-.160
x12 <--- x11	6.463	.148
x12 <--- x9	5.311	-.117
x21 <--- x16	5.365	.189
x22 <--- x11	5.458	-.119
x22 <--- x21	8.185	.128
x22 <--- x19	5.280	.117
x6 <--- x27	4.422	-.115
x8 <--- Responsiveness	4.322	.267
x8 <--- Usability	6.655	.216
x8 <--- x12	4.301	.174
x8 <--- x22	7.211	.228
x8 <--- x24	5.885	.184
x8 <--- x19	8.357	.228
x8 <--- x5	5.849	.192
x8 <--- x1	12.751	.295
x19 <--- x23	5.277	.120
x19 <--- x21	5.969	.114
x19 <--- x22	4.967	.127
x13 <--- Transaction quality	4.320	-.221
x13 <--- x23	5.379	-.119
x13 <--- x2	4.318	-.132
x13 <--- x24	4.160	-.102
x13 <--- x8	4.214	-.097
x13 <--- x9	5.956	.111
x9 <--- Content quality	4.219	-.174
x9 <--- Transaction quality	5.852	-.266
x9 <--- x23	5.783	-.128
x9 <--- x14	5.347	-.149
x9 <--- x27	4.859	-.121
x9 <--- x26	5.293	-.141
x9 <--- x11	6.073	-.133
x9 <--- x12	6.581	-.147
x9 <--- x13	7.306	.140
x5 <--- x14	4.570	.122
x5 <--- x7	9.027	.182
x5 <--- x6	4.900	.131

Minimization History (Default model)

Iteration	Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	12	-1.131	9999.000	2443.336	0	9999.000
1	e*	9	-1.342	5.456	1267.609	20	.158
2	e*	6	-.190	.355	1145.378	6	.801
3	e	3	-.140	.864	919.208	6	.927
4	e	2	-.098	.712	792.091	5	.824
5	e	1	-.019	.738	673.841	5	.851
6	e	0	1862.225	.983	618.374	7	.854
7	e	0	503.306	.854	608.949	2	.000
8	e	0	872.743	.265	604.730	1	1.123
9	e	0	1083.706	.119	604.422	1	1.093
10	e	0	1176.172	.025	604.412	1	1.031
11	e	0	1173.143	.002	604.412	1	1.002
12	e	0	1173.015	.000	604.412	1	.999

Bootstrap (Default model)

Summary of Bootstrap Iterations (Default model)

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0

2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	1	0
9	0	1	0
10	0	4	0
11	0	16	0
12	0	13	0
13	0	30	0
14	0	16	0
15	0	16	0
16	0	9	0
17	0	15	0
18	0	5	0
19	0	14	0
Total	0	140	0

0 bootstrap samples were unused because of a singular covariance matrix.

0 bootstrap samples were unused because a solution was not found.

140 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 139 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 1 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap $p = .014$

Bootstrap Distributions (Default model)

ML discrepancy (implied vs sample) (Default model)

	194.600	-----
	223.896	*
	253.191	****
	282.487	****
	311.783	*****
	341.079	*****
	370.375	*****
N = 140	399.670	*****
Mean = 372.865	428.966	*****
S. e. = 5.896	458.262	*****
	487.558	***
	516.853	*
	546.149	
	575.445	*
	604.741	*

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	56	604.412	269	.000	2.247
Saturated model	325	.000	0		
Independence model	25	2520.122	300	.000	8.400

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.138	.765	.716	.633
Saturated model	.000	1.000		
Independence model	.704	.178	.110	.165

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.760	.733	.851	.832	.849
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.897	.682	.761
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	335.412	267.926	410.623
Saturated model	.000	.000	.000
Independence model	2220.122	2063.615	2384.030

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	4.140	2.297	1.835	2.812
Saturated model	.000	.000	.000	.000
Independence model	17.261	15.206	14.134	16.329

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.092	.083	.102	.000
Independence model	.225	.217	.233	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	716.412	740.679	883.876	939.876
Saturated model	650.000	790.833	1621.891	1946.891
Independence model	2570.122	2580.956	2644.883	2669.883

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	4.907	4.445	5.422	5.073
Saturated model	4.452	4.452	4.452	5.417
Independence model	17.604	16.532	18.726	17.678

HOELTER

Model	HOELTER	HOELTER
	.05	.01
Default model	75	79
Independence model	20	21

Execution time summary

Minimization: .016
 Miscellaneous: .265
 Bootstrap: .733
 Total: 1.014

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\cfa model based on webqual 5-factor model.amw

Analysis Summary

Date and Time

Date: Sunday, 22 November 2009

Time: 5:24:17 p.m.

Title

cfa model based on webqual 5-factor model: Sunday, 22 November 2009 05:24 p.m.

Notes

Bootstrap confidence intervals are not available when the Bollen-Stine bootstrap is performed.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x7
x23
x27
x11
x6
x15
x3
x17
x26
x13
x16
x20
x22
x24
x5
x2
x14
x10
x9
x1
x18

Unobserved, exogenous variables

err-b
err-lod
err-twbd
err-pifs
err-tfs
err-p
err-rep
err-etu
err-af
err-etuse
err-attr
err-c
err-pe
err-cwo
information
empathy
trust
err-cau
err-a
err-r
err-t
usability
design
err-n
err-l
err-da

Variable counts (Group number 1)

Number of variables in your model:	47
Number of observed variables:	21
Number of unobserved variables:	26
Number of exogenous variables:	26
Number of endogenous variables:	21

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	26	0	0	0	0	26
Labeled	0	0	0	0	0	0
Unlabeled	15	10	26	0	0	51
Total	41	10	26	0	0	77

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 231
Number of distinct parameters to be estimated: 51
Degrees of freedom (231 - 51): 180

Result (Default model)

Minimum was achieved
Chi-square = 518.333
Degrees of freedom = 180
Probability level = .000

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

	Estimate	S.E.	C.R.	P	Label
x26 <--- information	.674	.094	7.140	***	
x24 <--- empathy	.940	.157	5.987	***	
x11 <--- trust	1.606	.296	5.428	***	
x6 <--- trust	1.000				
x3 <--- trust	1.506	.274	5.505	***	
x27 <--- trust	1.792	.309	5.798	***	
x15 <--- empathy	1.000				
x17 <--- information	.563	.089	6.342	***	
x7 <--- information	.773	.082	9.417	***	
x2 <--- information	.775	.086	9.034	***	
x23 <--- information	.860	.107	8.002	***	
x10 <--- information	1.000				
x1 <--- usability	.731	.063	11.684	***	
x9 <--- usability	1.000				
x13 <--- usability	.901	.062	14.585	***	
x5 <--- usability	.803	.062	12.858	***	
x16 <--- design	.688	.128	5.366	***	
x18 <--- design	1.000				
x20 <--- design	1.134	.109	10.399	***	
x22 <--- design	1.114	.118	9.405	***	

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
x26 <--- information	.597
x24 <--- empathy	.622
x11 <--- trust	.648
x6 <--- trust	.500
x3 <--- trust	.665
x27 <--- trust	.739
x15 <--- empathy	.558
x17 <--- information	.535
x7 <--- information	.766
x2 <--- information	.738
x23 <--- information	.663
x10 <--- information	.769
x1 <--- usability	.774
x9 <--- usability	.888
x13 <--- usability	.878
x5 <--- usability	.818
x16 <--- design	.457
x18 <--- design	.752
x20 <--- design	.848
x22 <--- design	.773

Covariances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
information <--> usability	.886	.167	5.319	***	

usability	<-->	design	.916	.152	6.010	***
empathy	<-->	design	.740	.147	5.036	***
empathy	<-->	trust	.513	.119	4.309	***
trust	<-->	design	.425	.090	4.694	***
information	<-->	design	.781	.132	5.913	***
trust	<-->	usability	.507	.114	4.470	***
information	<-->	trust	.513	.109	4.724	***
information	<-->	empathy	.835	.170	4.906	***
empathy	<-->	usability	1.039	.205	5.058	***

Correlations: (Group number 1 - Default model)

	Estimate
information <--> usability	.620
usability <--> design	.774
empathy <--> design	.918
empathy <--> trust	1.019
trust <--> design	.901
information <--> design	.856
trust <--> usability	.687
information <--> trust	.901
information <--> empathy	.857
empathy <--> usability	.822

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
information	1.102	.208	5.291	***	
empathy	.861	.273	3.157	.002	
trust	.294	.096	3.060	.002	
usability	1.855	.277	6.700	***	
design	.755	.147	5.156	***	
err-b	.463	.066	7.061	***	
err-lod	1.041	.135	7.737	***	
err-twbd	.786	.112	7.009	***	
err-pifs	1.051	.136	7.710	***	
err-tfs	.883	.108	8.189	***	
err-p	1.908	.263	7.260	***	
err-rep	.843	.111	7.615	***	
err-etu	.869	.107	8.132	***	
err-af	.903	.113	7.974	***	
err-etuse	.449	.075	5.998	***	
err-attr	1.353	.163	8.311	***	
err-c	.378	.062	6.069	***	
err-pe	.632	.088	7.177	***	
err-cwo	1.204	.187	6.450	***	
err-cau	.590	.084	7.028	***	
err-a	.553	.076	7.298	***	
err-r	1.270	.149	8.544	***	
err-t	.761	.108	7.031	***	
err-n	.499	.087	5.725	***	
err-l	.661	.089	7.436	***	
err-da	.582	.079	7.356	***	

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
x18	.565
x1	.599
x9	.788
x10	.592
x14	.000
x2	.545
x5	.670
x24	.387
x22	.597
x20	.720
x16	.209
x13	.770
x26	.357
x17	.287
x3	.442
x15	.311
x6	.250
x11	.419
x27	.546
x23	.439
x7	.587

Matrices (Group number 1 - Default model)**Residual Covariances (Group number 1 - Default model)**

	x18	x1	x9	x10	x14	x2	x5	x24	x22	x20	x16	x13	x26	x17	x3	x15	x6	x11	x27	x23
x18	.000																			
x1	-.012	.000																		
x9	-.099	-.040	.000																	
x10	-.034	-.049	.144	.000																
x14	.695	.595	.675	.921	.000															
x2	.013	.026	-.140	.036	.790	.000														
x5	.128	.057	-.037	.268	.843	.104	.000													
x24	-.011	-.150	-.127	.093	.776	-.046	-.014	.000												
x22	-.063	.045	-.016	.001	.622	-.101	.178	.185	.000											
x20	.050	-.011	-.116	-.149	.784	.060	.113	-.050	-.021	.000										
x16	.007	.022	.000	-.084	.417	.011	.134	-.035	-.034	.047	.000									
x13	.027	.001	.079	.024	.725	-.168	-.079	-.270	-.025	-.034	-.023	.000								
x26	-.006	-.020	-.121	-.128	.604	-.076	.163	.136	.072	.162	.023	-.008	.000							
x17	.017	.040	.037	.038	.613	-.023	.203	.169	.004	.132	.268	.171	.007	.000						
x3	.014	.266	.359	-.020	.539	-.117	.298	.025	.184	.026	.021	.206	-.107	-.100	.000					
x15	-.145	.167	.273	-.112	.650	-.193	.310	.000	.215	-.129	.089	.153	.141	-.087	.099	.000				
x6	-.074	-.025	-.103	-.043	.379	-.028	.127	.019	-.009	-.080	-.144	-.155	.219	-.203	-.001	.002	.000			
x11	.035	-.154	-.358	-.041	.773	-.040	-.036	.089	-.147	-.087	-.109	-.314	.086	-.084	-.020	-.194	.429	.000		
x27	-.047	.004	-.159	.048	.848	.024	.144	.061	.075	.061	-.011	-.042	.207	.009	-.061	-.170	-.156	-.009	.000	
x23	.031	-.004	-.141	.015	.833	-.062	.192	.194	.209	.013	-.139	-.155	.115	.021	-.152	-.193	-.033	.091	.160	.000
x7	.018	-.057	-.113	.073	.809	.090	.217	-.027	-.081	-.042	-.090	-.118	-.076	-.042	.005	-.026	.028	.058	-.037	-.098

Standardized Residual Covariances (Group number 1 - Default model)

	x18	x1	x9	x10	x14	x2	x5	x24	x22	x20	x16	x13	x26	x17	x3	x15	x6	x11	x27	x23
x18	.000																			
x1	-.091	.000																		
x9	-.601	-.202	.000																	
x10	-.230	-.319	.765	.000																
x14	6.440	4.968	4.718	7.237	.000															
x2	.114	.211	-.924	.254	7.685	.000														
x5	.906	.338	-.179	1.655	6.759	.796	.000													
x24	-.073	-.936	-.651	.543	5.937	-.335	-.083	.000												
x22	-.456	.308	-.090	.008	5.326	-.791	1.151	1.166	.000											
x20	.377	-.083	-.678	-.993	7.240	.500	.776	-.331	-.146	.000										
x16	.050	.153	.002	-.541	3.416	.092	.889	-.222	-.238	.349	.000									
x13	.179	.006	.352	.143	5.563	-1.220	-.415	-1.519	-.156	-.221	-.142	.000								
x26	-.053	-.152	-.761	-.867	5.470	-.646	1.190	.942	.542	1.304	.177	-.053	.000							
x17	.153	.334	.251	.283	5.952	-.216	1.606	1.266	.029	1.161	2.195	1.288	.058	.000						
x3	.107	1.915	2.132	-.129	4.699	-.954	2.053	.158	1.311	.194	.149	1.345	-.834	-.849	.000					
x15	-.847	.889	1.199	-.561	4.192	-1.201	1.575	.000	1.160	-.741	.479	.738	.834	-.552	.545	.000				
x6	-.679	-.210	-.715	-.329	3.748	-.268	1.020	.142	-.078	-.714	-1.203	-1.180	1.986	-1.992	-.013	.014	.000			
x11	.249	-1.016	-1.950	-.248	6.159	-.301	-.224	.525	-.959	-.601	-.721	-1.880	.612	-.656	-.133	-.982	3.378	.000		
x27	-.336	.026	-.865	.287	6.909	.182	.914	.360	.491	.421	-.073	-.255	1.488	.073	-.409	-.864	-1.240	-.058	.000	
x23	.215	-.024	-.764	.088	6.554	-.450	1.208	1.160	1.358	.091	-.913	-.928	.803	.161	-1.020	-.980	-.260	.562	.000	
x7	.157	-.475	-.773	.528	8.191	.813	1.723	-.203	-.662	-.362	-.753	-.888	-.666	-.401	.046	-.166	.275	.451	-.000	

Notes for Model (Group number 1 - Default model)

The following covariance matrix is not positive definite (Group number 1 - Default model)

	design	usability	trust	empathy	information
design	.755				
usability	.916	1.855			
trust	.425	.507	.294		
empathy	.740	1.039	.513	.861	
information	.781	.886	.513	.835	1.102

Notes for Group/Model (Group number 1 - Default model)

This solution is not admissible.

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

		M.I.	Par Change
err-t	<--> design	8.242	-.119
err-t	<--> usability	6.108	.195
err-t	<--> err-l	4.175	-.139
err-t	<--> err-n	10.075	.209
err-r	<--> information	22.837	.314
err-r	<--> err-n	4.905	-.173
err-cau	<--> usability	8.320	-.197
err-cau	<--> information	7.143	.130
err-cau	<--> err-r	9.410	.239
err-cwo	<--> usability	4.529	-.195
err-pe	<--> empathy	10.803	.215
err-pe	<--> err-a	4.479	-.119
err-pe	<--> err-cwo	4.934	.176
err-c	<--> err-t	12.039	-.192
err-c	<--> err-a	5.257	.106

err-etuse <--> err-n	6.816	.132
err-etuse <--> err-cau	5.185	-.119
err-etuse <--> err-cwo	4.371	-.151
err-af <--> information	7.507	-.154
err-af <--> err-t	4.311	-.159
err-af <--> err-c	6.315	.143
err-etu <--> trust	6.741	-.077
err-etu <--> err-attr	8.386	.269
err-etu <--> err-etuse	5.437	.143
err-rep <--> usability	13.548	.293
err-rep <--> information	4.551	-.118
err-rep <--> err-n	8.612	.195
err-p <--> usability	6.316	.290
err-p <--> err-pe	5.647	.236
err-tfs <--> err-af	7.327	.210
err-tfs <--> err-etu	8.117	-.215
err-pifs <--> usability	7.455	-.242
err-pifs <--> err-tfs	32.341	.479
err-twbd <--> err-tfs	6.047	-.184
err-lod <--> err-pe	9.277	.230
err-b <--> err-r	5.601	.163
err-b <--> err-a	6.107	.119
err-b <--> err-cau	6.622	.132

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
x9 <--- x14	4.905	-.136
x9 <--- x26	4.974	-.131
x9 <--- x27	4.110	-.107
x10 <--- x9	4.262	.106
x14 <--- design	70.952	.949
x14 <--- usability	42.805	.467
x14 <--- trust	78.305	1.607
x14 <--- empathy	71.356	.877
x14 <--- information	87.893	.881
x14 <--- x18	41.480	.519
x14 <--- x1	24.682	.361
x14 <--- x9	22.264	.287
x14 <--- x10	52.378	.494
x14 <--- x2	59.056	.650
x14 <--- x5	45.682	.472
x14 <--- x24	35.252	.395
x14 <--- x22	28.364	.396
x14 <--- x20	52.415	.581
x14 <--- x16	11.667	.244
x14 <--- x13	30.949	.371
x14 <--- x26	29.925	.431
x14 <--- x17	35.432	.503
x14 <--- x3	22.077	.356
x14 <--- x15	17.571	.235
x14 <--- x6	14.048	.322
x14 <--- x11	37.939	.427
x14 <--- x27	47.733	.490
x14 <--- x23	42.955	.449
x14 <--- x7	67.089	.721
x5 <--- design	5.657	.199
x5 <--- trust	8.340	.390
x5 <--- empathy	4.453	.163
x5 <--- information	11.853	.240
x5 <--- x10	6.310	.127
x5 <--- x14	9.410	.189
x5 <--- x2	6.125	.156
x5 <--- x22	4.144	.113
x5 <--- x26	5.055	.131
x5 <--- x6	7.703	.177
x5 <--- x27	4.674	.114
x5 <--- x23	7.957	.143
x5 <--- x7	16.137	.263
x22 <--- x24	4.211	.104
x22 <--- x15	5.090	.096
x22 <--- x23	4.876	.115
x20 <--- x10	4.004	-.086
x16 <--- x17	5.106	.200
x13 <--- x2	4.030	-.118
x13 <--- x24	4.737	-.100
x26 <--- x6	7.385	.203
x17 <--- x16	8.158	.172

x17 <--- x6	5.846	-.176
x3 <--- usability	8.289	.175
x3 <--- x1	7.821	.173
x3 <--- x9	12.905	.186
x3 <--- x13	6.263	.142
x6 <--- x26	4.155	.136
x6 <--- x17	5.832	-.173
x6 <--- x11	17.418	.246
x11 <--- usability	4.862	-.149
x11 <--- x9	6.259	-.144
x11 <--- x13	5.287	-.145
x11 <--- x6	23.481	.394
x27 <--- x6	4.400	-.152
x23 <--- x22	4.199	.144
x7 <--- x14	5.601	.128

Minimization History (Default model)

Iteration	Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	16	-.676	9999.000	1982.257	0	9999.000
1	e*	11	-.624	3.891	1028.740	20	.339
2	e	5	-.178	.542	823.573	6	.949
3	e*	2	-.077	.758	682.152	5	.675
4	e	0	3104.562	.617	582.188	5	.822
5	e	1	-.658	.769	577.545	3	.000
6	e	0	867.160	.347	533.252	5	.833
7	e	0	511.217	.493	520.841	1	1.057
8	e	0	860.767	.226	518.622	1	1.138
9	e	0	1201.181	.158	518.355	1	1.061
10	e	0	1457.390	.034	518.334	1	1.028
11	e	0	1517.068	.004	518.333	1	1.003
12	e	0	1518.404	.000	518.333	1	1.000

Bootstrap (Default model)**Summary of Bootstrap Iterations (Default model)**

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	2	0
11	0	4	0
12	0	7	0
13	0	14	0
14	0	11	0
15	0	15	0
16	0	13	0
17	0	9	0
18	0	9	0
19	0	56	0
Total	0	140	0

0 bootstrap samples were unused because of a singular covariance matrix.

0 bootstrap samples were unused because a solution was not found.

140 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 140 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 0 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .007

Bootstrap Distributions (Default model)**ML discrepancy (implied vs sample) (Default model)**

112.974	*
134.398	*
155.822	**
177.246	*****
198.670	*****
220.095	*****
241.519	*****

N = 140	262.943	*****
Mean = 234.359	284.367	****
S. e. = 3.874	305.791	***
	327.215	**
	348.639	
	370.063	*
	391.488	*
	412.912	*

Model Fit Summary**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	51	518.333	180	.000	2.880
Saturated model	231	.000	0		
Independence model	21	1961.618	210	.000	9.341

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.237	.777	.714	.605
Saturated model	.000	1.000		
Independence model	.686	.207	.128	.188

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.736	.692	.810	.775	.807
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.857	.631	.692
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	338.333	274.029	410.276
Saturated model	.000	.000	.000
Independence model	1751.618	1613.529	1897.118

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	3.550	2.317	1.877	2.810
Saturated model	.000	.000	.000	.000
Independence model	13.436	11.997	11.052	12.994

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.113	.102	.125	.000
Independence model	.239	.229	.249	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	620.333	638.430	772.845	823.845
Saturated model	462.000	543.968	1152.790	1383.790
Independence model	2003.618	2011.069	2066.417	2087.417

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	4.249	3.808	4.742	4.373
Saturated model	3.164	3.164	3.164	3.726
Independence model	13.723	12.778	14.720	13.774

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	60	64
Independence model	19	20

Execution time summary

Minimization:	.015
Miscellaneous:	.235
Bootstrap:	.687
Total:	.937

Appendix VIII: AMOS reports, formative model

This document provides output generated by the AMOS package after calculating the estimates for the formative model included in Chapter 7. Selected subsets of this report are included in the text.

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\formative model no fun.amw

Analysis Summary

Date and Time

Date: Thursday, 12 November 2009

Time: 4:16:25 p.m.

Title

formative model no fun: Thursday, 12 November 2009 04:16 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x5
x13
x8
x28
x22
x27

Observed, exogenous variables

x1
x9
x17
x18
x24
x2
x4
x10
x12
x14
x15
x23
x25
x3
x6
x11
x19
x20
x7
x26

Unobserved, endogenous variables

Ease of Use
Usefulness
Perceived online service quality
Trust

Unobserved, exogenous variables

err-x5
error x13
error x8
error x28
error x22
error ease of use
error POLSQ
error x 27
error Trust
error usefulness

Variable counts (Group number 1)

Number of variables in your model: 40
Number of observed variables: 26
Number of unobserved variables: 14
Number of exogenous variables: 30
Number of endogenous variables: 10

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	14	0	0	0	0	14
Labeled	0	0	0	0	0	0
Unlabeled	25	0	30	20	6	81

Total	39	0	30	20	6	95
-------	----	---	----	----	---	----

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x20	1.000	7.000	-1.022	-5.058	1.608	3.979
x19	1.000	7.000	-.638	-3.160	.219	.541
x11	1.000	7.000	-1.051	-5.201	1.040	2.575
x6	1.000	7.000	-1.329	-6.579	2.839	7.027
x3	1.000	7.000	-.670	-3.315	.664	1.643
x25	1.000	7.000	-.942	-4.662	.822	2.034
x23	1.000	7.000	-.789	-3.904	.144	.357
x15	1.000	7.000	-.282	-1.394	-.759	-1.877
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x12	1.000	7.000	-1.181	-5.847	1.673	4.141
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x4	1.000	7.000	-.820	-4.059	.331	.818
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
x24	1.000	7.000	-.710	-3.515	.228	.564
x18	1.000	7.000	-1.226	-6.070	2.271	5.621
x17	2.000	7.000	-.640	-3.168	.092	.227
x9	1.000	7.000	-.780	-3.860	.064	.159
x1	1.000	7.000	-.915	-4.531	.839	2.077
x27	1.000	7.000	-.944	-4.674	.866	2.143
x22	1.000	7.000	-.756	-3.742	.904	2.238
x28	1.000	7.000	-1.277	-6.320	2.257	5.586
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x13	1.000	7.000	-.871	-4.309	.278	.688
x5	1.000	7.000	-.988	-4.889	.864	2.138
Multivariate					149.290	23.718

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
86	66.238	.000	.003
116	62.944	.000	.000
27	62.512	.000	.000
17	56.943	.000	.000
77	56.631	.000	.000
74	56.106	.001	.000
65	54.675	.001	.000
97	53.967	.001	.000
30	53.531	.001	.000
55	52.471	.002	.000
50	51.242	.002	.000
48	50.754	.003	.000
64	50.649	.003	.000
41	47.519	.006	.000
46	46.934	.007	.000
144	45.423	.011	.000
103	45.075	.012	.000
2	45.054	.012	.000
80	43.679	.016	.000
145	42.832	.020	.000
11	41.790	.026	.000
121	40.597	.034	.000
38	40.357	.036	.000
47	40.029	.039	.000
40	38.953	.049	.000
10	38.374	.056	.000
5	38.313	.057	.000
12	37.716	.064	.000
101	37.664	.065	.000
118	37.662	.065	.000
111	37.595	.066	.000
58	37.214	.071	.000
68	36.415	.084	.000
69	36.402	.085	.000
98	36.216	.088	.000
22	35.200	.107	.000
52	34.355	.126	.000
110	34.322	.127	.000
43	33.738	.142	.000
35	33.700	.143	.000
109	32.373	.181	.002
123	32.209	.186	.002
87	32.163	.188	.001
91	32.058	.191	.001
8	31.913	.196	.001
81	31.833	.199	.001

60	31.038	.227	.006
136	30.655	.241	.012
135	30.655	.241	.007
72	29.978	.269	.033
107	29.778	.277	.038
119	29.039	.309	.141
6	29.033	.310	.107
113	29.024	.310	.080
125	27.770	.370	.488
16	27.644	.376	.484
147	27.632	.377	.422
56	27.401	.389	.472
137	26.917	.414	.649
49	26.756	.422	.665
45	25.721	.479	.948
127	25.700	.480	.932
131	25.467	.493	.949
92	25.433	.495	.936
78	25.385	.497	.922
23	24.593	.542	.990
21	24.534	.545	.988
141	24.526	.546	.982
90	23.769	.589	.999
20	23.680	.594	.999
66	23.588	.600	.998
128	23.301	.616	.999
124	23.247	.619	.999
122	23.247	.619	.998
28	23.064	.629	.999
18	22.962	.635	.999
67	22.616	.655	1.000
29	22.614	.655	.999
105	22.385	.667	1.000
70	22.283	.673	1.000
88	21.770	.701	1.000
95	21.744	.703	1.000
130	21.493	.716	1.000
126	21.181	.733	1.000
1	19.831	.800	1.000
108	19.772	.802	1.000
120	19.613	.810	1.000
51	19.452	.817	1.000
84	19.289	.824	1.000
34	18.895	.841	1.000
7	18.608	.853	1.000
57	18.447	.859	1.000
112	18.000	.876	1.000
117	17.711	.886	1.000
102	17.688	.887	1.000
134	17.454	.895	1.000
133	17.454	.895	1.000
15	17.302	.900	1.000
93	16.723	.917	1.000
82	16.493	.924	1.000

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 377
Number of distinct parameters to be estimated: 81
Degrees of freedom (377 - 81): 296

Result (Default model)

Minimum was achieved
Chi-square = 1987.854
Degrees of freedom = 296
Probability level = .000

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

	Estimate	S.E.	C.R.	P	Label

Ease of Use	<--- x1	.229	.036	6.352	***	par_1
Ease of Use	<--- x9	.427	.040	10.711	***	par_2
Ease of Use	<--- x17	.137	.039	3.516	***	par_3
Ease of Use	<--- x18	.173	.038	4.553	***	par_4
Ease of Use	<--- x24	-.060	.030	-2.015	.044	par_5
Usefulness	<--- x2	.037	.074	.507	.612	par_7
Usefulness	<--- x4	.143	.063	2.278	.023	par_8
Usefulness	<--- x10	-.160	.061	-2.610	.009	par_9
Usefulness	<--- x12	.187	.066	2.831	.005	par_10
Usefulness	<--- x14	-.049	.072	-.674	.501	par_11
Usefulness	<--- x15	.163	.051	3.201	.001	par_12
Usefulness	<--- x23	.206	.062	3.297	***	par_13
Usefulness	<--- x25	.123	.062	1.971	.049	par_14
Trust	<--- x3	.258	.052	4.938	***	par_19
Trust	<--- x6	.130	.051	2.568	.010	par_20
Trust	<--- x11	-.025	.039	-.648	.517	par_21
Trust	<--- x19	.080	.040	1.999	.046	par_22
Trust	<--- x20	.395	.065	6.108	***	par_23
Usefulness	<--- x7	.269	.080	3.343	***	par_24
Ease of Use	<--- x26	.092	.036	2.590	.010	par_25
Perceived online service quality	<--- Ease of Use	.110	.060	1.843	.065	par_16
Perceived online service quality	<--- Usefulness	.339	.090	3.742	***	par_17
Perceived online service quality	<--- Trust	.728	.139	5.254	***	par_18
x5	<--- Ease of Use	1.000				
x13	<--- Ease of Use	1.128	.124	9.070	***	par_6
x8	<--- Usefulness	1.000				
x28	<--- Perceived online service quality	1.000				
x22	<--- Perceived online service quality	1.045	.174	6.012	***	par_15
x27	<--- Trust	1.000				

Means: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
x24	4.966	.116	42.807	***	par_26
x18	5.537	.096	57.864	***	par_27
x17	5.340	.091	58.475	***	par_28
x9	4.986	.127	39.274	***	par_29
x1	5.170	.106	48.613	***	par_30
x2	5.735	.091	62.855	***	par_33
x4	5.429	.110	49.512	***	par_34
x10	5.204	.113	46.068	***	par_35
x12	5.293	.105	50.426	***	par_36
x14	5.619	.093	60.251	***	par_37
x15	3.993	.138	28.998	***	par_38
x23	5.150	.113	45.684	***	par_39
x25	5.340	.109	48.801	***	par_40
x3	5.075	.102	49.878	***	par_44
x6	5.728	.090	63.777	***	par_45
x11	5.565	.111	49.971	***	par_46
x19	4.973	.111	44.677	***	par_47
x20	5.327	.096	55.410	***	par_48
x7	6.034	.088	68.853	***	par_50
x26	5.551	.098	56.609	***	par_51

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
x5	.056	.564	.100	.920	par_31
x13	-.738	.551	-1.341	.180	par_32
x8	.566	1.148	.493	.622	par_41
x28	.017	.799	.021	.983	par_42
x22	-.849	.890	-.954	.340	par_43
x27	.777	.736	1.056	.291	par_49

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
x1	1.651	.193	8.544	***	par_52
x9	2.354	.275	8.544	***	par_53
x17	1.218	.143	8.544	***	par_54
x18	1.337	.156	8.544	***	par_55
x24	1.965	.230	8.544	***	par_56
x2	1.215	.142	8.544	***	par_57
x4	1.755	.205	8.544	***	par_58
x10	1.863	.218	8.544	***	par_59
x12	1.608	.188	8.544	***	par_60
x14	1.270	.149	8.544	***	par_61
x15	2.769	.324	8.544	***	par_62
x23	1.855	.217	8.544	***	par_63
x25	1.748	.205	8.544	***	par_64
x3	1.511	.177	8.544	***	par_65
x6	1.178	.138	8.544	***	par_66

x11	1.810	.212	8.544	***	par_67
x19	1.809	.212	8.544	***	par_68
x20	1.349	.158	8.544	***	par_69
x7	1.121	.131	8.544	***	par_70
x26	1.404	.164	8.544	***	par_71
error ease of use	.006	.042	.133	.894	par_72
error Trust	.059	.066	.896	.370	par_73
error usefulness	.078	.161	.481	.630	par_74
error POLSQ	.022	.057	.383	.702	par_75
err-x5	.655	.088	7.483	***	par_76
error x13	.515	.081	6.378	***	par_77
error x8	1.361	.227	6.007	***	par_78
error x28	.430	.068	6.347	***	par_79
error x22	.580	.084	6.899	***	par_80
error x 27	1.034	.138	7.510	***	par_81

