The translation and validation of the Postpartum Depression Screening Scale (PDSS): towards improving screening for postpartum depression in English- and Afrikaans-speaking South African women

MELONY STRUIK

Submitted in fulfillment of the requirements for the degree

PHD (PSYCHOLOGY)

In the

FACULTY OF HUMANITIES

At the

University of Pretoria

Supervisor: Prof D.J.F. Maree

2011

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ACKNOWLEDGEMENTS

I am thankful to my heavenly Father, my source of patience, hope, and strength. I am also thankful for the support and encouragement from my family and friends. I would also like to express my sincere appreciation to the following people who all had a significant input in this research in varying degrees

- To my friend, Marlé, the source of my inspiration for reaching out to postpartum mothers in distress;
- My supervisor, Prof. David Maree, for his prompt guidance, his encouragement and support throughout this research;
- To all the mothers who offered up some of their valuable time in an often hectic early postpartum period to participate in this research;
- To the obstetricians, general practitioners, and many nursing professionals for their interest in this study and for referring mothers for participation;
- Clare Huisamen at Your Baby magazine, Ruth Rehbock at Living and Loving magazine, and the editors of Kids Connection and Fit Pregnancy for the opportunity to recruit mothers for participation in this study;
- To all the dedicated people at PNDSA, the PNDSA affiliated psychologists, and other health practitioners who assisted with this research and accepted many referrals of mothers in distress.
- To the University of Pretoria for the bursary which helped to finance this research.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ABSTRACT

Postpartum depression is an illness that is frequently unreported and undetected for a variety of reasons and may be potentially devastating for the mother affected as well as her family. Routine screening of postpartum women enables health practitioners to detect symptoms of PPD early and provides an opportunity for early intervention which may improve the outcome and increase the mother's chances of an earlier recovery. It is therefore important that reliable and convenient screening tools are available to health practitioners who have contact with postpartum women.

The primary objective of this research was to make an Afrikaans version of an existing screening scale available – the Postpartum Depression Screening Scale (PDSS), designed specifically to encompass the multifaceted phenomenon of PPD. In accordance with this objective, the validity and reliability of the PDSS and its Afrikaans version was investigated in English- and Afrikaans-speaking South African mothers. A further objective of this study was to compare the performance of the PDSS with the Edinburgh Postnatal Depression Scale (EPDS) and the Quick Inventory of Depressive Symptomatology (QIDS-SR16).

Various factors have been reported to be associated with the development of PPD. The final objective of this study was to explore the relationship between known risk factors for PPD and high scores on the PDSS amongst women in South African.

A total of 365 South African mothers, between 4 and 16 weeks postpartum participated in this study. English-speaking mothers (n = 187) completed the PDSS, EPDS, QIDS, and a demographic and psychosocial questionnaire, while Afrikaans-

speaking mothers (n = 178) completed the respective Afrikaans versions of these questionnaires. A multiple translation method – Brislin's back-translation method and the committee approach – was used to translate the PDSS and the QIDS into Afrikaans.

An item response theory (IRT), Rasch analysis, was used to examine dimensionality, item difficulty, differential item functioning, and category functioning of the PDSS and the Afrikaans PDSS.

Results reveal excellent person reliability estimates for the Afrikaans PDSS as well as for the PDSS in a South African sample. Both language versions performed reasonably well and the majority of items in the PDSS dimensions and the Afrikaans PDSS dimensions demonstrated fit statistics that supported the underlying constructs of each dimension. Some items were identified as problematic, namely Item 2, Item 25, Item 28, and Item 30. The item person construct maps show reasonably good spread of items. There were, however, persons that scored higher than the items could measure and an overrepresentation of items at the mean level. The Likert response categories proved to be effective for all the Afrikaans PDSS items and almost all the PDSS items.

Results indicate that 49.7% of mothers screened positive for major PPD using the PDSS. A further 17.3% of mothers obtained scores indicating the presence of significant symptoms of PPD.

Statistically significant correlations were obtained between total scores on the PDSS, the EPDS, and the QIDS-SR16. Stepwise multiple regression analysis identified 11 variables that were significantly associated with a high PDSS total score. These were a history of psychiatric illness, postpartum blues, feeling negative or ambivalent about

expecting this baby, fearful of childbirth, infant temperament, antenatal depression in recent pregnancy, lack of support from the baby's father, concern about health related issues regarding the infant, lack of support from friends, difficulty conceiving, and life stress.

Key words:

Postpartum depression, screening, Postpartum Depression Screening Scale, Edinburgh Postnatal Depression Scale, Quick Inventory of Depressive Symptomatology, Item response theory, Rasch analysis, Multiple regression analysis, Risk factors, Afrikaans, Translation, Adaptation, Cross-cultural research.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 2
ABSTRACT 3
TABLE OF CONTENTS 6
LIST OF TABLES 16
LIST OF FIGURES 22

CHAPTER 1 23
INTRODUCTION 23
1.1 Aim of the Study 23
1.2 Contextualising the Research 24
1.3 An Overview of the Research Method 31
1.4 Orientation 33

CHAPTER 2 36
PERINATAL MOOD DISORDERS AND POSTPARTUM DEPRESSION 36
2.1 Chapter Preview 36
2.2 Introduction 36
2.3 Perinatal Mood Disorders 37
2.3.1 Antenatal mood and anxiety disorders. 38
2.3.2 Postpartum blues. 41
2.3.3 Postpartum depression. 43
2.3.4 Obsessive-compulsive disorder occurring in the postpartum period. 43
2.3.5 Postpartum onset of panic disorder. 46
2.3.6 Postpartum posttraumatic stress disorder. 49
2.3.7 Puerperal psychosis. 51
2.3.8 Anger in the postpartum period. 54
2.4 Postpartum Depression 55
2.4.1 Historical perspectives. 55

2.4.2 Diagnosing postpartum depression. 60
2.4.3 Symptoms of postpartum depression. 61
2.4.3.1 Symptom overlap between the postpartum period and postpartum depression. 61
2.4.3.2 Symptoms of postpartum depression versus depression. 61
2.4.3.3 Symptoms of postpartum depression. 62
2.4.4 Prevalence of postpartum depression. 66
2.4.5 Clinical course of postpartum depression. 70
2.4.6 Perspectives on the etiology of postpartum depression. 71
2.4.7 Risk factors for postpartum depression. 77
2.4.7.1 Antenatal depression and anxiety. 79
2.4.7.2 Past history of depression. 81
2.4.7.3 Postpartum blues. 83
2.4.7.4 Hormonal changes. 84
2.4.7.5 Obstetric risk factors. 89
2.4.7.6 Psychosocial adjustments. 95
2.4.7.7 Self-esteem. 96
2.4.7.8 Personality organization. 98
2.4.7.9 Infant temperament. 100
2.4.7.10 Sleep deprivation. 102
2.4.7.11 Lack of support. 105
2.4.7.12 Marital difficulties. 107
2.4.7.13 Single parenthood. 108
2.4.7.14 Adolescent age. 109
2.4.7.15 Unplanned pregnancy, ambivalence about having a child. 110
2.4.7.16 Maternal or paternal unemployment or poverty. 111
2.4.7.17 Childcare stress. 113
2.4.7.18 High stress levels and adverse life events. 114
2.4.8 Consequences of postpartum depression. 117
2.5 Conclusion 123

CHAPTER 3 124

SCREENING FOR POSTPARTUM DEPRESSION 124

3.1 Chapter Preview 124

3.2 Screening for Postpartum Depression 126

3.3 Screening Measures 129

3.3.1 The Beck Depression Inventory (BDI and BDI-II). 132

3.3.2 The Inventory of Depressive Symptomatology (IDS) and Quick Inventory of Depressive Symptomatology (QIDS). 133

3.3.3 The Bromley Postnatal Depression Scale (BPDS). 136

3.3.4 The Edinburgh Postnatal Depression Scale (EPDS). 137

3.3.5 The Postpartum Depression Screening Scale (PDSS). 143

3.4 Conceptual Basis of the PDSS 145

3.5 Development of the PDSS 148

3.5.1 Generation of items. 148

3.5.2 Item content validity. 149

3.6 Psychometric Properties of the PDSS 151

3.6.1 Reliability. 151

3.6.2 Validity. 155

3.6.2.1 Confirmatory factor analysis. 155

3.6.2.2 Item response theory. 156

3.7 Comparative Analysis of the Performance of the PDSS with Other Depression Instruments 161

3.8 Conclusion 166

CHAPTER 4 168

CROSS-CULTURAL ASSESSMENT 168

4.1 Chapter Preview 168

4.2 Cross-Cultural Assessment 168

4.2.1 Multicultural assessment in South Africa. 170

4.2.1.1 Instrument development versus translation and adaptation. 170

4.2.1.2 Progression of psychological assessment in South Africa. 171

4.3 Culture-Fair Tests ... 174
4.4 Factors Influencing Cross-Cultural Assessment ... 177
4.4.1 Schooling. ... 179
4.4.2 Language. ... 179
4.4.3 Culture. ... 181
4.4.4 Environmental factors. ... 182
4.4.4.1 The home environment. ... 183
4.4.4.2 Socio-economic status. ... 183
4.4.4.3 Urbanization. ... 183
4.5 Methodological Considerations in Cross-Cultural Assessment ... 184
4.5.1 Bias. ... 184
4.5.1.1 Construct bias. ... 185
4.5.1.2 Method bias. ... 185
4.5.1.3 Item bias ... 186
4.5.2 Equivalence. ... 189
4.5.2.1 Construct equivalence. ... 190
4.5.2.2 Measurement unit equivalence. ... 190
4.5.2.3 Scalar equivalence. ... 190
4.5.2.4 Linguistic equivalence. ... 191
4.6 Ethical Guidelines for Adaptation of Cross-Cultural Assessment Measures ... 191
4.7 Translating Assessment Measures ... 198
4.7.1 Techniques in translating instruments. ... 198
4.7.1.1 One way or bilingual translation. ... 200
4.7.1.2 Forward-translation. ... 200
4.7.1.3 Modified direct translation. ... 201
4.7.1.4 Parallel blind translation. ... 202
4.7.1.5 Committee approach. ... 202
4.7.1.6 Pilot-testing or pretest. ... 203
4.7.1.7 Field-testing. ... 205
4.7.1.8 Random probe. ... 205
4.7.1.9 Decentering. ... 205

4.7.1.10 Back-translation. 207
4.7.2 Translation procedure. 211
4.7.2.1 Application. 212
4.7.2.2 Adaptation. 212
4.7.2.3 Assembly. 213
4.8 Conclusion 213

CHAPTER 5 215
A CULTURAL APPROACH TO PERINATAL MOOD DISORDERS 215
5.1 Chapter Preview 215
5.2 Paradigms of Mental Illness 216
5.3 Prevalence of PPD Across Different Cultures 217
5.4 Environmental and Cultural Influence on PPD Prevalence 222
5.5 Symptom Definition and Expression Across Cultures 228
5.6 Cultural Factors, Beliefs, and Rituals Associated With Pregnancy and Childbirth in South Africa 232
5.7 Use of PPD Screening Measures Across Different Cultures 235
5.8 Conclusion 240

CHAPTER 6 241
AFRIKAANS-SPEAKING SOUTH AFRICANS 241
6.1 Chapter Preview 241
6.2 Definition of Terms 241
6.2.1 Afrikaner. 241
6.2.2 Culture. 243
6.2.3 Cultural group. 244
6.2.4 Ethnic group. 244
6.2.5 Racial group. 245
6.2.6 Classification group. 246
6.3 Historical Overview 246
6.4 The Development of Afrikaans 252

6.4.1 The history of the Afrikaans language. 252
6.4.2 The influence of other languages. 256
6.4.3 Landmarks in the extension of the functions of Afrikaans. 259
6.5 Linguistic Diversity in South Africa 260
6.6 Afrikaans-Speaking People: The Coloured – White Dichotomy 262
6.6.1 Classification and identification of Coloured and White Afrikaans-speakers. 263
6.6.2 Implications of classification. 265
6.7 Demographic Features 267
6.7.1 Geographical region. 268
6.7.2 Language. 268
6.8 Conclusion 268

CHAPTER 7 270
RESEARCH DESIGN AND METHODOLOGY 270
7.1 Introduction 270
7.2 Primary Objective of the Research 270
7.3 Research Methods and Designs Used in the Study 271
7.3.1 Multiple translation method: Brislin's back-translation method and the committee approach. 272
7.3.2 Item response theory and the Rasch measurement model. 273
7.4 Advantages of Item Response Theory and the Rasch Measurement Model over Classical Test Theory 282
7.4.1 Focus on item-level. 283
7.4.2 Better construct interpretation. 284
7.4.3 Better measurement precision across the continuum of the variable. 284
7.4.4 Test development. 285
7.4.5 Information on category functioning. 286
7.4.6 Scoring methods. 286
7.4.7 Differential item functioning. 287
7.4.8 Administrative efficiency and item banking. 289

7.4.9 Additivity. 290
7.4.10 Superior reliability estimates. 291
7.5 Participants and Sampling Procedures 298
7.5.1 Participants for the translating process. 298
7.5.2 Participants for the English PPD screening process. 300
7.5.3 Participants for the Afrikaans PPD screening process. 301
7.6 Measures 301
7.6.1 Demographic questionnaire. 301
7.6.2 The Postpartum Depression Screening Scale (PDSS). 302
7.6.3 The Edinburgh Postnatal Depression Scale (EPDS). 303
7.6.4 The Quick Inventory for Depressive Symptomatology – Self Report (QIDS-SR16). 304
7.7 Procedure 306
7.7.1 Procedure for the translation of the PDSS. 306
7.7.2 Procedure for the translation of the QIDS-SR. 308
7.7.3 Procedure for the screening process. 308
7.8 Ethical Considerations 311
7.9 Data Analysis 314
7.9.1 Descriptive statistics for the PDSS. 314
7.9.2 Qualitative data analysis. 314
7.9.3 Quantitative data analysis. 315
7.9.3.1 Rasch analysis. 315
7.9.3.2 Multiple regression analysis. 323
7.9.3.3 Correlation of PDSS, EPDS, and QIDS-SR16 total scores. 328

CHAPTER 8 329
RESULTS AND DISCUSSION 329
8.1 Introduction 329
8.2 Descriptive Statistics 330
8.3 Results of Rasch Analysis of the English PDSS 355
8.3.1 Summary of English Rasch analysis: persons and items. 357

8.3.2 Rating scale requirements: English PDSS. 360
8.3.3 Item person construct map: English PDSS. 365
8.3.4 Item fit: English PDSS. 366
8.3.5 Dimensionality: English PDSS. 372
8.3.6 Performance of English PDSS dimensions: Rasch analysis of persons and items. 376
8.3.6.1 Sleeping/Eating Disturbances (SLP) dimension. 376
8.3.6.2 Anxiety/Insecurity (ANX) dimension. 380
8.3.6.3 Emotional Lability (ELB) dimension. 382
8.3.6.4 Mental Confusion (MNT) dimension. 383
8.3.6.5 Loss of Self (LOS) dimension. 385
8.3.6.6 Guilt/Shame (GLT) dimension. 387
8.3.6.7 Suicidal Thoughts (SUI) dimension. 388
8.3.7 Item Fit Statistics for the PDSS Dimensions. 390
8.3.7.1 Sleeping/Eating Disturbances (SLP) dimension. 392
8.3.7.2 Anxiety/Insecurity (ANX) dimension. 393
8.3.7.3 Emotional Lability (ELB) dimension. 395
8.3.7.4 Mental Confusion (MNT) dimension. 396
8.3.7.5 Loss of Self (LOS) dimension. 397
8.3.7.6 Guilt/Shame (GLT) dimension. 399
8.3.7.7 Suicidal Thoughts (SUI) dimension. 400
8.3.8 Response category statistics: Item option and distractor frequencies for the PDSS dimensions. 402
8.4 Results of Rasch Analysis of the Afrikaans PDSS 403
8.4.1 Summary of Afrikaans Rasch analysis: persons and items. 403
8.4.2 Rating scale requirements: Afrikaans PDSS 406
8.4.3 Item person construct map: Afrikaans PDSS. 411
8.4.4 Item fit: Afrikaans PDSS. 415
8.4.5 Dimensionality: Afrikaans PDSS. 417
8.4.6 Performance of Afrikaans PDSS dimensions: Rasch analysis of persons and items. 421

8.4.6.1 Afrikaans Sleeping/Eating Disturbances (SLP) dimension. 421
8.4.6.2 Afrikaans Anxiety/Insecurity (ANX) dimension. 424
8.4.6.3 Afrikaans Emotional Lability (ELB) dimension. 426
8.4.6.4 Afrikaans Mental Confusion (MNT) dimension. 428
8.4.6.5 Afrikaans Loss of Self (LOS) dimension. 429
8.4.6.6 Afrikaans Guilt/Shame (GLT) dimension. 431
8.4.6.7 Afrikaans Suicidal Thoughts (SUI) dimension. 433
8.4.7 Item fit statistics for the Afrikaans PDSS dimensions. 435
8.4.7.1 Afrikaans Sleeping/Eating Disturbances (SLP) dimension. 435
8.4.7.2 Afrikaans Anxiety/Insecurity (ANX) dimension. 437
8.4.7.3 Afrikaans Emotional Lability (ELB) dimension. 438
8.4.7.4 Afrikaans Mental Confusion (MNT) dimension. 439
8.4.7.5 Afrikaans Loss of Self (LOS) dimension. 441
8.4.7.6 Afrikaans Guilt/Shame (GLT) dimension. 442
8.4.7.7 Afrikaans Suicidal Thoughts (SUI) dimension. 444
8.4.8 Response category statistics: Item option and distractor frequencies for the Afrikaans PDSS dimensions. 445
8.5 Items Marked as Difficult to Understand 447
8.6 Invariance and Differential Item Functioning 450
8.7 Results of the Analysis of Risk Factors for PPD 461
8.8 Results of the Comparison of the PDSS, the EPDS, and the QIDS-SR16 476
8.9 Discussion 483
8.9.1 Discussion of Rasch analysis. 483
8.9.2 Discussion of problematic items and items with differential item functioning. 486
8.9.3 Discussion of the risk factors for major PPD in this study. 494
8.9.4 Discussion of the correlation of the PDSS, the EPDS, and the QIDS-SR16. 500