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
Trust	.853
Usefulness	.839
Ease of Use	.991
Perceived online service quality	.926
x27	.280
x22	.360
x28	.409
x8	.261
x13	.598
x5	.479

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

		M.I.	Par Change
x7	<--> x26	23.027	.498
x20	<--> x26	44.389	.759
x20	<--> x7	39.772	.642
x19	<--> x26	16.692	.539
x19	<--> x7	19.311	.518
x19	<--> x20	50.688	.920
x11	<--> x26	23.622	.641
x11	<--> x7	34.747	.695
x11	<--> x20	28.160	.686
x11	<--> x19	16.808	.614
x6	<--> x26	28.181	.565
x6	<--> x7	19.898	.424
x6	<--> x20	14.833	.402
x6	<--> x19	6.686	.312
x6	<--> x11	55.706	.902
x3	<--> x26	11.825	.415
x3	<--> x7	31.314	.603
x3	<--> x20	40.390	.751
x3	<--> x19	45.242	.920
x3	<--> x11	25.583	.692
x3	<--> x6	16.034	.442
x25	<--> x26	25.609	.656
x25	<--> x7	25.669	.587
x25	<--> x20	40.342	.807
x25	<--> x19	30.952	.819
x25	<--> x11	20.405	.665
x25	<--> x6	11.271	.399
x25	<--> x3	31.652	.757
x23	<--> x26	31.893	.754
x23	<--> x7	28.244	.634
x23	<--> x20	34.969	.774
x23	<--> x19	32.271	.861
x23	<--> x11	27.810	.800
x23	<--> x6	11.128	.408
x23	<--> x3	13.683	.513
x23	<--> x25	27.320	.779
x15	<--> x26	18.639	.704
x15	<--> x7	18.036	.619
x15	<--> x20	19.687	.710
x15	<--> x19	32.822	1.061
x15	<--> x11	11.549	.630
x15	<--> x6	11.883	.515
x15	<--> x3	26.485	.871
x15	<--> x25	9.235	.553
x15	<--> x23	7.830	.525
x14	<--> x26	29.925	.604
x14	<--> x7	67.089	.809

x14	<--> x20	52.415	.784
x14	<--> x19	33.349	.724
x14	<--> x11	37.939	.773
x14	<--> x6	14.048	.379
x14	<--> x3	22.077	.539
x14	<--> x25	18.539	.531
x14	<--> x23	42.955	.833
x14	<--> x15	17.571	.650
x12	<--> x26	32.557	.710
x12	<--> x7	31.393	.623
x12	<--> x20	47.077	.836
x12	<--> x19	20.161	.634
x12	<--> x11	44.653	.944
x12	<--> x6	17.831	.481
x12	<--> x3	13.150	.468
x12	<--> x25	33.105	.798
x12	<--> x23	45.381	.963
x12	<--> x15	21.956	.818
x12	<--> x14	36.739	.717
x10	<--> x26	21.143	.615
x10	<--> x7	59.801	.925
x10	<--> x20	31.470	.736
x10	<--> x19	23.740	.740
x10	<--> x11	26.519	.783
x10	<--> x6	14.730	.470
x10	<--> x3	29.432	.753
x10	<--> x25	22.781	.713
x10	<--> x23	39.145	.963
x10	<--> x15	14.774	.722
x10	<--> x14	52.378	.921
x10	<--> x12	26.409	.736
x4	<--> x26	20.416	.587
x4	<--> x7	23.570	.564
x4	<--> x20	30.523	.704
x4	<--> x19	36.364	.889
x4	<--> x11	14.447	.561
x4	<--> x6	16.083	.477
x4	<--> x3	35.037	.798
x4	<--> x25	10.354	.466
x4	<--> x23	20.989	.684
x4	<--> x15	11.878	.629
x4	<--> x14	21.903	.578
x4	<--> x12	14.403	.528
x4	<--> x10	21.981	.702
x2	<--> x26	21.387	.500
x2	<--> x7	60.348	.751
x2	<--> x20	49.619	.746
x2	<--> x19	18.670	.530
x2	<--> x11	23.791	.599
x2	<--> x6	13.964	.370
x2	<--> x3	18.500	.482
x2	<--> x25	31.334	.675
x2	<--> x23	29.274	.672
x2	<--> x15	8.942	.454
x2	<--> x14	59.056	.790
x2	<--> x12	36.251	.697
x2	<--> x10	51.175	.891
x2	<--> x4	21.671	.563
x24	<--> x26	23.407	.665
x24	<--> x7	22.246	.579
x24	<--> x20	30.078	.739
x24	<--> x19	31.082	.870
x24	<--> x11	30.549	.863
x24	<--> x6	15.834	.501
x24	<--> x3	27.717	.751
x24	<--> x25	34.637	.903
x24	<--> x23	30.250	.869
x24	<--> x15	17.578	.809
x24	<--> x14	35.252	.776
x24	<--> x12	33.655	.853
x24	<--> x10	30.724	.878
x24	<--> x4	13.872	.572
x24	<--> x2	19.339	.562
x18	<--> x26	21.049	.520
x18	<--> x7	37.577	.621
x18	<--> x20	66.458	.906
x18	<--> x19	44.459	.858
x18	<--> x11	31.005	.717
x18	<--> x6	11.379	.350
x18	<--> x3	30.870	.654
x18	<--> x25	42.410	.824
x18	<--> x23	28.997	.702

	<--> x15	13.986	.595
x18	<--> x14	41.480	.695
x18	<--> x12	39.337	.761
x18	<--> x10	32.746	.747
x18	<--> x4	14.166	.477
x18	<--> x2	34.401	.619
x18	<--> x24	26.073	.685
x17	<--> x26	15.414	.425
x17	<--> x7	20.459	.437
x17	<--> x20	35.322	.630
x17	<--> x19	25.048	.615
x17	<--> x11	9.532	.379
x17	<--> x3	8.908	.335
x17	<--> x25	11.423	.408
x17	<--> x23	19.875	.555
x17	<--> x15	6.362	.383
x17	<--> x14	35.432	.613
x17	<--> x12	15.661	.458
x17	<--> x10	27.904	.658
x17	<--> x4	7.769	.337
x17	<--> x2	20.658	.458
x17	<--> x24	22.723	.610
x17	<--> x18	18.702	.457
x9	<--> x26	10.049	.477
x9	<--> x7	18.094	.572
x9	<--> x20	39.155	.923
x9	<--> x19	64.726	1.374
x9	<--> x11	7.145	.457
x9	<--> x6	8.617	.404
x9	<--> x3	51.804	1.123
x9	<--> x25	25.937	.855
x9	<--> x23	12.899	.621
x9	<--> x15	38.617	1.313
x9	<--> x14	22.264	.675
x9	<--> x12	12.470	.569
x9	<--> x10	35.323	1.030
x9	<--> x4	35.758	1.006
x9	<--> x2	15.295	.547
x9	<--> x24	22.804	.850
x9	<--> x18	30.957	.817
x9	<--> x17	14.595	.535
x1	<--> x26	10.925	.416
x1	<--> x7	15.488	.443
x1	<--> x20	36.586	.747
x1	<--> x19	59.130	1.100
x1	<--> x11	9.514	.441
x1	<--> x6	8.967	.346
x1	<--> x3	39.719	.824
x1	<--> x25	14.790	.541
x1	<--> x23	14.563	.553
x1	<--> x15	27.401	.926
x1	<--> x14	24.682	.595
x1	<--> x12	13.441	.494
x1	<--> x10	16.967	.598
x1	<--> x4	43.299	.927
x1	<--> x2	20.290	.528
x1	<--> x24	14.294	.564
x1	<--> x18	28.535	.657
x1	<--> x17	11.894	.405
x1	<--> x9	64.983	1.315
error Trust	<--> x25	4.303	.143
error Trust	<--> x23	7.547	.196
error Trust	<--> x10	6.435	.181
error Trust	<--> x24	6.762	.191
error usefulness	<--> x24	6.233	.284
error usefulness	<--> x1	5.668	.248
error x 27	<--> x26	9.693	.318
error x 27	<--> x7	4.152	.186
error x 27	<--> x11	5.880	.282
error x 27	<--> x25	6.287	.286
error x 27	<--> x23	11.694	.402
error x 27	<--> x14	10.288	.312
error x 27	<--> x12	11.991	.379
error x 27	<--> x10	10.064	.374
error x 27	<--> x2	6.233	.238
error x 27	<--> x24	6.716	.313
error x8	<--> x24	4.586	.298
error x8	<--> x1	4.949	.284
error x13	<--> x2	5.502	-.155
err-x5	<--> x7	11.755	.244
err-x5	<--> x6	6.779	.190
err-x5	<--> x23	4.766	.200

	<-->	x14	7.366	.206
err-x5	<-->	x12	5.989	.209
err-x5	<-->	x4	9.501	.275
err-x5	<-->	error Trust	4.842	.093
err-x5	<-->	error x28	4.282	.097

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

		M.I.	Par Change
x19 <---	x15	4.856	.057
x19 <---	x9	5.597	.050
x3 <---	x9	4.480	.041
x9 <---	x19	4.412	.052
x9 <---	x15	5.713	.070
x1 <---	x19	4.030	.041
x1 <---	x15	4.054	.049
x1 <---	x9	5.619	.048

Means: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Intercepts: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Minimization History (Default model)

Iteration		Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	8		-.664	9999.000	4618.186	0	9999.000
1	e	8		-.245	.509	2761.146	13	.888
2	e	9		-1.132	.920	2565.283	8	.797
3	e	7		-.255	.656	2402.254	6	.890
4	e	5		-.807	.765	2243.411	6	.820
5	e	3		-1.082	.556	2167.827	5	.631
6	e	4		-1.163	.484	2097.987	5	.637
7	e	2		-.235	.307	2051.501	5	.805
8	e	3		-.294	.335	2022.017	5	.845
9	e	1		-.032	.315	2005.143	7	.985
10	e	3		-.780	.361	1996.034	5	.681
11	e	0	10515.370		.102	1990.799	5	.839
12	e	1		-.008	.188	1988.274	2	.000
13	e	0	5926.339		.071	1987.893	5	.806
14	e	0	7044.053		.019	1987.855	1	.967
15	e	0	7054.466		.004	1987.854	1	.993
16	e	0	7047.680		.000	1987.854	1	1.001

Pairwise Parameter Comparisons (Default model)

Variance-covariance Matrix of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18	par_19	par_20	par_21	par_22	par_23	par_24	par_25	par_26
par_1	.001																									
par_2	.000	.002																								
par_3	.000	.000	.002																							
par_4	.000	.000	.000	.001																						
par_5	.000	.000	.000	.000	.001																					
par_6	-.002	-.004	-.001	-.001	.001	.015																				
par_7	.000	.000	.000	.000	.000	.000	.005																			
par_8	.000	.000	.000	.000	.000	.000	.000	.004																		
par_9	.000	.000	.000	.000	.000	.000	.000	.000	.004																	
par_10	.000	.000	.000	.000	.000	.000	.000	.000	.000	.004																
par_11	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.005															
par_12	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.003														
par_13	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.004													
par_14	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.004												
par_15	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.030											
par_16	.000	.000	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.001	.004									
par_17	.000	.000	.000	.000	.000	.000	.000	-.001	.001	-.001	.000	-.001	-.002	-.001	-.004	.000	.008									
par_18	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001									
par_19	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000								
par_20	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000							
par_21	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000						
par_22	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000					
par_23	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000				
par_24	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000			
par_25	.000	.000	.000	.000	.000	-.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000		
par_26	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	

[illegible]

Correlations of Estimates (Default model)

[illegible]

par_24	.000	.000	.000	.000	.000	.000	.014	.062	-.071	.077	-.018	.087	.089	.053	.000	.000	-.289
par_25	.077	.130	.043	.055	-.024	-.177	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.036	.000
par_26	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_27	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_28	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_29	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_30	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_31	-.542	-.626	-.499	-.533	-.172	.630	.000	.000	.000	.000	.000	.000	.000	.000	.000	.128	.000
par_32	-.117	.136	-.294	-.250	-.360	-.449	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_33	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_34	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_35	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_36	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_37	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_38	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_39	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_40	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_41	.000	.000	.000	.000	.000	.000	-.385	-.362	-.172	-.382	-.329	-.272	-.368	-.349	.000	.000	.366
par_42	-.008	.009	-.020	-.017	-.024	.000	-.163	-.068	-.208	-.051	-.192	.020	-.022	-.076	.506	-.442	-.471
par_43	-.007	.009	-.019	-.016	-.023	.000	-.153	-.063	-.195	-.048	-.180	.019	-.021	-.071	-.586	-.269	-.146
par_44	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_45	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_46	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_47	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_48	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_49	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_50	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_51	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_52	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_53	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_54	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_55	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_56	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_57	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_58	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_59	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_60	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_61	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_62	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_63	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_64	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_65	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_66	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_67	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_68	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_69	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_70	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_71	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_72	.007	.011	.004	.005	-.002	-.015	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_73	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
par_74	.000	.000	.000	.000	.000	.000	.006	.025	-.029	.031	-.007	.036	.037	.022	.000	.000	-.119
par_75	.000	.000	.000	.000	.000	.000	.002	.008	-.009	.010	-.002	.011	.012	.007	-.075	.009	-.017
par_76	-.003	-.005	-.002	-.002	.001	.007	.000	.000	.000	.000	.000	.000	.000	.000	.000	.002	.000
par_77	.004	.007	.002	.003	-.001	-.010	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000
par_78	.000	.000	.000	.000	.000	.000	-.004	-.018	.021	-.022	.005	-.025	-.026	-.016	.000	.000	.084
par_79	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.147	-.020	-.041
par_80	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.130	.018	.036
par_81	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Bootstrap (Default model)

Summary of Bootstrap Iterations (Default model)

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	2
14	0	0	0
15	0	0	2
16	0	1	0
17	0	0	2

18	0	0	1
19	0	190	2
Total	0	191	9

0 bootstrap samples were unused because of a singular covariance matrix.

0 bootstrap samples were unused because a solution was not found.

200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 200 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 0 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .005

Bootstrap Distributions (Default model)

ML discrepancy (implied vs sample) (Default model)

	191.235	-----
	222.113	*
	252.991	**
	283.869	***
	314.747	*****
	345.624	*****
	376.502	*****
N = 200	407.380	*****
Mean = 395.579	438.258	*****
S. e. = 4.674	469.136	*****
	500.014	*****
	530.892	***
	561.769	*
	592.647	*
	623.525	*

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	81	1987.854	296	.000	6.716
Saturated model	377	.000	0		
Independence model	52	2731.755	325	.000	8.405

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.272	.201	.305	.228	.297
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.911	.248	.271
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	1691.854	1554.418	1836.727
Saturated model	.000	.000	.000
Independence model	2406.755	2243.647	2577.254

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	13.615	11.588	10.647	12.580
Saturated model	.000	.000	.000	.000
Independence model	18.711	16.485	15.367	17.652

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.198	.190	.206	.000
Independence model	.225	.217	.233	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	2149.854	2186.611		

Saturated model	754.000	925.076
Independence model	2835.755	2859.352

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	14.725	13.784	15.717	14.977
Saturated model	5.164	5.164	5.164	6.336
Independence model	19.423	18.306	20.591	19.585

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	25	27
Independence model	20	21

Execution time summary

Minimization: .047
 Miscellaneous: .313
 Bootstrap: 3.310
 Total: 3.670

Appendix IX: AMOS reports, multi-indicator structural models

These documents provide output generated by the AMOS package after calculating the estimates for the seven multi-indicator structural models included in Chapter 8. Selected subsets of these reports are included in the text.

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\MISM-1.amw

Analysis Summary

Date and Time

Date: Monday, 23 November 2009

Time: 7:37:45 p.m.

Title

Mism-1: Monday, 23 November 2009 07:37 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x4
x8

Unobserved, endogenous variables

infQual
usefulness

Unobserved, exogenous variables

err-x2
believability
err-x7
err-x10
relevance
err-x14
timeliness
err-x4
err-x8
err-infQual
err-usefulness
accuracy

Variable counts (Group number 1)

Number of variables in your model: 20
Number of observed variables: 6
Number of unobserved variables: 14
Number of exogenous variables: 12
Number of endogenous variables: 8

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	14	0	6	0	0	20
Labeled	0	0	0	0	0	0
Unlabeled	5	6	6	0	0	17
Total	19	6	12	0	0	37

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x4	1.000	7.000	-.820	-4.059	.331	.818
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
Multivariate					17.188	10.635

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
17	29.741	.000	.006
27	28.137	.000	.000
97	23.369	.001	.000
30	22.361	.001	.000
50	20.732	.002	.000

144	18.599	.005	.000
74	18.265	.006	.000
64	17.435	.008	.000
12	17.237	.008	.000
78	15.892	.014	.000
69	15.500	.017	.000
145	15.148	.019	.000
86	13.485	.036	.003
46	12.809	.046	.008
113	12.716	.048	.005
109	12.309	.055	.008
110	12.146	.059	.006
28	11.847	.065	.008
121	11.563	.072	.010
1	11.269	.080	.014
55	10.641	.100	.061
87	10.613	.101	.040
101	10.550	.103	.029
68	10.252	.114	.047
147	10.222	.116	.031
124	9.934	.127	.052
122	9.934	.127	.032
47	9.420	.151	.115
95	9.404	.152	.083
80	9.017	.173	.183
141	8.631	.195	.349
91	8.188	.225	.612
38	7.961	.241	.709
65	7.956	.241	.643
29	7.790	.254	.700
41	7.709	.260	.694
84	7.631	.266	.687
120	7.569	.271	.667
103	7.390	.286	.741
118	7.260	.297	.776
49	7.197	.303	.764
52	6.872	.333	.904
67	6.781	.342	.911
45	6.763	.343	.888
61	6.326	.388	.984
2	6.264	.394	.983
5	6.253	.395	.976
100	6.203	.401	.974
137	6.198	.401	.963
102	6.193	.402	.948
99	5.790	.447	.995
58	5.772	.449	.992
93	5.663	.462	.995
88	5.424	.491	.999
111	5.325	.503	.999
107	5.314	.504	.999
31	4.923	.554	1.000
72	4.888	.558	1.000
35	4.837	.565	1.000
136	4.690	.584	1.000
135	4.690	.584	1.000
76	4.652	.589	1.000
54	4.652	.589	1.000
34	4.623	.593	1.000
85	4.515	.607	1.000
98	4.499	.610	1.000
14	4.396	.623	1.000
81	4.396	.623	1.000
15	4.350	.629	1.000
117	4.348	.630	1.000
143	4.300	.636	1.000
90	4.300	.636	1.000
77	4.168	.654	1.000
48	4.128	.659	1.000
75	4.043	.671	1.000
79	4.017	.674	1.000
83	4.017	.674	1.000
126	4.014	.675	1.000
70	4.010	.675	1.000
57	4.004	.676	1.000
26	3.990	.678	1.000
20	3.833	.699	1.000
66	3.833	.699	1.000
18	3.779	.707	1.000
112	3.779	.707	1.000
11	3.746	.711	1.000
89	3.648	.724	1.000

		3.648	.724	1.000
51		3.520	.741	1.000
21		3.517	.742	1.000
56		3.328	.767	1.000
132		3.193	.784	1.000
53		3.189	.785	1.000
16		3.189	.785	1.000
116		3.130	.792	1.000
22		3.127	.793	1.000
40		3.019	.806	1.000
82		2.997	.809	1.000
127		2.910	.820	1.000
140		2.895	.822	1.000

Sample Moments (Group number 1)**Sample Covariances (Group number 1)**

	x8	x4	x14	x10	x7	x2
x8	2.177					
x4	.660	1.755				
x14	.619	.578	1.270			
x10	.511	.702	.921	1.863		
x7	.659	.564	.809	.925	1.121	
x2	.529	.563	.790	.891	.751	1.215

Condition number = 14.009

Eigenvalues

5.088 1.671 1.182 .646 .453 .363

Determinant of sample covariance matrix = 1.066

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 21
Number of distinct parameters to be estimated: 17
Degrees of freedom (21 - 17): 4

Result (Default model)

Minimum was achieved

Chi-square = 4.917

Degrees of freedom = 4

Probability level = .296

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

		Estimate	S.E.	C.R.	P	Label
infQual	<--- believability	4.514	9.342	.483	.629	par_1
infQual	<--- timeliness	-.136	1.190	-.114	.909	par_2
infQual	<--- accuracy	1.000				
infQual	<--- relevance	1.509	3.811	.396	.692	par_5
usefulness	<--- infQual	.100	.182	.552	.581	par_4
x7	<--- believability	1.000				
x14	<--- relevance	1.000				
x10	<--- timeliness	1.000				
x4	<--- usefulness	1.000				
x8	<--- usefulness	.999	.204	4.898	***	par_3
x2	<--- accuracy	1.000				

Covariances: (Group number 1 - Default model)

		Estimate	S.E.	C.R.	P	Label
believability	<--> timeliness	.925	.142	6.513	***	par_6
believability	<--> relevance	.809	.119	6.780	***	par_7
relevance	<--> timeliness	.921	.148	6.209	***	par_8
believability	<--> accuracy	.751	.115	6.534	***	par_9
relevance	<--> accuracy	.790	.122	6.484	***	par_10
timeliness	<--> accuracy	.891	.145	6.156	***	par_11

Variances: (Group number 1 - Default model)

--	--	--	--	--	--	--

	Estimate	S.E.	C.R.	P	Label
believability	.921	.131	7.020	***	par_12
relevance	1.070	.149	7.198	***	par_13
timeliness	1.663	.218	7.627	***	par_14
accuracy	1.015	.142	7.138	***	par_15
err-infQual	1.000				
err-usefulness	.281	.125	2.252	.024	par_16
err-x2	.200				
err-x7	.200				
err-x10	.200				
err-x14	.200				
err-x4	1.000				
err-x8	1.477	.225	6.564	***	par_17

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
infQual	.976
usefulness	.599
x8	.321
x4	.412
x14	.843
x10	.893
x7	.822
x2	.835

Matrices (Group number 1 - Default model)

Implied (for all variables) Covariances (Group number 1 - Default model)

	accuracy	timeliness	relevance	believability	infQual	usefulness	x8	x4	x14	x10	x7	x2
accuracy	1.015											
timeliness	.891	1.663										
relevance	.790	.921	1.070									
believability	.751	.925	.809	.921								
infQual	5.474	6.230	5.930	6.004	41.671							
usefulness	.549	.625	.595	.602	4.180	.701						
x8	.549	.625	.594	.602	4.177	.700	2.177					
x4	.549	.625	.595	.602	4.180	.701	.700	1.701				
x14	.790	.921	1.070	.809	5.930	.595	.594	.595	1.270			
x10	.891	1.663	.921	.925	6.230	.625	.625	.625	.921	1.863		
x7	.751	.925	.809	.921	6.004	.602	.602	.602	.809	.925	1.121	
x2	1.015	.891	.790	.751	5.474	.549	.549	.549	.790	.891	.751	1.215

Implied Covariances (Group number 1 - Default model)

	x8	x4	x14	x10	x7	x2
x8	2.177					
x4	.700	1.701				
x14	.594	.595	1.270			
x10	.625	.625	.921	1.863		
x7	.602	.602	.809	.925	1.121	
x2	.549	.549	.790	.891	.751	1.215

Residual Covariances (Group number 1 - Default model)

	x8	x4	x14	x10	x7	x2
x8	.000					
x4	-.040	.054				
x14	.025	-.017	.000			
x10	-.113	.077	.000	.000		
x7	.057	-.039	.000	.000	.000	
x2	-.020	.014	.000	.000	.000	.000

Standardized Residual Covariances (Group number 1 - Default model)

	x8	x4	x14	x10	x7	x2
x8	.000					
x4	-.237	.273				
x14	.169	-.127	.000			
x10	-.649	.491	.000	.000		
x7	.412	-.310	.000	.000	.000	
x2	-.141	.106	.000	.000	.000	.000

Factor Score Weights (Group number 1 - Default model)

	x8	x4	x14	x10	x7	x2
accuracy	.011	.017	.096	.059	.096	.658
timeliness	.004	.006	.054	.790	.089	.059
relevance	.015	.022	.645	.054	.127	.096
believability	.025	.038	.127	.089	.558	.096
infQual	.192	.284	1.619	.432	2.761	1.215

usefulness	.142	.211	.110	.029	.188	.083
------------	------	------	------	------	------	------

Total Effects (Group number 1 - Default model)

	accuracy	timeliness	relevance	believability	usefulness
infQual	1.000	-.136	1.509	4.514	.000
usefulness	.100	-.014	.151	.453	.000
x8	.100	-.014	.151	.452	.999
x4	.100	-.014	.151	.453	1.000
x14	.000	.000	1.000	.000	.000
x10	.000	1.000	.000	.000	.000
x7	.000	.000	.000	1.000	.000
x2	1.000	.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	accuracy	timeliness	relevance	believability	usefulness
infQual	1.000	-.136	1.509	4.514	.000
usefulness	.000	.000	.000	.000	.000
x8	.000	.000	.000	.000	.999
x4	.000	.000	.000	.000	1.000
x14	.000	.000	1.000	.000	.000
x10	.000	1.000	.000	.000	.000
x7	.000	.000	.000	1.000	.000
x2	1.000	.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	accuracy	timeliness	relevance	believability	usefulness
infQual	.000	.000	.000	.000	.000
usefulness	.100	-.014	.151	.453	.000
x8	.100	-.014	.151	.452	.000
x4	.100	-.014	.151	.453	.000
x14	.000	.000	.000	.000	.000
x10	.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	.000
x2	.000	.000	.000	.000	.000

Modification Indices (Group number 1 - Default model)**Covariances: (Group number 1 - Default model)**

	M.I.	Par Change
--	------	------------

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Minimization History (Default model)

Iteration		Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	7		-.290	9999.000	363.092	0	9999.000
1	e	0	584.575		1.070	99.528	19	1.000
2	e	0	331.848		.740	63.735	5	.000
3	e	0	1210.711		.522	25.641	2	.000
4	e	0	2965.135		.682	15.819	2	.000
5	e	1		-.006	.716	6.749	1	1.074
6	e	0	7632.818		.267	5.280	10	.973
7	e	0	16117.350		.379	5.057	1	1.165
8	e	0	24545.624		.443	4.998	1	1.122
9	e	0	66378.116		.271	4.957	1	1.256
10	e	0	62456.585		.473	4.943	1	.752
11	e	0	230842.068		.170	4.925	1	1.107
12	e	0	142845.525		.446	4.924	1	.221
13	e	0	588853.096		.077	4.918	1	1.017
14	e	0	514129.854		.179	4.917	1	.826
15	e	0	771511.206		.012	4.917	1	1.004
16	e	0	768320.524		.005	4.917	1	1.002
17	e	0	793630.758		.000	4.917	1	1.000

Pairwise Parameter Comparisons (Default model)**Variance-covariance Matrix of Estimates (Default model)**

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	87.264																
par_2	-2.181	1.415															
par_3	.167	-.049	.042														

par_4	-1.646	.018	-.004	.033													
par_5	26.795	-.514	.039	-.602	14.525												
par_6	-.001	-.008	.000	.000	.000	.020											
par_7	-.009	.001	.000	.000	-.012	.012	.014										
par_8	.000	-.003	.000	.000	.000	.016	.013	.022									
par_9	.042	-.001	.000	-.001	.016	.012	.010	.010	.013								
par_10	.015	.000	.000	.000	.003	.010	.011	.013	.011	.015							
par_11	-.002	-.002	.000	.000	-.001	.015	.010	.016	.012	.012	.021						
par_12	-.036	.003	.000	.000	.000	.014	.012	.010	.012	.008	.010	.017					
par_13	-.002	.000	.000	.000	-.008	.010	.014	.016	.009	.014	.010	.009	.022				
par_14	.000	.000	.000	.000	.000	.024	.012	.024	.011	.011	.023	.012	.012	.048			
par_15	.020	.000	.000	.000	.007	.009	.008	.010	.012	.013	.015	.008	.009	.011	.020		
par_16	.221	.018	-.009	-.005	.111	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.016	
par_17	-.161	.047	-.016	.003	-.037	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.003	.051

Correlations of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	1.000																
par_2	-.196	1.000															
par_3	.088	-.202	1.000														
par_4	-.969	.083	-.103	1.000													
par_5	.753	-.113	.050	-.869	1.000												
par_6	.000	-.047	.000	.002	-.001	1.000											
par_7	-.008	.007	.000	.008	-.026	.720	1.000										
par_8	.000	-.015	.000	.000	.000	.767	.743	1.000									
par_9	.039	-.004	.000	-.046	.036	.711	.746	.583	1.000								
par_10	.013	-.001	.000	-.014	.006	.559	.750	.704	.771	1.000							
par_11	-.001	-.010	.000	.002	-.001	.741	.564	.731	.739	.708	1.000						
par_12	-.029	.017	.000	.012	.000	.762	.794	.526	.765	.520	.501	1.000					
par_13	-.001	.002	.000	.003	-.014	.484	.794	.727	.513	.759	.464	.460	1.000				
par_14	.000	.002	.000	.000	.000	.762	.449	.727	.451	.423	.720	.410	.359	1.000			
par_15	.015	-.002	.000	-.017	.013	.453	.479	.457	.765	.759	.720	.413	.404	.351	1.000		
par_16	.190	.121	-.339	-.220	.234	-.001	.000	.000	.015	.005	.000	.003	.000	.000	.005	1.000	
par_17	-.076	.176	-.352	.066	-.044	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.101	1.000

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	.000																
par_2	-.482	.000															
par_3	-.377	.910	.000														
par_4	-.464	.199	-3.132	.000													
par_5	-.433	.399	.134	.355	.000												
par_6	-.384	.881	-.298	3.577	-.153	.000											
par_7	-.397	.791	-.805	3.269	-.183	-1.161	.000										
par_8	-.384	.880	-.309	3.498	-.154	-.038	1.127	.000									
par_9	-.403	.741	-1.062	2.962	-.199	-1.731	-.699	-1.379	.000								
par_10	-.399	.774	-.880	3.131	-.189	-1.078	-.220	-1.229	.493	.000							
par_11	-.388	.856	-.433	3.404	-.162	-.331	.655	-.283	1.434	.969	.000						
par_12	-.384	.885	-.321	3.681	-.154	-.039	1.383	.000	1.991	1.056	.220	.000					
par_13	-.369	1.006	.280	4.134	-.115	.980	2.885	1.353	2.395	2.878	1.178	1.016	.000				
par_14	-.305	1.488	2.223	5.503	.040	5.155	4.357	4.941	4.671	4.370	5.090	3.649	2.755	.000			
par_15	-.375	.961	.065	3.931	-.129	.608	1.529	.621	2.883	2.405	1.160	.634	-.343	-3.019	.000		
par_16	-.454	.353	-2.630	.747	-.324	-3.403	-3.055	-3.300	-2.787	-2.923	-3.188	-3.539	-4.062	-5.499	-3.888	.000	
par_17	-.324	1.377	1.355	4.920	-.008	2.075	2.624	2.063	2.876	2.685	2.192	2.134	1.511	-.593	1.735	4.460	.000

Bootstrap (Default model)

Summary of Bootstrap Iterations (Default model)

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	1
11	0	0	0
12	0	0	0
13	0	0	0
14	0	3	0
15	0	1	0
16	0	3	0
17	0	2	0
18	0	2	0
19	0	177	11

Total	0	188	12
-------	---	-----	----

0 bootstrap samples were unused because of a singular covariance matrix.
 92 bootstrap samples were unused because a solution was not found.
 200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 106 bootstrap samples.
 It fit about equally well in 0 bootstrap samples.
 It fit worse or failed to fit in 94 bootstrap samples.
 Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .473

Bootstrap Distributions (Default model)

ML discrepancy (implied vs sample) (Default model)

	.237	*****
	1.961	*****
	3.685	*****
	5.409	*****
	7.133	*****
	8.857	*****
	10.581	****
N = 200	12.305	**
Mean = 5.949	14.029	**
S. e. = .323	15.753	**
	17.477	*
	19.201	
	20.925	*
	22.649	*
	24.373	*

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	17	4.917	4	.296	1.229
Saturated model	21	.000	0		
Independence model	6	357.290	15	.000	23.819

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.037	.989	.941	.188
Saturated model	.000	1.000		
Independence model	.601	.454	.236	.324

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.986	.948	.997	.990	.997
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.267	.263	.266
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	.917	.000	10.868
Saturated model	.000	.000	.000
Independence model	342.290	284.429	407.576

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	.034	.006	.000	.074
Saturated model	.000	.000	.000	.000
Independence model	2.447	2.344	1.948	2.792

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.040	.000	.136	.470
Independence model	.395	.360	.431	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	38.917	40.630	89.755	106.755
Saturated model	42.000	44.115	104.799	125.799
Independence model	369.290	369.895	387.233	393.233

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	.267	.260	.335	.278
Saturated model	.288	.288	.288	.302
Independence model	2.529	2.133	2.977	2.534

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	282	395
Independence model	11	13

Execution time summary

Minimization: .016
 Miscellaneous: .187
 Bootstrap: 1.046
 Total: 1.249

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\MISM-2.amw

Analysis Summary

Date and Time

Date: Monday, 23 November 2009

Time: 7:18:13 p.m.

Title

Mism-2: Monday, 23 November 2009 07:18 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x4
x8
x26
x23

Unobserved, endogenous variables

believability
relevance
timeliness
infQual
usefulness
accuracy

Unobserved, exogenous variables

err-x2
err-x7
err-x10
err-x14
err-x4
err-x8
err-infQual
err-usefulness
Perceived information adequacy
err-x26
err-x23
err-accuracy
err-believability
err-timeliness
err-relevance

Variable counts (Group number 1)

Number of variables in your model: 29
Number of observed variables: 8
Number of unobserved variables: 21
Number of exogenous variables: 15
Number of endogenous variables: 14

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	21	0	7	0	0	28
Labeled	0	0	0	0	0	0
Unlabeled	10	0	8	0	0	18
Total	31	0	15	0	0	46

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x23	1.000	7.000	-.789	-3.904	.144	.357
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x4	1.000	7.000	-.820	-4.059	.331	.818
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094

Multivariate	30.534	14.634
--------------	--------	--------

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
77	36.387	.000	.002
17	35.022	.000	.000
27	30.642	.000	.000
30	29.528	.000	.000
97	26.470	.001	.000
50	22.552	.004	.000
144	22.044	.005	.000
74	21.369	.006	.000
110	20.423	.009	.000
69	20.365	.009	.000
113	19.322	.013	.000
12	18.180	.020	.000
64	17.496	.025	.000
80	16.782	.032	.000
78	16.433	.037	.000
46	16.227	.039	.000
145	16.067	.041	.000
55	15.934	.043	.000
109	15.649	.048	.000
47	15.635	.048	.000
86	15.614	.048	.000
101	14.258	.075	.002
122	13.820	.087	.004
124	13.820	.087	.002
116	13.238	.104	.009
68	13.166	.106	.007
103	13.158	.107	.003
2	13.071	.109	.003
41	13.046	.110	.001
65	12.850	.117	.002
118	12.597	.126	.003
1	12.116	.146	.013
121	11.962	.153	.014
28	11.895	.156	.011
137	11.304	.185	.064
87	10.864	.210	.170
141	10.594	.226	.254
147	10.247	.248	.416
91	9.972	.267	.550
58	9.951	.268	.492
29	9.915	.271	.446
128	9.906	.272	.381
95	9.488	.303	.637
84	9.270	.320	.732
35	8.771	.362	.934
111	8.694	.369	.933
18	8.656	.372	.920
38	8.594	.378	.915
61	8.417	.394	.945
45	8.028	.431	.990
120	7.949	.438	.990
107	7.739	.459	.996
23	7.715	.462	.995
67	7.613	.472	.996
49	7.524	.481	.997
43	7.500	.484	.995
102	7.414	.493	.996
98	7.171	.518	.999
88	7.120	.524	.999
93	7.017	.535	.999
52	6.921	.545	.999
83	6.917	.546	.999
72	6.757	.563	1.000
5	6.603	.580	1.000
100	6.506	.591	1.000
34	6.306	.613	1.000
99	6.056	.641	1.000
81	5.880	.661	1.000
42	5.412	.713	1.000
14	5.399	.714	1.000
31	5.385	.716	1.000
85	5.322	.723	1.000
75	5.217	.734	1.000
125	5.166	.740	1.000
60	5.154	.741	1.000
51	5.050	.752	1.000
21	5.038	.754	1.000

11	5.013	.756	1.000
48	4.945	.763	1.000
143	4.858	.773	1.000
57	4.852	.773	1.000
135	4.756	.783	1.000
136	4.756	.783	1.000
129	4.714	.788	1.000
76	4.686	.791	1.000
54	4.686	.791	1.000
70	4.558	.804	1.000
40	4.525	.807	1.000
117	4.523	.807	1.000
112	4.513	.808	1.000
15	4.504	.809	1.000
90	4.498	.810	1.000
79	4.481	.811	1.000
59	4.353	.824	1.000
66	4.286	.830	1.000
26	4.219	.837	1.000
53	4.198	.839	1.000
126	4.148	.844	1.000
16	4.063	.851	1.000
92	4.014	.856	1.000

Sample Moments (Group number 1)**Sample Covariances (Group number 1)**

	x23	x26	x8	x4	x14	x10	x7	x2
x23	1.855							
x26	.754	1.404						
x8	.766	.417	2.177					
x4	.684	.587	.660	1.755				
x14	.833	.604	.619	.578	1.270			
x10	.963	.615	.511	.702	.921	1.863		
x7	.634	.498	.659	.564	.809	.925	1.121	
x2	.672	.500	.529	.563	.790	.891	.751	1.215

Condition number = 18.622

Eigenvalues

6.403 1.686 1.220 1.102 .844 .616 .445 .344

Determinant of sample covariance matrix = 1.155

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 36
Number of distinct parameters to be estimated: 18
Degrees of freedom (36 - 18): 18

Result (Default model)

Minimum was achieved

Chi-square = 34.477

Degrees of freedom = 18

Probability level = .011

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

			Estimate	S.E.	C.R.	P	Label
timeliness	<---	Perceived information adequacy	1.653	.264	6.253	***	par_6
relevance	<---	Perceived information adequacy	1.476	.226	6.542	***	par_7
believability	<---	Perceived information adequacy	1.362	.212	6.420	***	par_8
accuracy	<---	Perceived information adequacy	1.346	.214	6.277	***	par_9
infQual	<---	believability	2.273	2.273	1.000	.317	par_1
infQual	<---	timeliness	.356	.765	.465	.642	par_2
infQual	<---	accuracy	1.000				
infQual	<---	relevance	1.333	1.535	.868	.385	par_5
usefulness	<---	infQual	.128	.104	1.225	.221	par_4
x7	<---	believability	1.000				
x14	<---	relevance	1.000				
x10	<---	timeliness	1.000				

x4	<--- usefulness	1.000				
x8	<--- usefulness	.999	.204	4.898	***	par_3
x2	<--- accuracy	1.000				
x26	<--- Perceived information adequacy	1.000				
x23	<--- Perceived information adequacy	1.358	.238	5.711	***	par_10

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Perceived information adequacy	.397	.120	3.305	***	par_11
err-accuracy	.496	.071	6.948	***	par_12
err-believability	.384	.060	6.410	***	par_13
err-timeliness	.778	.112	6.960	***	par_14
err-relevance	.404	.066	6.130	***	par_15
err-infQual	1.000				
err-usefulness	.309	.124	2.500	.012	par_16
err-x2	.000				
err-x7	.000				
err-x10	.000				
err-x14	.000				
err-x4	1.000				
err-x26	1.000				
err-x8	1.477	.225	6.564	***	par_17
err-x23	1.123	.144	7.788	***	par_18

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
accuracy	.592
timeliness	.582
relevance	.682
believability	.657
infQual	.958
usefulness	.557
x23	.395
x26	.284
x8	.320
x4	.411
x14	1.000
x10	1.000
x7	1.000
x2	1.000

Matrices (Group number 1 - Default model)
Implied (for all variables) Covariances (Group number 1 - Default model)

	Perceived information adequacy	accuracy	timeliness	relevance	believability	infQual	usefulness	x23	x26	x8	x4	x14	x10	x7
Perceived information adequacy	.397													
accuracy	.535	1.215												
timeliness	.657	.883	1.863											
relevance	.586	.789	.969	1.270										
believability	.541	.728	.894	.799	1.121									
infQual	2.779	4.236	4.870	4.641	4.659	23.745								
usefulness	.356	.542	.623	.594	.596	3.039	.698							
x23	.539	.726	.891	.796	.735	3.773	.483	1.855						
x26	.397	.535	.657	.586	.541	2.779	.356	.539	1.397					
x8	.355	.542	.623	.593	.596	3.036	.697	.482	.355	2.174				
x4	.356	.542	.623	.594	.596	3.039	.698	.483	.356	.697	1.698			
x14	.586	.789	.969	1.270	.799	4.641	.594	.796	.586	.593	.594	1.270		
x10	.657	.883	1.863	.969	.894	4.870	.623	.891	.657	.623	.623	.969	1.863	
x7	.541	.728	.894	.799	1.121	4.659	.596	.735	.541	.596	.596	.799	.894	1.121
x2	.535	1.215	.883	.789	.728	4.236	.542	.726	.535	.542	.542	.789	.883	.728

Implied Covariances (Group number 1 - Default model)

	x23	x26	x8	x4	x14	x10	x7	x2
x23	1.855							
x26	.539	1.397						
x8	.482	.355	2.174					
x4	.483	.356	.697	1.698				
x14	.796	.586	.593	.594	1.270			
x10	.891	.657	.623	.623	.969	1.863		
x7	.735	.541	.596	.596	.799	.894	1.121	
x2	.726	.535	.542	.542	.789	.883	.728	1.215

Residual Covariances (Group number 1 - Default model)

	x23	x26	x8	x4	x14	x10	x7	x2
--	-----	-----	----	----	-----	-----	----	----

x23	.000							
x26	.215	.007						
x8	.283	.062	.003					
x4	.201	.231	-.037	.057				
x14	.036	.018	.026	-.016	.000			
x10	.071	-.041	-.112	.078	-.048	.000		
x7	-.100	-.043	.063	-.033	.010	.031	.000	
x2	-.053	-.035	-.013	.021	.001	.007	.022	.000

Standardized Residual Covariances (Group number 1 - Default model)

	x23	x26	x8	x4	x14	x10	x7	x2
x23	.000							
x26	1.530	.041						
x8	1.657	.418	.011					
x4	1.322	1.768	-.221	.289				
x14	.255	.151	.175	-.120	.000			
x10	.418	-.285	-.640	.503	-.318	.000		
x7	-.748	-.379	.456	-.262	.086	.219	.000	
x2	-.387	-.297	-.092	.162	.009	.051	.197	.000

Factor Score Weights (Group number 1 - Default model)

	x23	x26	x8	x4	x14	x10	x7	x2
Perceived information adequacy	.054	.044	.000	.000	.162	.094	.157	.120
accuracy	.000	.000	.000	.000	.000	.000	.000	1.000
timeliness	.000	.000	.000	.000	.000	1.000	.000	.000
relevance	.000	.000	.000	.000	1.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	1.000	.000
infQual	.000	.000	.056	.083	1.309	.349	2.233	.982
usefulness	.000	.000	.142	.211	.110	.029	.188	.083

Total Effects (Group number 1 - Default model)

	Perceived information adequacy	accuracy	timeliness	relevance	believability	usefulness
accuracy		1.346	.000	.000	.000	.000
timeliness		1.653	.000	.000	.000	.000
relevance		1.476	.000	.000	.000	.000
believability		1.362	.000	.000	.000	.000
infQual		6.997	1.000	.356	1.333	2.273
usefulness		.895	.128	.046	.171	.291
x23		1.358	.000	.000	.000	.000
x26		1.000	.000	.000	.000	.000
x8		.895	.128	.045	.170	.291
x4		.895	.128	.046	.171	.291
x14		1.476	.000	.000	1.000	.000
x10		1.653	.000	1.000	.000	.000
x7		1.362	.000	.000	.000	1.000
x2		1.346	1.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	Perceived information adequacy	accuracy	timeliness	relevance	believability	usefulness
accuracy		1.346	.000	.000	.000	.000
timeliness		1.653	.000	.000	.000	.000
relevance		1.476	.000	.000	.000	.000
believability		1.362	.000	.000	.000	.000
infQual		.000	1.000	.356	1.333	2.273
usefulness		.000	.000	.000	.000	.000
x23		1.358	.000	.000	.000	.000
x26		1.000	.000	.000	.000	.000
x8		.000	.000	.000	.000	.999
x4		.000	.000	.000	.000	1.000
x14		.000	.000	.000	1.000	.000
x10		.000	.000	1.000	.000	.000
x7		.000	.000	.000	.000	1.000
x2		.000	1.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	Perceived information adequacy	accuracy	timeliness	relevance	believability	usefulness
accuracy		.000	.000	.000	.000	.000
timeliness		.000	.000	.000	.000	.000
relevance		.000	.000	.000	.000	.000
believability		.000	.000	.000	.000	.000
infQual		6.997	.000	.000	.000	.000
usefulness		.895	.128	.046	.171	.291
x23		.000	.000	.000	.000	.000
x26		.000	.000	.000	.000	.000
x8		.895	.128	.045	.170	.291
x4		.895	.128	.046	.171	.291
x14		1.476	.000	.000	.000	.000

x10	1.653	.000	.000	.000	.000	.000
x7	1.362	.000	.000	.000	.000	.000
x2	1.346	.000	.000	.000	.000	.000

Modification Indices (Group number 1 - Default model)**Covariances: (Group number 1 - Default model)**

	M.I.	Par Change
err-x23 <--> err-believability	4.560	-.135
err-x23 <--> err-infQual	8.129	1.949
err-x23 <--> err-usefulness	8.129	.249
err-x26 <--> err-infQual	4.063	1.280
err-x26 <--> err-usefulness	4.063	.164
err-x26 <--> err-x23	6.757	.242
err-x8 <--> err-x23	4.482	.253
err-x4 <--> err-x26	5.491	.223
err-x7 <--> err-x23	7.156	-.165

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
x23 <--- x26	4.638	.166
x23 <--- x8	4.236	.127
x4 <--- x26	4.480	.167
x7 <--- x23	4.333	-.087

Minimization History (Default model)

Iteration		Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	2		-.503	9999.000	463.018	0	9999.000
1	e	1		-.063	2.026	193.093	20	.410
2	e	1		-.022	.665	99.567	5	.856
3	e	0	120.532		.929	54.515	6	.730
4	e	0	508.812		.741	39.497	1	1.126
5	e	0	1585.031		.552	35.696	1	1.119
6	e	0	2959.873		.631	35.200	1	.501
7	e	0	24524.046		.232	34.524	1	.945
8	e	0	16938.508		.317	34.485	2	.000
9	e	0	29637.400		.061	34.477	1	1.036
10	e	0	32597.072		.034	34.477	1	1.020
11	e	0	32905.388		.001	34.477	1	1.002
12	e	0	33892.666		.000	34.477	1	.996

Pairwise Parameter Comparisons (Default model)**Variance-covariance Matrix of Estimates (Default model)**

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	5.165																	
par_2	.657	.584																
par_3	.043	-.022	.042															
par_4	-.219	-.043	-.003	.011														
par_5	2.441	.415	.017	-.137	2.355													
par_6	.000	.000	.000	.000	.000	.070												
par_7	.000	.000	.000	.000	.000	.047	.051											
par_8	.000	.000	.000	.000	.000	.045	.039	.045										
par_9	.000	.000	.000	.000	.000	.044	.039	.036	.046									
par_10	.000	.000	.000	.000	.000	.043	.038	.035	.035	.057								
par_11	.000	.000	.000	.000	.000	-.026	-.023	-.021	-.021	-.020	.014							
par_12	.000	.000	.000	.000	.000	.000	.000	.000	-.002	.000	.000	.005						
par_13	.000	.000	.000	.000	.000	-.001	.000	-.002	.000	.000	.000	.000	.004					
par_14	.000	.000	.000	.000	.000	-.004	.000	.000	.000	-.001	.000	.000	.000	.012				
par_15	.000	.000	.000	.000	.000	.001	-.002	.000	.000	.000	.000	.000	.000	-.001	.004			
par_16	.049	.014	-.009	-.002	.033	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.015		
par_17	-.042	.021	-.016	.002	-.016	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.003	.051
par_18	.000	.000	.000	.000	.000	.000	.001	.001	.002	.001	-.002	-.001	.000	-.001	.000	.000	.000	.000

Correlations of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	1.000																	
par_2	.378	1.000																
par_3	.094	-.139	1.000															
par_4	-.923	-.537	-.142	1.000														
par_5	.700	.354	.053	-.857	1.000													
par_6	.000	.000	.000	.000	.000	1.000												
par_7	.000	.000	.000	.000	.000	.791	1.000											

par_8	.000	.000	.000	.000	.000	.796	.822	1.000										
par_9	.000	.000	.000	.000	.000	.774	.801	.800	1.000									
par_10	.000	.000	.000	.000	.000	.686	.710	.691	.679	1.000								
par_11	.000	.000	.000	.000	.000	-.811	-.842	-.837	-.815	-.717	1.000							
par_12	.000	.000	.000	.000	.000	-.013	-.005	-.018	-.128	.005	.024	1.000						
par_13	.000	.000	.000	.000	.000	-.034	-.018	-.154	-.033	.020	.040	-.018	1.000					
par_14	.000	.000	.000	.000	.000	-.127	.016	-.015	-.009	-.023	.019	-.029	-.014	1.000				
par_15	.000	.000	.000	.000	.000	.047	-.120	.024	.024	-.002	-.001	-.074	-.087	-.143	1.000			
par_16	.176	.152	-.361	-.166	.172	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000		
par_17	-.081	.121	-.352	.072	-.046	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	
par_18	.000	.000	.000	.000	.000	.027	.030	.056	.044	-.068	-.032	-.045	-.098	.011	.005	.000	.000	1

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	.000																
par_2	-.910	.000															
par_3	-.563	.787	.000														
par_4	-.905	-.276	-3.599	.000													
par_5	-.579	.673	.217	.741	.000												
par_6	-.271	1.604	1.958	5.365	.206	.000											
par_7	-.349	1.406	1.568	5.422	.092	-1.087	.000										
par_8	-.399	1.268	1.233	5.218	.019	-1.819	-.870	.000									
par_9	-.406	1.247	1.171	5.106	.008	-1.832	-.936	-.121	.000								
par_10	-.401	1.252	1.144	4.735	.016	-1.472	-.670	-.025	.066	.000							
par_11	-.824	.054	-2.542	1.691	-.608	-3.407	-3.238	-3.019	-2.964	-2.871	.000						
par_12	-.782	.183	-2.328	2.908	-.545	-4.212	-4.136	-3.848	-3.624	-3.476	.714	.000					
par_13	-.831	.037	-2.892	2.128	-.617	-4.646	-4.655	-4.266	-4.282	-3.988	-.097	-1.187	.000				
par_14	-.657	.547	-.951	4.248	-.360	-2.918	-2.791	-2.420	-2.339	-2.187	2.343	2.099	3.086	.000			
par_15	-.822	.064	-2.774	2.237	-.604	-4.634	-4.419	-4.340	-4.225	-3.862	.052	-.910	.215	-2.714	.000		
par_16	-.871	-.062	-2.519	1.036	-.674	-4.606	-4.537	-4.289	-4.190	-3.914	-.512	-1.311	-.549	-2.815	-.681	.000	
par_17	-.346	1.456	1.355	5.593	.093	-.505	.004	.373	.423	.366	4.233	4.156	4.692	2.783	4.575	4.361	.000
par_18	-.505	.986	.496	5.588	-.136	-1.780	-1.337	-.957	-.880	-.820	3.807	3.829	4.574	1.901	4.540	4.287	-1.326

Bootstrap (Default model)

Summary of Bootstrap Iterations (Default model)

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	1	0
9	0	0	0
10	0	0	0
11	0	2	0
12	0	3	0
13	0	13	0
14	0	13	0
15	0	7	0
16	0	20	0
17	0	22	0
18	0	22	0
19	0	91	6
Total	0	194	6

0 bootstrap samples were unused because of a singular covariance matrix.

35 bootstrap samples were unused because a solution was not found.

200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 154 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 46 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap $p = .234$

Bootstrap Distributions (Default model)

ML discrepancy (implied vs sample) (Default model)

7.056	**
12.460	*****
17.864	*****
23.269	*****
28.673	*****
34.078	*****
39.482	*****

N = 200	44.887	***
Mean = 27.367	50.291	***
S. e. = .810	55.695	*
	61.100	*
	66.504	
	71.909	
	77.313	
	82.718	*

Model Fit Summary**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	18	34.477	18	.011	1.915
Saturated model	36	.000	0		
Independence model	8	485.286	28	.000	17.332

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.089	.944	.889	.472
Saturated model	.000	1.000		
Independence model	.611	.403	.233	.314

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.929	.889	.965	.944	.964
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.643	.597	.620
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	16.477	3.631	37.108
Saturated model	.000	.000	.000
Independence model	457.286	389.544	532.461

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	.236	.113	.025	.254
Saturated model	.000	.000	.000	.000
Independence model	3.324	3.132	2.668	3.647

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.079	.037	.119	.111
Independence model	.334	.309	.361	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	70.477	72.842	124.305	142.305
Saturated model	72.000	76.730	179.656	215.656
Independence model	501.286	502.337	525.210	533.210

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	.483	.395	.624	.499
Saturated model	.493	.493	.493	.526
Independence model	3.433	2.969	3.948	3.441

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	123	148
Independence model	13	15

Execution time summary

Minimization:	.016
Miscellaneous:	.187
Bootstrap:	.656
Total:	.859

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\mism-3.amw

Analysis Summary

Date and Time

Date: Monday, 23 November 2009

Time: 7:22:11 p.m.

Title

mism-3: Monday, 23 November 2009 07:22 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x4
x8
x26
x23
x28
x22
x20

Unobserved, endogenous variables

believability
relevance
timeliness
usefulness
infQual
accuracy
overall SQ

Unobserved, exogenous variables

err-x2
err-x7
err-x10
err-x14
err-x4
err-x8
err-infQual
Perceived presentation adequacy
err-x26
err-x23
err-believability
err-timeliness
err-relevance
err-x28
er-x20
err-x22
err POLSQ
err-accuracy

Variable counts (Group number 1)

Number of variables in your model: 36
Number of observed variables: 11
Number of unobserved variables: 25
Number of exogenous variables: 18
Number of endogenous variables: 18

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	25	0	7	0	0	32
Labeled	0	0	0	0	0	0
Unlabeled	15	0	11	0	0	26
Total	40	0	18	0	0	58

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x20	1.000	7.000	-1.022	-5.058	1.608	3.979

x22	1.000	7.000	-.756	-3.742	.904	2.238
x28	1.000	7.000	-1.277	-6.320	2.257	5.586
x23	1.000	7.000	-.789	-3.904	.144	.357
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x4	1.000	7.000	-.820	-4.059	.331	.818
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
Multivariate					41.362	14.827

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
17	43.505	.000	.001
77	39.066	.000	.000
30	31.880	.001	.000
97	31.293	.001	.000
27	30.684	.001	.000
65	28.881	.002	.000
50	28.467	.003	.000
64	28.438	.003	.000
48	24.808	.010	.000
80	24.626	.010	.000
74	24.129	.012	.000
109	22.591	.020	.000
144	22.562	.020	.000
113	21.775	.026	.000
55	21.277	.031	.000
69	21.011	.033	.000
110	20.890	.035	.000
41	20.769	.036	.000
2	20.554	.038	.000
12	19.012	.061	.001
145	18.832	.064	.000
52	18.408	.073	.001
78	18.176	.078	.001
86	17.912	.084	.001
38	17.881	.084	.001
116	17.864	.085	.000
56	17.835	.085	.000
98	17.725	.088	.000
81	17.322	.099	.000
46	17.121	.104	.000
91	16.939	.110	.000
47	16.690	.117	.000
68	16.436	.126	.001
101	16.316	.130	.001
58	16.294	.131	.000
124	16.265	.132	.000
122	16.265	.132	.000
103	15.880	.146	.000
1	15.422	.164	.001
16	15.206	.173	.002
43	14.522	.205	.020
123	14.330	.215	.027
22	14.228	.221	.025
141	13.787	.245	.078
35	13.395	.268	.173
18	13.271	.276	.181
118	13.086	.288	.221
28	12.885	.301	.276
120	12.328	.339	.593
40	12.108	.356	.681
121	12.085	.357	.633
137	12.005	.363	.625
128	11.966	.366	.587
67	11.448	.407	.854
111	11.297	.419	.881
87	11.105	.434	.919
147	11.094	.435	.894
61	10.543	.482	.987
138	10.317	.502	.994
29	10.290	.505	.992
49	10.117	.520	.996
95	10.047	.526	.996
84	9.630	.564	1.000
60	9.336	.591	1.000
5	9.211	.602	1.000
45	9.036	.619	1.000
72	8.735	.646	1.000

83	8.585	.660	1.000
23	8.546	.664	1.000
107	8.411	.676	1.000
88	8.291	.687	1.000
85	8.196	.696	1.000
136	8.181	.697	1.000
135	8.181	.697	1.000
11	8.105	.704	1.000
126	7.676	.742	1.000
90	7.574	.751	1.000
143	7.518	.756	1.000
102	7.481	.759	1.000
21	7.328	.772	1.000
93	7.235	.780	1.000
54	7.163	.786	1.000
15	7.063	.794	1.000
42	6.999	.799	1.000
142	6.930	.805	1.000
53	6.902	.807	1.000
100	6.831	.813	1.000
14	6.743	.819	1.000
106	6.696	.823	1.000
108	6.675	.825	1.000
99	6.619	.829	1.000
51	6.614	.829	1.000
34	6.599	.831	1.000
134	6.586	.832	1.000
133	6.586	.832	1.000
125	6.547	.834	1.000
127	6.357	.849	1.000
112	6.306	.852	1.000
82	6.098	.867	1.000
75	6.061	.869	1.000

Sample Moments (Group number 1)

Sample Covariances (Group number 1)

	x20	x22	x28	x23	x26	x8	x4	x14	x10	x7	x2
x20	1.349										
x22	.933	1.570									
x28	.901	.947	1.336								
x23	.774	.957	.722	1.855							
x26	.759	.658	.644	.754	1.404						
x8	.704	.896	.678	.766	.417	2.177					
x4	.704	.753	.821	.684	.587	.660	1.755				
x14	.784	.622	.715	.833	.604	.619	.578	1.270			
x10	.736	.871	.718	.963	.615	.511	.702	.921	1.863		
x7	.642	.591	.679	.634	.498	.659	.564	.809	.925	1.121	
x2	.746	.574	.731	.672	.500	.529	.563	.790	.891	.751	1.215

Condition number = 28.159

Eigenvalues

8.739 1.719 1.301 1.148 .982 .833 .683 .439 .436 .325 .310

Determinant of sample covariance matrix = .242

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 66
Number of distinct parameters to be estimated: 26
Degrees of freedom (66 - 26): 40

Result (Default model)

Minimum was achieved
Chi-square = 74.698
Degrees of freedom = 40
Probability level = .001

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
timeliness	<---	Perceived presentation adequacy	1.547	.229	6.756	***	par_4
relevance	<---	Perceived presentation adequacy	1.380	.194	7.120	***	par_5
believability	<---	Perceived presentation adequacy	1.272	.183	6.969	***	par_6
accuracy	<---	Perceived presentation adequacy	1.252	.184	6.794	***	par_7
infQual	<---	relevance	3.079	3.146	.979	.328	par_3
infQual	<---	timeliness	2.276	2.256	1.009	.313	par_13
infQual	<---	accuracy	1.000				
infQual	<---	believability	2.584	2.840	.910	.363	par_15
usefulness	<---	infQual	-.077	.074	-1.034	.301	par_2
usefulness	<---	Perceived presentation adequacy	2.080	.491	4.238	***	par_14
overall SQ	<---	infQual	-.003	.025	-.133	.894	par_11
overall SQ	<---	usefulness	1.208	.345	3.500	***	par_12
x7	<---	believability	1.000				
x14	<---	relevance	1.000				
x10	<---	timeliness	1.000				
x4	<---	usefulness	1.000				
x8	<---	usefulness	.971	.172	5.647	***	par_1
x2	<---	accuracy	1.000				
x26	<---	Perceived presentation adequacy	1.000				
x23	<---	Perceived presentation adequacy	1.302	.210	6.205	***	par_8
x22	<---	overall SQ	1.017	.098	10.410	***	par_9
x20	<---	overall SQ	.995	.090	11.109	***	par_10
x28	<---	overall SQ	1.000				

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Perceived presentation adequacy	.451	.125	3.607	***	par_16
err-believability	.391	.061	6.378	***	par_17
err-timeliness	.784	.112	6.984	***	par_18
err-relevance	.410	.066	6.201	***	par_19
err-accuracy	.508	.071	7.125	***	par_20
err-infQual	1.000				
err POLSQ	.005	.103	.046	.963	par_21
err-x2	.000				
err-x7	.000				
err-x10	.000				
err-x14	.000				
err-x4	1.000				
err-x26	1.000				
err-x8	1.560	.200	7.783	***	par_22
err-x23	1.090	.134	8.162	***	par_23
err-x28	.418	.068	6.167	***	par_24
er-x20	.440	.070	6.299	***	par_25
err-x22	.620	.091	6.817	***	par_26

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
accuracy	.582
timeliness	.579
relevance	.677
believability	.651
infQual	.988
overall SQ	.995
x20	.673
x22	.604
x28	.686
x23	.412
x26	.311
x8	.283
x4	.395
x14	1.000
x10	1.000
x7	1.000
x2	1.000

Matrices (Group number 1 - Default model)

Implied (for all variables) Covariances (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	infQual	usefulness	overall SQ	x20	x22	x28	x23	x26	x8
Perceived presentation adequacy	.451													
accuracy	.565	1.215												
timeliness	.698	.874	1.863											
relevance	.623	.780	.963	1.270										
believability	.574	.719	.888	.793	1.121									
infQual	5.556	7.466	10.377	8.932	8.079	80.466								