CHAPTER 9 502
CONCLUSION, LIMITATIONS, AND RECOMMENDATIONS FOR FUTURE RESEARCH.......... 502

APPENDIX A 516
APPENDIX B 521
APPENDIX C 529
APPENDIX D 541
APPENDIX E 544
APPENDIX F 548

REFERENCES 573

LIST OF TABLES

Table 1 Signs and symptoms of PPD 64
Table 2 Item Analysis and Internal Consistency Estimates by Standardization Sample for 35-Item PDSS 153
Table 3 Confirmatory Factor Analysis: Maximum-Likelihood Dimensions and Loadings in the Development Sample (N = 525) 157
Table 4 Postpartum Depression Screening Scale: Likert Response Category Fit Statistics 160
Table 5 Sensitivity, Specificity, Positive and Negative Predictive Values of the PDSS, EPDS, and BDI-II 162
Table 6 Comparison of the Item Content of the PDSS' Seven Dimensions with the BDI-II and the EPDS 164
Table 7 Strategies for Identifying and Dealing with Bias in Cross-cultural Assessment 188
Table 8 Severity Thresholds for the QIDS-C16/QIDS-SR16 306
Table 9 Demographic Characteristics Stratified by Questionnaire Language: Home Language, Race/Ethnic Group, Marital Status and Age 333
Table 10 Demographic Characteristics Stratified by Questionnaire Language: Education Level and Employment Status 334
Table 11 Demographic Characteristics Stratified by Questionnaire Language: Number of Weeks Since Birth, Infant's Sex, Gestational Age at Birth, and Feeding Method 336
Table 12 Perceived Level of Support Obtained by Mothers, Stratified by Questionnaire Language 339
Table 13 Obstetric Profile of Mothers Stratified by Questionnaire Language 342
Table 14 Current PPD and Antenatal Depression Assessment and/or Treatment of Mothers, Stratified by Questionnaire Language 343
Table 15 Psychiatric History of Mothers Stratified by Questionnaire Language 345
Table 16 Self Evaluation PPD Statements Chosen by Mothers, Stratified by Questionnaire Language 346

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table 17 Peripartum and Psychological Profile of Mothers Stratified by Questionnaire Language 348

Table 18 Psychosocial Characteristics of Mothers Stratified by Questionnaire Language 350

Table 19 Profile of How Mothers Felt About Their Pregnancies, Stratified by Questionnaire Language 352

Table 20 Infant Temperament and Concerns Regarding Infant, Stratified by Questionnaire Language 354

Table 21a Summary Statistics of 182 Non-Extreme Persons and Items for the English PDSS. 359

Table 21b English PDSS: Summary of 35 Measured (Non-Extreme) PDSS 359

Table 22 Summary Statistics for the 5-Point Likert Response Categories Used for the PDSS 361

Table 23a Item-Person Distribution Map for the English PDSS (N = 187) 367

Table 23b Item Category-Person Distribution Map for the English PDSS (N = 187) .. 368

Table 24 Item Statistics for the English PDSS Total: Misfit Order (N = 187) 371

Table 25 Variance Decomposition of the Observations for the English PDSS Items (N = 187) 373

Table 26 Standardized Residual Loading for the English PDSS (Sorted by Loading) . 375

Table 27 Summary Statistics for the PDSS Dimensions 377

Table 28 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Sleeping/Eating Disturbances (SLP) Dimension (n=187) 393

Table 29 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Anxiety/Insecurity (ANX) Dimension (n=187) 395

Table 30 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Emotional Lability (ELB) Dimension (n=187) 396

Table 31 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Mental Confusion (MNT) Dimension (n=187) 397

Table 32 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Loss of Self (LOS) Dimension (n=187) 398

Table 33 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Guilt/Shame (GLT) Dimension (n=187)................................ 400

Table 34 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Suicidal Thoughts (SUI) Dimension (n=187) 401

Table 35a Summary Statistics of 170 Non-Extreme Persons and Items for the Afrikaans PDSS. 405

Table 35b Afrikaans PDSS: Summary of 35 Measured (Non-Extreme) PDSS 405

Table 36 Summary Statistics for the 5-Point Likert Response Categories Used for the Afrikaans PDSS 407

Table 37a Item Distribution Map for the Afrikaans PDSS (N=178) 412

Table 37b Item Category-Person Distribution Map for Afrikaans PDSS (N = 178)..... 413

Table 38 Item Statistics for the Afrikaans PDSS Total: Misfit Order (N = 178) 416

Table 39 Variance Decomposition of the Observations for the Afrikaans PDSS Items (n = 178) 418

Table 40 Standardized Residual Loading for the Afrikaans PDSS (Sorted by Loading) 420

Table 41 Summary Statistics for the Afrikaans PDSS Dimensions 422

Table 42 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Sleeping/Eating Disturbances (SLP) Dimension (n=178)... 436

Table 43 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Anxiety/Insecurity (ANX) Dimension (n=178) 438

Table 44 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Emotional Lability (ELB) Dimension (n=178) 439

Table 45 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Mental Confusion (MNT) Dimension (n=178) 440

Table 46 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Loss of Self (LOS) Dimension (n=178) 442

Table 47 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Guilt/Shame (GLT) Dimension (n=178) 443

Table 48 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Suicidal Thoughts (SUI) Dimension (n=178) 445

Table 49 Items Marked by Participants as Difficult to Understand after Completing English PDSS or Afrikaans PDSS 449
Table 50 Items that Exhibit Differential Item Functioning in the PDSS Total Item Rasch Analysis 456
Table 51 Items that Exhibit Differential Item Functioning in the PDSS Dimensions ... 457
Table 52 Demographic and Obstetric Variables by PDSS Screening Result (N = 365) 462
Table 53 Model Summary of the Dependent Variable (PDSS score) 465
Table 54 Analysis of Variance of the Dependent Variable (PDSS score).......... 466
Table 55 Multiple Regression Analysis of the Association between Demographic and Obstetric Variables and Scores on the PDSS (N = 365) 468
Table 56 Collinearity Diagnostics of the PDSS Scores 471
Table 57 Casewise Diagnostics of the PDSS Score 473
Table 58 Case Summaries 473
Table 59 Point Biserial Correlations of Psychiatric History and Life Stress Variables with Total PDSS Scores (N = 365) 475
Table 60 Descriptive Statistics for the PDSS, EPDS, and QIDS-SR16 478
Table 61 Cut-off Scores for Screening for the Diagnosis of Major Postpartum Depression for the PDSS, EPDS, and QIDS-SR16 479
Table 62 Cross Tabulation of the Participants According to Cut-off Scores for the PDSS and EPDS 479
Table 63 Cross Tabulation of the Participants According to Cut-off Scores for the PDSS and QIDS-SR16 480
Table 64 Cross Tabulation of the Participants According to Cut-off Scores for the EPDS and QIDS-SR16 481
Table 65 Pearson Correlations between the Total Scores of the PDSS, EPDS, and QIDS-SR16 (N=365) 482
Table 66 Infit and Outfit MNSQ Statistic for Misfit Items in the PDSS and Afrikaans PDSS Dimensions 484
Table 67 Association of sample characteristics with English and Afrikaans samples .. 548
Table 68a Crosstabulation of Support Recived from the Baby's Father and Questionnaire Language 551

Table 68a Chi-Square Statistics from Crosstabulation of Support Recived from the Baby's Father and Questionnaire Language 551
Table 69a Crosstabulation of Support Recived from Family and Questionnaire Language 552
Table 69b Chi-Square Statistics from Crosstabulation of Support Recived from Family and Questionnaire Language 552
Table 70 Summary Statistics of 187 Extreme and Non-Extreme Participants for the English PDSS... 553
Table 71 Summary Statistics of 178 Extreme and Non-Extreme Participants for the Afrikaans PDSS 553
Table 72 Item Option and Distractor Frequencies for English PDSS Sleeping/Eating Disturbances Content Scale: Measure Order (N = 187) 554
Table 73 Item Option and Distractor Frequencies for English PDSS Anxiety/Insecurity Content Scale: Measure Order (N = 187) 555
Table 74 Item Option and Distractor Frequencies for English PDSS Emotional Lability Content Scale: Measure Order (N = 187) 556
Table 75 Item Option and Distractor Frequencies for English PDSS Mental Confusion Content Scale: Measure Order (N = 187) 557
Table 76 Item Option and Distractor Frequencies for English PDSS Loss of Self Content Scale: Measure Order (N = 187)... 558
Table 77 Item Option and Distractor Frequencies for English PDSS Guilt/Shame Content Scale: Measure Order (N = 187)... 559
Table 78 Item Option and Distractor Frequencies for English PDSS Suicidal Thoughts Content Scale: Measure Order (N = 187) 560
Table 79 Item Option and Distractor Frequencies for Afrikaans PDSS Sleeping/Eating Disturbances Content Scale: Measure Order (N = 178) 561
Table 80 Item Option and Distractor Frequencies for Afrikaans PDSS Anxiety/Insecurity Content Scale: Measure Order (N = 178) 562
Table 81 Item Option and Distractor Frequencies for Afrikaans PDSS Emotional Lability Content Scale: Measure Order (N = 178) 563

Table 82 Item Option and Distractor Frequencies for Afrikaans PDSS Mental Confusion Content Scale: Measure Order (N = 178) 564
Table 83 Item Option and Distractor Frequencies for Afrikaans PDSS Loss of Self Content Scale: Measure Order (N = 178) 565
Table 84 Item Option and Distractor Frequencies for Afrikaans PDSS Guilt/Shame Content Scale: Measure Order (N = 178) 566
Table 85 Item Option and Distractor Frequencies for Afrikaans PDSS Suicidal Thoughts Content Scale: Measure Order (N = 178) 567
Table 86 Item Correlations with PDSS Dimensions (N = 365) 568

LIST OF FIGURES

Figure 1 Bio-psycho-socio-cultural model of the processes leading to postpartum disorders. (Halbreich, 2005) 74

Figure 2 Flow chart for examining the sources for differential item functioning. (Adapted from Allalouf & Sireci, 1998, p. 19). 322

Figure 3 Probability curves of observations in each category of the PDSS. 364

Figure 4 Standardized residual contrast of English PDSS items. 374

Figure 5 Probability curves of observations in each category. 410

Figure 6 Standardized residual contrast of Afrikaans PDSS items. 419

Figure 7 Differential Item Functioning of English and Afrikaans PDSS items. 455

Figure 8 Differential item functioning of items in the Sleeping/Eating Disturbances (SLP) dimension. 458

Figure 9. Differential item functioning of items in the Anxiety/Insecurity (ANX) dimension. 458

Figure 10. Differential item functioning of items in the Emotional Lability (ELB) dimension. 459

Figure 11. Differential item functioning of items in the Mental Confusion (MNT) dimension. 459

Figure 12. Differential item functioning of items in the Loss of Self (LOS) dimension. 460

Figure 13. Differential item functioning of items in the Guilt/Shame (GLT) dimension. 460

Figure 14. Differential item functioning of items in the Suicidal Thoughts (SUI) dimension. 461

Figure 15 Histogram showing the distribution of the regression standardized residuals. 570

Figure 16 Normal probability plot showing the distribution of the regression standardized residuals. 571

Figure 17 Scatterplot. 572

CHAPTER 1

INTRODUCTION

1.1 Aim of the Study

Postpartum depression (PPD) is a relatively common perinatal mental illness affecting, on average, approximately 13% of postpartum women. The prevalence of PPD is reportedly significantly higher in certain peri-urban areas of South Africa. Furthermore, it is estimated that up to 50% of mothers affected by this illness go undetected. Screening mothers for symptoms of postpartum depression after the birth of their babies is important for the wellbeing of the mother, her infant, and ultimately her entire family. This study aims to address the problem of the unavailability of an Afrikaans screening measure specifically for postpartum depression. The primary objective of this study is to provide an Afrikaans version of an existing postpartum depression screening measure – the Postpartum Depression Screening Scale (PDSS). Another objective of the study was to ascertain the level of agreement between the PDSS and two other self-report screening measures for depression, specifically whether all three screening measures identified the same subgroup of mothers as having major postpartum depression.

The etiology of perinatal mental illness is complex and likely to arise from the interaction of multiple risk factors: biological, psychological, social, and cultural. The final objective of the study is to determine the magnitude of the relationship between a number of known risk factors for PPD and a positive screen for major postpartum depression.

1.2 Contextualising the Research

The majority of women adapt well to having a new baby and the demands of motherhood. A significant percentage of women are, however, affected by perinatal mental illness. Postpartum depression is one of the more common perinatal mood disorders. The prevalence estimates vary widely and depend on a number of variables, namely, the assessment measure used, the sampling procedure, diagnostic criteria employed, and the location and cultural attributes of the population. In a very poor peri-urban settlement near Cape Town, South Africa, a 34.7% prevalence rate for PPD was reported (Cooper, Tomlinson, Swartz, Woolgar, Murray, & Molteno, 1999). This figure is roughly three times the expected rate internationally. High levels of social adversity were endemic in this South African population and maternal PPD was associated with disturbances in the mother-infant relationship and the absence of support from the woman's partner.

Numerous researchers have examined the risk factors for PPD. Meta-analyses have revealed that PPD develops from the interplay of multiple biopsychosocial and cultural factors (Beck, 1996a, 2001; O'Hara & Swain, 1996). Other researchers point out that biological, obstetric, psychosocial, and personality risk factors are significant (Kruckman & Smith, 2006).

Exposure to extreme societal stressors during the antenatal period, like being in danger of being murdered or witnessing a violent crime, is indicated as one of the strongest predictors of PPD in an urban South African cohort (Ramchandani, Richter, Stein, & Norris, 2009). This study aims to examine which of the known risk factors for PPD were present amongst mothers who screened positive for major PPD.

Research has shown that PPD is likely to have a negative impact on the mother, her infant and her family. In severe instances, it may be potentially devastating culminating in suicide or infanticide.

As a result there has been more focus in recent years on the early recognition of PPD. This is due to findings that early screening and intervention for PPD results in improved outcomes and increases the mother's chance for an earlier recovery (Hanna, Jarman, Savage, & Layton, 2004; Sobey, 2002). This is, however, often challenging as PPD shares certain physical symptoms that are considered normal in the postpartum period, like decreased libido, fatigue, lack of sleep, and appetite changes. PPD is also often experienced covertly making it difficult for health practitioners to identify. In many instances a general practitioner will only casually enquire about a new mother's mental status and is of the opinion that screening takes too much effort (Kumar & Robson, 1984; O'Hara, 1995, Seehusen, Baldwin, Runkle, & Clark, 2005). For these reasons missed diagnoses have been found to be frequent in situations which lack structured methods for evaluating mental health status (Evins, Theofrastous, & Galvin, 2000; Goldsmith, 2007; Reid et al., 1998).

The use of screening scales specifically for postpartum depression in the weeks following childbirth allows for the early detection of mothers who suffer from PPD and referral for appropriate treatment and support. The PDSS is a brief 35-item self-report questionnaire that was developed to help practitioners identify and respond to PPD at an early stage (Beck & Gable, 2000). It was designed to assess the presence, severity and type of PPD by identifying women who are likely to meet the diagnostic criteria for a depressive disorder with postpartum onset, as defined by the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000; Beck & Gable, 2002). The PDSS is composed of seven symptom content scales which were derived from C.T. Beck's qualitative research studies on the subjective experience of PPD (Beck, 1992, 1993, 1996c). The PDSS demonstrates excellent psychometric properties (Beck & Gable, 2000, 2001c). The internal consistency reliability for the content scales and the overall scale reliability were excellent. When screening for major or minor PPD, the PDSS demonstrated the highest combination of sensitivity and specificity compared with two other instruments depression screening scales that have been used to screen for PPD (Beck & Gable, 2001a).

To effectively identify women with PPD from different cultures and language groups, there should be no language barrier in the screening process. In any psychological measure, it is important that the respondents understand the language of the assessment measure. Respondents who are not proficient in the language of the measure may introduce construct irrelevant components to the assessment process (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999).

The need for psychometrically sound instruments available in an appropriate language for the population being assessed has resulted in substantially more adaptations of instruments for use in multiple cultures and languages (Hambleton, 1994). The considerable cost of test development and an increase in globalisation has led to the widespread use of tests in other countries and an increased interest in cross-cultural research (e.g., Van de Vijver, 2002; Van de Vijver & Poortinga, 1997; Van de Vijver & Lonner, 1995).

The International Test Commission (ITC), under the leadership of Ron Hambleton from the United States of America, released Guidelines for Adapting Educational and Psychological Tests (Hambleton, 1994, 2001; International Test Commission, 2010). They address issues pertaining to the construct equivalence in the target language groups, guidelines pertaining to the methodology employed in instrument development and adaptation, guidelines pertaining to the administration process and procedures, as well as guidelines for score interpretations.

A number of other professional bodies have subsequently also provided clear standards and guidelines that need to be adhered to when using psychological tests. These include the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) and the Guidelines for Computer-based Tests and Interpretations (APA, 1986). The ITC later also developed the "International Guidelines on Test Use – Version 2000" (ITC, 2000) and the "International Test Commission Guidelines for Translating and Adapting Tests – Version 2010" (International Test Commission, 2010) to address issues of fairness and bias in test use and set standards for the professional practice of assessment. These guidelines have become the benchmark for cross-cultural test adaptation around the world (Foxcroft, Roodt, & Abrahams, 2006).

Translation is an important component of the adaptation process. The goal of translating an instrument is to obtain another version of the instrument that is conceptually equivalent, with the same connotative meaning, to the original instrument. In cross-cultural research the linguistic translation as well as the cultural translation is important to produce an instrument that is equally valid in different languages and cultures.

In order to determine the cultural appropriateness of an instrument, it is imperative that item bias, differential item functioning, and construct equivalence be examined for the different groups (Foxcroft et al., 2006). Bias and equivalence are essential concepts in cross-cultural assessment. Equivalence (or the lack of bias) is a prerequisite for valid comparisons across cultural populations (Van de Vijver & Tanzer, 1997). The equivalence of an instrument's scores is challenged when bias is present (Van de Vijver, 2002). Attaining equivalence across different cultures and language versions of instruments is perhaps the central issue in cross-cultural comparative research (Van de Vijver, 2001; Van de Vijver & Leung, 1997b).

The translation methodology used is important as it has an impact on the equivalence of the different language versions and the instrument's cross-cultural validity. It must also be ensured that the instruments' testing instructions are translated using the same methodology as the items (Ægisdóttir, Gerstein, & Cinarbas, 2008).

A variety of translation techniques have been developed for adapting, translating, and re-norming psychological instruments for use in other cultures and languages (Ferraro, 2002; Fletcher-Janzen, Strickland, & Reynolds, 2000; Nell, 2000). These include back-translation, one-way translation, forward translation, parallel-blind translation, modified direct translation, pilot-testing, field testing, random probe, a committee approach, and decentering.