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	infQual	usefulness	overall SQ
accuracy	1.252	.000	.000	.000	.000	.000	.000	.000
timeliness	1.547	.000	.000	.000	.000	.000	.000	.000
relevance	1.380	.000	.000	.000	.000	.000	.000	.000
believability	1.272	.000	.000	.000	.000	.000	.000	.000
infQual	12.312	1.000	2.276	3.079	2.584	.000	.000	.000
usefulness	1.134	-.077	-.175	-.237	-.199	-.077	.000	.000
overall SQ	1.330	-.096	-.219	-.296	-.248	-.096	1.208	.000
x20	1.323	-.096	-.218	-.295	-.247	-.096	1.202	.995

x22	1.353	-.098	-.223	-.301	-.253	-.098	1.229	1.017
x28	1.330	-.096	-.219	-.296	-.248	-.096	1.208	1.000
x23	1.302	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000
x8	1.101	-.075	-.170	-.230	-.193	-.075	.971	.000
x4	1.134	-.077	-.175	-.237	-.199	-.077	1.000	.000
x14	1.380	.000	.000	1.000	.000	.000	.000	.000
x10	1.547	.000	1.000	.000	.000	.000	.000	.000
x7	1.272	.000	.000	.000	1.000	.000	.000	.000
x2	1.252	1.000	.000	.000	.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	infQual	usefulness	overall SQ
accuracy	1.252	.000	.000	.000	.000	.000	.000	.000
timeliness	1.547	.000	.000	.000	.000	.000	.000	.000
relevance	1.380	.000	.000	.000	.000	.000	.000	.000
believability	1.272	.000	.000	.000	.000	.000	.000	.000
infQual	.000	1.000	2.276	3.079	2.584	.000	.000	.000
usefulness	2.080	.000	.000	.000	.000	-.077	.000	.000
overall SQ	.000	.000	.000	.000	.000	-.003	1.208	.000
x20	.000	.000	.000	.000	.000	.000	.000	.995
x22	.000	.000	.000	.000	.000	.000	.000	1.017
x28	.000	.000	.000	.000	.000	.000	.000	1.000
x23	1.302	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000
x8	.000	.000	.000	.000	.000	.000	.971	.000
x4	.000	.000	.000	.000	.000	.000	1.000	.000
x14	.000	.000	.000	1.000	.000	.000	.000	.000
x10	.000	.000	1.000	.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	1.000	.000	.000	.000
x2	.000	1.000	.000	.000	.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	infQual	usefulness	overall SQ
accuracy	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	.000	.000	.000	.000	.000	.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000
infQual	12.312	.000	.000	.000	.000	.000	.000	.000
usefulness	-.946	-.077	-.175	-.237	-.199	.000	.000	.000
overall SQ	1.330	-.096	-.219	-.296	-.248	-.093	.000	.000
x20	1.323	-.096	-.218	-.295	-.247	-.096	1.202	.000
x22	1.353	-.098	-.223	-.301	-.253	-.098	1.229	.000
x28	1.330	-.096	-.219	-.296	-.248	-.096	1.208	.000
x23	.000	.000	.000	.000	.000	.000	.000	.000
x26	.000	.000	.000	.000	.000	.000	.000	.000
x8	1.101	-.075	-.170	-.230	-.193	-.075	.000	.000
x4	1.134	-.077	-.175	-.237	-.199	-.077	.000	.000
x14	1.380	.000	.000	.000	.000	.000	.000	.000
x10	1.547	.000	.000	.000	.000	.000	.000	.000
x7	1.272	.000	.000	.000	.000	.000	.000	.000
x2	1.252	.000	.000	.000	.000	.000	.000	.000

Modification Indices (Group number 1 - Default model)**Covariances: (Group number 1 - Default model)**

	M.I.	Par Change
er-x20 <--> err-relevance	4.355	.085
err-x22 <--> err-accuracy	6.603	-.133
err-x22 <--> err-timeliness	5.735	.154
err-x22 <--> err-relevance	4.677	-.102
err-x23 <--> err-believability	4.806	-.124
err-x23 <--> err-x22	7.232	.202
err-x26 <--> er-x20	4.360	.130
err-x14 <--> err-x22	4.296	-.107
err-x10 <--> err-x22	8.066	.194
err-x7 <--> err-x23	4.837	-.133
err-x2 <--> err-x22	6.520	-.136

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Minimization History (Default model)

	Negative	Smallest
--	----------	----------

Iteration	eigenvalues	Condition #	eigenvalue	Diameter	F	NTries	Ratio
0	e	5	-.780	9999.000	830.148	0	9999.000
1	e	4	-.370	2.986	390.055	22	.285
2	e	3	-.367	1.060	212.181	6	.656
3	e	2	-.072	.362	154.120	5	.751
4	e	1	-.081	.710	109.941	6	.634
5	e	1	-.001	.456	90.437	5	.734
6	e	1	-.017	.372	82.519	5	.709
7	e	0	5939.296	.328	79.409	5	.671
8	e	1	-.035	1.002	78.418	1	.249
9	e	0	15957.067	.186	75.818	5	.773
10	e	0	53122.701	.410	75.215	1	1.155
11	e	0	55959.156	.366	75.018	2	.000
12	e	0	112123.942	.382	74.859	1	1.294
13	e	0	178281.575	.377	74.776	1	1.243
14	e	0	361596.994	.283	74.731	1	1.293
15	e	0	475614.050	.307	74.710	1	1.182
16	e	0	965077.111	.170	74.701	1	1.231
17	e	0	1145210.501	.166	74.699	1	1.124
18	e	0	1606732.735	.051	74.698	1	1.098
19	e	0	1679908.368	.017	74.698	1	1.026
20	e	0	1660416.080	.001	74.698	1	1.001

Pairwise Parameter Comparisons (Default model)

Variance-covariance Matrix of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	.030																	
par_2	.001	.006																
par_3	-.004	.198	9.894															
par_4	.000	.000	-.015	.052														
par_5	.000	.000	.042	.033	.038													
par_6	.000	.000	-.007	.032	.028	.033												
par_7	.000	-.002	-.098	.031	.027	.026	.034											
par_8	.000	-.001	-.001	.032	.028	.026	.026	.044										
par_9	.001	.000	-.011	.000	.000	.000	.000	.000	.010									
par_10	.000	.000	-.001	.000	.000	.000	.000	.000	.004	.008								
par_11	.000	-.001	.005	.000	.000	.000	.000	.000	.000	.000	.001							
par_12	.014	.010	.032	-.001	-.001	.003	-.001	-.005	-.004	-.004	-.008	.119						
par_13	.000	.142	6.031	.035	.000	.001	-.066	.004	-.015	.002	.006	-.001	5.090					
par_14	-.024	-.013	.012	.056	.050	.043	.044	.047	-.001	-.002	.008	-.120	.004	.241				
par_15	-.006	.188	6.961	.003	.003	.052	-.073	-.021	-.006	.008	-.012	.243	5.036	-.180	8.065			
par_16	.000	.000	.006	-.022	-.020	-.018	-.018	-.018	.000	.000	.000	.000	.002	-.031	.000	.016		
par_17	.000	-.001	-.002	.000	.000	-.002	.000	.000	.000	.000	.000	-.004	-.004	.000	-.059	.000	.004	
par_18	.000	.000	.005	-.003	.000	.000	.000	-.001	.000	.000	.000	.001	-.055	-.004	-.003	.000	.000	
par_19	.000	.000	-.064	.001	-.002	.000	.000	.000	.000	.000	.000	.001	-.005	-.004	-.003	.000	.000	-
par_20	.000	.002	.100	.000	.000	.000	-.002	.000	.000	.000	.000	.001	.071	-.002	.083	.000	.000	
par_21	.000	-.002	.005	.000	.000	-.001	.000	.002	.000	.000	.002	-.028	.014	.022	-.059	.000	.001	
par_22	-.003	.002	.009	.000	.000	.000	.000	-.001	.000	.000	-.001	.018	-.001	-.016	.053	.000	-.001	
par_23	.000	.000	-.009	.000	.000	.001	.000	-.001	.000	.000	.000	.006	-.009	-.004	.025	.000	-.001	
par_24	.001	.000	-.006	.000	.000	.000	.000	.000	.001	.002	.000	.000	-.006	-.002	.002	.000	.000	
par_25	.000	.000	-.005	.000	.000	.000	.000	.000	.000	-.002	.000	.000	-.009	.002	-.012	.000	.000	
par_26	-.001	.000	.014	.000	.000	.000	.000	.000	-.002	.000	.000	.000	.021	.000	.013	.000	.000	

Correlations of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	1.000																	
par_2	.063	1.000																
par_3	-.008	.848	1.000															
par_4	.004	-.007	-.020	1.000														
par_5	.005	.001	.069	.752	1.000													
par_6	.005	.022	-.012	.756	.785	1.000												
par_7	.004	-.167	-.168	.732	.762	.758	1.000											
par_8	.009	-.033	-.002	.666	.694	.669	.662	1.000										
par_9	.034	-.033	-.034	-.003	-.003	-.002	-.005	.010	1.000									
par_10	.021	.030	-.002	-.014	-.015	-.012	-.015	-.011	.451	1.000								
par_11	-.002	-.318	.068	.015	.015	-.051	.000	.073	-.027	-.024	1.000							
par_12	.228	.399	.030	-.010	-.009	.050	-.014	-.073	-.109	-.114	-.880	1.000						
par_13	-.001	.845	.850	.068	.001	.001	-.158	.009	-.066	.009	.100	-.001	1.000					
par_14	-.282	-.350	.008	.496	.525	.479	.486	.461	-.018	-.049	.614	-.707	.003	1.000				
par_15	-.013	.890	.779	.005	.006	.100	-.140	-.035	-.022	.030	-.175	.248	.786	-.129	1.000			
par_16	-.005	.017	.016	-.776	-.810	-.803	-.779	-.701	.002	.014	-.001	.001	.008	-.500	-.001	1.000		
par_17	-.001	-.131	-.011	-.022	-.007	-.159	-.022	.028	.000	-.003	.179	-.174	-.030	.014	-.339	.031	1.000	
par_18	.002	-.025	.015	-.133	.020	-.012	-.009	-.025	.003	.003	-.039	.027	-.215	-.078	-.011	.018	-.019	1
par_19	-.001	-.066	-.310	.050	-.127	.027	.022	-.008	.003	.004	-.053	.029	-.031	-.113	-.018	-.002	-.085	-
par_20	.002	.440	.448	-.005	.002	-.007	-.117	-.005	.009	.005	.004	.040	.440	-.045	.413	.014	-.022	-
par_21	.000	-.239	.015	.010	.008	-.044	.011	.084	-.032	-.032	.760	-.783	.060	.431	-.201	-.002	.158	-
par_22	-.078	.111	.014	-.009	-.009	.009	-.009	-.035	-.019	.005	-.265	.262	-.002	-.163	.093	.007	-.052	
par_23	-.009	.033	-.022	.003	.005	.039	.014	-.039	-.021	-.001	-.134	.135	-.029	-.056	.067	-.010	-.101	
par_24	.043	.002	-.027	-.014	-.015	-.012	-.016	-.003	.205	.248	-.048	.010	-.041	-.056	.010	.013	-.002	

par_25	-.008	-.068	-.021	.019	.020	.017	.019	.023	-.014	-.244	.021	-.008	-.060	.062	-.061	-.019	.004	-
par_26	-.034	.068	.048	-.005	-.006	-.006	-.003	-.021	-.197	.010	.026	-.002	.101	-.008	.052	.006	-.002	-

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	.000																
par_2	-5.727	.000															
par_3	.669	1.024	.000														
par_4	2.015	6.730	-.485	.000													
par_5	1.585	7.021	-.541	-1.089	.000												
par_6	1.206	6.897	-.573	-1.826	-.873	.000											
par_7	1.120	6.332	-.574	-1.876	-.980	-.156	.000										
par_8	1.226	6.131	-.564	-1.357	-.496	.183	.302	.000									
par_9	.239	8.773	-.655	-2.124	-1.671	-1.230	-1.124	-1.234	.000								
par_10	.128	9.347	-.662	-2.232	-1.793	-1.355	-1.247	-1.338	-.223	.000							
par_11	-5.606	.859	-.981	-6.741	-7.092	-6.876	-6.750	-6.231	-10.054	-10.672	.000						
par_12	.681	3.982	-.593	-.814	-.434	-.168	-.113	-.225	.517	.581	3.299	.000					
par_13	.577	1.073	-.470	.324	.396	.444	.447	.431	.556	.568	1.012	.468	.000				
par_14	1.967	4.136	-.314	1.251	1.656	1.862	1.915	1.785	2.117	2.155	4.378	1.127	-.085	.000			
par_15	.567	.959	-.247	.364	.423	.464	.464	.449	.551	.560	.910	.496	.175	.171	.000		
par_16	-2.438	3.657	-.835	-3.266	-3.054	-2.806	-2.739	-2.739	-3.569	-3.559	3.563	-2.062	-.808	-2.888	-.750	.000	
par_17	-3.177	4.570	-.854	-4.850	-4.857	-4.372	-4.407	-4.200	-5.432	-5.562	6.368	-2.265	-.835	-3.421	-.767	-.439	.00
par_18	-.912	6.321	-.730	-2.846	-2.687	-2.268	-2.163	-2.155	-1.572	-1.475	6.790	-1.179	-.654	-2.532	-.633	1.996	3.04
par_19	-3.045	4.740	-.843	-4.835	-4.564	-4.481	-4.333	-4.046	-5.155	-5.269	5.749	-2.284	-.826	-3.323	-.765	-.292	.20
par_20	-2.491	7.582	-.826	-4.327	-4.229	-3.893	-3.629	-3.580	-4.234	-4.273	6.778	-2.004	-.795	-3.151	-.739	.393	1.22
par_21	-4.825	.581	-.977	-6.167	-6.291	-5.941	-5.941	-5.747	-7.032	-7.158	.095	-2.797	-1.009	-4.550	-.901	-2.756	-3.47
par_22	2.151	7.951	-.482	.045	.641	1.066	1.125	.875	2.416	2.577	7.500	1.004	-.316	-.929	-.362	4.707	5.45
par_23	.547	7.745	-.631	-1.723	-1.235	-.820	-.717	-.836	.437	.590	7.860	-.334	-.524	-1.919	-.527	3.474	4.58
par_24	-3.036	4.924	-.845	-4.708	-4.664	-4.369	-4.225	-4.005	-5.605	-5.887	5.745	-2.250	-.822	-3.329	-.763	-.234	.25
par_25	-2.853	4.903	-.838	-4.648	-4.593	-4.283	-4.147	-3.926	-4.774	-4.396	6.016	-2.178	-.812	-3.337	-.754	-.078	.52
par_26	-1.779	6.139	-.783	-3.755	-3.543	-3.191	-3.073	-2.959	-2.721	-2.955	6.654	-1.647	-.737	-2.921	-.692	1.094	2.08

Bootstrap (Default model)**Summary of Bootstrap Iterations (Default model)**

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	1
15	0	0	1
16	0	0	0
17	0	1	0
18	0	1	1
19	0	175	20
Total	0	177	23

0 bootstrap samples were unused because of a singular covariance matrix.

37 bootstrap samples were unused because a solution was not found.

200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 184 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 16 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .085

Bootstrap Distributions (Default model)**ML discrepancy (implied vs sample) (Default model)**

	23.873	*****
	30.345	*****
	36.817	*****
	43.289	*****
	49.761	*****
	56.233	*****
	62.705	*****
N = 200	69.177	*****

Mean = 52.663	75.650	****
S. e. = 1.124	82.122	**
	88.594	***
	95.066	*
	101.538	*
	108.010	
	114.482	*

Model Fit Summary**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	26	74.698	40	.001	1.867
Saturated model	66	.000	0		
Independence model	11	865.369	55	.000	15.734

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.067	.921	.870	.558
Saturated model	.000	1.000		
Independence model	.662	.295	.154	.246

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.914	.881	.958	.941	.957
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.727	.664	.696
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	34.698	14.187	63.024
Saturated model	.000	.000	.000
Independence model	810.369	718.815	909.342

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	.512	.238	.097	.432
Saturated model	.000	.000	.000	.000
Independence model	5.927	5.550	4.923	6.228

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.077	.049	.104	.054
Independence model	.318	.299	.337	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	126.698	131.355	204.449	230.449
Saturated model	132.000	143.821	329.369	395.369
Independence model	887.369	889.339	920.263	931.263

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	.868	.727	1.062	.900
Saturated model	.904	.904	.904	.985
Independence model	6.078	5.451	6.756	6.091

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	109	125
Independence model	13	14

Execution time summary

Minimization:	.016
Miscellaneous:	.219
Bootstrap:	1.360
Total:	1.595

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\MISM-4.amw

Analysis Summary

Date and Time

Date: Monday, 23 November 2009

Time: 7:25:13 p.m.

Title

Mism-4: Monday, 23 November 2009 07:25 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x26
x23
x4
x17
x13
x8

Unobserved, endogenous variables

believability
relevance
timeliness
usefulness
infQual
accuracy
ease of use

Unobserved, exogenous variables

err-x2
err-x7
err-x10
err-x14
err-infQual
Perceived presentation adequacy
err-x26
err-x23
err-accuracy
err-believability
err-timeliness
err-relevance
err-ease of use
err-x13
err-x4
err-x8
err-x17

Variable counts (Group number 1)

Number of variables in your model: 34
Number of observed variables: 10
Number of unobserved variables: 24
Number of exogenous variables: 17
Number of endogenous variables: 17

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	24	0	8	0	0	32
Labeled	0	0	0	0	0	0
Unlabeled	15	0	9	0	0	24
Total	39	0	17	0	0	56

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x13	1.000	7.000	-.871	-4.309	.278	.688
x17	2.000	7.000	-.640	-3.168	.092	.227

x4	1.000	7.000	-.820	-4.059	.331	.818
x23	1.000	7.000	-.789	-3.904	.144	.357
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
Multivariate					45.881	17.954

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
27	54.042	.000	.000
17	38.902	.000	.000
77	38.300	.000	.000
30	30.625	.001	.000
144	29.474	.001	.000
97	29.014	.001	.000
74	27.716	.002	.000
50	25.483	.005	.000
65	21.100	.020	.003
110	20.874	.022	.002
69	20.722	.023	.001
80	20.642	.024	.000
113	19.529	.034	.002
58	19.512	.034	.001
64	19.464	.035	.000
46	19.017	.040	.000
12	18.871	.042	.000
68	17.997	.055	.001
101	17.694	.060	.001
103	17.387	.066	.002
78	17.048	.073	.003
55	16.940	.076	.002
109	16.934	.076	.001
47	16.903	.077	.000
141	16.844	.078	.000
86	16.560	.085	.000
2	16.453	.087	.000
145	16.230	.093	.000
60	16.177	.095	.000
41	16.148	.095	.000
107	14.686	.144	.018
121	14.441	.154	.025
122	14.300	.160	.025
124	14.300	.160	.015
116	14.076	.170	.021
1	13.738	.185	.043
29	13.720	.186	.030
118	13.359	.204	.066
52	12.872	.231	.185
48	12.861	.232	.143
18	12.480	.254	.273
28	12.350	.262	.287
128	11.918	.291	.511
70	11.844	.296	.492
40	11.607	.312	.594
91	11.571	.315	.550
137	11.418	.326	.594
43	11.408	.327	.532
127	11.335	.332	.517
87	11.334	.332	.448
35	10.681	.383	.837
147	10.667	.384	.799
84	10.328	.412	.913
95	10.286	.416	.899
34	10.111	.431	.930
14	9.289	.505	.999
49	9.260	.508	.999
45	9.257	.508	.998
67	9.213	.512	.997
123	9.183	.515	.996
111	9.069	.526	.997
72	8.857	.546	.999
100	8.827	.549	.999
38	8.748	.556	.999
120	8.706	.560	.998
61	8.691	.562	.998
8	8.666	.564	.997
102	8.584	.572	.997
11	8.461	.584	.998
23	8.223	.607	1.000

20	7.998	.629	1.000
83	7.840	.644	1.000
98	7.769	.651	1.000
88	7.764	.652	1.000
90	7.717	.656	1.000
93	7.497	.678	1.000
92	7.249	.702	1.000
125	7.121	.714	1.000
5	7.097	.716	1.000
51	7.025	.723	1.000
81	6.567	.766	1.000
99	6.520	.770	1.000
119	6.478	.774	1.000
57	6.402	.780	1.000
54	6.364	.784	1.000
136	6.230	.796	1.000
135	6.230	.796	1.000
112	6.052	.811	1.000
75	5.878	.825	1.000
31	5.854	.827	1.000
42	5.709	.839	1.000
129	5.695	.840	1.000
85	5.623	.846	1.000
59	5.554	.851	1.000
21	5.339	.867	1.000
143	5.301	.870	1.000
15	5.281	.872	1.000
117	4.811	.903	1.000
26	4.805	.904	1.000
76	4.729	.909	1.000

Sample Moments (Group number 1)**Sample Covariances (Group number 1)**

	x8	x13	x17	x4	x23	x26	x14	x10	x7	x2
x8	2.177									
x13	.602	1.954								
x17	.330	.620	1.218							
x4	.660	.819	.337	1.755						
x23	.766	.531	.555	.684	1.855					
x26	.417	.530	.425	.587	.754	1.404				
x14	.619	.725	.613	.578	.833	.604	1.270			
x10	.511	.823	.658	.702	.963	.615	.921	1.863		
x7	.659	.499	.437	.564	.634	.498	.809	.925	1.121	
x2	.529	.451	.458	.563	.672	.500	.790	.891	.751	1.215

Condition number = 23.474

Eigenvalues

7.384 1.731 1.553 1.200 1.066 .844 .725 .590 .426 .315

Determinant of sample covariance matrix = 1.228

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 55
Number of distinct parameters to be estimated: 24
Degrees of freedom (55 - 24): 31

Result (Default model)

Minimum was achieved

Chi-square = 43.931

Degrees of freedom = 31

Probability level = .062

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

			Estimate	S.E.	C.R.	P	Label
timeliness	<---	Perceived presentation adequacy	1.660	.260	6.381	***	par_3
relevance	<---	Perceived presentation adequacy	1.488	.223	6.686	***	par_4
believability	<---	Perceived presentation adequacy	1.326	.205	6.470	***	par_5

accuracy	<---	Perceived presentation adequacy	1.315	.208	6.331	***	par_6
infQual	<---	accuracy	1.000				
infQual	<---	relevance	2.306	1.990	1.159	.247	par_2
infQual	<---	believability	-.293	.948	-.309	.757	par_14
infQual	<---	timeliness	2.133	1.679	1.270	.204	par_15
usefulness	<---	infQual	-.079	.065	-1.218	.223	par_1
ease of use	<---	Perceived presentation adequacy	1.112	.225	4.954	***	par_12
usefulness	<---	Perceived presentation adequacy	1.091	.643	1.696	.090	par_13
x7	<---	believability	1.000				
x14	<---	relevance	1.000				
x10	<---	timeliness	1.000				
x2	<---	accuracy	1.000				
x26	<---	Perceived presentation adequacy	1.000				
x23	<---	Perceived presentation adequacy	1.393	.235	5.919	***	par_7
x4	<---	ease of use	.598	.206	2.906	.004	par_8
x17	<---	Perceived presentation adequacy	.613	.269	2.279	.023	par_9
x17	<---	ease of use	.267	.188	1.416	.157	par_10
x13	<---	ease of use	1.000				
x4	<---	usefulness	1.000				
x8	<---	usefulness	2.650	1.610	1.646	.100	par_11

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Perceived presentation adequacy	.402	.120	3.356	***	par_16
err-accuracy	.521	.072	7.203	***	par_17
err-believability	.416	.061	6.799	***	par_18
err-timeliness	.756	.108	6.971	***	par_19
err-relevance	.381	.062	6.101	***	par_20
err-infQual	1.000				
err-ease of use	.500	.168	2.976	.003	par_21
err-x2	.000				
err-x7	.000				
err-x10	.000				
err-x14	.000				
err-x4	1.000				
err-x26	1.000				
err-x13	1.000				
err-x23	1.075	.135	7.973	***	par_22
err-x8	1.279	.278	4.600	***	par_23
err-x17	.850	.107	7.923	***	par_24

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
accuracy	.571
timeliness	.594
relevance	.700
believability	.629
infQual	.969
ease of use	.498
x8	.421
x13	.499
x17	.302
x4	.425
x23	.420
x26	.287
x14	1.000
x10	1.000
x7	1.000
x2	1.000

Matrices (Group number 1 - Default model)

Implied (for all variables) Covariances (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	infQual	ease of use	usefulness	x8	x13	x17	x4	x23	x26	x
Perceived presentation adequacy	.402														
accuracy	.528	1.215													
timeliness	.667	.877	1.863												
relevance	.598	.786	.992	1.270											
believability	.532	.700	.884	.792	1.121										
infQual	3.172	4.691	6.879	5.597	4.083	32.074									
ease of use	.447	.587	.742	.665	.592	3.528	.997								
usefulness	.189	.207	.187	.212	.260	.940	.210	.132							
x8	.501	.550	.495	.562	.689	2.491	.557	.351	2.208						
x13	.447	.587	.742	.665	.592	3.528	.997	.210	.557	1.997					
x17	.365	.480	.606	.543	.484	2.885	.540	.172	.455	.540	1.218				
x4	.456	.559	.630	.609	.614	3.049	.806	.258	.684	.806	.494	1.740			

x23	.560	.736	.929	.833	.742	4.420	.622	.263	.698	.622	.509	.635	1.855	
x26	.402	.528	.667	.598	.532	3.172	.447	.189	.501	.447	.365	.456	.560	1.402
x14	.598	.786	.992	1.270	.792	5.597	.665	.212	.562	.665	.543	.609	.833	.598
x10	.667	.877	1.863	.992	.884	6.879	.742	.187	.495	.742	.606	.630	.929	.667
x7	.532	.700	.884	.792	1.121	4.083	.592	.260	.689	.592	.484	.614	.742	.532
x2	.528	1.215	.877	.786	.700	4.691	.587	.207	.550	.587	.480	.559	.736	.528

Implied Covariances (Group number 1 - Default model)

	x8	x13	x17	x4	x23	x26	x14	x10	x7	x2
x8	2.208									
x13	.557	1.997								
x17	.455	.540	1.218							
x4	.684	.806	.494	1.740						
x23	.698	.622	.509	.635	1.855					
x26	.501	.447	.365	.456	.560	1.402				
x14	.562	.665	.543	.609	.833	.598	1.270			
x10	.495	.742	.606	.630	.929	.667	.992	1.863		
x7	.689	.592	.484	.614	.742	.532	.792	.884	1.121	
x2	.550	.587	.480	.559	.736	.528	.786	.877	.700	1.215

Residual Covariances (Group number 1 - Default model)

	x8	x13	x17	x4	x23	x26	x14	x10	x7	x2
x8	-.031									
x13	.045	-.043								
x17	-.125	.081	.000							
x4	-.024	.013	-.157	.015						
x23	.068	-.092	.046	.049	.000					
x26	-.084	.084	.060	.131	.195	.002				
x14	.057	.061	.069	-.031	.000	.007	.000			
x10	.016	.081	.052	.072	.034	-.051	-.071	.000		
x7	-.030	-.093	-.047	-.050	-.107	-.034	.017	.041	.000	
x2	-.021	-.136	-.023	.004	-.063	-.028	.005	.014	.051	.000

Standardized Residual Covariances (Group number 1 - Default model)

	x8	x13	x17	x4	x23	x26	x14	x10	x7	x2
x8	-.120									
x13	.248	-.183								
x17	-.887	.591	.000							
x4	-.138	.078	-1.236	.075						
x23	.385	-.548	.347	.309	.000					
x26	-.553	.584	.531	.973	1.378	.014				
x14	.390	.425	.615	-.235	.000	.057	.000			
x10	.094	.473	.388	.453	.196	-.354	-.466	.000		
x7	-.213	-.701	-.447	-.399	-.800	-.303	.142	.294	.000	
x2	-.147	-.989	-.209	.032	-.458	-.241	.037	.100	.449	.000

Factor Score Weights (Group number 1 - Default model)

	x8	x13	x17	x4	x23	x26	x14	x10	x7	x2
Perceived presentation adequacy	.068	.013	.026	.040	.040	.031	.162	.105	.094	.096
accuracy	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
timeliness	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
infQual	-.119	.020	.019	-.045	.022	.017	2.308	2.111	-.230	1.015
ease of use	.026	.296	.102	.189	.017	.013	.096	.071	.035	.052
usefulness	.083	.012	.027	.048	.042	.033	-.005	-.051	.121	.025

Total Effects (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	ease of use	usefulness
accuracy		1.315	.000	.000	.000	.000	.000
timeliness		1.660	.000	.000	.000	.000	.000
relevance		1.488	.000	.000	.000	.000	.000
believability		1.326	.000	.000	.000	.000	.000
infQual		7.898	1.000	2.133	2.306	-.293	.000
ease of use		1.112	.000	.000	.000	.000	.000
usefulness		.470	-.079	-.168	-.181	.023	.000
x8		1.247	-.208	-.444	-.480	.061	.000
x13		1.112	.000	.000	.000	.000	1.000
x17		.910	.000	.000	.000	.000	.267
x4		1.135	-.079	-.168	-.181	.023	.598
x23		1.393	.000	.000	.000	.000	.000
x26		1.000	.000	.000	.000	.000	.000
x14		1.488	.000	.000	1.000	.000	.000
x10		1.660	.000	1.000	.000	.000	.000
x7		1.326	.000	.000	.000	1.000	.000
x2		1.315	1.000	.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	ease of use	usefulness
accuracy	1.315	.000	.000	.000	.000	.000	.000
timeliness	1.660	.000	.000	.000	.000	.000	.000
relevance	1.488	.000	.000	.000	.000	.000	.000
believability	1.326	.000	.000	.000	.000	.000	.000
infQual	.000	1.000	2.133	2.306	-.293	.000	.000
ease of use	1.112	.000	.000	.000	.000	.000	.000
usefulness	1.091	.000	.000	.000	.000	.000	.000
x8	.000	.000	.000	.000	.000	.000	2.650
x13	.000	.000	.000	.000	.000	1.000	.000
x17	.613	.000	.000	.000	.000	.267	.000
x4	.000	.000	.000	.000	.000	.598	1.000
x23	1.393	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000
x14	.000	.000	.000	1.000	.000	.000	.000
x10	.000	.000	1.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	1.000	.000	.000
x2	.000	1.000	.000	.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	Perceived presentation adequacy	accuracy	timeliness	relevance	believability	ease of use	usefulness
accuracy	.000	.000	.000	.000	.000	.000	.000
timeliness	.000	.000	.000	.000	.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000
infQual	7.898	.000	.000	.000	.000	.000	.000
ease of use	.000	.000	.000	.000	.000	.000	.000
usefulness	-.621	-.079	-.168	-.181	.023	.000	.000
x8	1.247	-.208	-.444	-.480	.061	.000	.000
x13	1.112	.000	.000	.000	.000	.000	.000
x17	.297	.000	.000	.000	.000	.000	.000
x4	1.135	-.079	-.168	-.181	.023	.000	.000
x23	.000	.000	.000	.000	.000	.000	.000
x26	.000	.000	.000	.000	.000	.000	.000
x14	1.488	.000	.000	.000	.000	.000	.000
x10	1.660	.000	.000	.000	.000	.000	.000
x7	1.326	.000	.000	.000	.000	.000	.000
x2	1.315	.000	.000	.000	.000	.000	.000

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

	M.I.	Par Change
err-x13 <--> err-accuracy	4.113	-.152
err-x23 <--> err-believability	4.562	-.130
err-x26 <--> err-x23	5.617	.213
err-x7 <--> err-x23	4.662	-.130

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Minimization History (Default model)

Iteration	Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	3	-.723	9999.000	552.264	0	9999.000
1	e	2	-.133	2.203	221.323	20	.344
2	e	2	-.087	.883	121.238	4	.686
3	e	1	-.009	.814	76.142	6	.834
4	e	0	1326.805	.836	57.775	6	1.023
5	e	0	2727.303	.533	53.190	4	.000
6	e	1	-.060	1.410	50.865	1	.314
7	e	1	-.002	.163	45.465	7	.805
8	e	0	37883.478	.324	44.469	9	.936
9	e	1	-.001	.567	44.250	2	.000
10	e	0	207978.064	.194	43.985	5	.866
11	e	0	187468.189	.389	43.953	2	.000
12	e	0	291733.606	.147	43.934	1	.787
13	e	0	245781.454	.067	43.931	1	1.090
14	e	0	233272.815	.009	43.931	1	1.017
15	e	0	239782.951	.000	43.931	1	1.000

Pairwise Parameter Comparisons (Default model)