Researchers may opt to use one of three different translation procedures, namely, application, adaptation, and assembly (Van de Vijver & Leung, 1997a, 1997b). Application involves the literal translation of an instrument, assuming that the underlying construct is appropriate for each cultural group. Adaptations are appropriate when the translated version is open to changes in items and the addition of items to ensure construct equivalence is maintained and to eliminate bias (Van de Vijver & Tanzer, 1997). Assembly is used when an instrument requires dramatic adaptation from the original due to differential item functioning for the majority of items or aspects of the construct being measured, although salient for some cultures, are not covered by the instrument.

The various guidelines for adapting instruments together with the translation techniques and procedures for assessment across languages and cultures are important to eliminate bias and ensure construct equivalence. Careful consideration must also be given to the cultural applicability of the constructs being measured by an instrument (McGorry, 2000), in this case, the symptoms of PPD.

Childbirth and the transition to motherhood occur in a socio-cultural context, which is experienced and conceptualized in accordance with the mother's values, beliefs, and attitudes. Cultural factors, along with social, psychological, and biological perspectives must therefore be taken into account to fully comprehend PPD and its symptom definition and expression across different cultures (Bina, 2008; Cox, 1999; Leung, 2002). Although many studies across different countries have indicated that PPD is a universal experience, cultural attitudes, beliefs, ways of thinking, and cultural norms for behaviour and emotional responses influence how a mother experiences PPD, how she describes the symptoms, and how she seeks help. In different socio-cultural contexts the manner in which a mother's depression is confronted, discussed, and managed may vary. Some cultures have their own indigenous definitions for PPD along with explanations of its etiology (Bashiri & Spielvogel, 1999). Furthermore, the course of PPD is influenced, either positively or negatively, by cultural beliefs, meanings, and practices (Bina, 2008; Furnham & Malik, 1994; Kleinman, 2004). Cultural factors must therefore be taken into account when screening for PPD, and will be considered in this study.

This study aims to provide an Afrikaans version of the PDSS. The context within which this study was conducted concludes with a history of the Afrikaans-speaking people, the development of the Afrikaans language, and demographic features of the Afrikaans population in South Africa today.

It is believed that this study will make a contribution towards improving the screening of postpartum depression in Afrikaans-speaking South African women. This study will also provide valuable psychometric information for the Afrikaans version of the PDSS and provide information about the risk factors for PPD in a South African sample.

1.3 An Overview of the Research Method

Translation of the PDSS into Afrikaans was performed using a multiple translation technique: Brislin's back-translation method advocated by Brislin (1970) and the committee approach. A combination of the committee approach and the back translation technique has often been used by researchers (Van de Vijver & Leung, 1997b). The back-translation method was selected as it is regarded to be especially useful in cross-cultural research for checking the equivalence of the translations of measures in different languages (Bracken & Barona, 1991; Prieto, 1992). The committee approach has the advantage of a collaborative effort from a group of experts who have an input in the translation process. This improves the quality of the translation, reduces bias, and reduces misconceptions that a single person may bring (Ægisdóttir et al., 2008). This is especially true if the members have complimentary areas of expertise (Van de Vijver & Tanzer, 1997). Utilising a multiple translation method has been recommended to ensure semantic equivalence (Beck, Bernal, & Froman, 2003).

An IRT model, specifically the Rasch rating scale model is employed in this study as implemented by Winsteps (Linacre, 2009). Fundamental assumptions of the Rasch model are that the items assess a single or unidimensional construct and that the difference between person ability and item difficulty should determine the probability of any person being successful on any particular item. Person location (or person logit) and item location (or item logit) are the two parameter estimates within the Rasch model. The Rasch model places person and items on a common logit scale to provide equal-interval measures. This allows for more accurate determination of means, variances, and reliability (Schumacker, 2004; Smith, 2004).

Rasch analysis is performed on the 35-item PDSS and its Afrikaans translation to determine how well the items define the underlying construct of postpartum depression in a South African sample. Rasch analysis is also performed on each of the seven dimensions of the PDSS to determine how adequately the attitude continuum which underlies each dimension is assessed. The overall fit of the data to a one-dimensional model is determined. If the data demonstrates a good fit to the model then the responses from individuals should correspond well with the responses that are predicted by the model.

The assessment of unidimensionality is an important determinant of the scale's internal construct validity (Hong & Wong, 2005). Unidimensionality is ascertained by a Rasch principal components factorial analysis of the residuals as well as by analysis of fit statistics or indices (mean-square infit and mean-square outfit). Individual item-fit indices and Pearson item-total correlations (r_{it}) are examined as they indicate the degree to which the individual items define a unidimensional construct. Rasch analysis provides information on reliability estimates for the PDSS and Afrikaans PDSS in a South African sample. Invariance is determined through analysis of Differential Item Functioning (DIF). The category functioning of the PDSS and its Afrikaans translation is determined through Rasch analysis.

Convergent validity, an important aspect of construct validity, is examined to ascertain whether the PDSS correlates positively with other self-report screening scales for depression, namely the Edinburgh Postnatal Depression Scale (EPDS; Cox, Holden, & Sagovsky, 1987) and the 16 item Quick Inventory of Depressive Symptomatology (QIDS; Rush et al., 2003). The EPDS was developed to screen specifically for postpartum depression and is the most widely used screening questionnaire for PPD. The QIDS was designed to measure the severity of depressive symptoms, including all the criterion symptom domains required for the diagnosis of a major depressive episode as designated by the American Psychiatry Association Diagnostic and Statistical Manual of Mental Disorders - 4th edition (DSM-IV; APA, 1994). The QIDS has been used to examine the differences in the clinical features between postpartum and non-postpartum women and was considered a useful measure in the assessment of PPD (Bernstein et al., 2008).

The relationship between known risk factors for PPD and high scores on the PDSS amongst women in South African is determined through multiple regression analysis. Pearson's product-moment correlation is used to measure the associations among the scores on the PDSS, the EPDS and the QIDS.

1.4 Orientation

This study is presented in nine chapters. Chapter one has covered the aim of this study, the contextualisation of the research, given an overview of the research methodology employed, and an outline of the orientation.

Chapter 2 provides an overview of the types of perinatal mood disorders with a focus on the symptoms, prevalence and clinical course of PPD. Furthermore, the perspectives on the etiology of PPD as well as the risk factors and consequences of PPD are outlined.

Chapter 3 describes the importance of screening postpartum women for PPD in light of the fact that PPD may have serious implications for the mother, her infant, and the entire family. This chapter provides a review of screening measures available that assist in assessing, identifying and treating mothers who present with PPD. The PDSS is discussed in more detail with respect to its conceptual basis, psychometric properties, and comparative analysis of the performance of the PDSS with two other depression instruments.

Chapter 4 addresses cross-cultural assessment with a brief discussion of multicultural assessment in South Africa. Factors that influence cross-cultural assessment and methodological considerations for cross-cultural assessment are outlined. A review of the ethical guidelines for adapting cross-cultural assessment follows. Finally the different techniques and procedures that researchers use to prepare target language versions of existing instruments are described.

Chapter 5 explores some cultural approaches to the understanding of childbirth and related mental disorders as well as how these impact on adapting a postpartum depression screening measure cross-culturally.

Chapter 6 focuses on the history of the Afrikaans-speaking people, the development of the Afrikaans language, and demographic features of the Afrikaans population in South Africa today.

Chapter 7 presents the primary objectives of the study and the methods employed in conducting the study and analysing the results. The main features of the Rasch model are presented along with the advantages of using item response theory (IRT) techniques, like Rasch analysis, as opposed to classical test theory (CTT) based statistical models.

Chapter 8 begins with an overview of the descriptive statistics. The results of the Rasch analysis for the PDSS and the Afrikaans PDSS are presented and discussed next. This is followed by a presentation of the risk factors for PPD in this sample. Finally the comparative analyses of the PDSS, the QIDS-SR and the EPDS are presented.

Chapter 9 considers the major insights gained in the study as well as limitations and recommendations for future research.

CHAPTER 2

PERINATAL MOOD DISORDERS AND POSTPARTUM DEPRESSION

2.1 Chapter Preview

This chapter provides an overview of the types of perinatal mood disorders, focussing on postpartum depression (PPD) – which has been reported to affect 10 to 20% of postpartum women (Dalton & Holton, 2001; O'Hara & Swain, 1996; Wilkinson, 2001). Lee, Yip, Chiu and Chung (2000) maintain that PPD is the most common disorder following childbirth. The symptoms, prevalence and clinical course of PPD are presented and perspectives on the etiology of PPD as well as the risk factors and consequences of PPD are outlined.

2.2 Introduction

The transition to motherhood can be a stressful time in the lives of women and often brings with it a number of major life changes. Apart from the physiological changes, the birth of a baby also has an emotional impact on the mother. Many women adapt well to these changes, but a significant percentage of mothers are affected by the development of a psychological disorder, particularly depression, during the postpartum period (Dalton, 1996).

2.3 Perinatal Mood Disorders

Perinatal mood disorders affect a large number of women both during and after pregnancy and are potentially devastating. Perinatal mood disorders are often broadly termed maternal depression. The term "postpartum depression" has also been used as a catchall phrase for many postpartum emotional symptoms (Beck, 1998a). This may lead to misdiagnosis of PPD when in actuality the mother is suffering from another perinatal mood disorder.

Researchers and clinicians have struggled to come to a consensus regarding the definition, onset and course of perinatal mood disorders (Boyd, Pearson, & Blehar, 2002). It has been debated whether they should fall within the category of mood disorders in the Diagnostic and Statistical Manual, Fourth Edition, Text Revision (DSM-IV-TR; APA, 2000) or whether they represent a constellation of symptoms unique to the antenatal and postpartum period (Brockington, Macdonald, & Wainscott, 2006; Beck & Indman, 2005; Halbreich, 2005). A recent study by Marrs, Durette, Ferraro and Cross (2009) to determine the underlying factor structure of a broad range of emotional experiences and psychiatric symptoms which occur within 30 days after childbirth, indicate that postpartum symptoms are more diverse than has currently been recognized, and may not fit within the current classification system.

A review of literature seems to indicate, however, that perinatal mood disorders may be classified into groups which are considered to be interrelated but potentially separate conditions. These disorders may be classified into 7 separate conditions (Beck, 1998a; Dalton, 1996; Kumar, 1994; Perinatal Mood Disorders, 2004; Roan, 1997;

Bennett & Indman, 2003; Spinelli & Endicott, 2003). Each of these 7 conditions has its own symptomatology and ability to disrupt the family unit affected by it. Although they may share some symptoms, they are considered different conditions, each requiring different treatment (Kirschenbaum, 1995).

The perinatal mood disorders that are classified separately in this section do not appear as separate diagnosable conditions in the DSM-IV-TR. The manner in which the DSM-IV-TR recognizes some of the postpartum mental illness is discussed in each section. Anger in the postpartum period is also included in this section, but postpartum anger is not a syndrome or pathological condition, and as such is not listed as a separate condition of perinatal disorders.

2.3.1 Antenatal mood and anxiety disorders.

Antenatal mood and anxiety disorders can occur any time during pregnancy. Some symptoms of antenatal mood disorders may be overlooked by clinicians because they are also common symptoms which women may experience as a result of the pregnancy. A woman may report frequent changes in energy, appetite and sleep (neurovegetative symptoms), for example, and these may not be considered unusual symptoms during the first and third trimesters (Bennett & Indman, 2003; Sugawara, Sakamoto, Kitamura, Toda, & Shima, 1999). These neurovegetative symptoms are very common in pregnancy. This is frequently a reason why perinatal mood disorders may go undetected – some clinicians don't pay much attention to pregnant patients' complaints about neurovegetative symptoms (Hoffman & Hatch, 2000; Kendler, Walters, & Kessler, 1997) and postpartum women easily dismiss these symptoms as just part of the pregnancy (Bennett & Indman, 2003).

Even though the symptoms of depression and normal pregnancy do overlap and pose a dilemma to clinicians, labelling antenatal women's alterations in appetite, sleep and energy as "normal" may be problematic. Antenatal women with depression are more likely to complain of fatigue and sleep deprivation than antenatal women without depression (Kelly, Zatzick, & Anders, 2001). Depressed patients may be more willing to mention behavioural changes to their health-care provider than to disclose their depressed mood. In addition, depressed patients often attribute their symptoms to being tired, overworked, or having a cold and fail to recognize that they're mentally ill (Smith, Brunetto, & Yonkers, 2004). Cognitive as well as behavioural and somatic symptoms should therefore be explored.

Depressive disorder during pregnancy varies in length and time of onset. Rates of depression are higher for pregnant women with inadequate social support, chronic stressors such as marital dysfunction, a personal history or family history of mood disorder, a history of child abuse, and financial and housing problems. Gonadal hormones have also been blamed for provoking uncertain mood effects (Wisner & Stowe, 1997). Demographic variables such as young age, minimal education, poverty and a large number of children have been found to put women at a greater risk for antenatal depression (Barnett, Joffe, Duggan, Wilson, & Repke, 1996; Evans, Heron, Francomb, Oke, & Golding, 2001).

According to Bennett and Indman (2003) approximately 15-20 percent of all antenatal women experience depression. Around 15 percent of these women are so severely depressed that they attempt suicide. It is therefore essential that health care providers adequately explore all symptoms that antenatal women report so that intervention may be given for the percentage of women whose symptoms are not merely pregnancy related. According to the American College of Obstetricians and Gynaecologists (ACOG, 2002) and Sichel and Driscoll (1999) antenatal depression necessitates careful monitoring to ensure a healthy outcome for both the mother and foetus. Timely and appropriate treatment is imperative to avoid depression-associated appetite and weight loss. Women suffering from antenatal depression are more vulnerable to nicotine, drug, and alcohol abuse and failure to obtain adequate antenatal care – factors that compromise foetal development (Zuckerman, Bauchner, Parker, & Cabral, 1990).

In studies by ACOG (2002) and Sichel and Driscoll (1999) it is reported that depression which is not monitored or which is left in pregnant women may lead to premature labour and delivery. Chung, Lau, Alexander, Chiu, & Lee (2001) also report that antenatal depression and stress is associated with lower gestational age, and also found it to be associated with lower birth weight, delivery by caesarean section, and admittance of infants to a neonatal care unit. Furthermore, Matthey, Barnett, Ungerer, & Waters (2000) have found that antenatal depression and PPD are linked by as much as 75%. This stresses the importance of identifying women with antenatal depression for intervention and treatment.

Some antidepressant medications have demonstrated relative safety during pregnancy, but warn that absolute safety cannot be ensured as the infant's developing brain is vulnerable to adverse events (Wisner, Gelenberg, Leonard, Zarin, & Frank, 1999). A study by Oren et al. (2002) suggests that light therapy is beneficial for the treatment of antenatal depression. Spinelli and Endicott (2003) are but some researchers who found interpersonal psychotherapy to be an effective method of treatment during pregnancy and recommended it as a first-line treatment in the hierarchy of antenatal depression. They found that it showed significant improvement compared to a parenting education program. Wisner et al. (2000) compiled a set of guidelines for choosing appropriate interventions. Their risk-benefit analysis regarding decision making for the treatment of childbearing women has been regarded as the most appropriate method for intervention.

2.3.2 Postpartum blues.

The first type of postpartum disturbance is termed postpartum blues, which commonly occurs around day 3 to 5 postpartum in 50% to 80% of mothers (Henshaw, 2003; Postpartum depression consensus statement, 2002). The DSM-IV-TR uses the term “baby blues” and states that it can affect up to 70% of women in the first 10 days postpartum, that the symptoms are transient and do not impair functioning (APA, 2000). Bennett and Indman (2003) point out that the “Baby Blues” is commonly experienced by a majority of mothers and should not be considered a disorder. They suggest that it may be more accurate to consider the blues as a normal experience following childbirth rather than a disorder or psychiatric illness.

It is assumed that postpartum blues are a result of fluctuating hormone levels that result from the expulsion of the placenta in the third stage of labour (Halbreich, 2005). Bennett and Indman (2003) add a number of factors that contribute to the etiology of postpartum blues which include the physical and emotional stress of birth, physical discomfort after birth, the emotional letdown experienced after pregnancy and the birth, an awareness and anxiety about the increase in responsibility that having a baby brings, fatigue and lack of sleep, as well as disappointments around the birth, spousal support, breastfeeding and the baby.

Women with postpartum blues commonly report mood swings, anxiety or irritability, feeling tearful, sadness, lack of concentration and feelings of dependency. Typically, these symptoms reach a peak on the fourth or fifth day after delivery and may last from a few hours to a few days. They usually disappear spontaneously within two weeks after the delivery. These symptoms do not interfere with a women’s ability to function, but they may be unpredictable and often unsettling. No specific treatment is usually required. Women who experience this form of depression seldom pose any significant physical threat to themselves or to their babies (Postpartum depression consensus statement, 2002). However, in some women, particularly women with a history of depression, postpartum blues may herald the development of a more significant mood disorder. An evaluation to rule out a more serious mood disorder is warranted if the symptoms of postpartum blues last for a period longer than two weeks.

2.3.3 Postpartum depression.

The second type of postpartum disturbance is termed postpartum depression (PPD) which may range from moderate to severe. This can occur as a gradual onset from postpartum blues, it may start when breastfeeding is discontinued, or it can manifest itself at any point in the first year after childbirth or up to the return of normal menstruation (Dalton & Holton, 2001). The majority of studies indicate that most cases of PPD occur within the first 3 months postpartum. Up to 20% of mothers develop PPD, although O'Hara and Swain (1996) reported that the average rate of this mood disorder based on findings of 59 studies was 13%. PPD will be discussed in further detail in the next section.

2.3.4 Obsessive-compulsive disorder occurring in the postpartum period.

Obsessive-compulsive disorder (OCD) is an anxiety disorder characterized by "(a) recurrent, unwelcome thoughts, ideas, or doubts that seem senseless, yet give rise to anxiety or distress (obsessions), and (b) urges to perform excessive behavioural or mental acts (compulsive rituals) to suppress or neutralize the obsessional distress" (Abramowitz, Schwartz, Moore and Luenzmann, 2003, p. 462) . An adult would recognise that the obsessions or compulsions are excessive or unreasonable, yet would try in most cases to avoid situations related to obsessional fears. The obsessions or compulsions cause marked distress and interfere with the person's normal functioning (APA, 2000).

Limited research exists on OCD during pregnancy and the puerperium. Abramowitz et al. (2003) reviewed literature on OCD in pregnancy and suggest that obsessional phenomena in postpartum women may occur at higher than expected rates.