Variance-covariance Matrix of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	.004																	
par_2	.081	3.959																
par_3	.000	-.017	.068															
par_4	.000	.016	.046	.050														
par_5	.000	-.012	.042	.037	.042													
par_6	-.001	-.048	.041	.037	.033	.043												
par_7	.000	-.003	.043	.038	.034	.034	.055											
par_8	.005	-.077	.002	.001	.001	.002	.001	.042										
par_9	.002	-.001	.018	.017	.014	.014	.016	.010	.072									
par_10	-.002	.000	.000	.000	.001	.001	.000	-.007	-.041	.035								
par_11	.044	-.785	.022	.017	.018	.019	.015	.265	.054	-.045	2.591							
par_12	.001	-.004	.035	.031	.027	.027	.028	-.003	.008	.000	.043	.050						
par_13	-.022	.282	.034	.032	.027	.025	.026	-.099	-.014	.022	-.902	.011	.414					
par_14	-.013	-.265	.015	.012	.031	.016	.004	-.003	-.021	.016	.070	.007	.139	.900				
par_15	.081	2.762	.009	-.008	-.008	-.040	.000	-.032	.010	-.009	-.394	-.001	.074	-.287	2.820			
par_16	.000	.005	-.025	-.022	-.020	-.020	-.021	-.001	-.009	.000	-.010	-.017	-.016	-.008	.004	.014		
par_17	.001	.041	.000	.000	.000	-.002	.000	.000	.000	.000	-.003	.000	-.004	.035	.000	.000	.005	
par_18	.000	.003	.000	.000	-.002	.000	.000	.000	.001	.000	-.002	.000	-.002	-.019	.001	.000	.000	
par_19	.000	.008	-.004	.001	.000	.000	-.001	.000	.000	.000	-.001	-.001	-.002	.001	-.025	.000	.000	
par_20	.000	-.030	.001	-.002	.000	.000	.000	.000	.000	.000	.002	.000	-.003	.003	-.001	.000	.000	
par_21	-.001	.017	.000	.000	.000	.000	.000	-.013	.004	-.004	-.074	-.001	.027	.001	.006	.000	.000	
par_22	.000	-.007	.001	.001	.001	.001	-.001	.000	.000	.000	.003	.001	.001	.011	-.008	.000	.000	-
par_23	.004	.152	-.007	-.006	-.006	-.006	-.004	-.012	.003	-.002	-.160	-.005	.009	-.098	.146	.003	.001	
par_24	.000	-.004	.000	.000	.000	.000	.000	.002	.004	-.004	.014	.000	-.004	.004	-.001	.000	.000	

Correlations of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	1.000																	
par_2	.630	1.000																
par_3	-.011	-.032	1.000															
par_4	-.010	.036	.796	1.000														
par_5	-.016	-.029	.790	.818	1.000													
par_6	-.097	-.115	.768	.796	.785	1.000												
par_7	.001	-.006	.703	.730	.704	.690	1.000											
par_8	.378	-.189	.034	.027	.035	.040	.027	1.000										
par_9	.111	-.002	.263	.278	.254	.254	.245	.174	1.000									
par_10	-.132	.001	.007	.006	.020	.015	.002	-.168	-.818	1.000								
par_11	.425	-.245	.052	.048	.055	.058	.038	.800	.126	-.147	1.000							
par_12	.042	-.009	.599	.621	.596	.580	.537	-.061	.133	-.003	.119	1.000						
par_13	-.535	.220	.202	.221	.209	.191	.172	-.745	-.083	.180	-.871	.076	1.000					
par_14	-.209	-.141	.062	.057	.160	.082	.018	-.018	-.084	.088	.046	.032	.228	1.000				
par_15	.750	.827	.022	-.020	-.024	-.116	-.001	-.091	.022	-.029	-.146	-.002	.069	-.180	1.000			
par_16	.017	.022	-.810	-.844	-.824	-.803	-.733	-.033	-.275	-.009	-.052	-.622	-.205	-.068	.022	1.000		
par_17	.245	.287	-.010	-.002	-.024	-.111	.004	-.026	.014	-.022	-.027	.017	.008	-.054	.288	.016	1.000	
par_18	-.002	.028	-.030	-.017	-.132	-.038	.011	-.011	.037	-.041	-.016	.010	-.040	-.336	.011	.031	.028	1
par_19	-.016	.036	-.124	.021	-.021	-.014	-.021	-.008	-.007	.003	-.007	-.027	-.023	.008	-.137	.021	-.017	
par_20	-.029	-.245	.050	-.121	.011	.014	.001	.027	-.019	.009	.019	-.010	-.072	.047	-.007	.006	-.061	-
par_21	-.130	.052	-.008	-.001	.008	.006	.001	-.368	.097	-.132	-.272	-.032	.250	.008	.020	.003	-.028	-
par_22	-.030	-.026	.022	.028	.045	.035	-.044	.005	-.006	.010	.015	.022	.016	.082	-.037	-.025	-.036	-
par_23	.241	.275	-.091	-.094	-.099	-.107	-.062	-.213	.034	-.039	-.356	-.087	.050	-.372	.313	.092	.064	
par_24	.030	-.021	.009	.004	.013	.009	.004	.096	.156	-.220	.078	.000	-.059	.035	-.008	-.008	-.004	-

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	.000																	
par_2	1.223	.000																
par_3	6.470	-.320	.000															
par_4	6.743	-.410	-1.087	.000														
par_5	6.508	-.488	-2.095	-1.248	.000													
par_6	6.238	-.489	-2.070	-1.255	-.081	.000												
par_7	6.031	-.455	-1.387	-.560	.395	.447	.000											
par_8	3.543	-.838	-3.257	-2.977	-2.552	-2.503	-2.579	.000										
par_9	2.565	-.843	-3.260	-2.940	-2.427	-2.379	-2.511	.048	.000									
par_10	1.668	-1.020	-4.353	-4.201	-3.842	-3.766	-3.741	-1.099	-.792	.000								
par_11	1.723	.121	.612	.720	.822	.829	.777	1.415	1.275	1.446	.000							
par_12	5.155	-.595	-2.498	-1.930	-1.101	-1.019	-1.268	1.640	1.529	2.881	-.962	.000						
par_13	1.720	-.622	-.884	-.627	-.370	-.351	-.468	.610	.667	1.294	-.711	-.032	.000					
par_14	-.223	-1.119	-2.019	-1.852	-1.726	-1.685	-1.733	-.915	-.900	-.589	-1.608	-1.453	-1.361	.000				
par_15	1.356	-.154	.279	.380	.476	.477	.436	.898	.897	1.101	-.208	.603	.593	1.171	.000			
par_16	3.557	-.956	-3.458	-3.294	-2.972	-2.926	-2.976	-.813	-.654	.602	-1.388	-2.269	-1.017	.721	-1.030	.000		
par_17	7.109	-.906	-4.207	-4.128	-3.675	-3.490	-3.545	-.349	-.330	1.252	-1.320	-2.519	-.881	.853	-.971	.862	.000	
par_18	5.554	-.950	-4.626	-4.626	-4.110	-4.111	-4.031	-.847	-.721	.743	-1.386	-3.002	-1.042	.730	-1.023	.105	-1.131	
par_19	6.569	-.779	-3.073	-2.979	-2.435	-2.369	-2.439	.679	.494	2.255	-1.173	-1.413	-.511	1.101	-.811	2.219	1.789	
par_20	5.043	-.959	-4.837	-4.646	-4.424	-4.324	-4.159	-1.017	-.837	.576	-1.410	-3.131	-1.092	.712	-1.043	-.154	-1.426	
par_21	3.083	-.908	-3.732	-3.541	-3.128	-3.059	-3.091	-.316	-.372	.869	-1.293	-2.151	-.949	.825	-.970	.478	-.115	
par_22	7.629	-.616	-2.014	-1.605	-1.042	-.982	-1.150	1.946	1.534	3.507	-.976	-.142	-.024	1.445	-.626	3.692	3.568	
par_23	5.031	-.531	-.958	-.561	-.129	-.098	-.305	1.795	1.752	2.960	-.793	.447	.273	1.452	-.529	3.000	2.679	

par_24	7.518	-.730	-2.888	-2.587	-2.068	-1.996	-2.104	1.132	.866	2.467	-1.122	-1.055	-.367	1.202	-.762	2.779	2.535
--------	-------	-------	--------	--------	--------	--------	--------	-------	------	-------	--------	--------	-------	-------	-------	-------	-------

Bootstrap (Default model)**Summary of Bootstrap Iterations (Default model)**

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	1
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	0
15	0	0	1
16	0	2	0
17	0	0	1
18	0	1	2
19	0	160	32
Total	0	163	37

0 bootstrap samples were unused because of a singular covariance matrix.

57 bootstrap samples were unused because a solution was not found.

200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 116 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 84 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .423

Bootstrap Distributions (Default model)**ML discrepancy (implied vs sample) (Default model)**

	11.847	*
	17.802	***
	23.757	*****
	29.712	*****
	35.667	*****
	41.621	*****
	47.576	*****
N = 200	53.531	*****
Mean = 41.754	59.486	*****
S. e. = .971	65.441	**
	71.396	*
	77.350	**
	83.305	
	89.260	
	95.215	*

Model Fit Summary**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	24	43.931	31	.062	1.417
Saturated model	55	.000	0		
Independence model	10	602.990	45	.000	13.400

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.065	.947	.905	.534
Saturated model	.000	1.000		
Independence model	.588	.379	.241	.310

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.927	.894	.977	.966	.977
Saturated model	1.000		1.000		1.000

Independence model	.000	.000	.000	.000	.000
--------------------	------	------	------	------	------

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.689	.639	.673
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	12.931	.000	34.584
Saturated model	.000	.000	.000
Independence model	557.990	482.408	641.010

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	.301	.089	.000	.237
Saturated model	.000	.000	.000	.000
Independence model	4.130	3.822	3.304	4.390

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.053	.000	.087	.408
Independence model	.291	.271	.312	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	91.931	95.842	163.702	187.702
Saturated model	110.000	118.963	274.474	329.474
Independence model	622.990	624.620	652.895	662.895

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	.630	.541	.778	.656
Saturated model	.753	.753	.753	.815
Independence model	4.267	3.749	4.836	4.278

HOELTER

Model	HOELTER	HOELTER
	.05	.01
Default model	150	174
Independence model	15	17

Execution time summary

Minimization: .015
 Miscellaneous: .217
 Bootstrap: 1.609
 Total: 1.841

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\MISM-5.amw

Analysis Summary

Date and Time

Date: Friday, 13 November 2009

Time: 10:40:20 a.m.

Title

Mism-5: Friday, 13 November 2009 10:40 a.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x26
x23
x17
x13
x20
x22
x28
x8
x4

Unobserved, endogenous variables

believability
relevance
timeliness
infQual
usefulness
accuracy
ease of use
overall SQ

Unobserved, exogenous variables

err-x2
err-x7
err-x10
err-x14
Perceived presentation fit
err-x26
err-x23
err-accuracy
err-believability
err-timeliness
err-relevance
err-ease of use
err-x13
err-x17
err-x20
err-x22
err-x28
err-x8
err-x4
err-OSQ
err-infQual
err-usefulness

Variable counts (Group number 1)

Number of variables in your model: 43
Number of observed variables: 13
Number of unobserved variables: 30
Number of exogenous variables: 22
Number of endogenous variables: 21

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	30	0	5	0	0	35
Labeled	0	0	0	0	0	0
Unlabeled	17	0	17	0	0	34

Total	47	0	22	0	0	69
-------	----	---	----	---	---	----

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x4	1.000	7.000	-.820	-4.059	.331	.818
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x28	1.000	7.000	-1.277	-6.320	2.257	5.586
x22	1.000	7.000	-.756	-3.742	.904	2.238
x20	1.000	7.000	-1.022	-5.058	1.608	3.979
x13	1.000	7.000	-.871	-4.309	.278	.688
x17	2.000	7.000	-.640	-3.168	.092	.227
x23	1.000	7.000	-.789	-3.904	.144	.357
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
Multivariate					57.903	17.774

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
27	54.560	.000	.000
17	44.386	.000	.000
77	40.646	.000	.000
65	34.175	.001	.000
30	33.969	.001	.000
97	32.818	.002	.000
144	31.178	.003	.000
74	30.742	.004	.000
64	29.479	.006	.000
50	29.056	.006	.000
58	27.756	.010	.000
48	27.059	.012	.000
52	26.551	.014	.000
80	26.420	.015	.000
40	25.546	.020	.000
2	24.632	.026	.000
55	22.837	.044	.000
109	22.598	.047	.000
68	22.337	.050	.000
113	21.862	.058	.000
41	21.340	.066	.001
69	21.281	.068	.000
110	21.132	.070	.000
103	20.557	.082	.001
46	20.287	.088	.001
12	20.062	.094	.001
81	19.898	.098	.001
86	19.665	.104	.001
107	19.573	.106	.001
101	19.537	.107	.000
78	19.034	.122	.002
145	18.950	.125	.001
141	18.885	.127	.001
56	18.660	.134	.001
116	18.611	.136	.001
60	18.543	.138	.000
127	18.472	.140	.000
98	18.317	.146	.000
123	18.263	.148	.000
38	18.063	.155	.000
43	17.848	.163	.000
47	17.724	.168	.000
91	17.353	.184	.001
35	17.246	.188	.001
122	16.756	.211	.004
124	16.756	.211	.002
1	16.353	.231	.008
18	16.325	.232	.006
22	16.105	.243	.009
16	15.913	.254	.012
121	14.767	.322	.287
128	14.260	.356	.551
29	13.762	.391	.798
118	13.709	.395	.776
28	13.584	.404	.792
67	13.223	.431	.904
70	12.562	.482	.991
120	12.384	.496	.995
137	12.212	.510	.997

111	12.009	.527	.999
147	11.665	.555	1.000
49	11.584	.562	1.000
87	11.507	.568	1.000
84	11.158	.598	1.000
90	10.791	.628	1.000
61	10.584	.646	1.000
72	10.562	.647	1.000
95	10.520	.651	1.000
45	10.453	.657	1.000
138	10.395	.661	1.000
34	10.282	.671	1.000
11	9.902	.702	1.000
20	9.898	.702	1.000
5	9.692	.719	1.000
8	9.545	.731	1.000
14	9.535	.731	1.000
83	9.504	.734	1.000
136	9.371	.744	1.000
135	9.371	.744	1.000
100	9.266	.753	1.000
88	9.231	.755	1.000
92	8.913	.779	1.000
102	8.843	.785	1.000
85	8.687	.796	1.000
23	8.676	.797	1.000
54	8.448	.813	1.000
15	8.198	.830	1.000
143	7.912	.849	1.000
126	7.872	.852	1.000
142	7.832	.854	1.000
57	7.799	.856	1.000
125	7.748	.860	1.000
51	7.632	.867	1.000
21	7.627	.867	1.000
112	7.594	.869	1.000
106	7.561	.871	1.000
93	7.544	.872	1.000
130	7.453	.877	1.000
108	7.146	.894	1.000
42	7.113	.896	1.000

Sample Moments (Group number 1)

Sample Covariances (Group number 1)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
x4	1.755												
x8	.660	2.177											
x28	.821	.678	1.336										
x22	.753	.896	.947	1.570									
x20	.704	.704	.901	.933	1.349								
x13	.819	.602	.878	.894	.901	1.954							
x17	.337	.330	.432	.493	.630	.620	1.218						
x23	.684	.766	.722	.957	.774	.531	.555	1.855					
x26	.587	.417	.644	.658	.759	.530	.425	.754	1.404				
x14	.578	.619	.715	.622	.784	.725	.613	.833	.604	1.270			
x10	.702	.511	.718	.871	.736	.823	.658	.963	.615	.921	1.863		
x7	.564	.659	.679	.591	.642	.499	.437	.634	.498	.809	.925	1.121	
x2	.563	.529	.731	.574	.746	.451	.458	.672	.500	.790	.891	.751	1.215

Condition number = 34.661

Eigenvalues

9.814 1.774 1.615 1.223 1.136 .974 .813 .704 .637 .426 .374 .313 .283

Determinant of sample covariance matrix = .196

Sample Correlations (Group number 1)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
x4	1.000												
x8	.338	1.000											
x28	.536	.398	1.000										
x22	.454	.485	.654	1.000									
x20	.457	.411	.671	.641	1.000								
x13	.442	.292	.543	.510	.555	1.000							
x17	.231	.203	.339	.357	.492	.402	1.000						
x23	.379	.381	.459	.561	.489	.279	.369	1.000					
x26	.374	.238	.470	.443	.551	.320	.325	.467	1.000				
x14	.387	.372	.549	.441	.599	.460	.493	.542	.453	1.000			
x10	.388	.254	.455	.510	.464	.431	.437	.518	.381	.599	1.000		
x7	.402	.422	.554	.445	.522	.337	.374	.440	.397	.678	.640	1.000	
x2	.385	.325	.574	.416	.583	.293	.376	.448	.383	.636	.592	.643	1.000

Condition number = 33.101

Eigenvalues

6.490 1.054 .889 .767 .735 .610 .538 .475 .407 .331 .273 .235 .196

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 91
 Number of distinct parameters to be estimated: 34
 Degrees of freedom (91 - 34): 57

Result (Default model)

Iteration limit reached
 The results that follow are therefore incorrect.
 Chi-square = 114.067
 Degrees of freedom = 57
 Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
timeliness	<---	Perceived presentation fit	1.563	.226	6.900	***	par_5
relevance	<---	Perceived presentation fit	1.407	.193	7.279	***	par_6
believability	<---	Perceived presentation fit	1.248	.178	7.025	***	par_7
accuracy	<---	Perceived presentation fit	1.206	.174	6.939	***	par_8
infQual	<---	believability	-49.575	44.097	-1.124	.261	par_1
infQual	<---	timeliness	-55.684	40.050	-1.390	.164	par_2
infQual	<---	accuracy	1.000				
infQual	<---	relevance	-131.234	98.345	-1.334	.182	par_4
ease of use	<---	Perceived presentation fit	1.198	.205	5.843	***	par_11
usefulness	<---	infQual	.005	.004	1.460	.144	par_3
usefulness	<---	ease of use	2.456	.896	2.741	.006	par_17
overall SQ	<---	ease of use	-.267	.502	-.532	.595	par_14
overall SQ	<---	usefulness	1.508	.578	2.608	.009	par_15
x7	<---	believability	1.000				
x14	<---	relevance	1.000				
x10	<---	timeliness	1.000				
x2	<---	accuracy	1.000				
x26	<---	Perceived presentation fit	1.000				
x23	<---	Perceived presentation fit	1.305	.205	6.369	***	par_9
x17	<---	ease of use	.603	.118	5.123	***	par_10
x13	<---	ease of use	1.000				
x20	<---	overall SQ	1.000				
x22	<---	overall SQ	.980	.093	10.501	***	par_12
x28	<---	overall SQ	.966	.084	11.562	***	par_13
x4	<---	usefulness	1.000				
x8	<---	usefulness	.966	.178	5.419	***	par_16

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
timeliness	<---	Perceived presentation fit	.814
relevance	<---	Perceived presentation fit	.913
believability	<---	Perceived presentation fit	.873
accuracy	<---	Perceived presentation fit	.804
infQual	<---	believability	-.203
infQual	<---	timeliness	-.306
infQual	<---	accuracy	.004
infQual	<---	relevance	-.578
ease of use	<---	Perceived presentation fit	.968
usefulness	<---	infQual	1.569
usefulness	<---	ease of use	2.523
overall SQ	<---	ease of use	-.224
overall SQ	<---	usefulness	1.231
x7	<---	believability	.906
x14	<---	relevance	.918
x10	<---	timeliness	.945
x2	<---	accuracy	.914
x26	<---	Perceived presentation fit	.567
x23	<---	Perceived presentation fit	.643

x17	<--- ease of use	.454
x13	<--- ease of use	.595
x20	<--- overall SQ	.845
x22	<--- overall SQ	.769
x28	<--- overall SQ	.821
x4	<--- usefulness	.609
x8	<--- usefulness	.528

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Perceived presentation fit	.451	.123	3.666	***	par_18
err-accuracy	.359	.072	5.007	***	par_19
err-believability	.219	.062	3.564	***	par_20
err-timeliness	.562	.109	5.155	***	par_21
err-relevance	.178	.063	2.843	.004	par_22
err-infQual	1.000				
err-ease of use	.044	.026	1.693	.090	par_23
err-usefulness	-.358	.293	-1.223	.221	par_24
err-OSQ	.029	.092	.309	.757	par_25
err-x2	.200				
err-x7	.200				
err-x10	.200				
err-x14	.200				
err-x26	.953	.113	8.427	***	par_26
err-x23	1.088	.132	8.244	***	par_27
err-x13	1.263	.143	8.824	***	par_28
err-x17	.966	.118	8.218	***	par_29
err-x20	.393	.062	6.297	***	par_30
err-x22	.652	.091	7.154	***	par_31
err-x28	.444	.067	6.673	***	par_32
err-x8	1.577	.194	8.137	***	par_33
err-x4	1.113	.141	7.872	***	par_34

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
accuracy	.646
timeliness	.662
relevance	.833
believability	.762
ease of use	.936
infQual	1.000
usefulness	1.547
overall SQ	.971
x4	.371
x8	.279
x28	.674
x22	.591
x20	.714
x13	.354
x17	.206
x23	.413
x26	.321
x14	.843
x10	.893
x7	.822
x2	.835

Matrices (Group number 1 - Default model)

Implied (for all variables) Covariances (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	infQual	usefulness	overall SQ	x4	x8	x28	x22	x20
Perceived presentation fit	.451													
accuracy	.544	1.015												
timeliness	.704	.850	1.663											
relevance	.634	.765	.991	1.071										
believability	.563	.679	.879	.792	.921									
ease of use	.540	.651	.844	.760	.674	.691								
infQual	-149.794	-180.324	-265.409	-234.157	-197.846	-179.497	55137.344							
usefulness	.516	.625	.638	.600	.586	.727	-142.651	.655						
overall SQ	.634	.768	.736	.702	.703	.911	-167.200	.793	.982					
x4	.516	.625	.638	.600	.586	.727	-142.651	.655	.793	1.767				
x8	.499	.604	.616	.580	.566	.702	-137.792	.633	.766	.633	2.188			
x28	.613	.742	.711	.678	.679	.880	-161.504	.766	.948	.766	.740	1.360		
x22	.622	.753	.721	.688	.689	.893	-163.822	.777	.962	.777	.751	.929	1.594	
x20	.634	.768	.736	.702	.703	.911	-167.200	.793	.982	.793	.766	.948	.962	1.360
x13	.540	.651	.844	.760	.674	.691	-179.497	.727	.911	.727	.702	.880	.893	.962

x17	.326	.393	.509	.458	.407	.417	-108.249	.438	.550	.438	.423	.531	.539	.531
x23	.588	.709	.919	.827	.734	.705	-195.414	.674	.828	.674	.651	.800	.811	.811
x26	.451	.544	.704	.634	.563	.540	-149.794	.516	.634	.516	.499	.613	.622	.622
x14	.634	.765	.991	1.071	.792	.760	-234.157	.600	.702	.600	.580	.678	.688	.688
x10	.704	.850	1.663	.991	.879	.844	-265.409	.638	.736	.638	.616	.711	.721	.721
x7	.563	.679	.879	.792	.921	.674	-197.846	.586	.703	.586	.566	.679	.689	.689
x2	.544	1.015	.850	.765	.679	.651	-180.324	.625	.768	.625	.604	.742	.753	.753

Implied (for all variables) Correlations (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	infQual	usefulness	overall SQ	x4	x8	x28	x22	x20
Perceived presentation fit	1.000													
accuracy	.804	1.000												
timeliness	.814	.654	1.000											
relevance	.913	.734	.743	1.000										
believability	.873	.702	.710	.797	1.000									
ease of use	.968	.778	.787	.883	.845	1.000								
infQual	-.950	-.762	-.876	-.964	-.878	-.919	1.000							
usefulness	.950	.766	.611	.717	.754	1.080	-.751	1.000						
overall SQ	.954	.770	.576	.685	.739	1.106	-.719	.990	1.000					
x4	.579	.466	.372	.436	.459	.658	-.457	.609	.602	1.000				
x8	.502	.405	.323	.379	.398	.571	-.397	.528	.523	.322	1.000			
x28	.783	.632	.473	.562	.607	.908	-.590	.812	.821	.494	.429	1.000		
x22	.733	.592	.443	.526	.569	.851	-.553	.761	.769	.463	.402	.631	1.000	
x20	.806	.650	.487	.579	.625	.935	-.607	.836	.845	.509	.442	.694	.650	1.000
x13	.575	.463	.468	.525	.502	.595	-.547	.642	.658	.391	.339	.540	.506	.556
x17	.440	.353	.358	.401	.384	.454	-.418	.491	.503	.299	.259	.413	.386	.425
x23	.643	.517	.523	.587	.561	.622	-.611	.611	.613	.372	.323	.503	.472	.518
x26	.567	.455	.461	.517	.495	.548	-.538	.539	.540	.328	.285	.444	.416	.457
x14	.838	.674	.682	.918	.732	.811	-.885	.658	.628	.400	.348	.516	.483	.531
x10	.769	.618	.945	.702	.671	.744	-.828	.577	.544	.351	.305	.447	.418	.460
x7	.791	.636	.644	.722	.906	.766	-.796	.684	.670	.416	.361	.550	.515	.566
x2	.735	.914	.598	.671	.641	.711	-.697	.700	.703	.426	.370	.577	.541	.594

Implied Covariances (Group number 1 - Default model)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
x4	1.767												
x8	.633	2.188											
x28	.766	.740	1.360										
x22	.777	.751	.929	1.594									
x20	.793	.766	.948	.962	1.375								
x13	.727	.702	.880	.893	.911	1.954							
x17	.438	.423	.531	.539	.550	.417	1.218						
x23	.674	.651	.800	.811	.828	.705	.425	1.855					
x26	.516	.499	.613	.622	.634	.540	.326	.588	1.404				
x14	.600	.580	.678	.688	.702	.760	.458	.827	.634	1.271			
x10	.638	.616	.711	.721	.736	.844	.509	.919	.704	.991	1.863		
x7	.586	.566	.679	.689	.703	.674	.407	.734	.563	.792	.879	1.121	
x2	.625	.604	.742	.753	.768	.651	.393	.709	.544	.765	.850	.679	1.215

Implied Correlations (Group number 1 - Default model)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
x4	1.000												
x8	.322	1.000											
x28	.494	.429	1.000										
x22	.463	.402	.631	1.000									
x20	.509	.442	.694	.650	1.000								
x13	.391	.339	.540	.506	.556	1.000							
x17	.299	.259	.413	.386	.425	.270	1.000						
x23	.372	.323	.503	.472	.518	.370	.283	1.000					
x26	.328	.285	.444	.416	.457	.326	.249	.364	1.000				
x14	.400	.348	.516	.483	.531	.482	.368	.539	.475	1.000			
x10	.351	.305	.447	.418	.460	.442	.338	.494	.436	.644	1.000		
x7	.416	.361	.550	.515	.566	.455	.348	.509	.448	.663	.608	1.000	
x2	.426	.370	.577	.541	.594	.423	.323	.472	.416	.616	.565	.581	1.000

Residual Covariances (Group number 1 - Default model)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
x4	-.012												
x8	.027	-.011											
x28	.055	-.062	-.024										
x22	-.024	.145	.018	-.025									
x20	-.090	-.063	-.047	-.029	-.026								
x13	.092	-.100	-.003	.001	-.010	.000							
x17	-.101	-.093	-.099	-.045	.081	.204	.000						

x23	.011	.115	-.078	.146	-.053	-.174	.130	.000						
x26	.071	-.082	.031	.037	.124	-.010	.099	.166	.000					
x14	-.022	.040	.037	-.065	.082	-.035	.154	.005	-.030	-.001				
x10	.064	-.105	.007	.150	.000	-.021	.149	.044	-.089	-.070	.000			
x7	-.022	.093	-.001	-.098	-.061	-.175	.031	-.100	-.064	.017	.046	.000		
x2	-.062	-.075	-.011	-.179	-.022	-.200	.065	-.037	-.044	.025	.041	.072	.000	

Standardized Residual Covariances (Group number 1 - Default model)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
x4	-.059												
x8	.160	-.044											
x28	.382	-.399	-.151										
x22	-.159	.871	.124	-.133									
x20	-.621	-.401	-.344	-.198	-.160								
x13	.560	-.556	-.017	.006	-.066	.000							
x17	-.798	-.666	-.856	-.366	.695	1.539	.000						
x23	.066	.657	-.527	.931	-.359	-1.035	1.003	.000					
x26	.515	-.543	.247	.274	.984	-.067	.889	1.170	.000				
x14	-.163	.271	.305	-.500	.665	-.239	1.406	.036	-.243	-.005			
x10	.402	-.599	.047	.971	-.001	-.124	1.136	.255	-.609	-.460	-.001		
x7	-.176	.675	-.007	-.790	-.520	-1.303	.301	-.744	-.566	.146	.327	-.002	
x2	-.471	-.520	-.087	-1.365	-.175	-1.448	.612	-.269	-.374	.208	.288	.643	.001

Factor Score Weights (Group number 1 - Default model)

	x4	x8	x28	x22	x20	x13	x17	x23	x26	x14	x10	x7	x2
Perceived presentation fit	.029	.020	.115	.079	.134	-.039	-.031	.002	.002	.185	.122	.080	.002
accuracy	.013	.009	.050	.035	.059	-.018	-.014	.000	.000	.080	.053	.034	.642
timeliness	.001	.000	-.014	-.010	-.017	.027	.021	.029	.026	.071	.774	.064	.053
relevance	.005	.003	-.007	-.005	-.008	.035	.028	.044	.039	.601	.071	.107	.080
believability	.010	.007	.030	.020	.035	.004	.003	.019	.017	.107	.064	.590	.034
ease of use	.063	.043	.261	.180	.305	-.103	-.082	-.059	-.052	.221	.168	.023	-.111
infQual	-1.150	-.784	.340	.235	.398	-6.272	-4.944	-8.395	-7.346	-88.056	-55.532	-46.838	-14.456
usefulness	.058	.040	.147	.101	.172	.056	.044	.039	.034	.026	.004	.057	.071
overall SQ	.061	.041	.182	.126	.213	.095	.075	.063	.055	-.017	-.033	.068	.116

Total Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	usefulness	overall SQ
accuracy		1.206	.000	.000	.000	.000	.000	.000
timeliness		1.563	.000	.000	.000	.000	.000	.000
relevance		1.407	.000	.000	.000	.000	.000	.000
believability		1.248	.000	.000	.000	.000	.000	.000
ease of use		1.198	.000	.000	.000	.000	.000	.000
infQual		-332.334	1.000	-55.684	-131.234	-49.575	.000	.000
usefulness		1.146	.005	-.301	-.710	-.268	2.456	.000
overall SQ		1.408	.008	-.454	-1.070	-.404	3.435	1.508
x4		1.146	.005	-.301	-.710	-.268	2.456	1.000
x8		1.107	.005	-.291	-.685	-.259	2.372	.966
x28		1.360	.008	-.438	-1.033	-.390	3.318	1.456
x22		1.379	.008	-.445	-1.048	-.396	3.366	1.477
x20		1.408	.008	-.454	-1.070	-.404	3.435	1.508
x13		1.198	.000	.000	.000	.000	1.000	.000
x17		.723	.000	.000	.000	.000	.603	.000
x23		1.305	.000	.000	.000	.000	.000	.000
x26		1.000	.000	.000	.000	.000	.000	.000
x14		1.407	.000	.000	1.000	.000	.000	.000
x10		1.563	.000	1.000	.000	.000	.000	.000
x7		1.248	.000	.000	.000	1.000	.000	.000
x2		1.206	1.000	.000	.000	.000	.000	.000

Standardized Total Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	usefulness	overall SQ
accuracy		.804	.000	.000	.000	.000	.000	.000
timeliness		.814	.000	.000	.000	.000	.000	.000
relevance		.913	.000	.000	.000	.000	.000	.000
believability		.873	.000	.000	.000	.000	.000	.000
ease of use		.968	.000	.000	.000	.000	.000	.000
infQual		-.950	.004	-.306	-.578	-.203	.000	.000
usefulness		.950	.007	-.480	-.907	-.318	2.523	.000
overall SQ		.954	.008	-.591	-1.117	-.392	2.882	1.231
x4		.579	.004	-.292	-.552	-.194	1.536	.609
x8		.502	.004	-.254	-.479	-.168	1.333	.528
x28		.783	.007	-.485	-.917	-.321	2.366	1.011
x22		.733	.006	-.454	-.859	-.301	2.216	.947
x20		.806	.007	-.499	-.944	-.331	2.436	1.040
x13		.575	.000	.000	.000	.000	.595	.000
x17		.440	.000	.000	.000	.000	.454	.000
x23		.643	.000	.000	.000	.000	.000	.000
x26		.567	.000	.000	.000	.000	.000	.000

x14	.838	.000	.000	.918	.000	.000	.000	.000
x10	.769	.000	.945	.000	.000	.000	.000	.000
x7	.791	.000	.000	.000	.906	.000	.000	.000
x2	.735	.914	.000	.000	.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	usefulness	overall SQ
accuracy	1.206	.000	.000	.000	.000	.000	.000	.000
timeliness	1.563	.000	.000	.000	.000	.000	.000	.000
relevance	1.407	.000	.000	.000	.000	.000	.000	.000
believability	1.248	.000	.000	.000	.000	.000	.000	.000
ease of use	1.198	.000	.000	.000	.000	.000	.000	.000
infQual	.000	1.000	-55.684	-131.234	-49.575	.000	.000	.000
usefulness	.000	.000	.000	.000	.000	2.456	.000	.000
overall SQ	.000	.000	.000	.000	.000	-.267	1.508	.000
x4	.000	.000	.000	.000	.000	.000	1.000	.000
x8	.000	.000	.000	.000	.000	.000	.966	.000
x28	.000	.000	.000	.000	.000	.000	.000	.966
x22	.000	.000	.000	.000	.000	.000	.000	.980
x20	.000	.000	.000	.000	.000	.000	.000	1.000
x13	.000	.000	.000	.000	.000	1.000	.000	.000
x17	.000	.000	.000	.000	.000	.603	.000	.000
x23	1.305	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000
x14	.000	.000	.000	1.000	.000	.000	.000	.000
x10	.000	.000	1.000	.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	1.000	.000	.000	.000
x2	.000	1.000	.000	.000	.000	.000	.000	.000

Standardized Direct Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	usefulness	overall SQ
accuracy	.804	.000	.000	.000	.000	.000	.000	.000
timeliness	.814	.000	.000	.000	.000	.000	.000	.000
relevance	.913	.000	.000	.000	.000	.000	.000	.000
believability	.873	.000	.000	.000	.000	.000	.000	.000
ease of use	.968	.000	.000	.000	.000	.000	.000	.000
infQual	.000	.004	-.306	-.578	-.203	.000	.000	.000
usefulness	.000	.000	.000	.000	.000	2.523	.000	.000
overall SQ	.000	.000	.000	.000	.000	-.224	1.231	.000
x4	.000	.000	.000	.000	.000	.000	.609	.000
x8	.000	.000	.000	.000	.000	.000	.528	.000
x28	.000	.000	.000	.000	.000	.000	.000	.821
x22	.000	.000	.000	.000	.000	.000	.000	.769
x20	.000	.000	.000	.000	.000	.000	.000	.845
x13	.000	.000	.000	.000	.000	.595	.000	.000
x17	.000	.000	.000	.000	.000	.454	.000	.000
x23	.643	.000	.000	.000	.000	.000	.000	.000
x26	.567	.000	.000	.000	.000	.000	.000	.000
x14	.000	.000	.000	.918	.000	.000	.000	.000
x10	.000	.000	.945	.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	.906	.000	.000	.000
x2	.000	.914	.000	.000	.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	usefulness	overall SQ
accuracy	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	.000	.000	.000	.000	.000	.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000
ease of use	.000	.000	.000	.000	.000	.000	.000	.000
infQual	-.332.334	.000	.000	.000	.000	.000	.000	.000
usefulness	1.146	.005	-.301	-.710	-.268	.000	.000	.000
overall SQ	1.408	.008	-.454	-1.070	-.404	3.702	.000	.000
x4	1.146	.005	-.301	-.710	-.268	2.456	.000	.000
x8	1.107	.005	-.291	-.685	-.259	2.372	.000	.000
x28	1.360	.008	-.438	-1.033	-.390	3.318	1.456	.000
x22	1.379	.008	-.445	-1.048	-.396	3.366	1.477	.000
x20	1.408	.008	-.454	-1.070	-.404	3.435	1.508	.000
x13	1.198	.000	.000	.000	.000	.000	.000	.000
x17	.723	.000	.000	.000	.000	.000	.000	.000
x23	.000	.000	.000	.000	.000	.000	.000	.000
x26	.000	.000	.000	.000	.000	.000	.000	.000
x14	1.407	.000	.000	.000	.000	.000	.000	.000
x10	1.563	.000	.000	.000	.000	.000	.000	.000
x7	1.248	.000	.000	.000	.000	.000	.000	.000
x2	1.206	.000	.000	.000	.000	.000	.000	.000

Standardized Indirect Effects (Group number 1 - Default model)

--	--	--	--	--	--	--	--	--

	Perceived presentation fit	accuracy	timeliness	relevance	believability	ease of use	usefulness	overall SQ
accuracy	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	.000	.000	.000	.000	.000	.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000
ease of use	.000	.000	.000	.000	.000	.000	.000	.000
infQual	-.950	.000	.000	.000	.000	.000	.000	.000
usefulness	.950	.007	-.480	-.907	-.318	.000	.000	.000
overall SQ	.954	.008	-.591	-1.117	-.392	3.106	.000	.000
x4	.579	.004	-.292	-.552	-.194	1.536	.000	.000
x8	.502	.004	-.254	-.479	-.168	1.333	.000	.000
x28	.783	.007	-.485	-.917	-.321	2.366	1.011	.000
x22	.733	.006	-.454	-.859	-.301	2.216	.947	.000
x20	.806	.007	-.499	-.944	-.331	2.436	1.040	.000
x13	.575	.000	.000	.000	.000	.000	.000	.000
x17	.440	.000	.000	.000	.000	.000	.000	.000
x23	.000	.000	.000	.000	.000	.000	.000	.000
x26	.000	.000	.000	.000	.000	.000	.000	.000
x14	.838	.000	.000	.000	.000	.000	.000	.000
x10	.769	.000	.000	.000	.000	.000	.000	.000
x7	.791	.000	.000	.000	.000	.000	.000	.000
x2	.735	.000	.000	.000	.000	.000	.000	.000

Notes for Model (Group number 1 - Default model)

The following variances are negative. (Group number 1 - Default model)

err-usefulness
-.358

Notes for Group/Model (Group number 1 - Default model)

This solution is not admissible.

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

	M.I.	Par Change
err-x22 <--> err-accuracy	7.417	-.145
err-x22 <--> err-relevance	5.248	-.071
err-x22 <--> err-x8	4.375	.191
err-x20 <--> err-relevance	4.135	.050
err-x13 <--> err-accuracy	4.199	-.136
err-x13 <--> err-believability	4.498	-.109
err-x17 <--> err-relevance	5.714	.084
err-x17 <--> err-ease of use	4.136	-.030
err-x17 <--> err-infQual	4.413	-11.830
err-x17 <--> err-usefulness	4.413	-.064
err-x17 <--> err-x28	5.739	-.139
err-x17 <--> err-x13	4.270	.175
err-x23 <--> err-x22	5.847	.180
err-x14 <--> err-x22	4.046	-.102
err-x10 <--> err-x22	8.951	.201
err-x2 <--> err-x22	7.436	-.145
err-x2 <--> err-x13	4.219	-.136

Variances: (Group number 1 - Default model)

M.I.	Par Change
------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
x28 <--- x17	4.457	-.116
x10 <--- x22	4.959	.137
x2 <--- x22	5.127	-.111

Minimization History (Default model)

Iteration		Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	9		-.555	9999.000	1058.981	0	9999.000
1	e	9		-.145	2.427	558.150	20	.491
2	e*	2		-.097	1.522	284.080	4	.700
3	e	1		-.160	.558	191.539	4	.922
4	e	1		-.050	.525	151.552	6	.889
5	e	0	1439.243	.650	135.145		7	.868
6	e	1		-.003	.647	130.957	3	.000
7	e	1		.000	.324	126.794	5	.836
8	e	1		-.032	.478	125.066	6	.713
9	e	1		.000	.157	124.145	6	.688
10	e	1		-.012	.574	122.831	7	.736

11	e	0	77122.717		.179	122.121	6	.685
12	e	1		-.005	.607	121.267	5	.000
13	e	1		-.001	.281	120.437	5	.730
14	e	1		-.002	.474	119.590	6	.852
15	e	1		-.002	.276	119.189	5	.586
16	e	1		-.003	.438	118.648	6	.757
17	e	1		.000	.271	118.312	5	.656
18	e	1		.000	.494	117.847	6	.905
19	e	0	421776.638		.259	117.613	5	.806
20	e	1		-.014	1.278	117.590	2	.000
21	e	0	562672.460		.175	116.985	4	.855
22	e	1		-.016	1.205	116.805	1	.427
23	e	0	4384704.624		.078	116.337	5	.950
24	e	1		-.005	.913	116.228	4	.000
25	e	0	5947766.086		.114	115.998	8	.985
26	e	0	5826165.530		.635	115.858	4	.000
27	e	1		-.001	1.302	115.784	2	.000
28	e	0	5950784.287		.276	115.601	5	.855
29	e	1		-.004	1.539	115.533	1	.448
30	e	0	205722898.829		.083	115.346	5	.926
31	e	1		-.002	1.287	115.323	6	.000
32	e	0	82310728.960		.087	115.182	7	.971
33	e	1		.000	1.181	115.127	4	.000
34	e	0	88049755.737		.314	115.046	7	1.031
35	e	0	40771210.315		.794	114.993	4	.000
36	e	0	74375782.947		1.905	114.918	1	1.105
37	e	1		.000	.283	114.883	1	.540
38	e	0	81214458.104		.739	114.837	5	.684
39	e	1		.000	2.817	114.787	1	.510
40	e	0	156519843.799		.306	114.706	6	.879
41	e	1		.000	2.204	114.687	1	.395
42	e	0	5515985847.831		.055	114.625	7	.952
43	e	1		.000	1.822	114.614	6	.000
44	e	0	4447672356.776		.291	114.568	10	.952
45	e	0	409154023.765		1.059	114.545	6	.000
46	e	0	950832607.137		2.145	114.526	1	.713
47	e	0	1393231429.689		.615	114.486	1	1.165
48	e	0	662760587.999		1.335	114.468	3	.000
49	e	0	878503389.622		1.595	114.443	1	1.342
50	e	0	944043410.508		1.617	114.423	1	1.131
51	e	0	1786835738.041		1.269	114.399	1	1.324
52	e	0	1303423430.828		1.443	114.386	2	.000
53	e	0	2052627440.214		1.341	114.366	1	1.372
54	e	0	2038514281.483		2.270	114.355	1	.747
55	e	0	5910496866.197		.977	114.332	1	1.196
56	e	0	2968699774.546		1.405	114.322	3	.000
57	e	0	3162185920.985		1.913	114.309	1	1.249
58	e	0	5024804818.263		1.539	114.294	1	1.373
59	e	0	4217268934.821		2.567	114.287	1	.677
60	e	0	15434963262.826		1.060	114.269	1	1.181
61	e	0	6442464113.193		1.701	114.261	3	.000
62	e	0	7548638465.075		1.938	114.250	1	1.336
63	e	0	8866677342.359		2.059	114.239	1	1.317
64	e	0	11031656305.937		1.990	114.229	1	1.350
65	e	0	11347659633.115		2.257	114.220	1	1.271
66	e	0	17632317060.982		1.890	114.211	1	1.375
67	e	0	13065982732.117		2.859	114.205	1	.906
68	e	0	45413116928.750		1.385	114.195	1	1.247
69	e	0	16331365743.033		3.126	114.192	2	.000
70	e	0	86759505051.990		1.217	114.181	1	1.160
71	e	0	27715682939.503		2.352	114.176	3	.000
72	e	0	43728400758.268		2.067	114.169	1	1.375
73	e	0	28255259423.351		3.399	114.165	1	.738
74	e	0	135663723969.646		1.409	114.156	1	1.199
75	e	0	53488142634.302		2.158	114.152	3	.000
76	e	0	55712907614.490		2.780	114.147	1	1.279
77	e	0	84424555701.905		2.373	114.141	1	1.374
78	e	0	60050045445.833		3.438	114.137	1	.982
79	e	0	196245151444.046		1.781	114.131	1	1.276
80	e	0	73425455481.072		3.479	114.129	2	.000
81	e	0	294408699448.015		1.714	114.123	1	1.238
82	e	0	83220857434.795		4.207	114.122	2	.000
83	e	0	620075102656.745		1.413	114.115	1	1.130
84	e	0	123317262015.605		3.567	114.112	3	.000
85	e	0	397257439057.677		2.015	114.108	1	1.289
86	e	0	153827112871.310		3.684	114.106	2	.000
87	e	0	550616386296.665		2.030	114.101	1	1.275
88	e	0	187858835487.314		4.039	114.099	2	.000
89	e	0	800799039870.837		1.919	114.095	1	1.230
90	e	0	198956446972.144		5.010	114.095	2	.000
91	e	0	1799917903436.970		1.510	114.089	1	1.109
92	e	0	270117752583.167		4.661	114.087	3	.000

par_32	-.031	.019	-.009	.011	-.013	-.012	-.011	-.022	-.005	.035	-.017	-.051	-.203	-.044	.064	.032	-.016
par_33	-.010	-.011	.014	-.012	-.004	-.006	-.009	-.004	-.010	.010	.003	-.019	.000	.062	-.050	-.075	.026
par_34	.000	.005	-.004	.005	-.002	.000	-.004	-.001	-.004	.038	.007	-.011	-.027	.042	.017	.080	-.034

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	.000																
par_2	-.172	.000															
par_3	1.124	1.391	.000														
par_4	-1.012	-1.049	-1.335	.000													
par_5	1.159	1.429	6.880	1.350	.000												
par_6	1.156	1.425	7.259	1.348	-1.040	.000											
par_7	1.152	1.421	6.999	1.347	-2.092	-1.287	.000										
par_8	1.151	1.420	6.905	1.347	-2.239	-1.514	-.325	.000									
par_9	1.154	1.423	6.342	1.348	-1.435	-.646	.351	.614	.000								
par_10	1.138	1.406	5.072	1.341	-3.733	-3.533	-2.999	-2.850	-2.945	.000							
par_11	1.151	1.420	5.818	1.347	-1.939	-1.257	-.297	-.046	-.551	2.246	.000						
par_12	1.146	1.415	10.436	1.344	-2.386	-1.998	-1.341	-1.147	-1.456	2.499	-.973	.000					
par_13	1.146	1.414	11.490	1.344	-2.482	-2.104	-1.444	-1.252	-1.538	2.483	-1.055	-.147	.000				
par_14	1.117	1.379	-.543	1.330	-3.368	-3.160	-2.848	-2.776	-2.938	-1.666	-2.742	-2.460	-2.445	.000			
par_15	1.160	1.433	2.594	1.352	-.088	.163	.428	.499	.327	1.546	.504	.886	.909	1.670	.000		
par_16	1.146	1.415	5.379	1.344	-2.070	-1.677	-1.123	-.966	-1.250	1.706	-.859	-.069	.000	2.276	-.966	.000	
par_17	1.173	1.445	2.738	1.354	.989	1.190	1.352	1.370	1.258	2.075	1.320	1.640	1.658	3.781	.695	1.557	.000
par_18	1.135	1.402	3.617	1.339	-3.361	-3.167	-2.800	-2.714	-2.817	-.903	-2.487	-3.414	-3.447	1.381	-1.798	-2.377	-2.15
par_19	1.132	1.399	4.949	1.338	-5.133	-5.186	-4.701	-4.469	-4.351	-1.770	-3.927	-5.334	-5.489	1.245	-1.958	-3.153	-2.37
par_20	1.130	1.396	3.461	1.337	-5.682	-5.829	-5.246	-5.296	-5.091	-2.905	-4.598	-6.821	-7.207	.969	-2.203	-3.955	-2.47
par_21	1.137	1.405	5.098	1.340	-3.780	-3.838	-3.252	-3.109	-3.167	-.253	-2.714	-2.907	-2.934	1.603	-1.617	-1.937	-2.08
par_22	1.128	1.394	2.726	1.337	-5.944	-5.782	-5.662	-5.565	-5.251	-3.157	-4.742	-7.117	-7.533	.872	-2.306	-4.177	-2.46
par_23	1.125	1.391	1.437	1.335	-6.600	-6.852	-6.611	-6.649	-6.126	-4.349	-5.595	-9.653	-10.553	.617	-2.537	-5.122	-2.65
par_24	1.117	1.384	-1.236	1.332	-4.988	-4.737	-4.523	-4.560	-4.566	-3.196	-4.458	-4.345	-4.326	-.122	-5.045	-4.360	-2.41
par_25	1.125	1.391	.249	1.335	-6.257	-6.403	-6.129	-5.995	-5.682	-3.715	-5.152	-7.026	-7.288	.526	-2.741	-4.663	-2.62
par_26	1.146	1.414	8.385	1.344	-2.441	-2.056	-1.424	-1.224	-1.505	2.134	-1.054	-.183	-.092	2.385	-.937	-.061	-1.68
par_27	1.148	1.418	8.208	1.345	-1.832	-1.386	-.741	-.541	-.883	2.753	-.457	.660	.780	2.592	-.713	.549	-1.53
par_28	1.153	1.422	8.785	1.347	-1.125	-.604	.066	.254	-.166	3.340	.264	1.655	1.784	2.906	-.415	1.305	-1.31
par_29	1.146	1.414	8.176	1.344	-2.352	-1.963	-1.341	-1.149	-1.438	2.143	-.954	-.090	.002	2.413	-.908	.002	-1.67
par_30	1.133	1.400	6.203	1.338	-5.000	-5.018	-4.561	-4.418	-4.294	-1.557	-3.778	-5.733	-6.155	1.313	-1.901	-3.025	-2.30
par_31	1.139	1.407	7.089	1.341	-3.725	-3.524	-2.982	-2.836	-2.882	.327	-2.431	-2.322	-2.497	1.798	-1.467	-1.554	-2.00
par_32	1.134	1.401	6.578	1.339	-4.724	-4.694	-4.224	-4.067	-3.991	-1.196	-3.483	-4.569	-4.467	1.396	-1.842	-2.773	-2.23
par_33	1.160	1.430	8.110	1.350	.048	.620	1.245	1.421	.961	4.314	1.344	2.756	2.895	3.502	.112	2.239	-.96
par_34	1.149	1.418	7.830	1.346	-1.685	-1.230	-.597	-.418	-.770	2.823	-.345	.780	.883	2.675	-.667	.671	-1.47

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	34	114.067	57	.000	2.001
Saturated model	91	.000	0		
Independence model	13	1022.879	78	.000	13.114

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.080	.899	.838	.563
Saturated model	.000	1.000		
Independence model	.644	.280	.160	.240

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.888	.847	.941	.917	.940
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.731	.649	.687
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	57.067	30.455	91.461
Saturated model	.000	.000	.000
Independence model	944.879	845.365	1051.818

FMIN

Model	FMIN	F0	LO 90	HI 90
-------	------	----	-------	-------

Default model	.781	.391	.209	.626
Saturated model	.000	.000	.000	.000
Independence model	7.006	6.472	5.790	7.204

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.083	.060	.105	.010
Independence model	.288	.272	.304	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	182.067	189.279	283.742	317.742
Saturated model	182.000	201.303	454.129	545.129
Independence model	1048.879	1051.636	1087.754	1100.754

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	1.247	1.065	1.483	1.296
Saturated model	1.247	1.247	1.247	1.379
Independence model	7.184	6.502	7.917	7.203

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	97	109
Independence model	15	16

Execution time summary

Minimization: .110
 Miscellaneous: .328
 Bootstrap: .000
 Total: .438

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\MISM-6.amw

Analysis Summary

Date and Time

Date: Monday, 23 November 2009

Time: 7:28:39 p.m.

Title

Mism-6: Monday, 23 November 2009 07:28 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x26
x23
x13
x4
x8
x17
x18
x16

Unobserved, endogenous variables

believability
relevance
timeliness
infQual
accuracy
Perceived presentation fit
ease of use
usefulness
perceived design fit

Unobserved, exogenous variables

err-x2
err-x7
err-x10
err-x14
err-x26
err-x23
err-accuracy
err-believability
err-timeliness
err-relevance
err-ease of use
err-x13
err-x4
err-x8
err-usefulness
err-x17
err-x18
err-x16
err per des fit
err presentation fit

Variable counts (Group number 1)

Number of variables in your model: 41
Number of observed variables: 12
Number of unobserved variables: 29
Number of exogenous variables: 20
Number of endogenous variables: 21

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	29	0	9	0	0	38
Labeled	0	0	0	0	0	0
Unlabeled	15	0	11	0	0	26
Total	44	0	20	0	0	64

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x16	1.000	7.000	-.604	-2.987	.037	.092
x18	1.000	7.000	-1.226	-6.070	2.271	5.621
x17	2.000	7.000	-.640	-3.168	.092	.227
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x4	1.000	7.000	-.820	-4.059	.331	.818
x13	1.000	7.000	-.871	-4.309	.278	.688
x23	1.000	7.000	-.789	-3.904	.144	.357
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
Multivariate					59.799	19.777

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
27	62.390	.000	.000
17	43.535	.000	.000
77	40.169	.000	.000
30	39.689	.000	.000
97	35.063	.000	.000
74	30.873	.002	.000
65	29.990	.003	.000
144	29.777	.003	.000
48	26.628	.009	.000
50	26.324	.010	.000
110	25.224	.014	.000
46	25.076	.014	.000
80	23.608	.023	.000
60	22.434	.033	.000
64	22.377	.034	.000
101	21.930	.038	.000
69	21.104	.049	.001
113	20.919	.052	.001
121	20.871	.052	.000
103	19.939	.068	.002
2	19.828	.070	.002
68	19.760	.072	.001
58	19.610	.075	.001
40	19.537	.076	.000
12	19.215	.083	.001
47	18.998	.089	.001
52	18.963	.089	.000
109	18.801	.093	.000
145	18.637	.098	.000
55	18.421	.104	.000
78	18.009	.115	.001
141	17.632	.127	.002
1	16.831	.156	.018
86	16.722	.160	.016
107	16.514	.169	.020
41	16.428	.172	.016
118	15.210	.230	.296
29	14.856	.249	.430
147	14.828	.251	.375
91	14.669	.260	.400
122	14.395	.276	.502
124	14.395	.276	.429
116	14.362	.278	.379
28	14.085	.295	.489
8	13.600	.327	.733
43	13.561	.330	.695
38	13.381	.342	.742
18	13.312	.347	.725
127	12.824	.382	.904
128	12.694	.392	.915
70	12.156	.433	.987
84	11.930	.451	.993
87	11.701	.470	.997
137	11.520	.485	.998
35	11.459	.490	.998
120	11.430	.492	.997
10	11.322	.502	.998
131	11.313	.502	.997
95	11.216	.510	.997
14	11.151	.516	.997
45	11.014	.528	.998
34	10.937	.534	.998

92	10.913	.536	.997
49	10.795	.547	.997
90	10.765	.549	.996
135	10.674	.557	.997
136	10.674	.557	.995
111	10.655	.559	.992
123	10.397	.581	.998
23	10.334	.587	.997
5	10.110	.606	.999
67	10.016	.615	.999
72	9.975	.618	.999
83	9.507	.659	1.000
125	9.506	.659	1.000
93	9.262	.680	1.000
61	9.164	.689	1.000
100	9.131	.692	1.000
119	8.662	.732	1.000
102	8.597	.737	1.000
11	8.472	.747	1.000
54	8.396	.753	1.000
20	8.380	.755	1.000
98	8.059	.781	1.000
88	7.958	.788	1.000
99	7.873	.795	1.000
7	7.823	.799	1.000
21	7.785	.802	1.000
112	7.753	.804	1.000
6	7.665	.811	1.000
22	7.463	.826	1.000
57	7.251	.841	1.000
51	7.198	.844	1.000
82	7.186	.845	1.000
85	7.068	.853	1.000
126	7.034	.855	1.000
56	6.984	.859	1.000
66	6.958	.860	1.000
75	6.845	.868	1.000
81	6.832	.869	1.000

Sample Moments (Group number 1)**Sample Covariances (Group number 1)**

	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x16	1.710											
x18	.526	1.337										
x17	.570	.457	1.218									
x8	.101	.470	.330	2.177								
x4	.400	.477	.337	.660	1.755							
x13	.545	.852	.620	.602	.819	1.954						
x23	.323	.702	.555	.766	.684	.531	1.855					
x26	.386	.520	.425	.417	.587	.530	.754	1.404				
x14	.417	.695	.613	.619	.578	.725	.833	.604	1.270			
x10	.454	.747	.658	.511	.702	.823	.963	.615	.921	1.863		
x7	.325	.621	.437	.659	.564	.499	.634	.498	.809	.925	1.121	
x2	.428	.619	.458	.529	.563	.451	.672	.500	.790	.891	.751	1.215

Condition number = 27.255

Eigenvalues

8.241 2.011 1.601 1.266 1.212 1.067 .844 .798 .633 .487 .420 .302

Determinant of sample covariance matrix = 1.142

Models**Default model (Default model)****Notes for Model (Default model)****Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 78
Number of distinct parameters to be estimated: 26
Degrees of freedom (78 - 26): 52

Result (Default model)

Minimum was achieved
Chi-square = 152.700
Degrees of freedom = 52
Probability level = .000

Group number 1 (Group number 1 - Default model)**Estimates (Group number 1 - Default model)**

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
timeliness	<---	Perceived presentation fit	1.638	.256	6.386	***	par_1
relevance	<---	Perceived presentation fit	1.470	.220	6.694	***	par_2
believability	<---	Perceived presentation fit	1.312	.202	6.485	***	par_3
accuracy	<---	Perceived presentation fit	1.301	.205	6.337	***	par_4
infQual	<---	accuracy	1.000				
infQual	<---	believability	4.137	5.841	.708	.479	par_11
infQual	<---	relevance	1.671	2.803	.596	.551	par_12
infQual	<---	timeliness	-.756	1.239	-.610	.542	par_13
ease of use	<---	perceived design fit	.604	.177	3.420	***	par_9
ease of use	<---	Perceived presentation fit	.657	.217	3.034	.002	par_10
usefulness	<---	infQual	.109	.139	.784	.433	par_14
x7	<---	believability	1.000				
x14	<---	relevance	1.000				
x10	<---	timeliness	1.000				
x2	<---	accuracy	1.000				
x26	<---	Perceived presentation fit	1.000				
x23	<---	Perceived presentation fit	1.347	.231	5.821	***	par_5
x13	<---	ease of use	1.000				
x4	<---	usefulness	.327	.145	2.259	.024	par_6
x8	<---	usefulness	1.000				
x17	<---	Perceived presentation fit	.944	.180	5.234	***	par_7
x18	<---	perceived design fit	1.000				
x16	<---	perceived design fit	.649	.143	4.530	***	par_8
x4	<---	ease of use	.614	.150	4.082	***	par_15

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
err presentation fit	.411	.122	3.365	***	par_16
err-accuracy	.519	.073	7.155	***	par_17
err-believability	.414	.061	6.755	***	par_18
err-timeliness	.761	.109	6.960	***	par_19
err-relevance	.382	.063	6.067	***	par_20
err per des fit	.831	.156	5.335	***	par_21
err-ease of use	.327	.147	2.220	.026	par_22
err-usefulness	.751	.201	3.745	***	par_23
err-x2	.000				
err-x7	.000				
err-x10	.000				
err-x14	.000				
err-x4	1.000				
err-x8	1.000				
err-x26	1.000				
err-x13	1.000				
err-x18	.500				
err-x23	1.109	.142	7.832	***	par_24
err-x17	.852	.105	8.093	***	par_25
err-x16	1.360	.180	7.541	***	par_26

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
Perceived presentation fit	.000
accuracy	.573
timeliness	.592
relevance	.699
believability	.631
perceived design fit	.000
usefulness	.357
ease of use	.595
x16	.205
x18	.624
x17	.301
x8	.539
x4	.344
x13	.447
x23	.402
x26	.291
x14	1.000
x10	1.000
x7	1.000
x2	1.000

Matrices (Group number 1 - Default model)

Implied (for all variables) Covariances (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	infQual	usefulness	ease of use	x16	x18	x17	x8	x4
Perceived presentation fit	.411													
accuracy	.535	1.215												
timeliness	.673	.876	1.863											
relevance	.604	.786	.989	1.270										
believability	.539	.702	.883	.793	1.121									
perceived design fit	.000	.000	.000	.000	.000	.831								
infQual	3.266	4.769	4.773	5.439	5.997	.000	35.055							
usefulness	.356	.520	.521	.593	.654	.000	3.824	1.169						
ease of use	.270	.351	.442	.397	.354	.502	2.145	.234	.807					
x16	.000	.000	.000	.000	.000	.539	.000	.000	.326	1.710				
x18	.000	.000	.000	.000	.000	.831	.000	.000	.502	.539	1.331			
x17	.388	.505	.635	.570	.509	.000	3.082	.336	.255	.000	.000	1.218		
x8	.356	.520	.521	.593	.654	.000	3.824	1.169	.234	.000	.000	.336	2.169	
x4	.282	.386	.442	.438	.432	.308	2.569	.526	.572	.200	.308	.267	.526	1.524
x13	.270	.351	.442	.397	.354	.502	2.145	.234	.807	.326	.502	.255	.234	.572
x23	.554	.721	.907	.814	.726	.000	4.400	.480	.364	.000	.000	.523	.480	.380
x26	.411	.535	.673	.604	.539	.000	3.266	.356	.270	.000	.000	.388	.356	.282
x14	.604	.786	.989	1.270	.793	.000	5.439	.593	.397	.000	.000	.570	.593	.438
x10	.673	.876	1.863	.989	.883	.000	4.773	.521	.442	.000	.000	.635	.521	.442
x7	.539	.702	.883	.793	1.121	.000	5.997	.654	.354	.000	.000	.509	.654	.432
x2	.535	1.215	.876	.786	.702	.000	4.769	.520	.351	.000	.000	.505	.520	.386

Implied Covariances (Group number 1 - Default model)

	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x16	1.710											
x18	.539	1.331										
x17	.000	.000	1.218									
x8	.000	.000	.336	2.169								
x4	.200	.308	.267	.526	1.524							
x13	.326	.502	.255	.234	.572	1.807						
x23	.000	.000	.523	.480	.380	.364	1.855					
x26	.000	.000	.388	.356	.282	.270	.554	1.411				
x14	.000	.000	.570	.593	.438	.397	.814	.604	1.270			
x10	.000	.000	.635	.521	.442	.442	.907	.673	.989	1.863		
x7	.000	.000	.509	.654	.432	.354	.726	.539	.793	.883	1.121	
x2	.000	.000	.505	.520	.386	.351	.721	.535	.786	.876	.702	1.215

Residual Covariances (Group number 1 - Default model)