Bennett and Indman (2003) indicate that symptoms usually begin at about two to six weeks after birth. They describe the postpartum obsessive thoughts as intrusive, repetitive, and persistent thoughts or mental pictures which usually centre on harming or killing the baby. The thoughts are conscious, usually intensify and frighten the mother to the extent that she may start to avoid being alone with her baby for fear that she may lose control and act out the obsessive ruminations. The mother is usually horrified and disgusted by the thoughts. The obsessions may be accompanied by behaviours to reduce the anxiety she experiences (for example, hiding dangerous items like knives). The mother may also experience compulsive behaviour like counting, checking or cleaning. Abramowitz et al. (2003) found a significantly consistent pattern regarding the content of obsessions and compulsions in perinatal women with OCD symptoms. Antenatal women report being obsessed by fear of contamination, followed by cleaning and washing rituals. Postpartum women tended to report experiencing intrusive unwanted obsessional thoughts of harming their babies, accompanied by phobic avoidance of fear cues. Symptoms of perfectionism, hoarding, and symmetry or ordering which often is present in OCD, were not prominent. Typically, the onset of OCD is gradual yet Abramowitz et al. (2003) found that in postpartum OCD, clinical reports emphasized a rapid onset of obsessive symptoms. They further state that there is evidence of a relationship between PPD and OCD symptoms, particularly unwanted intrusive thoughts of hurting the newborn. Furthermore, postpartum obsessive thoughts (regardless of how horrendous the content) were not associated with an increased risk of harming the infant. They state that this is due to the obsessive thoughts being experienced as unwanted, senseless and ego-dystonic.

Abramowitz et al. (2003) provide a distinction between the symptoms of postpartum psychosis and postpartum OCD given that either of these disorders may give rise to thoughts or ideas of harming the infant. The postpartum OCD patient differs from the postpartum psychotic patient in that she fears participating in unacceptable behaviour, and also fears merely thinking about it (unlike the delusional thinking typically found in a postpartum psychotic patient). Furthermore, excessive avoidance behaviour and rituals can be seen in postpartum OCD patients as they attempt to control their thoughts and to ensure that they refrain from committing the frightful acts featured in their obsessive thoughts. Severe anxiety complaints are typical in mothers with postpartum OCD. The anxiety may, for example, have to do with concern over whether they will harm their infant or not. In postpartum psychosis, general psychotic symptoms are more prominent, such as losing touch with reality and unpredictable, aggressive behaviour.

According to Bennett and Indman (2003), three to five percent of new mothers develop obsessive symptoms. Women at risk of OCD in the postpartum period may have a personal or family history of OCD. This is a condition that seems to recur and women at-risk should therefore be monitored closely and be given prompt treatment after a subsequent pregnancy (ACOG, 2002; Sichel & Driscoll, 1999).

2.3.5 Postpartum onset of panic disorder.

Metz, Sichel, and Goff (1988) reported the initial onset of panic disorder during the postpartum period. They recommended that clinicians differentiate between postpartum panic disorder and PPD. According to Roan (1997), panic disorder is common among women of childbearing age and is twice as common among women as among men. Prevalence rates vary, with figures ranging from 0.5% to 1.5% at 6 weeks postpartum (Matthey, Barnett, Howie, & Kavanagh, 2003). The emergence of the disorder for the first time in the postpartum period could be coincidental, but is likely to be triggered by the birth. Stressful life events can precipitate panic attacks, and childbirth, although considered a positive event by most people, is stressful.

The symptoms of panic disorder include fear, episodes of extreme anxiety, and a number of physical sensations like shortness of breath, a sense of being smothered or choking sensations, chest pain, palpitations, hot or cold flushes, dizziness, trembling, and tingling sensations or numbness. The mother may be restless, agitated or irritated. During an attack the mother may fear she is losing control, going crazy, or even dying. The panic attack can be so intense that it may wake her up. Typically the attack has no identifiable trigger. It is often accompanied by excessive worry or fears, including fear of having another panic attack (ACOG, 2002; APA, 1994; Bennett & Indman, 2003).

Beck (1998a) found that mothers experienced considerable impairment in their quality of life due to the panic attacks and accompanying fear and anxiety, to the extent that fulfilling maternal responsibilities became a struggle. Her phenomenological study of panic disorder in postpartum mothers indicated that recurring panic attacks led to impairment in quality of life, feelings of disappointment and guilt, a decrease in self-esteem, feeling exhausted, and concern about the residual effects it would have on their children.

A woman who has a personal or family history of anxiety or panic disorder may trigger its onset in the stressful postpartum period. Thyroid dysfunction has also been described as a risk factor (Bennett & Indman, 2003), and Roan (1997) reports that the female hormone progesterone, which is approximately 170 times higher than before pregnancy, may trigger the onset of panic disorder. Panic attacks have also been found to be precipitated by certain times of day such as sunset, on awakening, feeding time, by being confined indoors, being alone, being away from the infant, the infant crying or by multiple demands on the mother's time (Beck, 1998a; Matthey et al., 2003). Beck's phenomenological study revealed six themes that describe the experiences of panic during the postpartum period (Beck, 1998a, p. 133-134):

Theme 1. The terrifying physical and emotional components of panic paralyzed the women, leaving them feeling totally out of control;

Theme 2. During panic attacks, women's cognitive functioning abruptly diminished while between these attacks women experienced a more insidious decrease in their cognitive functioning;

Theme 3. During the panic attacks, women feverishly struggled to maintain their composure, leading to exhaustion;

Theme 4. Because of the terrifying nature of panic, preventing further panic attacks was paramount in the lives of the women;

Theme 5. Due to recurring panic attacks, negative changes in women's lifestyles ensued, lowering their self-esteem and leaving them to bear the burden of disappointing not only themselves, but also their families;

Theme 6. Mothers were haunted by the prospect that their panic could have residual effects on themselves and their families.

There is a potential adverse effect on foetal well-being when stress hormones are released into the blood stream (Diego et al., 2004). Stress hormones can cause contraction of the blood vessels to the placenta which may induce abruptio placentae. The Postpartum depression consensus statement (2002) emphasises that early identification and treatment of anxiety may prevent pregnancy complications. Women with a history of anxiety or panic attacks prior to pregnancy warrant medical investigation to prevent maternal and foetal problems during the pregnancy.

Beck (1998a) discusses a number of specific interventions for nursing practice that can be formulated based on each theme to facilitate the correct treatment of mothers experiencing panic attacks in the postpartum period. For example, an intervention for Theme 2 is reassuring mothers that it is not unusual to fear insanity and feel a sense of impending doom during a panic attack. The fears are transient and disappear as soon as the panic attack is over. Beck (1998a) further advocates a multidisciplinary treatment plan specifically for postpartum onset of panic disorder that promotes healthy development of the woman’s maternal role and family integration.

2.3.6 Postpartum posttraumatic stress disorder.

This disorder is characterised by symptoms of re-experiencing a trauma, avoidance of stimuli that are associated with and remind the person of the trauma, numbing of general responsiveness and increased arousal. The DSM-IV (APA, 1994) describes posttraumatic stress disorder (PTSD) as a response of “intense fear, helplessness or horror” (p. 424) to an extreme traumatic stressor that the person experienced, witnessed, or was confronted with. The extreme stressor may be an “event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others” (p. 427).

Childbirth experiences with associated high levels of fear and increased risk of injury and mortality may well include some of the specified features described above. In a study by Arizmendi and Affonso (1987) it was found that the experience of labour continues to impact after the birth. Schreiber and Galai-Gat (1993) found that the experience of intense pain itself may act as a traumatic event.

A number of studies have identified women who experience posttraumatic stress symptoms following labour and childbirth. These studies found the posttraumatic stress symptoms to be associated with long or complicated labour and feelings of lack of control over the situation (Ballard, Stanley, & Brockington, 1995; Fones, 1996; Ichida, 1996; Moleman, Van der Hart, & Van der Kolk, 1992). Another potential aetiological factor is a previous experience of an extreme traumatic event, in particular, sexual abuse (Watson, Juba, Manifold, Kucala, & Anderson, 1991). Watson et al. (1991) also identified contributing factors which include levels of control, attitude of the doctor, degree to which patients' views were listened to, the level of information given during the procedure and if consent was perceived to have been given.

Czarnocka and Slade (2000) and Soderquist, Wijma, and Wijma, (2006) researched the potential predictors and prevalence of posttraumatic stress type symptoms following labour. In both studies three percent (n = 264 and n = 1224) presented with symptoms that suggested clinically significant levels on all three posttraumatic stress dimensions of intrusions, avoidance and hyper arousal. A further 24% of mothers presented with symptoms on at least one of these dimensions (Czarnocka & Slade, 2000).

Soderquist et al. (2006) assessed posttraumatic stress in early and late pregnancy, and up to 11 months postpartum. They report that during the pregnancy, pre-traumatic stress, severe fear of childbirth, depression, previous counselling related to the pregnancy or childbirth, as well as self-reported prior psychological problems were associated with an increased risk of having posttraumatic stress within the first 11 months postpartum. A decrease in perceived social support was also reported in postpartum women who had posttraumatic stress.

According to Czarnocka and Slade (2000), the potential predictors of posttraumatic stress type symptoms following labour are a) the partner not being present at the birth; b) perceptions of low levels of support from the attending partner or relative or staff member; c) self-blame and particularly blaming staff for difficulties experienced during the labour and delivery; d) fear and amount of distress experienced; and e) perceptions of low control during labour and delivery. Furthermore, they found that a history of mental health problems and trait anxiety were significant predictors for depression and anxiety and were also related to posttraumatic stress symptoms.

2.3.7 Puerperal psychosis.

Puerperal psychosis is the final and most extreme form of perinatal mental illness and is regarded a medical emergency. The DSM-IV-TR terms this condition "postpartum-onset mood episode with psychotic features" and reports it to be more common in primiparous women (APA, 2000).

Puerperal psychosis is typically characterised by severe behavioural changes and psychotic episodes. In many cases puerperal psychosis signifies an episode or a variant of bipolar disorder triggered by childbirth (Brockington et al., 1981; Jones & Craddock, 2001; Kendell, Chalmers, & Platz, 1987; Perinatal Mood Disorders, 2004). Puerperal psychosis may present with mostly depressive symptoms, yet differs from PPD due to the presence of hallucinations, delusions, perplexity, confusion, and the psychotic symptoms that appear after the delivery tend to resemble those of a manic or mixed episode (Brockington, 2004). Hypomanic symptoms are particularly characteristic in mothers who develop puerperal psychosis in the initial days after childbirth with symptoms like irritability, restlessness, and insomnia (Heron, McGuinness, Blackmore, Craddock, & Jones, 2008). Mothers with this disorder show signs of disorientation or confusion, rapid mood variations from depressed to elated, and disorganized or erratic behaviour.

Delusions are common and often centre on the infant and may include religious themes (Heron et al., 2008). The mother may also experience auditory hallucinations that instruct her to harm herself or her infant. For these reasons, there is often a suicidal risk as well as a risk that the mother may harm her newborn (Bennett & Indman, 2003; Craig, 2004; Spinelli, 2004;). This disorder has a 5 percent suicide and a 4 percent infanticide rate (Bennett & Indman, 2003). King, Slaytor, and Sullivan (2004) suggest that figures could be much higher if risk events and near misses were taken into account.

The DSM-IV-TR states infanticide is most often associated with postpartum onset mood episode with psychotic features "that are characterized by command hallucinations to kill the infant, or delusions that the infant is possessed" (APA, 2000, p. 422). These psychotic features have, however, also been known to occur in severe postpartum mood episodes that do not have such specific hallucinations or delusions. Researchers agree that infanticide usually occurs when a woman is psychotic (Spinelli, 2004) or involved in the act of committing suicide, to avoid abandoning her children (Jennings, Ross, Popper, & Elmore, 1999; Spinelli, 2005).

This form of psychological disturbance probably has the highest detection rate owing to the severe nature of its manifestation, which typically occurs whilst the mother is still undergoing a period of hospitalisation. Puerperal psychosis (or postpartum mood episode with psychotic feature, as it is referred to in the DSM-IV-TR) is comparatively rare at around 1 to 2 in 1000 (0.1% – 0.2%) women afflicted with this condition (APA, 2000; Perinatal Mood Disorders, 2004; Munk-Olsen, Laursen, Pedersen, Mors, & Mortensen, 2006).

The onset is usually sudden and within a few days postpartum. It has been reported that the first 48 hours postpartum are symptom-free (Brockington and Hamilton as cited in Doucet, Dennis, Letourneau, & Blackmore, 2009, p.270), however, more recent research revealed that about one half of mothers present with mild hypomanic symptoms within the first three days after childbirth (Heron et al., 2008). After the initial mood symptoms, puerperal psychosis progresses rapidly (Heron et al., 2008). Onset typically occurs within the first three months postpartum and 80% of all incidences present within 3-14 days postpartum (Kumar, 1994; Kruckman & Smith, 2006).

Risk factors for puerperal psychosis include a personal or family history of psychosis, alcoholism, depression, premenstrual symptoms, stressful life events, bipolar disorder, or schizophrenia, and a previous postpartum psychotic or bipolar episode - especially bipolar I disorder according to the DSM-IV-TR (APA, 2000; Bennett & Indman, 2003; Heron et al., 2008). Women who have had a postpartum episode with psychotic features have a greater risk of recurrence with each subsequent delivery. The risk of recurrence is reportedly between 30% and 50% (APA, 2000). The DSM-IV-TR also reports that among women without a history of mood disorders, there is still evidence of an increased risk of postpartum psychotic mood episodes if they have a family history of bipolar disorder (APA, 2000).

Higher rates of postpartum mania, delirium, and psychosis were reportedly associated with postpartum thyroiditis (PPT; e.g. Bokhari, Bhatara1, Bandettini, & McMillin, 1998). PPT is the postpartum occurrence of transient hypothyroidism or transient hyperthyroidism. The majority of women return to the euthyroid state by 1 year postpartum. PPT occurs, on average, in 7.5% of women (Stagnaro-Green, 2004).

Spinelli (2009) recommends prompt treatment after delivery to prevent psychosis in women with mood swing disorders. The management of puerperal psychosis should include a physical examination, a clinical evaluation with complete blood chemistry, thyroid functioning tests, and calcium, vitamin B12 and folate levels (Sit, Rothschild, & Wisner, 2006). The treatment of puerperal psychosis is dependent on the outcome of the evaluations and the symptom profile but usually requires hospitalization (Sharma, 2003). Acute treatment may include mood stabilizing medication, antipsychotics, benzodiazepines, and aggressive treatment of insomnia. Electroconvulsive therapy may be a treatment to consider if the illness is unresponsive to conventional therapy (Sharma, 2003). Furthermore, pending the outcome of the clinical evaluation, the nueuroendocrine role in the pathophysiology of puerperal psychosis may warrant hormone replacement therapy if indicated (Spinelli, 2009).

2.3.8 Anger in the postpartum period.

Graham, Lobel and DeLuca (2002) explored state anger as a likely emotional response in the postpartum period and found that a considerable number of women reported angry feelings at approximately six weeks postpartum. Thirty-five percent of their sample reported moderate to high levels of anger. They also determined that anger and depressed mood were associated but relatively independent. Over 80% of women in their study who reported high levels of anger reported low levels of depressed mood. Their findings suggest that there is a group of women, separated from those who experience PPD, who experience anger after delivery. Their research does not suggest that postpartum anger is a syndrome or pathological condition, and as such was not listed as a separate condition of perinatal disorders above. They also do not state that postpartum anger is qualitatively different from that which occurs at other times in a women's life. Rather, they advocate that the longstanding focus on PPD is too narrow and further investigations should look more closely at a more comprehensive range of postpartum emotional experiences.

2.4 Postpartum Depression

The term "postpartum depression" (PPD) is widely used but varies considerably in its definition as the diagnosis of PPD is often erroneously used as a general term to incorporate many of the other postpartum mood disorders mentioned earlier. The phenomenon of what is now termed "postpartum depression" has been the subject of some debate in the past century.

2.4.1 Historical perspectives.

Historically, the connection between psychiatric illness and childbirth has been well-documented. Hippocrates described the emotional problems and psychotic behaviour of postpartum women as a severe case of insomnia and restlessness that began on the sixth day in a woman who bore twins. This condition was referred to as "peurperal fever", and it was theorised that suppressed lochial discharge was transported to the brain where it produced symptoms of "agitation, delirium and attacks of mania" (Thurtle, 1995). An 11th century gyneacologist, Trotula of Salerno, speculated that when a woman's womb was too moist, then her brain was filled with water which would spill over her eyes and cause her to shed tears involuntarily (Steiner, 1990). The writings of Galen, Celsus and others also documented the problems and behaviour of postpartum women. Greater systematic efforts were made in the mid-19th century to describe and classify postpartum mental illness when Esquirol wrote about how nursing women and those recently confined suffered from mental alienation (Steiner, 1990).

The first thorough scrutiny of postpartum disorders took place in 1858 when a French physician, Louis Victor Marcé, published a definitive study, Traits de la Folie des Femmes Enceintes (Insanity in Pregnant and Lactating Women). This study linked negative emotional reactions with childbirth and the development of postpartum psychiatric illness. Marcé noted melancholy, anaemia, weight loss, constipation, and menstrual abnormalities. He also described the presence of confusion, faulty memory, and fogginess which are now recognized as hallmark symptoms in postpartum illness (Roan, 1997; Steiner, 1990; Stern & Kruckman, 1983).

During the first half of the 20th century relatively few studies of maternal mental health were done and there was much disagreement about postpartum psychiatric illness. This changed, however, during the latter half of the 20th century when research among the interrelated, albeit diverse, disciplines of biology, psychology, sociology, and anthropology increased. Many of these have focused on the etiology and treatment of PPD.

The APA published the first edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM) in 1952. This was done in an effort to give all psychiatric ailments names and definitions that would be agreed upon worldwide. The leaders of this reform did not, however, include a category for PPD. Women who were afflicted with this condition, it seemed, were suffering from “dementia praecox”, “neurotic states” or “toxic confusion”, or they were “manic depressive”. Numerous early psychiatrists, including Kraepelin, therefore concluded that “postpartum psychosis” did not exist as a separate syndrome (Brockington, Schofield, Donnelly, & Hyde, 1978; Roan 1997). According to Jacobs, puerperal psychosis as a clinical entity did not exist because every reaction type may have occurred during the puerperium (as cited in Stern & Kruckman, 1983, p. 1030). Foundeur, Fixsen, Triebel, and White were even more insistent, stating that “the results would not appear to justify terming the postpartum illness as a separate illness any more than one might term those young patients who react unfavourably to college as sufferers from a ‘college psychosis’ ” (as cited in Stern & Kruckman, 1983, p. 1030).