	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x16	.000											
x18	-.013	.006										
x17	.570	.457	.000									
x8	.101	.470	-.006	.008								
x4	.200	.169	.071	.134	.231							
x13	.219	.350	.366	.368	.247	.147						
x23	.323	.702	.032	.286	.304	.167	.000					
x26	.386	.520	.037	.061	.305	.261	.201	-.007				
x14	.417	.695	.042	.026	.140	.328	.019	.000	.000			
x10	.454	.747	.023	-.010	.260	.380	.056	-.058	-.068	.000		
x7	.325	.621	-.071	.005	.132	.145	-.092	-.041	.016	.042	.000	
x2	.428	.619	-.047	.008	.177	.100	-.048	-.035	.004	.015	.049	.000

Standardized Residual Covariances (Group number 1 - Default model)

	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x16	.000											
x18	-.099	.040										
x17	4.775	4.335	.000									
x8	.634	3.346	-.043	.033								
x4	1.489	1.403	.615	.852	1.297							
x13	1.483	2.597	2.935	2.228	1.700	.695						
x23	2.189	5.398	.243	1.674	2.129	1.081	.000					
x26	2.999	4.587	.327	.410	2.465	1.944	1.417	-.043				
x14	3.416	6.456	.375	.176	1.163	2.535	.130	.003	.000			
x10	3.071	5.736	.172	-.056	1.802	2.436	.327	-.396	-.450	.000		
x7	2.836	6.144	-.677	.033	1.158	1.191	-.689	-.361	.138	.300	.000	
x2	3.586	5.879	-.433	.059	1.509	.791	-.350	-.298	.032	.104	.433	.000

Factor Score Weights (Group number 1 - Default model)

	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
Perceived presentation fit	-.002	-.008	.048	-.001	.011	.018	.053	.043	.166	.093	.136	.108
accuracy	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
timeliness	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000

relevance	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
perceived design fit	.124	.519	-.005	-.009	.063	.107	-.005	-.004	-.018	-.008	-.019	-.012
infQual	.000	.000	.000	.000	.000	.000	.000	.000	1.671	-.756	4.137	1.000
usefulness	-.004	-.018	-.002	.412	.121	-.023	-.002	-.001	.094	-.048	.243	.056
ease of use	.051	.213	.020	-.023	.164	.279	.022	.018	.063	.041	.043	.042

Total Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	usefulness	ease of use
accuracy	1.301	.000	.000	.000	.000	.000	.000	.000
timeliness	1.638	.000	.000	.000	.000	.000	.000	.000
relevance	1.470	.000	.000	.000	.000	.000	.000	.000
believability	1.312	.000	.000	.000	.000	.000	.000	.000
infQual	7.947	1.000	-.756	1.671	4.137	.000	.000	.000
usefulness	.867	.109	-.082	.182	.451	.000	.000	.000
ease of use	.657	.000	.000	.000	.000	.604	.000	.000
x16	.000	.000	.000	.000	.000	.649	.000	.000
x18	.000	.000	.000	.000	.000	1.000	.000	.000
x17	.944	.000	.000	.000	.000	.000	.000	.000
x8	.867	.109	-.082	.182	.451	.000	1.000	.000
x4	.687	.036	-.027	.060	.148	.371	.327	.614
x13	.657	.000	.000	.000	.000	.604	.000	1.000
x23	1.347	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000
x14	1.470	.000	.000	1.000	.000	.000	.000	.000
x10	1.638	.000	1.000	.000	.000	.000	.000	.000
x7	1.312	.000	.000	.000	1.000	.000	.000	.000
x2	1.301	1.000	.000	.000	.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	usefulness	ease of use
accuracy	1.301	.000	.000	.000	.000	.000	.000	.000
timeliness	1.638	.000	.000	.000	.000	.000	.000	.000
relevance	1.470	.000	.000	.000	.000	.000	.000	.000
believability	1.312	.000	.000	.000	.000	.000	.000	.000
infQual	.000	1.000	-.756	1.671	4.137	.000	.000	.000
usefulness	.000	.000	.000	.000	.000	.000	.000	.000
ease of use	.657	.000	.000	.000	.000	.604	.000	.000
x16	.000	.000	.000	.000	.000	.649	.000	.000
x18	.000	.000	.000	.000	.000	1.000	.000	.000
x17	.944	.000	.000	.000	.000	.000	.000	.000
x8	.000	.000	.000	.000	.000	.000	1.000	.000
x4	.000	.000	.000	.000	.000	.000	.327	.614
x13	.000	.000	.000	.000	.000	.000	.000	1.000
x23	1.347	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000
x14	.000	.000	.000	1.000	.000	.000	.000	.000
x10	.000	.000	1.000	.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	1.000	.000	.000	.000
x2	.000	1.000	.000	.000	.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	usefulness	ease of use
accuracy	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	.000	.000	.000	.000	.000	.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000
infQual	7.947	.000	.000	.000	.000	.000	.000	.000
usefulness	.867	.109	-.082	.182	.451	.000	.000	.000
ease of use	.000	.000	.000	.000	.000	.000	.000	.000
x16	.000	.000	.000	.000	.000	.000	.000	.000
x18	.000	.000	.000	.000	.000	.000	.000	.000
x17	.000	.000	.000	.000	.000	.000	.000	.000
x8	.867	.109	-.082	.182	.451	.000	.000	.000
x4	.687	.036	-.027	.060	.148	.371	.000	.000
x13	.657	.000	.000	.000	.000	.604	.000	.000
x23	.000	.000	.000	.000	.000	.000	.000	.000
x26	.000	.000	.000	.000	.000	.000	.000	.000
x14	1.470	.000	.000	.000	.000	.000	.000	.000
x10	1.638	.000	.000	.000	.000	.000	.000	.000
x7	1.312	.000	.000	.000	.000	.000	.000	.000
x2	1.301	.000	.000	.000	.000	.000	.000	.000

Modification Indices (Group number 1 - Default model)**Covariances: (Group number 1 - Default model)**

	M.I.	Par Change
err per des fit <--> err presentation fit	54.115	.454

err-x18	<-->	err presentation fit	35.618	.341
err-x17	<-->	err-x16	9.590	.294
err-x4	<-->	err-x18	4.090	-.182
err-x13	<-->	err-accuracy	4.480	-.160
err-x13	<-->	err-x17	4.459	.194
err-x23	<-->	err-usefulness	8.269	.335
err-x23	<-->	err-x8	6.003	.289
err-x26	<-->	err-x23	5.938	.225
err-x7	<-->	err-x23	6.232	-.154
err-x2	<-->	err-x13	4.994	-.169

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

		M.I.	Par Change
perceived design fit	<--- Perceived presentation fit	54.115	1.104
perceived design fit	<--- accuracy	32.909	.473
perceived design fit	<--- timeliness	34.508	.392
perceived design fit	<--- relevance	41.832	.522
perceived design fit	<--- believability	34.024	.501
perceived design fit	<--- infQual	41.174	.099
x16	<--- x17	11.199	.305
x18	<--- Perceived presentation fit	35.618	.831
x18	<--- accuracy	24.912	.382
x18	<--- timeliness	19.730	.275
x18	<--- relevance	27.257	.391
x18	<--- believability	29.888	.436
x18	<--- infQual	34.342	.084
x18	<--- usefulness	17.849	.410
x18	<--- ease of use	9.142	.336
x18	<--- x17	5.474	.179
x18	<--- x8	7.939	.162
x18	<--- x13	5.671	.150
x18	<--- x23	22.574	.294
x18	<--- x26	10.407	.229
x18	<--- x14	27.257	.391
x18	<--- x10	19.730	.275
x18	<--- x7	29.888	.436
x18	<--- x2	24.912	.382
x17	<--- x16	10.483	.194
x17	<--- x13	4.591	.125
x8	<--- x23	4.110	.161
x23	<--- x8	4.101	.124
x23	<--- x26	4.040	.153

Minimization History (Default model)

Iteration	Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	3	-.589	9999.000	709.587	0	9999.000
1	e	1	-.078	2.409	336.865	20	.364
2	e	1	-.055	.690	233.427	5	.844
3	e	0	180.042	.975	174.625	6	.735
4	e	0	332.627	1.028	158.153	1	1.131
5	e	0	2122.394	.954	157.261	1	.212
6	e	1	-.012	.917	154.476	1	.502
7	e	0	14227.176	.142	152.840	7	.926
8	e	0	30871.079	.460	152.737	1	1.157
9	e	0	57561.515	.382	152.712	1	1.245
10	e	0	109107.401	.272	152.703	1	1.238
11	e	0	159582.936	.187	152.700	1	1.177
12	e	0	204130.378	.076	152.700	1	1.100
13	e	0	219852.494	.015	152.700	1	1.022
14	e	0	221066.892	.000	152.700	1	1.001

Pairwise Parameter Comparisons (Default model)

Variance-covariance Matrix of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	.066																	
par_2	.045	.048																
par_3	.041	.036	.041															
par_4	.040	.036	.033	.042														
par_5	.041	.036	.032	.032	.054													
par_6	.000	.000	.000	.000	.000	.021												
par_7	.029	.026	.023	.023	.023	.000	.033											
par_8	.000	.000	.000	.000	.000	.000	.000	.021										
par_9	.000	.001	.001	.001	.000	.007	.000	.003	.031									

par_10	.020	.017	.015	.015	.015	-.008	.012	-.001	-.020	.047								
par_11	-.001	-.002	-.002	.002	-.001	-.066	-.001	.002	-.023	.030	34.121							
par_12	.000	-.001	.000	.001	.000	-.035	.000	.001	-.009	.012	13.471	7.857						
par_13	.000	.001	.001	.000	.000	.029	.000	.000	.013	-.015	-4.786	-2.238	1.536					
par_14	.000	.000	.000	.000	.000	.001	.000	.000	.000	-.001	-.790	-.350	.104	.019				
par_15	.000	.000	.000	.000	.000	-.014	.000	.000	-.011	.004	.052	.020	-.026	-.001	.023			
par_16	-.025	-.023	-.020	-.020	-.020	.000	-.014	.000	.000	-.010	.001	.000	.000	.000	.000	.015		
par_17	.000	.000	.000	-.002	.000	.000	.000	.000	.000	.001	-.003	-.001	.000	.000	.000	.000	.005	
par_18	-.001	.000	-.002	-.001	.000	.000	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000	
par_19	-.004	.000	.000	.000	-.001	.000	.000	.000	.000	-.001	.000	-.001	.001	.000	.000	.000	.000	
par_20	.001	-.002	.000	.000	.000	.000	.000	.000	.000	.000	.001	.001	-.001	.000	.000	.000	.000	
par_21	.000	.000	.000	.000	.000	.000	.000	-.007	-.007	.000	-.001	.000	.000	.000	.000	.000	.000	
par_22	.000	.000	.000	.000	.000	-.001	.000	.000	-.007	.005	-.009	-.006	.000	.000	-.001	.000	.000	
par_23	.000	.000	.000	.000	.000	-.007	.000	.000	-.002	.002	.021	.012	-.009	.000	.004	.000	.000	
par_24	.001	.001	.001	.001	-.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	-
par_25	.000	.000	.000	.000	.000	.000	-.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
par_26	.000	.000	.000	.000	.000	.000	.000	-.007	-.001	.001	-.002	-.001	.000	.000	.000	.000	.000	

Correlations of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	1.000																	
par_2	.793	1.000																
par_3	.788	.817	1.000															
par_4	.766	.795	.784	1.000														
par_5	.686	.712	.686	.673	1.000													
par_6	.007	.010	.009	.010	.004	1.000												
par_7	.624	.654	.626	.613	.556	.007	1.000											
par_8	.001	.001	.002	.001	.001	.003	-.002	1.000										
par_9	.009	.015	.023	.023	.010	.267	.002	.123	1.000									
par_10	.356	.365	.341	.330	.309	-.252	.296	-.019	-.535	1.000								
par_11	-.001	-.001	-.002	.001	.000	-.078	-.001	.002	-.022	.024	1.000							
par_12	.000	-.002	.000	.002	-.001	-.085	-.001	.001	-.018	.020	.823	1.000						
par_13	-.001	.004	.003	.001	.001	.159	.001	-.002	.058	-.056	-.661	-.644	1.000					
par_14	.001	.002	.001	-.002	.001	.056	.001	-.002	.018	-.021	-.972	-.896	.602	1.000				
par_15	-.007	-.013	-.009	-.009	-.004	-.623	-.010	.007	-.398	.113	.059	.047	-.141	-.050	1.000			
par_16	-.808	-.843	-.823	-.802	-.716	-.008	-.653	-.001	-.015	-.364	.001	.000	-.002	-.001	.009	1.000		
par_17	-.015	-.006	-.028	-.119	.003	-.007	.007	.000	-.025	.036	-.006	-.007	.003	.007	.001	.021	1.000	
par_18	-.032	-.018	-.137	-.040	.013	-.002	.014	-.003	-.031	.034	.004	-.001	-.005	-.003	.002	.033	.028	1
par_19	-.127	.021	-.021	-.015	-.022	.003	-.013	.000	.016	-.033	.000	-.002	.008	-.001	-.005	.021	-.016	
par_20	.047	-.127	.008	.012	-.001	-.007	-.018	.000	.003	-.012	.003	.008	-.009	-.004	.016	.010	-.063	-
par_21	.000	.000	.000	.000	.000	-.001	.001	-.310	-.237	.005	-.001	.000	.000	.001	-.002	.000	.000	
par_22	.000	.003	.006	.006	.003	-.029	.000	-.022	-.279	.144	-.011	-.015	-.002	.017	-.034	-.002	-.013	-
par_23	-.001	-.001	-.001	-.002	-.002	-.235	-.001	.002	-.047	.056	.018	.021	-.035	-.013	.117	.001	.002	
par_24	.024	.030	.047	.038	-.065	.005	.016	-.001	.006	.010	.000	.000	.002	.000	-.008	-.027	-.038	-
par_25	.006	.001	.019	.015	.001	-.001	-.074	.003	.013	-.017	.000	.001	.000	.000	.004	-.005	-.031	-
par_26	-.001	-.001	-.001	-.001	-.001	-.002	.001	-.276	-.019	.014	-.002	-.001	.001	.002	-.005	.001	.000	

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	.000																	
par_2	-1.067	.000																
par_3	-2.062	-1.227	.000															
par_4	-2.037	-1.234	-.079	.000														
par_5	-1.492	-.715	.203	.258	.000													
par_6	-4.461	-4.362	-3.972	-3.893	-3.741	.000												
par_7	-3.445	-3.093	-2.208	-2.089	-2.026	2.673	.000											
par_8	-3.366	-3.131	-2.675	-2.605	-2.566	1.580	-1.278	.000										
par_9	-3.335	-3.096	-2.668	-2.605	-2.566	1.408	-1.348	-.212	.000									
par_10	-3.626	-3.308	-2.722	-2.639	-2.620	1.138	-1.209	.029	.154	.000								
par_11	.427	.456	.483	.485	.477	.651	.546	.597	.604	.596	.000							
par_12	.012	.071	.128	.131	.115	.476	.259	.364	.379	.361	-.636	.000						
par_13	-1.891	-1.769	-1.647	-1.638	-1.668	-.884	-1.357	-1.126	-1.095	-1.112	-.727	-.652	.000					
par_14	-5.241	-5.238	-4.902	-4.802	-4.586	-1.118	-3.666	-2.701	-2.220	-2.108	-.674	-.533	.745	.000				
par_15	-3.433	-3.197	-2.757	-2.689	-2.652	1.077	-1.398	-.170	.037	-.172	-.604	-.377	1.079	2.405	.000			
par_16	-3.385	-3.218	-2.900	-2.854	-2.837	.439	-1.930	-1.265	-.892	-.864	-.638	-.449	.937	1.630	-1.052	.000		
par_17	-4.179	-4.102	-3.655	-3.463	-3.416	1.180	-2.189	-.808	-.439	-.609	-.619	-.411	1.027	2.621	-.567	.770	.00	
par_18	-4.608	-4.611	-4.096	-4.096	-3.911	.549	-2.795	-1.508	-1.007	-1.090	-.637	-.448	.942	2.003	-1.233	.022	-1.12	
par_19	-3.010	-2.915	-2.375	-2.309	-2.272	2.391	-.862	.620	.762	.423	-.578	-.324	1.220	3.682	.788	2.158	1.82	
par_20	-4.809	-4.610	-4.400	-4.295	-4.024	.343	-2.926	-1.708	-1.185	-1.216	-.643	-.460	.916	1.783	-1.432	-.213	-1.38	
par_21	-2.689	-2.374	-1.884	-1.826	-1.851	2.365	-.474	.750	.867	.654	-.566	-.299	1.270	3.457	1.000	2.122	1.81	
par_22	-4.431	-4.327	-3.946	-3.866	-3.723	-.002	-2.648	-1.550	-1.066	-1.353	-.652	-.478	.867	1.085	-1.340	-.438	-1.16	
par_23	-2.721	-2.414	-1.966	-1.914	-1.944	1.548	-.713	.415	.539	.329	-.580	-.328	1.194	2.614	.581	1.450	1.08	
par_24	-1.822	-1.400	-.840	-.784	-.853	3.867	.727	2.282	2.239	1.756	-.518	-.200	1.495	5.036	2.388	3.685	3.65	
par_25	-2.842	-2.541	-2.035	-1.961	-1.950	2.924	-.428	1.140	1.212	.804	-.562	-.292	1.292	4.255	1.297	2.727	2.56	
par_26	-.885	-.387	.177	.215	.044	4.458	1.634	2.739	2.968	2.513	-.475	-.111	1.690	5.496	3.169	4.359	4.32	

Bootstrap (Default model)

Summary of Bootstrap Iterations (Default model)

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	1	0
14	0	4	0
15	0	2	0
16	0	5	0
17	0	5	0
18	0	14	0
19	0	152	17
Total	0	183	17

0 bootstrap samples were unused because of a singular covariance matrix.

48 bootstrap samples were unused because a solution was not found.

200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 200 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 0 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap $p = .005$

Bootstrap Distributions (Default model)

ML discrepancy (implied vs sample) (Default model)

	30.235	*
	37.237	***
	44.239	*****
	51.241	*****
	58.242	*****
	65.244	*****
	72.246	*****
N = 200	79.248	*****
Mean = 73.319	86.250	*****
S. e. = 1.358	93.251	*****
	100.253	*****
	107.255	**
	114.257	***
	121.259	*
	128.260	*

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	26	152.700	52	.000	2.937
Saturated model	78	.000	0		
Independence model	12	734.347	66	.000	11.126

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.273	.873	.809	.582
Saturated model	.000	1.000		
Independence model	.566	.355	.237	.300

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.792	.736	.852	.809	.849
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.788	.624	.669
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	100.700	67.424	141.614
Saturated model	.000	.000	.000
Independence model	668.347	584.910	759.229

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	1.046	.690	.462	.970
Saturated model	.000	.000	.000	.000
Independence model	5.030	4.578	4.006	5.200

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.115	.094	.137	.000
Independence model	.263	.246	.281	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	204.700	209.783	282.451	308.451
Saturated model	156.000	171.248	389.254	467.254
Independence model	758.347	760.693	794.232	806.232

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	1.402	1.174	1.682	1.437
Saturated model	1.068	1.068	1.068	1.173
Independence model	5.194	4.623	5.817	5.210

HOELTER

Model	HOELTER	HOELTER
	.05	.01
Default model	67	76
Independence model	18	20

Execution time summary

Minimization: .016
 Miscellaneous: .216
 Bootstrap: 1.439
 Total: 1.671

\\staff\home\fca\tatema\My Documents\PhD\joerg commented version\apendices\MISM-7.amw

Analysis Summary

Date and Time

Date: Monday, 23 November 2009

Time: 7:31:51 p.m.

Title

Mism-7: Monday, 23 November 2009 07:31 p.m.

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.

Sample size = 147

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x2
x7
x10
x14
x26
x23
x13
x4
x8
x17
x18
x16
x20
x22
x28

Unobserved, endogenous variables

believability
relevance
timeliness
infQual
usefulness
accuracy
Perceived presentation fit
ease of use
perceived design fit
overall SQ

Unobserved, exogenous variables

err-x2
err-x7
err-x10
err-x14
err-x26
err-x23
err-accuracy
err-believability
err-timeliness
err-relevance
err-ease of use
err-x13
err-x4
err-x8
err-x17
err-x18
err-x16
err per des fit
err presentation fit
err-x20
err-x22
err-x28
err-OSQ
err-usefulness

Variable counts (Group number 1)

Number of variables in your model:	49
Number of observed variables:	15
Number of unobserved variables:	34
Number of exogenous variables:	24
Number of endogenous variables:	25

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	34	0	5	0	0	39
Labeled	0	0	0	0	0	0
Unlabeled	18	0	19	0	0	37
Total	52	0	24	0	0	76

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
x28	1.000	7.000	-1.277	-6.320	2.257	5.586
x22	1.000	7.000	-.756	-3.742	.904	2.238
x20	1.000	7.000	-1.022	-5.058	1.608	3.979
x16	1.000	7.000	-.604	-2.987	.037	.092
x18	1.000	7.000	-1.226	-6.070	2.271	5.621
x17	2.000	7.000	-.640	-3.168	.092	.227
x8	1.000	7.000	-1.042	-5.158	.883	2.185
x4	1.000	7.000	-.820	-4.059	.331	.818
x13	1.000	7.000	-.871	-4.309	.278	.688
x23	1.000	7.000	-.789	-3.904	.144	.357
x26	1.000	7.000	-1.091	-5.400	1.658	4.103
x14	1.000	7.000	-1.438	-7.119	3.131	7.750
x10	1.000	7.000	-1.062	-5.256	1.036	2.563
x7	1.000	7.000	-1.958	-9.694	5.865	14.514
x2	1.000	7.000	-1.199	-5.933	2.058	5.094
Multivariate					74.834	20.088

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
27	62.692	.000	.000
30	47.210	.000	.000
17	45.908	.000	.000
77	42.805	.000	.000
40	38.095	.001	.000
97	38.066	.001	.000
65	37.733	.001	.000
48	37.605	.001	.000
74	33.386	.004	.000
64	32.372	.006	.000
144	31.182	.008	.000
80	29.569	.014	.000
50	29.061	.016	.000
52	28.970	.016	.000
58	27.874	.022	.000
2	26.966	.029	.000
46	26.279	.035	.000
55	26.111	.037	.000
110	25.492	.044	.000
68	25.073	.049	.000
60	24.579	.056	.000
109	24.577	.056	.000
101	24.134	.063	.000
103	23.734	.070	.000
113	22.496	.095	.003
43	22.021	.107	.007
41	21.863	.111	.006
38	21.836	.112	.004
69	21.651	.117	.003
145	21.395	.125	.004
12	21.103	.134	.006
121	21.076	.134	.004
123	20.891	.140	.004
47	20.361	.159	.013
81	20.328	.160	.009
56	19.805	.180	.029
35	19.790	.180	.019
141	19.771	.181	.012
107	19.756	.181	.008
86	19.667	.185	.006
98	19.563	.189	.005
78	19.478	.193	.004
91	19.468	.193	.002
127	18.911	.218	.013
120	18.747	.225	.015
1	18.676	.229	.012
116	18.618	.232	.009
10	18.378	.243	.014
16	17.931	.266	.043
22	17.199	.307	.217
18	17.092	.313	.214
124	17.069	.315	.176
122	17.069	.315	.135

147	16.242	.366	.519
118	15.706	.402	.779
28	15.246	.434	.916
29	14.949	.455	.958
128	14.804	.466	.965
8	14.486	.489	.987
45	13.978	.527	.999
111	13.832	.538	.999
135	13.541	.561	1.000
136	13.541	.561	1.000
67	13.441	.568	1.000
84	13.386	.573	.999
70	13.035	.600	1.000
49	12.996	.603	1.000
92	12.662	.628	1.000
5	12.658	.629	1.000
90	12.636	.630	1.000
23	12.537	.638	1.000
137	12.394	.649	1.000
87	12.196	.664	1.000
131	11.937	.684	1.000
34	11.664	.704	1.000
138	11.526	.715	1.000
14	11.485	.717	1.000
95	11.368	.726	1.000
72	11.306	.731	1.000
6	11.222	.737	1.000
61	11.029	.751	1.000
125	10.772	.769	1.000
83	10.351	.797	1.000
119	10.337	.798	1.000
11	10.145	.811	1.000
85	10.069	.815	1.000
126	10.048	.817	1.000
20	9.914	.825	1.000
93	9.847	.829	1.000
88	9.836	.830	1.000
82	9.669	.840	1.000
54	9.648	.841	1.000
100	9.466	.852	1.000
115	9.246	.864	1.000
21	9.151	.869	1.000
102	8.893	.883	1.000
66	8.882	.884	1.000
99	8.882	.884	1.000
112	8.775	.889	1.000
104	8.744	.891	1.000

Sample Moments (Group number 1)

Sample Covariances (Group number 1)

	x28	x22	x20	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x28	1.336														
x22	.947	1.570													
x20	.901	.933	1.349												
x16	.669	.545	.636	1.710											
x18	.772	.778	.906	.526	1.337										
x17	.432	.493	.630	.570	.457	1.218									
x8	.678	.896	.704	.101	.470	.330	2.177								
x4	.821	.753	.704	.400	.477	.337	.660	1.755							
x13	.878	.894	.901	.545	.852	.620	.602	.819	1.954						
x23	.722	.957	.774	.323	.702	.555	.766	.684	.531	1.855					
x26	.644	.658	.759	.386	.520	.425	.417	.587	.530	.754	1.404				
x14	.715	.622	.784	.417	.695	.613	.619	.578	.725	.833	.604	1.270			
x10	.718	.871	.736	.454	.747	.658	.511	.702	.823	.963	.615	.921	1.863		
x7	.679	.591	.642	.325	.621	.437	.659	.564	.499	.634	.498	.809	.925	1.121	
x2	.731	.574	.746	.428	.619	.458	.529	.563	.451	.672	.500	.790	.891	.751	1.215

Condition number = 38.874

Eigenvalues

10.772 2.042 1.703 1.290 1.235 1.132 .955 .892 .707 .569 .503 .421 .354 .283 .277

Determinant of sample covariance matrix = .136

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 120

Number of distinct parameters to be estimated: 37
 Degrees of freedom (120 - 37): 83

Result (Default model)

Minimum was achieved
 Chi-square = 260.082
 Degrees of freedom = 83
 Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

		Estimate	S.E.	C.R.	P	Label
timeliness	<--- Perceived presentation fit	1.588	.244	6.507	***	par_5
relevance	<--- Perceived presentation fit	1.432	.209	6.861	***	par_6
believability	<--- Perceived presentation fit	1.286	.193	6.663	***	par_7
accuracy	<--- Perceived presentation fit	1.276	.196	6.505	***	par_8
infQual	<--- believability	1.078	.781	1.380	.168	par_1
infQual	<--- timeliness	-.243	.268	-.907	.365	par_2
infQual	<--- accuracy	1.000				
infQual	<--- relevance	.344	.524	.657	.511	par_4
usefulness	<--- infQual	.344	.140	2.448	.014	par_3
ease of use	<--- perceived design fit	1.044	.257	4.055	***	par_12
ease of use	<--- Perceived presentation fit	.401	.186	2.162	.031	par_18
overall SQ	<--- ease of use	.444	.074	5.959	***	par_15
overall SQ	<--- usefulness	.727	.144	5.034	***	par_17
x7	<--- believability	1.000				
x14	<--- relevance	1.000				
x10	<--- timeliness	1.000				
x2	<--- accuracy	1.000				
x26	<--- Perceived presentation fit	1.000				
x23	<--- Perceived presentation fit	1.343	.223	6.016	***	par_9
x13	<--- ease of use	1.000				
x8	<--- usefulness	1.000				
x17	<--- Perceived presentation fit	.921	.173	5.313	***	par_10
x18	<--- perceived design fit	1.000				
x16	<--- perceived design fit	.732	.172	4.248	***	par_11
x20	<--- overall SQ	1.000				
x22	<--- overall SQ	.983	.095	10.387	***	par_13
x28	<--- overall SQ	.989	.084	11.705	***	par_14
x4	<--- usefulness	.930	.177	5.253	***	par_16

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
err presentation fit	.431	.125	3.456	***	par_19
err-accuracy	.311	.073	4.283	***	par_20
err-believability	.204	.062	3.271	.001	par_21
err-timeliness	.577	.110	5.257	***	par_22
err-relevance	.185	.063	2.913	.004	par_23
err per des fit	.719	.208	3.461	***	par_24
err-ease of use	.295	.208	1.420	.156	par_25
err-usefulness	.241	.093	2.583	.010	par_26
err-OSQ	.001	.050	.014	.989	par_27
err-x2	.200				
err-x7	.200				
err-x10	.200				
err-x14	.200				
err-x13	.500				
err-x26	.973	.120	8.089	***	par_28
err-x23	1.078	.138	7.794	***	par_29
err-x4	1.146	.156	7.364	***	par_30
err-x8	1.473	.198	7.448	***	par_31
err-x17	.852	.105	8.101	***	par_32
err-x18	.618	.171	3.623	***	par_33
err-x16	1.325	.176	7.536	***	par_34
err-x20	.410	.063	6.499	***	par_35
err-x22	.662	.092	7.214	***	par_36
err-x28	.418	.064	6.575	***	par_37

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
Perceived presentation fit	.000
accuracy	.693
timeliness	.653

relevance	.827
believability	.778
perceived design fit	.000
ease of use	.743
usefulness	.657
overall SQ	.999
x28	.622
x22	.507
x20	.632
x16	.225
x18	.538
x17	.300
x8	.323
x4	.346
x13	.696
x23	.419
x26	.307
x14	.842
x10	.893
x7	.821
x2	.835

Matrices (Group number 1 - Default model)

Implied (for all variables) Covariances (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	infQual	ease of use	usefulness	overall SQ	x28	x22	x20	x16
Perceived presentation fit	.431													
accuracy	.549	1.012												
timeliness	.684	.873	1.664											
relevance	.617	.787	.979	1.068										
believability	.554	.707	.879	.793	.916									
perceived design fit	.000	.000	.000	.000	.000	.719								
infQual	1.192	1.832	1.753	1.771	1.753	.000	3.905							
ease of use	.173	.220	.274	.247	.222	.750	.478	1.147						
usefulness	.410	.630	.602	.608	.602	.333	1.187	.628	.702					
overall SQ	.374	.555	.560	.552	.536	.333	1.187	.628	.583	.703				
x28	.370	.549	.553	.546	.530	.329	1.174	.621	.576	.695	1.106			
x22	.368	.546	.550	.543	.527	.327	1.168	.618	.573	.692	.684	1.342		
x20	.374	.555	.560	.552	.536	.333	1.187	.628	.583	.703	.695	.692	1.114	
x16	.000	.000	.000	.000	.000	.526	.000	.549	.000	.244	.241	.240	.244	1.710
x18	.000	.000	.000	.000	.000	.719	.000	.750	.000	.333	.329	.327	.333	.520
x17	.397	.506	.630	.568	.510	.000	1.098	.159	.377	.345	.341	.339	.345	.000
x8	.410	.630	.602	.608	.602	.000	1.342	.164	.702	.583	.576	.573	.583	.000
x4	.381	.585	.560	.566	.560	.000	1.248	.153	.653	.542	.536	.533	.542	.000
x13	.173	.220	.274	.247	.222	.750	.478	1.147	.164	.628	.621	.618	.628	.540
x23	.578	.738	.919	.828	.744	.000	1.602	.232	.550	.503	.497	.495	.503	.000
x26	.431	.549	.684	.617	.554	.000	1.192	.173	.410	.374	.370	.368	.374	.000
x14	.617	.787	.979	1.068	.793	.000	1.771	.247	.608	.552	.546	.543	.552	.000
x10	.684	.873	1.664	.979	.879	.000	1.753	.274	.602	.560	.553	.550	.560	.000
x7	.554	.707	.879	.793	.916	.000	1.753	.222	.602	.536	.530	.527	.536	.000
x2	.549	1.012	.873	.787	.707	.000	1.832	.220	.630	.555	.549	.546	.555	.000