The Diagnostic and Statistical Manual, Second Edition (DSM-II, APA, 1968) had described a separate entity: “294.4 Psychosis with Childbirth”. The DSM Third Edition (DSM-III, 1980) eliminated this category, however, arguing that there was no compelling evidence that postpartum psychosis could be classified as a distinct entity (APA, 1980). Furthermore, many physicians believed that if the name of an illness was removed, then it meant that the illness did not exist (Roan, 1997). With the connection between childbirth and psychiatric illness ignored in the DSM, research in the field of postpartum psychiatric illness diminished until the 1980’s.

Dr. James Hamilton was dedicated to bringing attention to the postpartum psychiatric illness. In 1962 he wrote the book entitled “Postpartum Psychiatric Problems” and thirty years later, in 1992, he co-edited “Postpartum Psychiatric Illness: A Picture Puzzle” (Hamilton & Harberger, 1992).

In 1980, Dr Ian Brockington of Great Britain held an international meeting on postpartum psychiatric illness. Dr James Hamilton and other physicians who attended this pioneering meeting founded the Marcé Society, named after Louis Marcé. This scientific organization comprised a group of professionals dedicated to advancing the understanding and treatment of postpartum psychiatric illness (Roan, 1997). The Marcé Society has held biennial international conferences on PPD and related disorders since 1984. The formation of the Marcé Society at this time did not however offset the confusion that resulted from the omission of the link between childbirth and psychiatric illness in the DSM-III.

This uncertainty regarding PPD as a clinical entity continued. According to Roan (1997), the revision of the DSM-III published in 1987, heralded for its improvements over past editions due to its more detailed definitions and information, only briefly mentioned postpartum illness and practically dismissed it for its complexity.

According to Walther (1997), the DSM-IV does not have a useful category for psychiatric disorders of the puerperium period. This edition published in 1994, was preceded by an intense discussion on the topic of postpartum psychiatric illness. Although the DSM-IV excludes PPD, psychosis, anxiety, or any of the other observed variations as separate and distinct illnesses, it does contain a few additions that are helpful in the recognition and diagnosis of postpartum psychiatric disorders. The DSM-IV also cautions mental health practitioners about the risk of suicide and infanticide in severe cases of psychosis, of the risk of recurrence in subsequent pregnancies, and that healthy development of the mother-infant relationship is dependent upon prompt treatment (APA, 2000).

In 2001 a symposium was held in London to discuss contemporary issues of diagnosis and classification in psychiatry (Cox, 2002). At this symposium, the policy guidelines from Community Mental Health in the United Kingdom, specifically recommended, amongst others, that there be new funding for some mental disorders like PPD and puerperal psychosis. According to the International Classification of Diseases and Related Health Problems, tenth edition (ICD-10) and DSM-IV these disorders do not exist as distinct psychiatric illnesses, but only as an optional 4-week postpartum onset specifier in the DSM-IV or 6 weeks for ICD-10. Cox (2002) further reports that “these and other anomalies were encouraging a new look at international classification” (p. 195). Cox (2002) stressed that a common classification language is essential so that researchers and clinicians can communicate.

While more common than other pregnancy related conditions like gestational diabetes and preeclampsia, and preterm delivery, PPD has received less attention in medical literature, clinical practice, and training. In recent years, there has been an increase in academic and lay press that focus on PPD, yet this condition remains frequently overlooked despite its potentially devastating consequences. The cause, definition, diagnostic criteria, and even the existence of PPD as a distinct entity is still a topic of debate among some clinicians today.

2.4.2 Diagnosing postpartum depression.

PPD is defined in different ways depending on the source. The diagnosis of PPD often depends on the severity of the depression as well as the duration of time between onset of depression and delivery. A number of related conditions should be differentiated from PPD when assessing the patient:

- Postpartum blues
- Postpartum psychosis
- Anxiety disorders
- Medical conditions

The latest edition of the Diagnostic and Statistical Manual of Mental Disorders, the DSM-IV-TR (APA, 2000), currently uses the term "with postpartum onset" as a specifier to describe the current or most recent major depressive, manic, or mixed episode in major depressive disorder, bipolar I or II disorder, or brief psychotic disorder that has its onset within 4 weeks postpartum. It further states that the symptoms in postpartum-onset major depressive, manic, or mixed episode do not differ from the symptoms in non-postpartum mood episodes. The DSM-IV-TR also mentions symptoms that are common in postpartum-onset episodes, though not specific to postpartum onset. These symptoms are mentioned in the next section.

2.4.3 Symptoms of postpartum depression.

2.4.3.1 Symptom overlap between the postpartum period and postpartum depression.

Many symptoms of mood disorders are similar to those that naturally follow childbirth, such as lack of sleep, appetite changes, fatigue, decreased libido, and mood lability (O'Hara, Neunaber, & Zekoski, 1984). Women also tend to lose weight postpartum in an attempt to regain their pre-pregnancy figures, and many lose weight naturally due to breastfeeding. Furthermore, depressed patients often fail to recognize that they are mentally ill. Their symptoms are often attributed to being tired, having a cold, or feeling overworked (Smith et al., 2004). Cognitive symptoms should therefore be monitored closely along with behavioural and somatic symptoms during the antenatal as well as the postpartum period.

2.4.3.2 Symptoms of postpartum depression versus depression.

Symptoms of PPD may be similar to depression experienced at other times (Yonkers, 2003), however, a number of studies have indicated that perinatal mood disorders are quite different from other mood disorders. According to Fowles (1998) the difference between PPD and depression experienced at other times in a woman's life rests in the postpartum mother's feelings of guilt about being an inadequate and incompetent parent. Wilkinson (2001) further states that the significance of PPD relates to the time of onset and the impact it may have on the family as well as on the woman herself. Depression after childbirth may be considered unique due to the presence of an infant and the stress the mother experiences as she adapts to motherhood (Weinberg et al., 2001). Bennett and Indman (2003) consider the influence of hormonal fluctuations. Roan (1997) points out that rapidly changing symptoms and poor interaction with the baby are particular symptoms that occur commonly among women experiencing postpartum illness that do not typically occur among other psychiatric patients.

2.4.3.3 Symptoms of postpartum depression.

According to the DSM-IV-TR, symptoms frequently found in postpartum-onset episodes, although they are not limited to postpartum onset, include fluctuations in mood, mood lability, and a preoccupation with the infant's well-being. The intensity of these symptoms may range from over-concern with infant well-being to frank delusions. The DSM-IV-TR mentions that the presence of delusional thoughts or severe ruminations concerning the infant is associated with a notably increased risk of causing harm to the infant (APA, 2000).

Postpartum-onset mood episodes can, according to the DSM-IV-TR, present either with or without psychotic features. Although infanticide can occur in severe postpartum mood episodes without specific hallucinations or delusions, it is most often associated with postpartum psychotic episodes (APA, 2000).

Women with postpartum major depressive episodes often have, according to the DSM-IV-TR, severe anxiety, and some present with panic attacks. The DSM-IV-TR further recognizes that maternal attitudes toward the infant are highly variable and states that the symptoms may include "disinterest, fearfulness of being alone with the infant, or over intrusiveness that inhibits adequate infant rest" (APA, 2000, p. 243).

Table 1 lists signs and symptoms of PPD that occur with varying degrees of severity, frequency, and extremity in women with PPD and are based on research findings in recent years (Dalton & Holton, 2001; Kruckman & Smith, 2006; Mehta & Sheth, 2006; Bennett & Indman, Roan, 1997; Smith et al., 2004). Symptoms needed to make a clinical diagnosis according to the DSM-IV-TR are listed with an asterix.

Researchers who study PPD agree that, in addition to the typical symptoms of depression, women with PPD may experience feelings of inadequacy, severe anxiety related to feeling incompetent in the care of their infant and worrying about the infant's welfare, feelings of hostility towards others, including the baby, thoughts of harming their infants, obsessive symptoms, unprovoked tearfulness, unexplained mood swings, feelings of abandonment, feelings of hopelessness, and suicidal thoughts. These symptoms can occur with varying degrees in women with PPD (Dalton & Holton, 2001; Mehta & Sheth, 2006; Bennett & Indman, 2003; Wilkinson, 2001).

Mothers with PPD may also worry about the involvement of child protection services and possible separation from their baby. Women also may feel reluctant to confide their distress, as childbirth is usually expected to be a joyful event. These issues raise special concerns for women with suicidal thoughts or thoughts of harming their babies.

Table 1 Signs and symptoms of PPD

Symptoms
Emotional state
• Depressed or low mood most of the day, nearly every day* *
• Markedly diminished interest or pleasure in most, if not all, activities* *
• Severe anxiety related to feeling incompetent in the care of the infant and worrying about the infant's welfare
• Feelings of inadequacy
• Unexplained mood swings
• Unprovoked tearfulness
• Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day *
• Low self-esteem
• Feelings of hopelessness
• Feeling of unreality and of not being one's usual self
• Feeling emotionally detached from loved ones, in particular from the infant
• Feelings of hostility towards others, including the infant
• Feelings of ambivalence about the infant
Cognitive functioning
• Diminished ability to think or concentrate, or difficulty making decisions nearly every day *
• Poor short-term memory
• Recurrent thoughts of death, suicidal ideation, suicide attempt or a specific plan *
• Bizarre, strange or obsessive thoughts
• Over-concern for baby's health
• Thoughts of harming the baby
• Misinterpretation of baby's cues
Behavioural symptoms
• Insomnia or hypersomnia nearly every day with sleep disturbance unrelated to

Symptoms
the new baby, struggling to fall asleep, frequent waking and waking up unusually early * • Psychomotor agitation or retardation nearly every day * • Fatigue or loss of energy nearly every day * • Complaints of lack of social support • Extreme Behaviour • Panic Attacks • Hostility • Nightmares • Unresponsiveness towards the baby • Over-concern for the baby **Physical Symptoms** • Significant weight loss when not dieting, weight gain, or an increase or decrease in appetite nearly every day * • Loss of libido • Headaches • Numbness, Tingling in Limbs • Chest Pains, Heart Palpitations • Hyperventilation

* In order to diagnose a major depressive episode with postpartum onset, five or more
** of the items marked with a single asterix (*)must have been present during the same 2-week period and represent a change from previous functioning, and at least one of the symptoms marked with a double asterix (**). Onset of the episode must be within 4 weeks postpartum.

In the following section the prevalence and clinical course of PPD is presented along with a review of explanations concerning the cause of PPD in medical and psychological literature is presented. This is followed by an examination of the consequences and treatment of PPD.

2.4.4 Prevalence of postpartum depression.

Publications report that PPD effects up to 20% of women (Dalton & Holton, 2001; Josefsson, Berg, Nordin, & Sydsjö, 2001; Stuart, Couser, Schilder, O'Hara, & Gorman, 1998; Wilkinson, 2001). Halbreich (2005) discusses the diversified epidemiology of pregnant and postpartum symptoms and disorders. Halbreich reports that most publications estimate that PPD affects 10–15% of women. Prevalence estimates vary widely depending on the diagnostic criteria, the measures used in assessment, the sampling procedures, and the location of populations.

A meta-analysis by O'Hara and Swain (1996) reported that the rate of PPD in developed countries was approximately 13%. Research by Righetti-Veltema, Conne-Perreard, Bousquet, and Manzano, 1998, and Whitton, Warner and Appleby (1996), states that the incidence of PPD, without psychotic features, is roughly 10 - 15% for first time mothers. Women with a previous history of depression have a 2-fold rate of recurrence of a depressive disorder in the perinatal period (Banti et al., 2011; Sichel & Driscoll, 1999), while women with a previous history of PPD have an estimated 10-35% rate of recurrence (Kruckman & Smith, 2006).

Priest, Henderson, Evans, & Hagan (2003) compared the prevalence rate in the first few weeks postpartum to the rate in the first year postpartum. They found that in industrialised countries, rates for PPD varied between 13% in the first few weeks after delivery, to 20% in the first year postpartum.

A large study comprising 6,000 postpartum women estimated that the 2-month prevalence for postpartum-onset of major depressive disorder was 15% (Cooper, Murray, Hooper, & West, 1996). Transculturally, the rates were estimated at 10% to 15%, with a higher rate in adolescent mothers (Kumar, 1994). Higher rates of PPD were also reported in some developing countries (Patel, Rahman, Jacob, & Hughes, 2004). Halbreich's and Karkun's (2006) detailed review of the literature reveals that the reported prevalence of PPD varies among countries between 0.5% to over 60% of new mothers. Even in the USA, reports vary between 3.7% and 48.6% (Halbreich, 2005). This is despite the fact that most surveys applied the same instruments – the Edinburgh Postnatal Depression Scale (EPDS) or the Beck Depression Inventory (BDI).

Halbreich and Karkun (2006) found that in several countries like Denmark, Austria Singapore, Malaysia and Malta, PPD or postpartum depressive symptoms are seldom reported, unlike other countries (e.g. South Africa, Brazil, Costa Rica, Guyana, Italy, Taiwan, Chile, and Korea) where postpartum depressive symptoms were very prevalent. They believe that, due to the varying reports, the broadly cited mean prevalence of PPD (10-15%) is not truly representative of the real global prevalence and magnitude of the problem.

Affonso, De, Horowitz, and Mayberry (2000) attribute the diversity across countries and cultures to cultural, socio-economic, genetic, and reporting style differences. Halbreich and Karkun (2006) agree that these factors may contribute to the variability in reported PPD. They further state that factors such as cross cultural differences in the perception and stigma of mental health, differences in socio-economic environments (for example levels of social support or its perception, poverty, stress and nutrition), and factors due to biological vulnerability may be significant too. Halbreich (2004) attributes the diversity in reported prevalence to factors such as sampling and assessment methods. Halbreich (2004) found that most reports, especially those on minority women or developing countries, were based on relatively small samples, did not include a control group, have been based on self reports of symptoms – mostly with a short dimensional screening instrument (e.g. EPDS), and were not necessarily based on structured clinical interviews to formulate a DSM-IV, ICD-10 diagnoses or both. Greater insight into the underlying processes impacting on the varied prevalence of PPD along with insight into the range of normal postpartum versus abnormal postpartum expressions of symptoms may lead to a better understanding of the diversified phenomena in perinatal mental illness (Halbreich & Karkun, 2006).

Some published reviews assert that the prevalence of mental disorders during pregnancy and postpartum is not higher than during other periods of a woman's reproductive life. Despite inconsistent findings there are other researchers who suggest that after childbirth women are at a 12% to 15% higher risk for serious depressive illness than are non-childbearing women (Whiffen, 1992).

A major challenge in dealing with PPD has been early recognition, partly because PPD is covertly experienced. Researchers found that not many women with PPD seek assistance of their own accord (Murray, Woolgar, Murray, & Cooper, 2003). Furthermore, missed diagnosis is frequent in settings where mental health status does not undergo a structured method of review (Reid et al., 1998; Evins et al., 2000). It has been reported that up to 50% of mothers affected by postpartum depression go undetected (Ramsay, 1993). According to Kruckman and Smith (2006) and Milgrom, Mendelsohn, and Gemmill (2011), the use of depression scales specifically aimed at screening for perinatal mental illness will benefit future research by providing a more accurate picture of the incidence of PPD.

Clinicians tend to trivialise the seriousness of PPD and equate it simply with maternity blues (Huysman, 1998). Furthermore, many symptoms are similar to those that naturally follow childbirth, such as lack of sleep, appetite changes, fatigue, decreased libido, and mood lability (O'Hara et al., 1984). As a result only a small percentage of these women are identified by health practitioners as depressed. Mothers often suffer in silence, fear, and confusion before PPD is diagnosed. In a study by Hearn et al (1998), it was reported that up to 50% of cases go unreported. This may be due to the mother's concern about the stigma associated with mental health issues, or concern that the custody of her baby may be jeopardized if she were to report her mood swings and emotional state.

2.4.5 Clinical course of postpartum depression.

PPD can be mild, moderate, or so severe that it includes suicidal thoughts and requires hospitalisation (Roan, 1997). The onset may be gradual and insidious or sudden, but it commonly occurs within two to four weeks after delivery. Depression may occur at any time after childbirth, but more commonly sets in after the woman has returned home from hospital (Kruckman & Smith, 2006). According to Roan (1997), in some cases PPD may have started as postpartum blues that lingered and developed into a serious condition. According to Bennett and Indman (2003) and the Marcé Society, an international organization for the study of psychiatric illness related to childbearing, the onset is usually gradual, but it can be rapid and begin any time in the first year. The DSM-IV-TR stipulates that the initial episode of postpartum-onset depression begins within the first 4 weeks after delivery (APA, 2000). Many clinician's and researchers agree, however, that PPD symptoms are insidious and may occur at anytime up to a year after childbirth, but more commonly occurs within the first three months (Beck, 2006).

Early research by Gelder indicates that the symptoms of PPD may last anything from a few weeks to several months with approximately 4% of incidences persisting for as long as a year (as cited in Kruckman & Smith, 2006) and that the majority of women who have PPD recover within 6 months (Kumar & Robson, 1984). Subsequent research, however, suggests otherwise. England, Ballard, and George, (1994) claim that 20% of women will have chronic depression lasting longer than two years. Beck (2006) states that recent evidence reported by the Agency for Healthcare Research and Quality (Gaynes et al. as cited in Beck, 2006) indicates that up to 19.2% of new mothers may have either major or minor depression in the first three months postpartum, of which as many as 7.1% have major depression.

Researchers agree that the duration of PPD may vary. For some women it may be mild and short-lived (a matter of weeks), vanishing on its own, but for most women it may languish for several months or a year (Beck, 2006; Roan, 1997). Some women who experienced PPD have depressive episodes that persist throughout life (Roan 1997). Women who have had a severe episode of PPD may continue to suffer from depression for up to two years (Smith et al., 2004). Philipps and O'Hara (1991) found substantial recurrence in the long-term follow up of women with PPD. Half of the mothers with PPD either felt the need to continue or again sought treatment over four years. Smith et al (2004) indicate, based on research findings, that a woman has a significant risk for developing chronic depression as well as lifetime recurrence of depression regardless whether the postpartum episode was the first depressive event or whether it was a recurrence.

2.4.6 Perspectives on the etiology of postpartum depression.

A number of possible hormonal, biological, cultural and psycho-social theories have attempted to explain the onset of perinatal disorders. The role that certain hormones may play in the development of PPD has attracted substantial scientific research. Yet, a clear link between hormones and PPD has not been found. This has led some researchers to conclude that causality may be found in psychological or social factors. Researchers do, however, seem to agree PPD develops from an interplay of multiple factors.

Kruckman and Smith's (2006) review of journal research from the past five years revealed that a majority of articles focussed on biological cause or pharmacotherapy linked to a biological etiological view. Researchers who strictly examined psychological factors as dominant causes, and related research on predictions, risk, and screening scales also comprised a large percentage of publications.