Implied Covariances (Group number 1 - Default model)

	x28	x22	x20	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x28	1.106														
x22	.684	1.342													
x20	.695	.692	1.114												
x16	.241	.240	.244	1.710											
x18	.329	.327	.333	.526	1.337										
x17	.341	.339	.345	.000	.000	1.218									
x8	.576	.573	.583	.000	.000	.377	2.175								
x4	.536	.533	.542	.000	.000	.351	.653	1.753							
x13	.621	.618	.628	.549	.750	.159	.164	.153	1.647						
x23	.497	.495	.503	.000	.000	.533	.550	.512	.232	1.855					
x26	.370	.368	.374	.000	.000	.397	.410	.381	.173	.578	1.404				
x14	.546	.543	.552	.000	.000	.568	.608	.566	.247	.828	.617	1.268			
x10	.553	.550	.560	.000	.000	.630	.602	.560	.274	.919	.684	.979	1.864		
x7	.530	.527	.536	.000	.000	.510	.602	.560	.222	.744	.554	.793	.879	1.116	
x2	.549	.546	.555	.000	.000	.506	.630	.585	.220	.738	.549	.787	.873	.707	1.212

Residual Covariances (Group number 1 - Default model)

	x28	x22	x20	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x28	.230														
x22	.263	.228													
x20	.206	.242	.236												

x16	.428	.305	.392	.000											
x18	.443	.451	.573	.000	.000										
x17	.091	.154	.286	.570	.457	.000									
x8	.102	.323	.120	.101	.470	-.047	.002								
x4	.285	.220	.161	.400	.477	-.014	.007	.002							
x13	.256	.276	.273	-.004	.102	.461	.437	.666	.307						
x23	.225	.463	.271	.323	.702	.022	.215	.172	.299	.000					
x26	.274	.290	.384	.386	.520	.028	.007	.206	.358	.176	.000				
x14	.170	.080	.232	.417	.695	.045	.011	.012	.478	.004	-.012	.002			
x10	.165	.321	.177	.454	.747	.029	-.091	.141	.548	.044	-.068	-.058	-.001		
x7	.148	.064	.106	.325	.621	-.073	.057	.004	.277	-.110	-.055	.016	.046	.006	
x2	.182	.028	.191	.428	.619	-.048	-.101	-.023	.231	-.066	-.050	.003	.018	.044	.003

Standardized Residual Covariances (Group number 1 - Default model)

	x28	x22	x20	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
x28	1.779														
x22	2.279	1.450													
x20	1.898	2.078	1.808												
x16	3.704	2.404	3.383	.000											
x18	4.252	3.953	5.476	.000	.000										
x17	.914	1.409	2.842	4.775	4.325	.000									
x8	.744	2.163	.876	.633	3.333	-.339	.009								
x4	2.310	1.637	1.301	2.794	3.766	-.110	.041	.009							
x13	2.086	2.071	2.207	-.029	.739	3.911	2.781	4.720	1.592						
x23	1.791	3.383	2.153	2.189	5.385	.165	1.250	1.111	2.046	.000					
x26	2.544	2.468	3.559	3.006	4.588	.250	.049	1.542	2.825	1.239	.000				
x14	1.572	.681	2.143	3.419	6.446	.396	.072	.094	3.940	.030	-.099	.015			
x10	1.294	2.318	1.380	3.071	5.722	.212	-.525	.903	3.735	.256	-.471	-.384	-.003		
x7	1.458	.576	1.032	2.843	6.145	-.689	.409	.029	2.434	-.817	-.489	.136	.326	.043	
x2	1.720	.243	1.793	3.591	5.873	-.445	-.700	-.175	1.949	-.476	-.423	.027	.126	.390	.024

Factor Score Weights (Group number 1 - Default model)

	x28	x22	x20	x16	x18	x17	x8	x4	x13	x23	x26	x14	x10	x7	x2
Perceived presentation fit	.017	.011	.017	-.004	-.012	.048	.006	.007	.005	.055	.046	.162	.093	.133	.101
accuracy	.034	.021	.035	-.004	-.012	.022	.016	.020	-.014	.025	.021	.070	.048	.046	.638
timeliness	-.001	.000	-.001	-.001	-.003	.020	-.002	-.002	.007	.023	.019	.070	.781	.060	.048
relevance	.019	.012	.020	-.004	-.011	.035	.008	.009	-.001	.040	.033	.598	.070	.093	.070
believability	.033	.021	.034	-.005	-.013	.029	.016	.019	-.012	.033	.027	.093	.060	.571	.046
perceived design fit	.070	.044	.072	.110	.321	-.008	-.019	-.023	.227	-.009	-.008	-.034	-.010	-.041	-.038
infQual	.077	.048	.079	-.010	-.030	.060	.037	.044	-.029	.069	.057	.359	-.053	.679	.700
ease of use	.146	.092	.151	.063	.184	.003	-.037	-.044	.458	.003	-.002	.018	-.030	-.036	
usefulness	.172	.108	.178	-.016	-.046	.009	.091	.108	-.108	.010	.009	.058	-.013	.116	.121
overall SQ	.191	.120	.197	.016	.048	.008	.049	.059	.124	.009	.007	.041	-.001	.071	.071

Total Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	ease of use	usefulness	overall SQ
accuracy	1.276	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	1.588	.000	.000	.000	.000	.000	.000	.000	.000
relevance	1.432	.000	.000	.000	.000	.000	.000	.000	.000
believability	1.286	.000	.000	.000	.000	.000	.000	.000	.000
infQual	2.769	1.000	-.243	.344	1.078	.000	.000	.000	.000
ease of use	.401	.000	.000	.000	.000	1.044	.000	.000	.000
usefulness	.952	.344	-.084	.118	.371	.000	.000	.000	.000
overall SQ	.870	.250	-.061	.086	.269	.463	.444	.727	.000
x28	.860	.247	-.060	.085	.266	.458	.439	.718	.989
x22	.855	.246	-.060	.085	.265	.455	.436	.715	.983
x20	.870	.250	-.061	.086	.269	.463	.444	.727	1.000
x16	.000	.000	.000	.000	.000	.732	.000	.000	.000
x18	.000	.000	.000	.000	.000	1.000	.000	.000	.000
x17	.921	.000	.000	.000	.000	.000	.000	.000	.000
x8	.952	.344	-.084	.118	.371	.000	.000	1.000	.000
x4	.885	.320	-.078	.110	.345	.000	.000	.930	.000
x13	.401	.000	.000	.000	.000	1.044	1.000	.000	.000
x23	1.343	.000	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000	.000
x14	1.432	.000	.000	1.000	.000	.000	.000	.000	.000
x10	1.588	.000	1.000	.000	.000	.000	.000	.000	.000
x7	1.286	.000	.000	.000	1.000	.000	.000	.000	.000
x2	1.276	1.000	.000	.000	.000	.000	.000	.000	.000

Direct Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	ease of use	usefulness	overall SQ
accuracy	1.276	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	1.588	.000	.000	.000	.000	.000	.000	.000	.000
relevance	1.432	.000	.000	.000	.000	.000	.000	.000	.000
believability	1.286	.000	.000	.000	.000	.000	.000	.000	.000
infQual	.000	1.000	-.243	.344	1.078	.000	.000	.000	.000
ease of use	.401	.000	.000	.000	.000	1.044	.000	.000	.000

usefulness	.000	.000	.000	.000	.000	.000	.000	.000	.000
overall SQ	.000	.000	.000	.000	.000	.000	.444	.727	.000
x28	.000	.000	.000	.000	.000	.000	.000	.000	.989
x22	.000	.000	.000	.000	.000	.000	.000	.000	.983
x20	.000	.000	.000	.000	.000	.000	.000	.000	1.000
x16	.000	.000	.000	.000	.000	.732	.000	.000	.000
x18	.000	.000	.000	.000	.000	1.000	.000	.000	.000
x17	.921	.000	.000	.000	.000	.000	.000	.000	.000
x8	.000	.000	.000	.000	.000	.000	.000	1.000	.000
x4	.000	.000	.000	.000	.000	.000	.000	.930	.000
x13	.000	.000	.000	.000	.000	.000	1.000	.000	.000
x23	1.343	.000	.000	.000	.000	.000	.000	.000	.000
x26	1.000	.000	.000	.000	.000	.000	.000	.000	.000
x14	.000	.000	.000	1.000	.000	.000	.000	.000	.000
x10	.000	.000	1.000	.000	.000	.000	.000	.000	.000
x7	.000	.000	.000	.000	1.000	.000	.000	.000	.000
x2	.000	1.000	.000	.000	.000	.000	.000	.000	.000

Indirect Effects (Group number 1 - Default model)

	Perceived presentation fit	accuracy	timeliness	relevance	believability	perceived design fit	ease of use	usefulness	overall SQ
accuracy	.000	.000	.000	.000	.000	.000	.000	.000	.000
timeliness	.000	.000	.000	.000	.000	.000	.000	.000	.000
relevance	.000	.000	.000	.000	.000	.000	.000	.000	.000
believability	.000	.000	.000	.000	.000	.000	.000	.000	.000
infQual	2.769	.000	.000	.000	.000	.000	.000	.000	.000
ease of use	.000	.000	.000	.000	.000	.000	.000	.000	.000
usefulness	.952	.344	-.084	.118	.371	.000	.000	.000	.000
overall SQ	.870	.250	-.061	.086	.269	.463	.000	.000	.000
x28	.860	.247	-.060	.085	.266	.458	.439	.718	.000
x22	.855	.246	-.060	.085	.265	.455	.436	.715	.000
x20	.870	.250	-.061	.086	.269	.463	.444	.727	.000
x16	.000	.000	.000	.000	.000	.000	.000	.000	.000
x18	.000	.000	.000	.000	.000	.000	.000	.000	.000
x17	.000	.000	.000	.000	.000	.000	.000	.000	.000
x8	.952	.344	-.084	.118	.371	.000	.000	.000	.000
x4	.885	.320	-.078	.110	.345	.000	.000	.000	.000
x13	.401	.000	.000	.000	.000	1.044	.000	.000	.000
x23	.000	.000	.000	.000	.000	.000	.000	.000	.000
x26	.000	.000	.000	.000	.000	.000	.000	.000	.000
x14	1.432	.000	.000	.000	.000	.000	.000	.000	.000
x10	1.588	.000	.000	.000	.000	.000	.000	.000	.000
x7	1.286	.000	.000	.000	.000	.000	.000	.000	.000
x2	1.276	.000	.000	.000	.000	.000	.000	.000	.000

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

		M.I.	Par Change
err per des fit	<--> err presentation fit	46.956	.392
err-usefulness	<--> err per des fit	4.889	.134
err-OSQ	<--> err-ease of use	7.551	-.122
err-x22	<--> err-timeliness	5.249	.156
err-x22	<--> err-relevance	6.336	-.129
err-x22	<--> err-usefulness	4.540	.109
err-x20	<--> err-ease of use	5.499	-.136
err-x16	<--> err-OSQ	7.477	.154
err-x16	<--> err-x28	5.504	.167
err-x18	<--> err presentation fit	31.369	.306
err-x18	<--> err-believability	4.859	.122
err-x18	<--> err-OSQ	5.450	.104
err-x18	<--> err-x20	8.074	.160
err-x17	<--> err per des fit	4.746	.170
err-x17	<--> err-x20	5.496	.131
err-x17	<--> err-x16	7.546	.258
err-x8	<--> err-x22	5.168	.207
err-x8	<--> err-x16	5.592	-.299
err-x4	<--> err-ease of use	14.980	.359
err-x13	<--> err-accuracy	6.093	-.151
err-x13	<--> err-OSQ	6.646	-.107
err-x13	<--> err-x16	4.131	-.192
err-x13	<--> err-x4	14.667	.342
err-x23	<--> err-usefulness	13.442	.241
err-x23	<--> err-OSQ	7.014	.134
err-x23	<--> err-x22	15.025	.300
err-x23	<--> err-x13	4.571	-.183
err-x26	<--> err-usefulness	11.505	.209
err-x26	<--> err-OSQ	11.684	.162
err-x26	<--> err-x20	7.383	.162
err-x26	<--> err-x23	4.849	.198
err-x14	<--> err-x22	7.507	-.138

err-x10	<-->	err-x22	6.447	.174
err-x7	<-->	err-ease of use	4.570	-.120
err-x7	<-->	err-x23	7.812	-.169
err-x2	<-->	err-ease of use	4.682	-.132
err-x2	<-->	err-x22	6.590	-.137
err-x2	<-->	err-x13	7.007	-.156

Variances: (Group number 1 - Default model)

	M.I.	Par Change
err-x13	8.787	.320

Regression Weights: (Group number 1 - Default model)

		M.I.	Par Change
perceived design fit	<--- Perceived presentation fit	46.956	.909
perceived design fit	<--- accuracy	35.362	.522
perceived design fit	<--- timeliness	35.416	.400
perceived design fit	<--- relevance	44.411	.565
perceived design fit	<--- believability	37.795	.566
perceived design fit	<--- infQual	39.975	.275
infQual	<--- perceived design fit	4.889	.541
usefulness	<--- perceived design fit	4.889	.186
usefulness	<--- ease of use	4.289	.130
x28	<--- x16	4.079	.092
x22	<--- x23	7.087	.140
x20	<--- x18	5.427	.119
x20	<--- x17	4.791	.117
x20	<--- x26	6.171	.124
x16	<--- x28	6.276	.237
x16	<--- x17	9.633	.279
x18	<--- Perceived presentation fit	31.369	.712
x18	<--- accuracy	29.236	.454
x18	<--- timeliness	20.269	.290
x18	<--- relevance	28.150	.430
x18	<--- believability	32.673	.504
x18	<--- infQual	33.909	.243
x18	<--- usefulness	28.271	.556
x18	<--- overall SQ	20.073	.448
x18	<--- x28	11.050	.250
x18	<--- x22	9.681	.212
x18	<--- x20	24.066	.367
x18	<--- x8	5.122	.121
x18	<--- x23	19.332	.255
x18	<--- x26	7.312	.180
x18	<--- x14	22.498	.333
x18	<--- x10	16.484	.235
x18	<--- x7	28.903	.402
x18	<--- x2	24.370	.354
x17	<--- perceived design fit	4.746	.236
x17	<--- ease of use	5.233	.187
x17	<--- x20	4.025	.149
x17	<--- x16	10.230	.191
x17	<--- x13	5.215	.139
x4	<--- ease of use	8.160	.278
x4	<--- x13	13.667	.269
x13	<--- x16	4.274	-.126
x13	<--- x4	7.717	.167
x23	<--- x22	10.887	.254
x26	<--- x20	8.546	.232
x10	<--- x22	4.278	.141
x7	<--- x23	4.646	-.089
x2	<--- x22	4.670	-.115

Minimization History (Default model)

Iteration	Negative eigenvalues	Condition #	Smallest eigenvalue	Diameter	F	NTries	Ratio
0	e	8	-.638	9999.000	1225.064	0	9999.000
1	e	9	-.141	2.452	724.819	20	.444
2	e*	2	-.105	1.540	429.647	4	.741
3	e	0	1451.590	.622	331.541	4	.880
4	e	0	556.904	.747	310.888	5	.000
5	e	1	-.067	.622	279.073	2	.000
6	e	0	489.835	.484	263.934	8	.921
7	e	0	717.941	.426	260.558	1	1.136
8	e	0	968.866	.182	260.118	1	1.125
9	e	0	1229.807	.078	260.084	1	1.111
10	e	0	1353.025	.023	260.082	1	1.051
11	e	0	1383.687	.002	260.082	1	1.005
12	e	0	1382.853	.000	260.082	1	1.000

Pairwise Parameter Comparisons (Default model)

Variance-covariance Matrix of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	.610																	
par_2	-.084	.072																
par_3	-.078	.006	.020															
par_4	.015	-.041	-.039	.275														
par_5	.001	-.001	.000	-.001	.060													
par_6	-.001	.001	.000	.000	.040	.044												
par_7	.003	-.001	.000	-.003	.037	.033	.037											
par_8	-.001	.000	.000	-.003	.036	.032	.029	.038										
par_9	.000	.000	.000	.000	.037	.033	.029	.029	.050									
par_10	-.001	.000	.000	.000	.026	.023	.020	.020	.021	.030								
par_11	.003	-.001	.000	-.001	.000	.000	.000	.000	.000	.030	.030							
par_12	.001	.000	.000	-.002	.000	.000	.001	.001	.001	.000	.019	.066						
par_13	.003	.001	-.001	-.001	.000	.000	.000	.000	.000	.000	.000	.000	.009					
par_14	.002	.000	.000	-.002	.000	.000	.000	.000	.000	.000	.000	.000	.004	.007				
par_15	.004	-.003	.001	-.004	.000	.000	.000	.000	.000	.000	.001	.000	-.002	-.002	.006			
par_16	-.008	.005	-.006	.000	.000	.000	.000	.000	.000	.000	-.001	-.001	.000	.000	-.001	.031		
par_17	-.020	.005	-.003	.001	.000	.000	.000	.000	.000	.000	-.001	.000	-.003	-.002	-.004	.014	.021	
par_18	.004	-.002	.000	.000	.011	.010	.008	.008	.009	.007	.002	-.016	.000	.000	.000	.000	-.004	
par_19	.000	.000	.000	.001	-.024	-.022	-.020	-.019	-.020	-.014	.000	.000	.000	.000	.000	.000	.000	-
par_20	.007	-.001	-.002	.004	.000	.000	.000	-.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	
par_21	-.011	.002	.001	.004	.000	.000	-.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
par_22	-.001	.003	.000	-.001	-.003	.000	-.001	.000	-.001	.000	.000	.000	.000	.000	.000	.000	.000	-
par_23	.002	.000	.000	-.004	.001	-.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
par_24	-.003	.001	.000	.001	.000	.000	.000	.000	.000	.000	-.020	-.039	.000	.000	-.001	.001	.001	-
par_25	-.004	.002	.000	.002	.000	.000	.000	.000	.000	.000	-.009	-.039	.001	.000	-.006	.001	.005	
par_26	.001	.001	.001	.004	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000	.001	-.009	-.007	
par_27	.006	.000	-.001	.002	.000	.000	.000	.000	.000	.000	.000	-.001	.000	.000	.000	.000	-.002	
par_28	.000	.000	.000	-.001	.003	.003	.002	.002	.002	.002	.000	.000	.000	.000	.000	.000	.000	
par_29	.000	.000	.000	-.001	.001	.001	.001	.001	-.002	.000	.000	.000	.000	.000	.000	.000	.000	
par_30	-.003	-.003	.001	-.002	.000	.000	.000	.000	.000	.000	.000	.001	.000	.000	.000	-.005	.001	-
par_31	-.016	.004	.000	-.003	.000	.000	.000	.000	.000	.000	-.001	-.001	-.001	.000	-.001	.006	.007	-
par_32	.001	.000	.000	.000	.000	.000	.000	.000	.000	-.001	.000	.001	.000	.000	.000	.000	.000	-
par_33	.003	-.001	.000	-.001	.000	.000	.000	.000	.000	.000	.014	.030	.000	.000	.001	-.001	-.001	-
par_34	-.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.007	.000	.000	.000	.000	.000	.001	-
par_35	.003	.001	.000	-.001	.000	.000	.000	.000	.000	.000	.000	.000	.001	.001	-.001	.000	.000	
par_36	-.003	-.002	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.001	.000	.000	.000	.001	
par_37	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	-.001	.000	.000	.000	

Correlations of Estimates (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17	par_18
par_1	1.000																	
par_2	-.400	1.000																
par_3	-.710	.162	1.000															
par_4	.037	-.288	-.531	1.000														
par_5	.003	-.021	.008	-.006	1.000													
par_6	-.008	.009	-.004	.000	.782	1.000												
par_7	.018	-.014	-.009	-.033	.778	.808	1.000											
par_8	-.007	-.002	.003	-.025	.755	.786	.777	1.000										
par_9	.000	-.005	.002	.001	.679	.709	.684	.671	1.000									
par_10	-.006	.005	.003	-.002	.609	.641	.612	.600	.547	1.000								
par_11	.020	-.019	.001	-.010	.001	.000	.000	-.002	.002	-.005	1.000							
par_12	.007	-.006	.006	-.012	.006	.006	.017	.014	.009	-.008	.437	1.000						
par_13	.043	.054	-.044	-.017	.006	.003	.005	.003	.004	.001	-.001	-.014	1.000					
par_14	.024	.009	-.009	-.035	.004	.002	.005	.003	.002	-.001	.013	-.001	.445	1.000				
par_15	.071	-.140	.062	-.103	.003	-.001	.001	.001	.002	-.001	.065	.019	-.244	-.289	1.000			
par_16	-.059	.103	-.223	-.004	-.002	-.002	-.003	-.002	-.004	-.001	-.023	-.022	-.006	.024	-.076	1.000		
par_17	-.175	.125	-.156	.010	-.007	-.005	-.004	.001	-.004	-.003	-.045	-.009	-.198	-.182	-.391	.531	1.000	
par_18	.028	-.037	-.009	.002	.239	.251	.219	.212	.208	.209	.051	-.336	.016	.008	.015	.000	-.157	1
par_19	.000	-.002	-.002	.009	-.796	-.833	-.816	-.795	-.713	-.640	.000	-.009	-.004	-.003	-.002	.002	.005	-
par_20	.123	-.037	-.167	.115	-.015	-.001	-.024	-.123	.002	.009	.006	-.018	.004	.000	.005	-.003	-.019	
par_21	-.223	.102	.068	.132	-.032	-.007	-.142	-.034	.012	.020	-.003	-.029	-.004	-.006	.003	.005	.001	
par_22	-.009	.091	-.018	-.009	-.130	.008	-.030	-.022	-.023	-.019	-.002	.005	-.006	-.003	-.003	.000	.007	-
par_23	.044	-.009	.047	-.131	.038	-.130	.014	.015	-.003	-.023	-.001	.014	.006	.007	.018	-.003	-.004	-
par_24	-.016	.016	-.001	.008	.000	.000	.000	.002	-.002	.004	-.555	-.723	.000	-.011	-.053	.019	.037	-
par_25	-.027	.039	-.002	.016	-.003	-.002	-.005	-.003	-.003	.004	-.261	-.721	.027	.023	-.356	.029	.176	
par_26	.019	.031	.093	.074	.016	.020	.011	.012	.003	.014	.019	.012	.074	.008	.082	-.521	-.525	
par_27	.152	-.025	-.094	.081	.007	.009	.006	.000	.003	.005	-.039	-.094	.031	.004	-.044	.017	-.266	
par_28	.000	.002	.002	-.009	.104	.103	.107	.102	.072	.074	.000	.009	.004	.003	.002	-.002	-.005	
par_29	-.001	.012	-.001	-.016	.026	.031	.048	.039	-.066	.018	-.005	-.002	-.001	.002	.000	.003	.000	
par_30	-.028	-.062	.068	-.026	-.001	-.002	.000	.001	.001	-.001	.011	.019	-.027	-.028	.026	-.175	.033	-
par_31	-.105	.083	.010	-.029	-.004	-.004	-.004	-.001	-.004	-.003	-.021	-.013	-.033	.007	-.080	.179	.237	-
par_32	.008	-.004	-.002	-.006	.005	.000	.020	.014	.001	-.074	.007	.019	.002	.005	.003	.000	-.001	-
par_33	.020	-.019	.001	-.010	.001	.000	.000	-.002	.002	-.005	.465	.679	-.001	.013	.064	-.023	-.045	
par_34	-.010	.010	-.001	.005	.000	.000	.000	.001	-.001	.003	-.242	.007	.000	-.007	-.034	.012	.023	-
par_35	.052	.046	-.039	-.044	.008	.004	.007	.005	.004	.000	.012	-.010	.158	.204	-.113	.018	-.032	
par_36	-.047	-.072	.057	.003	-.006	-.003	-.004	-.002	-.005	-.003	.009	.019	-.158	-.019	.036	.024	.058	-

par_37	-.008	.024	-.017	.043	-.002	-.001	-.004	-.003	.000	.003	-.022	-.009	-.047	-.202	.088	-.042	-.018	-
--------	-------	------	-------	------	-------	-------	-------	-------	------	------	-------	-------	-------	-------	------	-------	-------	---

Critical Ratios for Differences between Parameters (Default model)

	par_1	par_2	par_3	par_4	par_5	par_6	par_7	par_8	par_9	par_10	par_11	par_12	par_13	par_14	par_15	par_16	par_17
par_1	.000																
par_2	-1.433	.000															
par_3	-.829	2.081	.000														
par_4	-.794	.898	.001	.000													
par_5	.624	4.998	4.436	2.146	.000												
par_6	.436	4.949	4.318	1.927	-1.021	.000											
par_7	.259	4.595	3.933	1.668	-1.971	-1.162	.000										
par_8	.245	4.566	3.872	1.651	-1.946	-1.170	-.075	.000									
par_9	.326	4.533	3.795	1.754	-1.302	-.534	.343	.391	.000								
par_10	-.196	3.651	2.593	1.044	-3.418	-3.096	-2.249	-2.132	-2.180	.000							
par_11	-.435	3.031	1.749	.701	-2.867	-2.586	-2.141	-2.081	-2.170	-.771	.000						
par_12	-.042	3.451	2.394	1.192	-1.540	-1.174	-.758	-.723	-.883	.395	1.304	.000					
par_13	-.121	4.384	3.704	1.196	-2.316	-1.960	-1.411	-1.346	-1.487	.315	1.277	-.220	.000				
par_14	-.115	4.388	3.922	1.207	-2.325	-1.970	-1.414	-1.348	-1.487	.350	1.343	-.204	.056	.000			
par_15	-.814	2.382	.647	.185	-4.489	-4.458	-4.073	-3.969	-3.824	-2.529	-1.574	-2.251	-4.029	-4.267	.000		
par_16	-.183	3.834	2.353	1.057	-2.182	-1.833	-1.358	-1.310	-1.449	.036	.791	-.361	-.265	-.302	2.466	.000	
par_17	-.429	3.364	1.770	.705	-3.029	-2.772	-2.316	-2.256	-2.315	-.860	-.023	-1.070	-1.367	-1.454	1.518	-1.284	.000
par_18	-.849	1.941	.246	.102	-4.413	-4.259	-3.741	-3.651	-3.638	-2.302	-1.342	-1.764	-2.813	-2.890	-.214	-2.062	-1.291
par_19	-.819	2.276	.463	.160	-3.294	-3.129	-2.820	-2.775	-2.816	-1.813	-1.418	-2.137	-3.526	-3.702	-.090	-2.309	-1.557
par_20	-.990	1.974	-.195	-.064	-4.997	-5.073	-4.693	-4.442	-4.400	-3.258	-2.259	-2.728	-5.651	-6.088	-1.282	-3.234	-2.555
par_21	-1.097	1.660	-.935	-.271	-5.455	-5.629	-5.128	-5.161	-4.932	-3.920	-2.882	-3.152	-6.869	-7.459	-2.477	-3.877	-3.325
par_22	-.634	2.924	1.300	.434	-3.606	-3.636	-3.151	-3.079	-3.051	-1.661	-.757	-1.670	-2.791	-2.963	1.006	-1.692	-.826
par_23	-1.144	1.549	-1.049	-.297	-5.617	-5.520	-5.442	-5.316	-4.986	-3.958	-2.979	-3.250	-7.024	-7.631	-2.668	-3.958	-3.430
par_24	-.443	2.857	1.496	.666	-2.713	-2.422	-2.000	-1.952	-2.046	-.749	-.040	-.752	-1.159	-1.198	1.227	-.780	-.032
par_25	-.963	1.617	-.195	-.088	-4.032	-3.860	-3.490	-3.431	-3.436	-2.321	-1.446	-1.736	-3.051	-3.122	-.610	-2.363	-1.871
par_26	-1.067	1.721	-.637	-.196	-5.184	-5.249	-4.896	-4.788	-4.560	-3.475	-2.527	-2.944	-5.802	-5.966	-1.770	-2.880	-2.324
par_27	-1.390	.890	-2.237	-.657	-6.382	-6.684	-6.458	-6.301	-5.872	-5.109	-4.035	-3.910	-9.301	-10.090	-4.845	-5.075	-4.405
par_28	-.133	4.138	3.410	1.167	-2.359	-1.994	-1.446	-1.380	-1.505	.257	1.147	-.249	-.065	-.104	3.746	.203	1.308
par_29	.000	4.397	3.726	1.348	-1.840	-1.434	-.897	-.841	-.982	.714	1.562	.117	.565	.553	4.038	.661	1.757
par_30	.085	4.362	3.966	1.456	-1.527	-1.096	-.564	-.519	-.725	.966	1.793	.343	.884	.880	4.113	.847	2.005
par_31	.477	5.364	4.678	1.994	-.368	.141	.674	.705	.432	2.095	2.794	1.312	2.204	2.257	4.747	2.255	3.463
par_32	-.287	3.795	2.898	.949	-2.774	-2.479	-1.989	-1.915	-1.990	-.328	.598	-.693	-.925	-1.011	3.176	-.376	.705
par_33	-.578	2.685	1.244	.495	-3.258	-3.018	-2.593	-2.528	-2.584	-1.242	-.642	-2.252	-1.870	-1.956	.961	-1.253	-.475
par_34	-.307	4.910	4.361	1.776	-.876	-.392	.149	.185	-.065	1.638	2.161	.904	1.711	1.720	4.561	1.593	2.660
par_35	-.856	2.395	.427	.124	-4.682	-4.691	-4.323	-4.208	-4.026	-2.769	-1.761	-2.385	-5.451	-6.115	-.325	-2.781	-1.986
par_36	-.526	3.123	1.950	.597	-3.546	-3.374	-2.916	-2.834	-2.818	-1.320	-.361	-1.406	-2.265	-2.594	1.881	-1.357	-.390
par_37	-.842	2.411	.482	.141	-4.636	-4.644	-4.265	-4.156	-3.984	-2.725	-1.697	-2.354	-4.850	-4.935	-.271	-2.684	-1.942

Bootstrap (Default model)**Summary of Bootstrap Iterations (Default model)**

(Default model)

Iterations	Method 0	Method 1	Method 2
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	1	0
13	0	4	0
14	0	4	1
15	0	4	0
16	0	10	0
17	0	9	0
18	0	18	0
19	0	145	4
Total	0	195	5

0 bootstrap samples were unused because of a singular covariance matrix.

4 bootstrap samples were unused because a solution was not found.

200 usable bootstrap samples were obtained.

Bollen-Stine Bootstrap (Default model)

The model fit better in 200 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 0 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap p = .005

Bootstrap Distributions (Default model)**ML discrepancy (implied vs sample) (Default model)**

	52.253	-----
	61.241	*
	70.229	****
	79.217	*****
	88.205	*****
	97.193	*****
	106.181	*****
N = 200	115.168	*****
Mean = 108.951	124.156	*****
S. e. = 1.722	133.144	*****
	142.132	*****
	151.120	***
	160.108	****
	169.096	*
	178.084	*

Model Fit Summary**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	37	260.082	83	.000	3.134
Saturated model	120	.000	0		
Independence model	15	1196.727	105	.000	11.397

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.286	.823	.744	.569
Saturated model	.000	1.000		
Independence model	.625	.263	.158	.230

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.783	.725	.841	.795	.838
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.790	.619	.662
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	177.082	132.205	229.580
Saturated model	.000	.000	.000
Independence model	1091.727	984.124	1206.759

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	1.781	1.213	.906	1.572
Saturated model	.000	.000	.000	.000
Independence model	8.197	7.478	6.741	8.265

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.121	.104	.138	.000
Independence model	.267	.253	.281	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	334.082	343.190	444.728	481.728
Saturated model	240.000	269.538	598.852	718.852
Independence model	1226.727	1230.420	1271.584	1286.584

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	2.288	1.981	2.648	2.351
Saturated model	1.644	1.644	1.644	1.846
Independence model	8.402	7.665	9.190	8.428

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	60	66
Independence model	16	18

Execution time summary

Minimization: .016
Miscellaneous: .249
Bootstrap: .919
Total: 1.184