According to Kruckman and Smith (2006), the majority of researchers agree that studies of hormonal influence in PPD have not produced a direct link to PPD. Since psychological stimuli affect the neuroendocrine systems it has been recommended that research on hormonal impact should be performed in conjunction with psychosocial research. Gelder reviewed the hormonal link to PPD over 2 decades ago, and concluded that psychological and social factors were responsible (as cited in Kruckman and Smith, 2006). Two decades later, Hendrick, Altshuler and Suri (1998) came to a similar conclusion: "The literature to date does not consistently support any single biological etiology for postpartum depression." (p. 98). Hendrick et al. (1998) recommend that future research which investigates biological factors as triggers for postpartum mood disorders ought to control for psychosocial variables, as they believe that these variables are likely to confound the data.

The socio-cultural context of childbirth has also been considered. Childbirth may be a similar physiological experience universally, but it occurs in a socio-cultural context and is conceptualised and experienced according to people's specific values, attitudes, and beliefs. Anthropological perspectives view postpartum disorders from a bio-cultural approach and examine the influence of cultural patterns such as family values, structure, roles, and beliefs. Anthropologists believe that while objective measures of underlying physiological processes may explain symptoms, it is important to take into account that the experience of those mechanisms are filtered, mediated and directed by culturally constituted frameworks (Kruckman & Smith, 2006; Stern & Kruckman, 1983).

Much research on PPD has focussed on biological and psycho-social etiologies such as hormonal changes, psychiatric history, maternal age, marital relationship, and so forth. Although these are important contributing factors, the influence that cultural patterning of the postpartum period has in the etiology of PPD needs consideration. This relates to factors like the social context, structure, and organization of the family. Furthermore, the role expectations of the new mother and father also need consideration.

In the following section Halbreich's model that explains the evolving etiology and pathology of postpartum disorders is discussed. Halbreich's model is comprehensive and considers the influence of numerous factors in the onset of postpartum disorders. Furthermore, Halbreich (2005) takes into account that diversified postpartum disorders may have different predictors for the different underlying processes. The processes leading to postpartum disorders are, according to Halbreich (2003), multifaceted on several levels. The Bio-Psycho-Socio-Cultural Model by Halbreich (2005) of the processes leading to postpartum disorders is presented in Figure 1. According to this model, symptoms are a consequence of a process starting from a genetic predisposition to dysregulation and impaired ability to adapt. The model takes into account a person's dynamically evolving vulnerability that is shaped along the individual's life. Symptoms and disorders may surface in response to biological and social triggers. Halbreich (2005) explains that the individual's response depends on the perinatal and postpartum environment at the time.

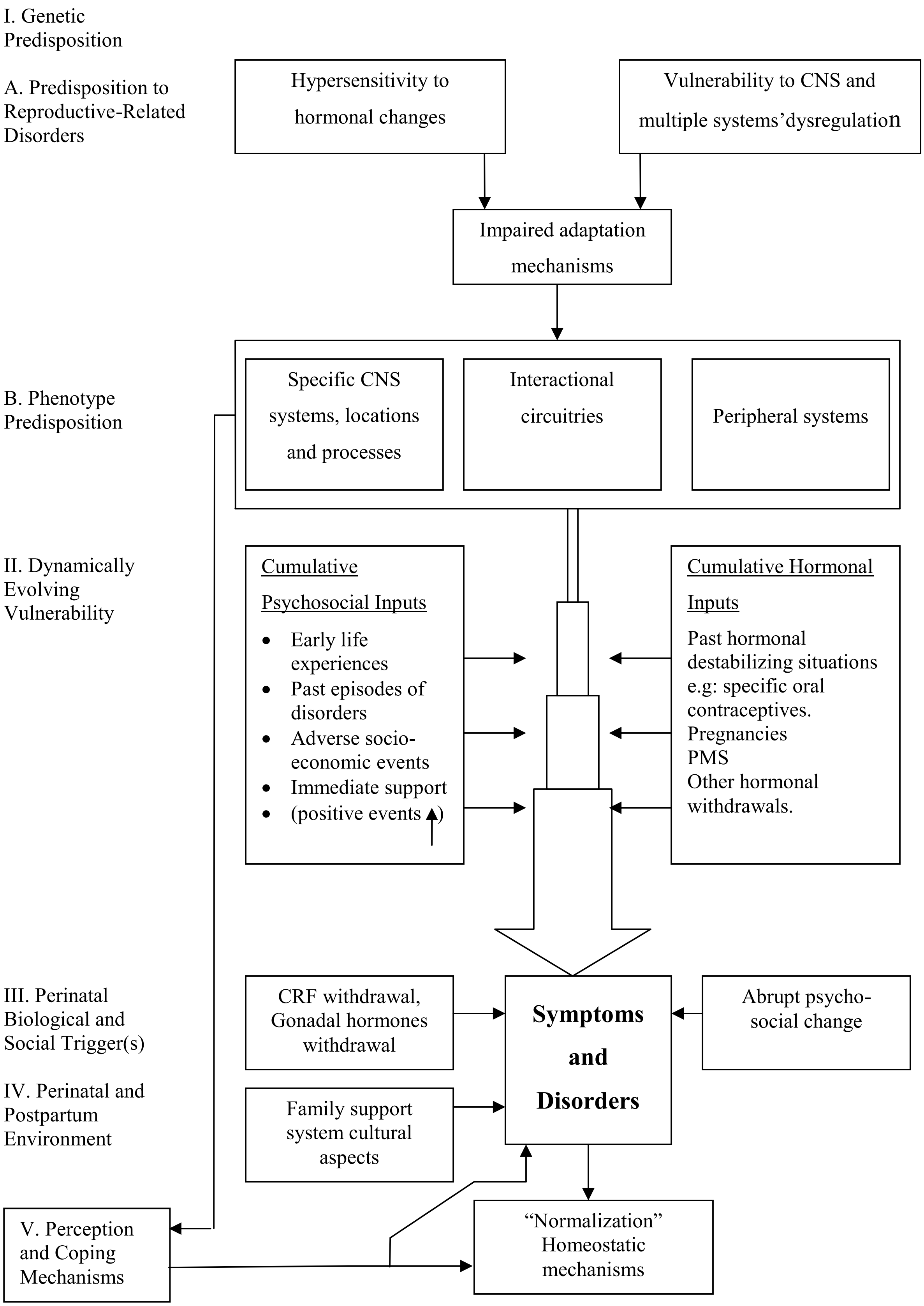

Figure 1 Bio-psycho-socio-cultural model of the processes leading to postpartum disorders. (Halbreich, 2005)

According to Halbreich (2005), the main etiological factor may be a genetically determined predisposition to reproductive-related disorders. He states that this vulnerability is likely to be due to a combination of two factors: (a) an individual's hypersensitivity to changes in gonadal hormones and possibly also to other steroids. Such hypersensitivity would also bring about symptoms a woman may experience during other periods of hormonal change or instability; and (b) vulnerability to these factors is compounded by the tendency towards central nervous system dysregulation as well as adaptation mechanisms that are impaired.

Central nervous system and other peripheral systems' vulnerability may be due to changes in the activity of steroids as well as impaired adaptation to other external situations of change causing stress, such as abrupt drug and hormonal withdrawal and other psychological, biological, and social changes. This dysregulation leads to a disruption of homeostasis or circuitry that is hypersensitive to change (Halbreich, 2005).

The second stage of genetic vulnerability in Halbreich's (2005) model involves a phenotype predisposition. Vulnerability would depend on an individual's predisposition to phenomena resulting from abnormalities in their own central nervous system systems – although which central nervous system may be responsible for impaired adaptation is unclear as it depends on the individual's sensitivity to these systems. Some women may show symptoms due to two or more systems that are out of balance. The vulnerable system or systems will determine the nature of symptoms and their clusters. Halbreich (2005) points out that in some women, the vulnerable system may be peripheral, leading to diversity of postpartum mood, behaviour, and physical disorders.

Halbreich (2005) refers to the genetic vulnerability as “dynamically evolving vulnerability” as it is continually shaped by internal as well as external environmental inputs, and it constantly changes according to cumulative life experiences, both negative and positive. Cumulative hormonal inputs and their influence on many physiological processes may further increase the vulnerability an individual has to disorders. Furthermore, the kindling effect of psychosocial factors such as repeated dysphoric states, repeated episodes of disorders and adverse socio-economic circumstances may cause an increase in the dynamically evolving vulnerability.

The influence of hormonal disturbance in the perinatal period is shown in level III in Figure 1. Halbreich points out that the most powerful trigger of postpartum symptoms is the abrupt delivery of the placenta. Levels of progesterone and some other steroids reach highest levels during pregnancy. In addition, corticotrophin releasing factor (CRF) is in its highest peak just prior to delivery that results in hyperactivity of the hypothalamo–pituitary–adrenal (HPA) system of both the mother and her pre-born baby. When the placenta is delivered its hormonal secretions are abruptly withdrawn. This causes immediate changes in every system influenced by its regulating hormones. Halbreich (2005) points out vulnerable women may experience dysregulation within the affected systems and, as a result, may develop symptoms.

Halbreich (2005) explains that the culprit for postpartum disorders cannot be found by focusing on the functioning of a single neurotransmitter. He discusses the possibility that a large number of systems may be functioning abnormally or that the multiple interactions among these systems may be in a state of imbalance or may be impaired in women with postpartum disorders. He proposes that such a dysregulated state can be a consequence of the inividual's impaired homeostatic mechanism and may ultimately be the cause of the development of a disorder.

Halbreich's model takes into account the effect of psychosocial change from pregnancy to motherhood, along with its demands and stresses, which may contribute to the onset of postpartum disorders. Halbreich also discusses the influence of socio-cultural aspects and the environment at the time of delivery and thereafter, the amount of support the mother received, the mother's perceptions and coping ability – which are shaped by past experiences and influenced by the functioning of the central nervous system – as contributing factors to consider in the onset of postpartum disorders.

Halbreich (2005) regards the interactions between trait, state, environment and culture as especially noticeable in women with postpartum disorders as well as other reproductive-related disorders. He does, however, advocate a diverse and open-minded approach to explain how and why an external event is regarded as a challenge, as something pleasant or something negative and anxiety provoking. How our perception of the environment is translated to the biological mechanisms of the central nervous system is, no doubt, an intricate and complex process.

2.4.7 Risk factors for postpartum depression.

PPD develops from the interplay of multiple biopsychosocial and cultural factors. A number of researchers have examined specific factors and their contribution to the onset of postpartum disorders. Beck (1996a) conducted a meta-analysis of 44 studies to determine the magnitude of the relationships between various predictor variables and

PPD. The strongest predictor of this mood disorder was antenatal depression. Moderate effect sizes were revealed for the relationships between PPD and the following predictors: social support, life stress, childcare stress, marital satisfaction, antenatal anxiety, and maternity blues. Lastly, history of previous depression was shown to have a small effect size when determining its relationship with PPD. In a different meta-analysis (Beck, 1996b) infant temperament was also revealed to be a significant predictor of PPD.

In addition to Beck's (1996a, 1996b) meta-analyses, one other meta-analysis of predictors of PPD had been conducted. O'Hara and Swain (1996) determined the effect sizes of a number of risk factors for PPD that had been measured during pregnancy. They reported that the strongest predictors of PPD were antenatal depression, antenatal anxiety, social support, life events, and mother's history of psychopathology. The meta-analysis revealed the following three predictors that had small, significant relationships with PPD: neuroticism, negative cognitive attributional style, and obstetric variables.

Since 1996 the amount of research on risk factors for PPD has dramatically increased. Beck (2001) conducted another meta-analysis to update the findings of these earlier meta-analyses of PPD predictors. Results confirmed findings of the earlier meta-analyses and also revealed four additional predictors of PPD: marital status, self-esteem, unwanted or unplanned pregnancy, and socio-economic status.

In these studies, a total of thirteen significant predictors of PPD were revealed. The risk factors were as follows: antenatal anxiety, antenatal depression, a history of depression, unplanned or unwanted pregnancy, postpartum blues, infant temperament, childcare stress, social support, life stress, marital status, marital relationship, self-esteem, and socio-economic status.

Kruckman and Smith (2006) state that the etiology of perinatal mental illness is complex and also likely diverse, with some distinct and shared risk factors. These risk factors may be biological variables, personality variables, psychological variables, demographic variables, interpersonal variables, or obstetric variables. The significant predictors of PPD identified by Beck (1996a, 1996b, 2001) as well as significant biological, obstetric risk, and other psychosocial and personality factors identified by other researchers are discussed in more detail below.

2.4.7.1 Antenatal depression and anxiety.

A number of studies have since demonstrated an association between anxiety or depression during pregnancy with PPD (e.g. Josefsson, 2003; O'Hara, Zekoski, Philipps, & Wright, 1990; Laizner & Jeans, 1990; Orr, James, & Blackmore Prince, 2002; Robertson, Grace, Wallington, & Stewart, 2004; Sutter-Dallay, Giaconne-Marcesche, Glatigny-Dallay, & Verdoux, 2004; Verkerk, Pop, Van Son, & Van Heck, 2003). O'Hara and Swain's (1996) results from their meta-analysis of the rates and risk of PPD, demonstrated a strong association between antenatal depression and PPD. This was particularly the case when a self-report measure was used to determine the presence of symptoms. Consistent with the literature, Da Costa, Larouche, Dritsa, and Brender (2000) reported that the best predictor of postpartum depressed mood was antenatal depression. Beck (2001), in a replicated meta-analysis, validated some findings from O'Hara and

Swain's (1996) meta-analysis, and also confirmed findings from her own earlier meta-analysis (Beck's 1996a), including that amongst the strongest predictors of PPD were antenatal depression and anxiety.

Orr, et al. (2002) established that the risk for PPD in women who had the most depressive symptoms during pregnancy was doubled. Matthey et al., (2000) noted a link ranging from between 18% and 75% and stated that the rate is dependent upon the population group. Leung, Martinson and Arthur's (2005) study aimed to identify correlations between demographic variables and PPD, and psychosocial variables and antenatal depression in Honk Kong Chinese women. One of the major three predictors in this group was depression during pregnancy. Rich-Edwards et al. (2006) also found that the strongest risk for postpartum depressive symptoms was antenatal depressive symptoms. The DSM-IV-TR also specifies that having the "baby blues" in addition to mood and anxiety symptoms during pregnancy, increases the risk for a postpartum major depressive episode (APA, 2000).

Verkerk et al. (2003) investigated, amongst other things, whether the occurrence of depression during the first year after childbirth can be predicted in around the mid trimester of pregnancy. Their findings led them to conclude that it was possible to detect during pregnancy those women who were at high risk as well as women who were at low risk for PPD during the initial months after childbirth. Women who were high-risk were only at particular risk during the first 3 months after delivery. A personal history of depression and high depressive symptomatology during mid-pregnancy were found to be independently predictive risk factors of PPD. Bloch, Rotenberg, Koren and Klein (2005) noted a strong trend for a significant effect of mood symptoms during the 3rd trimester and the development of PPD.

The severity as well as the duration of depressive symptoms impact on a woman during pregnancy. The physiological impact, for example, can be seen in poor maternal weight gain, or even weight loss due to poor appetite. Depression during pregnancy has been associated with low birth weight (less than 2,500 grams) and preterm delivery (less than 37 weeks; Bennett & Indman, 2003). Severe anxiety during pregnancy may cause harm to a growing foetus due to constriction of the placental blood vessels and higher cortisol levels". ACOG (2002) agree and report that unmonitored and untreated depression in antenatal women may initiate premature labour and delivery. They further stipulate that careful monitoring is required of the depressed mother during pregnancy to ensure a healthy outcome for both the mother and her foetus.

2.4.7.2 Past history of depression.

Bender (2003) reports that women with a history of major depression were 5 times more likely to present with depressive symptoms in the peripartum period. Ryan, Milis, & Misri (2005) found that a history of depression or any other psychiatric disorder may increase the risk of developing PPD. According to the DSM-IV-TR a family history of Mood Disorders on top of a personal history of non-postpartum Mood Disorder increases the risk for developing a postpartum Mood Disorder (APA, 2000). A family history of mental health problems increases the risk of PPD significantly. Genetic predisposition and psychosocial variables related to having a family member with psychiatric illness may be responsible for the increased risk (Freeman et al., 2005).

O'Hara and Swain's (1996) meta-analysis of the risk factors for PPD found that past history of psychopathology is a significant risk factor for PPD, although the kind of PPD assessment that was used influences the magnitude of the relation between PPD and previous psychiatric history. Forman, Videbech, Hedegaard, Salvig and Secher (2000) conducted a large study to identify and test the predictive power of risk factors of PPD. A history of pre-pregnant psychiatric disease was among the strongest identified risk factors. Beck's (1996a; 2001) meta-analyses confirmed that a mother's history of depression is a strong and significant predictor of PPD.

A mother's prior psychiatric history, especially the occurrence of previous depressive episodes, has emerged as one of the most salient predictors of PPD (e.g. Baker & Oswalt, 2008; Bloch et al., 2005; Dennis, Janssen, & Singer, 2004; Freeman et al., 2005; Rich-Edwards et al., 2006; Robertson et al., 2004). As mentioned in the previous section, the occurrence of depressive symptoms during pregnancy was found to be a risk factor in the development of PPD. Swendsen and Mazure (2000) point out that, taken together, these findings signal that PPD may, in some cases, constitute an exacerbation or recurrence of illness, rather than the onset of a depressive disorder that is only due to the state-related changes of motherhood. ACOG (2002) encourages counselling women prior to conception about their risk for recurrent depression during their pregnancy and also during the postpartum period.

2.4.7.3 Postpartum blues.

The prevailing perception that postpartum blues is inevitable and self-limiting has led to the condition receiving comparatively little attention from perinatal researchers (Kruckman & Smith, 2006). The exact mechanisms responsible for the development of postpartum blues or psychotic depression have been debated, but have not been clearly identified. Kruckman and Smith (2006) suggest that it may be that postpartum blues is "simply the milder end of a biologically based continuum in which the severe end is psychosis".

Henshaw (2003) did a comprehensive review of postpartum blues and described an earlier investigation, which took place over a period of 6 months, of 103 women with severe postpartum blues and their controls with no postpartum blues. It was found that severe postpartum blues was an independent predictor of depression. Depressive episodes in the severe postpartum blues group had onset earlier in the puerperium, lasted longer and were more likely to be major than minor depression. Henshaw (2003) concludes that the most convincing relationships with early mood disturbance are dysphoria during pregnancy, a personal history of depression, premenstrual depression, neuroticism, and depression later in the postpartum period suggesting that postpartum blues is a predictor of subsequent PPD and appears to be an index of affective vulnerability.

Lane et al. (1997) investigated the predictors of PPD and found that amongst the factors associated with PPD, mothers' mood state at 3 days postpartum (symptomatology related to the "blues") was the best predictor of psychopathology at 6 weeks. In this study, EPDS scores at day 3 postpartum were similar to EPDS scores at week 6 postpartum. O'Hara and Swain's (1996) and Beck's (1996a; 2001) meta-analyses of predictors of PPD have shown that of the 13 significant risk factors for PPD that they identified, postpartum blues was one of 10 of these predictors that had moderate *r* effect sizes. More recent studies by Bloch and Klein (2005) and Bloch et al. (2005) also found that mood symptoms during the first 2-4 days postpartum were amongst the significant risk factors for postpartum mood disorders.

Postpartum blues is, however, more prevalent than PPD, affecting up to 70% percent of postpartum women (APA, 2000). All women who experience postpartum blues will not necessarily develop PPD. The results of this study do, however, highlight the importance of screening prior to discharge from hospital and the need for early intervention for women at risk.

2.4.7.4 Hormonal changes.

2.4.7.4.1 Neuroendocrine alterations.

Hormonal changes are dramatic during pregnancy and shortly after delivery. Numerous studies have explored how reproductive events may contribute to the development of postpartum mood disorders. The DSM-IV-TR recognizes that neuroendocrine alterations render the postpartum period unique (APA, 2000). The levels of progesterone, estrogens, human chorionic gonadotropin, beta-endorphin, cortisol and prolactin increase during pregnancy and reach a maximum level near term and then decline rapidly after delivery. A topic of intense debate amongst researchers is whether postpartum mood disorders have a distinct pathophysiology.

In some studies that explored biological factors, a specific etiologic link between postpartum mood disorders and reproductive changes has not been identified (Hendrick et al., 1998; Ross, Sellers, & Romach, 2003). Ross, Sellers, and Romach (2003) investigated the interactions between psychosocial and biological risk factors in PPD and anxiety. They reached the conclusion that hormonal variables may not have a direct impact on women's moods during pregnancy and the postpartum period. Their results did, however, indicate that hormonal variables do play an important role in perinatal anxiety as they may mediate sensitivity to psychosocial stressors. Their results also emphasize the importance of examining the effect of biological variables in PPD in addition to examining demographic and psychosocial risk factors.

The theory of hormonal influence as a risk factor for PPD is supported by researchers like Bloch et al. (2000) who attribute PPD to hormonal changes. They report that women had a greater risk of recurrence of depressive symptoms during a pseudopregnancy and parturition if they have a history of PPD.

Epperson et al. (2003; 2006) are of the opinion that neuroactive steroids play a definite role in postpartum mood disorders considering the temporal relationship between hormonal changes associated with parturition and the onset of symptoms. Their examination of the GABA levels of postpartum women led them to conclude that some postpartum women are more vulnerable to the fluctuations in sex steroids and the onset of postpartum affective disorders.

Altemus et al. (2004) examined the changes that occur in the neurochemistry of cerebrospinal fluid during pregnancy. They report that levels of prolactin, but not oxytocin, in CSF and plasma were correlated in pregnant women. These results suggest that pregnancy alters regulation of brain GABA, norepinephrine, and prolactin, which may play a role in changes in vulnerability to anxiety and depression during pregnancy and postpartum.

Studies have indicated that postpartum hormone withdrawal may contribute to depressive symptoms experienced after giving birth (e.g. Ahokas, Kuakoranta, & Aito, 1999; Bloch, Daly, & Rubinow, 2003).According to Suri (2004) women who develop PPD may be particularly sensitive to these dramatic hormonal fluctuations that take place in the immediate postpartum period. Halbreich (2005) states that hormonal changes in conjunction with genetic predisposition, causing hypersensitivity to these changes, places women at an increased risk of developing PPD.

2.4.7.4.2 Premenstrual dysphoric disorder.

Halbreich and Halbreich and Endicott (as cited in Halbreich, 2005) demonstrated and suggested a statistical association between PPD and depressions during other reproductive-related situations, like premenstrual dysphoric disorder (PMDD) and puberty. The statistical association reflects common underlying mechanisms, most likely hormonal withdrawal, changes, or instability. A kindling effect was also suggested. This implies that repeated hormonally-related episodes have a cumulative effect resulting in increased sensitivity or vulnerability to develop symptoms in response to future situations of change.

Sugawara et al. (1999) reported that high postpartum depressive scores were associated with a history of PMS. Bloch and Klein (2005) found a clear association between having a history of PMDD and the development of either PPD or the blues. While their study was limited by the retrospective report of PMDD symptoms, the subsequent diagnosis of postpartum mood disorders and comparison to a control group strongly supported considering a history of PMDD as a risk factor for postpartum mood disorders.

Bloch et al. (2005) later also report that significant risk factors for postpartum mood disorders were a history of PMDD and a history of mood symptoms in prior oral contraceptive use. These studies provide evidence that putatively hormone-related phenomena are related to the occurrence of postpartum mood disorders. The results go some way to support the hypothesis that the etiology for postpartum mood disorders may be related to differential hormonal sensitivity. Such risk factors should be included in any assessment of the risk for these disorders.

2.4.7.4.3 Thyroid dysfunction.

It has been suggested that abnormalities in thyroid functioning in the postpartum period contribute to postpartum mood disorders (Bokhari et al., 1998; Pop et al., 1991; Pop et al., 1993). The percentage of women with postpartum hypothyroidism is fairly high in the first six months after delivery. The rate of thyroiditis in postpartum women has been found to reach 9%, compared to 3% to 4% in the general population (Goldman, 1986). The relationship between PPD and postpartum thyroid dysfunction may substantiate a hormonal theory for the development of PPD in a small number of women (Pop et al., 1991; Harris et al., 1996), but it does not account for most cases of PPD. Nevertheless, Pop et al. (1991) study shows that a significant fraction (7%) of euthyroid women developed postpartum thyroid dysfunction after childbirth. Thirty-eight percent of these women had PPD that resolved when the thyroid abnormality was treated. Thyroid dysfunction should therefore be given consideration in the assessment of women who present with PPD. Stronger associations have, however, been found with factors like social support and infant variables, and PPD also occurs in fathers. Therefore, it would be faulty to assume a strictly hormonal etiology for most cases of PPD.

2.4.7.4.4 Serum n-3 polyunsaturated fatty acid levels.

Alterations in serum fatty acid composition accompany major depression (De-Vriese, Christophe, & Maes, 2003). Maternal serum 22:6n-3 is depleted due to pregnancy. This serum level gradually declines further after childbirth. De-Vriese et al. (2003) investigated whether cholesterol esters and the postpartum fatty acid profile of maternal serum phospholipids differs in women who develop PPD. They found that abnormalities in fatty acid status were also observed in PPD just as it had previously been observed in major depression. Their results further show that antenatal women may benefit from prophylactic treatment with serum n-3 polyunsaturated fatty acids if they are at risk of developing PPD.

2.4.7.5 Obstetric risk factors.

2.4.7.5.1 Preterm infants.

Depression in mothers of pre-term infants is not uncommon. This mood disorder may impact on the health of the infant (Kruckman & Smith, 2006). Elevated depression scores after childbirth were significantly more frequent among mothers whose infants were born preterm. These findings were evident even when antenatal depression scores were controlled (Drewett, Blair, Emmett, & Emond, 2004). Locke et al. (1997) found that the severity of the initial neonatal illness was associated with maternal depression in mothers of preterm infants.

Halbreich (2005) proposes that the risk factors for PPD may be similar to the risk factors for low birth weight or preterm delivery. This may suggest that the 3 situations - low birth weight, preterm delivery and PPD - may be an outcome of similar or partially overlapping pregnancy processes. Halbreich (2005) states that it may be that low birth weight and preterm delivery are predictive factors for PPD, particularly when the infant's special needs severely affect the mother.

2.4.7.5.2 Perinatal loss.

Depression and anxiety are not uncommon after a pregnancy is terminated, either through own choice or in miscarriage. Furthermore, a bereaved mother typically experiences depressive symptomatology when a stillbirth or neonatal death occurs (Bennett & Indman, 2003).

Turton, Hughes, Evans, and Fainman (2001) demonstrated that women, who in stillbirth have suffered the double psychological burdens of trauma and bereavement, are at a significant risk of developing PTSD and comorbid symptoms of depression and state-anxiety during and after the pregnancy following stillbirth. Stowe, Levy, and Nemeroff (1997) caution that patients may be deprived of adequate support and treatment if either the patients or the professionals consider severe depression to be normal after a significant loss. These studies highlight the need for education about PPD and PTSD and the importance of careful ongoing diagnostic, pharmacological, and psychotherapeutic treatment of patients who suffer from perinatal loss. Furthermore, bereaved mothers ought to be carefully monitored for symptoms of depression and anxiety in subsequent pregnancies and in the postpartum period.

2.4.7.5.3 Care during labour and delivery.

The quality of care the mother receives during labour and delivery has been reported to be a risk factor for PPD. Studies have shown that the emotional and psychological care a woman receives during labour and delivery, as well as the physical care provided determine her satisfaction with childbirth. How all these needs are met is considered an important factor in postpartum outcomes like PPD (Baker, Henshaw, & Choi, 2003).

2.4.7.5.4 Delivery complications.

Birth complications have been investigated as possible risk factors in the development of PPD, and results have varied. In a study by Warner, Appleby, Whitton, and Faragher (1996) where obstetric risk factors for postpartum psychiatric morbidity were examined, there was no association reported by subjects. O'Hara and Swain (1996), however, found a moderate correlation between women with higher levels of obstetrical complications and those with higher levels of self-reported symptoms of depression during the postpartum period.

2.4.7.5.5 Unplanned caesarean delivery.

Concern has been expressed since the 1970s that caesarean delivery and PPD may be linked. A broad range of findings have been reported. Carter, Frampton, and Mulder (2006) point out that this may be partly due to methodological factors employed. Most commonly, however, studies have found no association between PPD and caesarean delivery. Carter et al. (2006) performed a meta-analysis of suitable studies and report that methodologically superior studies were more likely to find no significant association.

A recent study by Patel, Murphy, and Peters (2005) examined the association between PPD and elective caesarean delivery compared with planned vaginal delivery. Patel et al. (2005) further explored whether assisted vaginal delivery or an emergency caesarean section is associated with PPD compared with vaginal delivery that proceeds spontaneously. Their results show that women who plan vaginal delivery and due to complications require an emergency caesarean section or assisted vaginal delivery are not at increased risk of PPD. Variables such as whether the caesarean delivery is planned or unplanned do not appear to significantly increase the risk of PPD (Patel, Murphy, & Peters, 2005).

The studies reviewed in the meta-analysis by Carter et al. (2006) suggest an association between a variety of other risk factors and PPD. They suggest that caesarean delivery operates as a risk factor for PPD only if women are vulnerable to PPD for some other reason. Gottlieb and Barrett (1986) found that lack of experience with children was a moderating variable between caesarean delivery and PPD. Two studies are of indirect relevance to this issue. Green (1990) found that low antenatal mood and negative experiences of labour had independent and cumulative effects on PPD. Murray and Cartwright (1993) found that the mode of delivery was only associated with PPD if women had a history of depressive disorder.

2.4.7.5.6 Tokophobia.

Hofberg (2003) studied the profound dread and avoidance some women have of childbirth. Fear of childbirth is not uncommon in pregnant women. It may, however, be disabling in up to 10% of parous women (Saisto & Halmesmäki, 2003; Waldenström, Hildingsson, & Ryding, 2006). Fear of childbirth is equally common in nulliparous as in parous women. In up to 13% of women who are not pregnant, the fear is so intense that they prefer to postpone or avoid pregnancy altogether.

Profound fear of childbirth is not a modern day phenomenon. In 1858, Marcé described fear of parturition (as cited in Hofberg, 2003). Despite advances in medicine and the types of assistance offered to women in childbirth, many women still fear pain and death during childbirth. When this fear precedes pregnancy and is so intense that the woman avoids pregnancy, and hence childbirth, it is a phobic state termed tokophobia. Wijma (2003) refers to this phenomenon as “clinical fear of childbirth” or “clinical FOC”. He agrees that the fear may be so intense that it meets the criteria for a specific phobia. The fear is specific in some instances, only concerning the process of labour and childbirth, but in others it is coupled with various other anxiety problems. Furthermore, Wijma (2003) reports that it may be so disabling that it interferes with the woman’s academic or occupational functioning, with her social and domestic activities or with her relationships. Fear of childbirth may manifest as physical complaints, nightmares and difficulty in concentrating.

According to Wijma (2003), fear of childbirth can be experienced by women during their pregnancy, during the delivery and in the postpartum period. Their fear usually reaches a phobic level after they have become pregnant. Furthermore, the woman’s fear of childbirth often leads to a request for an elective cesarean section without any obvious medical reason (Saisto & Halmesmäki, 2003; Wijma, 2003). They may also request to be sterilised so that they can avoid a subsequent pregnancy – with the fear being parturition and not parenting. Various studies (e.g. Waldenström et al., 2006) have indicated that women with greater fear of childbirth antenatally are prone to more intense fear during the delivery and are more likely to report a negative birth experience. They may also suffer the most from it in the postpartum period, regardless of the type of delivery they had. Soderquist, Wijma, Thorbert, & Wijma (2009) found that antenatal women with pre-traumatic stress or women with severe fear of childbirth in late pregnancy were more likely to have depression and post-traumatic stress one month after childbirth.

Women with PTSD after childbirth often have had symptoms of PTSD prior to delivery. According to Hofberg (2003), women who suffer from tokophobia may be more vulnerable to PTSD and PPD. A large percentage of women who have had emergency caesarean sections or instrumental deliveries have PTSD after childbirth. Not all women will, however, develop PTSD after a problematic delivery.

2.4.7.5.7 Primiparity.

Birth order has frequently been suggested as a factor related to the development of PPD. In earlier studies Davidson, Yalom et al., and Jackson and Laymeyer, as cited in Kruckman and Smith (2006, section 4, paragraph 5), suggest that the birth of the first child brings about a unique stress as the woman adopts the role and identity of a mother. This was found to correlate more strongly with depression than the birth of the second or third child. Tamaki, Murata, and Okano (1997) also suggest a possible association between first childbirth and PPD.

Studies looking at the possible effect of pregnancy number on PPD are, however, controversial. Posner, Unterman, Williams, and Williams (1997) did not find an association between number of deliveries and PPD. Righetti-Veltema et al., (1998) also found no association indicating that high parity is associated with PPD. Munk-Olsen et al., (2006) found that a higher risk of postpartum mental disorders was evident among primiparous women for several months after childbirth. Bloch and Klein (2005) found that in the order of pregnancies, earlier ones entail a higher risk of PPD. They report that this result may reflect the possibility that women who develop PPD are less inclined to become pregnant again. Women with multiple pregnancies may therefore represent a group of women with a relatively lower vulnerability for PPD. Alternatively, Bloch and Klein (2005) suggest it is possible that the vulnerability to PPD diminishes with multiple deliveries due to a non-specific decrease in stress associated with the pregnancy and delivery, or for other yet unexplained reasons.

2.4.7.6 Psychosocial adjustments.

Halbreich (2005) cites that the abrupt psychosocial change from pregnancy to motherhood and its demands and stresses may be a risk factor in PPD. This factor is related to the environment at time of delivery and the immediate postpartum period. It is well documented that in cultures where the new mother is provided with a higher level of care and family support during the first month after childbirth, reported rates of PPD are low and may be delayed until this period of pampering ends and the new mother is faced with the reality of day-to-day life (e.g. Harkness as cited in Bina, 2008).

The occurrence of symptoms and their perceived severity is also dependent on the individual's ability to cope with them. According to Halbreich (2005), these perception and coping mechanisms are shaped by past experiences as well as by the individual's central nervous system functioning. Kruckman and Smith (2006) point out that psychological functioning plays a major role in PPD.

It has been suggested that PPD is partly associated with the explicitness of role expectations for females and mothers (Tentoni & High, 1980). Kruckman and Smith (2006) report on some findings that identify role conflict as a psychosocial risk factor for emotional problems. They indicate that attitude towards pregnancy, especially ambivalence and sexual identity, are concepts related to role conflict, which may be associated with the development of psychological symptoms following childbirth.

2.4.7.7 Self-esteem.

There appears to be a strong relationship between depression and self-esteem. Mothers with low self-esteem are 39 times more likely to have depressive symptoms than mothers with high self-esteem (Hall, Kotch, Browne, & Rayens, 1996). A number of studies have found that self-esteem is related to depression after childbirth. Fontaine and Jones (1997) found a significant relationship with moderate depressive symptomatology at two weeks postpartum. Beck's (2001) meta-synthesis shows that, based on research in the 1990s, self-esteem has emerged not only as a new, significant predictor of PPD but also as one of the strongest predictors.

Srisaeng's (2004) study focused on the relationships between self-esteem and stressful life events with PPD in adolescent mothers in Thailand. When controlling for maternal characteristics, only self-esteem and negative stressful life events were significant predictors of PPD. Adolescent mothers who have been subjected to a high level of negative stressful life events and who have low self-esteem should be identified as they are at increased risk for PPD.

According to Hall et al. (1996) self-esteem, with its emphasis on feelings of self worth, buffers the negative effects of stressful life events. Mothers with high self-esteem are better able to withstand stressors that may impact on their sense of self-worth and be a factor in the development of PPD. Clinicians should still be wary though, even if a mother does possess a high level of self-esteem. Sichel and Driscoll (1999, p. 198) warn in their model of women's mental health, that the postpartum period "is a fragile time for the self-esteem of the ablest of women and is made much worse by the occurrence of a depression." Logsdon and Usui (2003) recommend that social support interventions for postpartum women should include assistance with building self-esteem, maintaining or improving relations with her partner, and providing support in areas that are important to her.

Researchers have explored how the weight retained after childbirth influences a mother's self-esteem, her weight satisfaction, and mood. Jenkin and Tiggemann (1997) conclude that a mother's postpartum weight determined her psychological well-being. Their findings show that the negative response to weight gain is not uncommon after childbirth. Women who gained more weight report depressive symptoms more often than women who gained less weight (Walker, 1997). Furthermore, mothers who reported having low self-esteem were found to have higher body mass indexes, greater weight gains, and more symptoms of depression. A study conducted by Carter, Baker, and Brownell (2000) found a strong association among BMI, eating attitudes, and depressive and anxiety symptoms during the postpartum period that are not present during pregnancy. Morgan, Lacey, and Chung (2006) investigated whether active bulimia nervosa affects obstetric outcome. They found that active bulimia nervosa during pregnancy was associated with postnatal depression, miscarriage, and preterm delivery.

2.4.7.8 Personality organization.

Personality has been associated with clinical depression. A number of researchers have explored personality in relation to PPD. Traits such as neuroticism have often been found to be associated with postnatal depression (Dennis et al., 2004; Matthey et al., 2000; O'Hara & Zekoski, 1988; Thio, 2004). O'Hara and Swain (1996) found that a negative cognitive attributional style was found to be related to PPD when assessed through self-report. Matthey et al. (2000) who examined the course of postnatal depression in first-time mothers and fathers with an emphasis on the role of personality as one possible major risk factor, further indicate the mother's level of interpersonal sensitivity is associated with depressed mood postpartum. Sved-Williams (2003) found associations between antenatal perfectionism and mood changes both during the antenatal and postnatal period. Highly self-critical women's risk for depression was lowered if they became strongly attached to the foetus during pregnancy (Priel & Besser, 1999). Boyce and Hickey (2005) confirm that psychosocial risk factors, predominantly in the areas of social support and personality style, are closely associated with postnatal depression.

Verkerk, Denollet, Van-Heck, Van-Son, and Pop, (2005) investigated introversion and neuroticism as predictors of PPD. They conclude that a person's personality traits may be constant and important determinants of PPD. Furthermore, their findings show that the risk estimates for clinical depression in the first year postpartum are considerably higher when both neuroticism and introversion scores are high.

Mazzeo et al. (2006) explored how perfectionism in women is related to antenatal and postpartum symptoms of depression and eating disorders. Their results propose that the particular aspect of perfectionism, namely "concern over mistakes", may contribute to the severity of PPD symptomatology.

The properties of antenatal screening instruments, developed specifically to determine a mother's risk of PPD, are described by Austin and Lumley (2003). They report that certain factors may have influenced poor sensitivity and positive predictive values of antenatal screening measures. The exclusion of key domains in predicting risk, particularly personality, is one such factor. They believe that the influence personality traits have may be under-estimated in studies where measures of risk prediction are evaluated.

The Vulnerability Personality Style Questionnaire (VPSQ) was developed to identify women at-risk for PPD due to personality vulnerability. Preliminary research with this 9-item self-report scale suggests it has satisfactory psychometric properties. Dennis and Boyce (2004) report that this measure will aid in the identification of women who are at-risk of developing PPD thereby allowing for appropriate secondary preventive or treatment interventions.

2.4.7.9 Infant temperament.

According to Beck (1995), a source of stress that contributes to the development of PPD can be difficult infant temperament. Beck's (1996b) meta-analysis investigated the relationship between infant temperament and PPD. In this study the confidence interval, which was calculated at 95%, ranged from 0.261 to 0.369, indicating a significant relationship between infant temperament and PPD. A subsequent updated meta-analysis conducted by Beck (2001) revealed that infant temperament was a significant predictor of PPD.

A relationship between PPD and infant temperament has been indicated in a number of studies (Austin, Hadzi-Pavlovic, Leader, Saint, & Parker, 2005; Aydin, Inandi, & Karabulut, 2005; Coplan, O'Neil, Arbeau, 2005; Edhborg, Seimyr, Lundh, Widstroem, 2000; Pesonen, Raikkonen, Strandberg, Kelitikangas, & Jarvenpaa, 2004; Whiffen & Gotlib, 1989). Murray, Stanley, Hooper, King, & Fiori-Cowley, (1996) found that high levels of irritability in infants were strongly predictive of the onset of maternal depression by 8 wks postpartum.

Whiffen and Gotlib (1989) state that a depressed postpartum woman's ability to mother effectively may be impaired by non-affective symptoms of depression, such as self-preoccupation, withdrawal and passivity, which may further contribute to an infant's difficult temperament. Irritable infants can make caretaking efforts largely ineffective and raise doubts in mothers' minds about their competence, resulting in feelings of inadequacy and depression. Rowe, Fisher and Feekery (2003) agree that difficult infant temperament has an effect on the quality of the relationship the mother has with her baby and also contributes to her diminished maternal confidence. Furthermore, high rates of co-incidental maternal psychological distress, particularly clinically significant anxiety and exhaustion, were found to be related to very high rates of difficult infant temperament. Sheinkopf et al (2006) furthermore report that a mother's psychological distress has an effect on the extent to which her baby's behavioural characteristics were experienced as difficult or stressful.

Maxted et al. (2005) performed a study on infant colic and maternal depression. Their sample included 93 consecutive patients seen at an outpatient clinic for colic, and results show that 45.2% of these mothers reported moderate to severe depressive symptoms. They report that factors like fussy or difficult infant temperament, lower parental self-esteem, more parenting stress, and more family-functioning problems were associated with more severe symptoms of depression in mothers whose infants suffered from colic. Howell, Mora, & Leventhal (2006) report that patients reporting depressive symptoms were more likely to have infants that suffered from colic. Akman et al (2006) also report that the mean EPDS score of mothers whose infants suffered from colic were significantly higher in comparison to mothers whose infants did not have colic.

Murray (2001) identified women at risk for PPD prior to the birth of their babies. She found that women were three times more likely to be depressed postpartum when their infant was difficult and had poorly organised motor behaviour – characteristically either jerky and strung up or else flat and sluggish. The influence that an infant's early behaviour has on the mother's mood was seen regardless of whether her perception of her baby's behaviour was difficult and whether or not she had postpartum blues. These factors did contribute to her risk of PPD, but her baby's behaviour added significantly to that risk.

An infant's early behaviour is important because it contributes to the risk of depression in the mother, which can cause relationship problems between mother and baby. Interventions during the postpartum period that focus on assisting mothers who have infants with difficult temperaments or colic, may prevent PPD. Conversely, postpartum psychological interventions aimed at minimising maternal depression and anxiety may optimise infant temperament outcomes and are likely to impact positively upon maternal perceptions of their infants, with implications for improving child behavioural development and health.

2.4.7.10 Sleep deprivation.

Infant sleep problems and PPD are highly prevalent in the postpartum period and both have adverse sequelae. It has been suggested that changes in sleep physiology and sleep deprivation plays a role in perinatal psychiatric disorders. Hiscock and Wake (2001; 2002) investigated the relationships between infant sleep problems and maternal well-being and found that there is a significant relationship between the two, even when known depression risk factors are taken into account. They report that both PPD and infant sleep problems can negatively impact, in apparently similar ways, on the infant, the mother, and the mother-infant relationship. Disrupted sleep in the infant can result in maternal sleep disruption, which in turn has an adverse effect on motor function, cognition, and mood. Lavigne et al., 1999) found that infants with sleep disruption are more likely to be irritable, inattentive, and tired, and find it more difficult to modulate their emotions and impulses. Similar problems along with poorer behavioural and cognitive outcomes and difficulty in forming attachments are seen in children of depressed mothers (Beck, 1998b; Murray, Hipwell, & Hooper, 1996).

Hiscock and Wake's (2002) results confirm that there is a strong association between maternal report of depression symptoms and infant sleep problems, even when already determined risk factors for PPD, like a past history of depression, are taken into account. According to Hiscock and Wake (2002), an important mediator in the relationship between infant sleep problems and depression may be maternal sleep quality, the reason being that mothers who reported good or very good sleep quality were less likely to report symptoms of depression, even when they regarded their infant's sleep to be problematic. Their findings suggest that there are other factors which contribute to good maternal sleep quality and protect mothers from depression if they have an infant with a sleep disturbance.

Mothers who reported sleep disturbance in their infants noted that they were significantly more likely to sleep in their parent's bed, would wake frequently and for typically for longer periods, and would need an adult to settle them back to sleep (Hiscock & Wake, 2002). According to Ferber (1995), these behaviours are typically learned behaviours and are therefore amenable to change through behaviour modification techniques. Hiscock and Wake (2002) found that night waking was related to a high EPDS score. They suggest that in order to decrease maternal report of depression symptoms, assistance should be offered to parents in teaching their infants to settle independently.

Parry et al. (2003) hypothesized that underlying chronobiological abnormalities may be associated with depression. They examined the relationship between endocrine measures and sleep in women with onset of a major depressive episode during their pregnancy or within the first year after childbirth. Their findings revealed that disruptions in the timing relationships of endocrine and sleep rhythms may play a role in antenatal and PPD.

Ross, Murray and Steiner (2006) provide a review about changes in antenatal and postpartum behaviour and sleep physiology. Their review particularly focuses on the association between sleep and postpartum "blues," depression and psychosis as well as on sleep-based interventions for the prevention and treatment of perinatal mood disorders. Their review suggests that there is a significant relationship between perinatal mood disorders and sleep. They recommend that studies employ objective measurement tools to measure both mood and sleep during the perinatal period in order to gain important information about the etiology, treatment, and prevention of perinatal mood disorders.

Mothers who report sleep problems in their infants are likely to be experiencing symptoms of depression and should be carefully monitored by their practitioners. Appropriate anticipatory guidance, which addresses problems with infant sleep patterns, has the potential to greatly reduce the number of maternal reports of depressive symptoms. It may also improve the infant's sleep and consequently have a positive impact on the well-being of the infant, the mother, and her family.

2.4.7.11 Lack of support.

According to Kruckman and Smith (2006) the relationships between social variables such as role conflict, stress and support have frequently been correlated. This indicates the likelihood that a more complex causal pattern is involved in the etiology of PPD than merely biologically-based theories can encompass.

O'Hara and Swain (1996) report that social support, as it is manifest during pregnancy, is a significant risk factor for the development of PPD – even more so when the mother has high levels of antenatal depressive symptomatology and lacks support from the baby's father. Morton (2000) observed a link between a mother's prenatal perceived lack of personal support and PPD. The meta-analyses of 44 studies by Beck (1996a), the meta-analyses of 84 studies by Beck (2001) and a meta-analysis (Robertson et al., 2004) that included subsequent studies of nearly 10 000 additional subjects reveal that a low level of social support is one of the strongest predictors of PPD.

Dennis and Ross (2006) found that women with postpartum depressive symptoms had significantly lower perceptions of postpartum-specific partner support. The significant relationship between social support and postpartum depressive symptomatology has been documented in numerous studies. Some researchers assert that measures of social support are the strongest predictors of postpartum outcome (e.g. Bennett & Indman, 2003; Boyce & Hickey, 2005; Dennis et al., 2004; Kruckman & Smith, 2006; Martinez-Schallmoser, Telleen, & MacMullen, 2003; Misri, Kostaras, Fox, & Kostaras, 2000; Nath, 2005; Pierce, Strauman, & Lowe-Vandell, 1999; Seguin, Potvin, St-Denis, & Loiselle, 1999). Forman et al (2000) determined that one in three women with perceived social isolation who suffer from psychological distress in late pregnancy will develop PPD. Logsdon and Usui (2003) state that social support as a predictor of PPD is the same across diverse samples of women.

In relation to support, studies have shown that early discharge from the hospital increases a mother's risk for developing PPD. This was found to be the case even when psychosocial, obstetric, and socio-demographic risk factors are controlled for (Dennis et al., 2004; Hickey, Boyce, Ellwood, & Morris-Yates, 1997). Hospitals that have an early postpartum discharge policy are likely, unless planned effectively, to leave the new mother at risk for emotional stress due to lack of social support.

Cooper et al. (1999) found that the pattern of socio-demographic variables associated with PPD in Khayelitsha, South Africa, was somewhat different from that found in Western samples. They report that in Western studies social adversity was a major risk factor for postnatal depression (Cooper & Murray, 1998). High levels of social adversity were endemic in Khayelitsha however, and it was, therefore, not possible to examine usefully the role of adversity. The only socio-demographic factor examined which they found related to maternal depression in Khayelitsha was the absence of support from the woman's partner.

It is essential that health care providers enquire about the presence of depressive symptoms and assess the mother's available social support during the antenatal period. Vulnerable mothers need to be identified and targeted for assistance so that they may receive additional support and assistance in dealing with everyday stressors.

2.4.7.12 Marital difficulties.

O'Hara and Swain (1996) examined the relationship between the mother's antenatal relationship with her spouse and PPD. Their findings show that a comparatively clear risk factor for PPD is the state of a woman's marital relationship during pregnancy. The meta-analyses conducted by Beck (1996a; 2001) also point out that marital satisfaction is a significant predictor of PPD.

Parents in a stable marital relationship are better able to adapt to the demands of marriage, family and parenthood. In contrast, a number of studies indicate that a poor marital relationship is a consistent psychosocial risk factor for the development of PPD (Alkar & Gencoz, 2005; Crockenberg & Leerkes, 2003; Ghubash & Abou-Saleh, 1997; Martinez-Schallmoser et al., 2003; Matthey et al, 2000; Merchant, Alfonso, & Mayberry 1995). Women with postpartum depressive symptoms are more likely to report conflict in their relationship with their partner (Dennis & Ross, 2006). Partner violence has been found to be significantly associated with PPD. Beydoun, Al-Sahab, Beydoun, & Tamim (2010) found that the odds of PPD were 60% greater among mothers who experienced physical or sexual abuse by their partners in comparison to mothers who had not. Fisher, Feekery, and Rowe-Murray (2002) found that the severity of PPD was associated most consistently with the quality of a woman's relationship with her partner and with her infant if classified as "difficult to settle".

Significant complications in both family and marital relationships may result from the presence of maternal depression (Larsen & O'Hara, 2002). Furthermore, existing depression may worsen after childbirth in a troubled environment (Robertson et al., 2004). Women with a history of mood disorders are more prone to experience a relapse after childbirth if they are dissatisfied with their partners. A lack of communication is the most common complaint among these women. Conversely, there is evidence that if a psychologically vulnerable woman is in relationship, within which she is appreciated by her partner, this appreciation may actually protect her from PPD (Marks, Wieck, Checkley, & Kumar, 1996).

Numerous women find handling both marital and maternal roles stressful. Researchers found that significant psychosocial stresses arose in postpartum marital adjustment when partners were not involved in child-rearing and were not supportive (Boyd-Bragadeste, 1998; Misri et al., 2000). In addition, in a normal postpartum marital adjustment the lack of support after the birth of a child acts as a source of psychosocial stress. Furthermore, how a woman perceives her partner's support influences her sense of well-being as a wife, a mother, and a woman (Misri et al., 2000).

The amount and type of support a partner gives is an important factor in the treatment of PPD as it has a significant positive effect on women experiencing PPD. Husbands or partners should be routinely included in women's visits with both primary care physicians and psychiatrists.

2.4.7.13 Single parenthood.

Warner et al. (1996) and Wickberg and Hwang (1997) found a significantly increased risk of PPD among single women. According to Kruckman and Smith, (2006) the lack of a natural support system and marital intimacy that a marital relationship may provide has been correlated with mental health problems.

Carter, Garrity-Rokous, Chazan-Cohen, Little, and Briggs-Gowan (2001) report that when maternal depression is combined with single parenting, the risk to the parent–infant system may be amplified and developmental progress disrupted. Lane et al. (1997) found that amongst the factors associated with high EPDS scores were single status.

2.4.7.14 Adolescent age.

Research reveals that adolescent pregnancy is associated with PPD. Warner et al. (1996) found a significant association between PPD and a younger age.

Lesser, Koniak-Griffin, and Anderson (1999) examined depressed adolescent mothers' perceptions of their own maternal role. Many adolescent mothers in their study had engaged in impulsive high risk activities prior to their pregnancies. Their findings propose that the experience of motherhood may help some adolescent mothers improve their previously self-destructive lives. Furthermore, establishing maternal identity and simultaneously developing a sense of maternal protectiveness led to realistic, future-oriented decision making. Some adolescent mothers, however, who experienced chronic depressive mood along with social isolation after childbirth were found to be at increased risk for developing problematic maternal behaviours.

According to Srisaeng (2004), in Thailand, premarital relations and pregnancy out of marriage are considered dishonourable and bring great shame upon the family. This research indicated that PPD was common among Thai adolescent mothers. This risk was increased if they had low self-esteem and had experienced high negative stressful life events.

Adolescent mothers experience distinct social and personal challenges that may determine their postpartum functioning. Research indicates that maternal competence, social isolation, and shape or weight concerns contribute to the unique variance that may predict their depression level (Birkeland, Thompson, & Phares, 2005).

In a study of 1662 participants, Rich-Edwards et al. (2006) investigated whether age was a factor that may be related to antenatal and postpartum depressive symptoms. They concluded that young maternal age was related to an increased risk of antenatal and postpartum depressive symptoms. Unwanted pregnancy, lack of partner support, and financial hardship largely contributed to the risk in this age group.

2.4.7.15 Unplanned pregnancy, ambivalence about having a child.

An unplanned pregnancy has been shown to be associated with PPD, which in turn, may lead to difficulty in adjusting to parenthood and feelings of entrapment. Furthermore, an unplanned pregnancy may result in ambivalence towards the child antenatally or lack of commitment to the infant (Chee et al., 2005; Warner et al., 1996). Warner et al. (1996) suggest that reducing unwanted pregnancies and, perhaps, better opportunities to return to employment postnatally would have a substantial effect on the rate of postnatal depression.

Lane et al. (1997) report that amongst the factors associated with high EPDS scores were unplanned pregnancy. Ghubash and Abou-Saleh (1997) studied postpartum psychiatric illness in an Arab culture. Their study identified four major risk factors for PPD, of which unplanned pregnancy was one. Rich-Edwards et al. (2006) investigated socio-demographic risk factors for antenatal and postpartum depressive symptoms among women. They conclude that unwanted pregnancy and financial hardship are significant factors associated with antenatal and postpartum depressive symptoms.

Beck's meta-analysis (2001) revealed that an unplanned or unwanted pregnancy was found to be another new predictor of PPD. She indicates that even if unplanned pregnancies were a welcome surprise, the mothers still had to cope with the ramifications of this unplanned event that would impact on the rest of their lives.

Cooper et al. (1999) researched PPD and the mother-infant relationship in a South African peri-urban settlement. They found that depression, whether with antenatal or postpartum onset, was strongly related to whether or not the pregnancy was planned. An unplanned pregnancy was reported by 69% of mothers with PPD. In this study, an unplanned pregnancy was strongly related to the woman reporting that the pregnancy was also unwanted.

2.4.7.16 Maternal or paternal unemployment or poverty.

Epidemiological evidence indicates a high rate of depression in women in conditions of socio-economic hardship (Harpham, 1994). It is likely that a similarly high rate would be found in puerperal samples.

Unemployment in both the mother and in the head of the household is a significant risk factor for PPD. The association between maternal unemployment and PPD is thought to reflect the isolation and low self-esteem of unemployed mothers, or the substantial role change for women who were previously employed but who, following childbirth, have no future employment planned. Alternatively, women who are vulnerable to depression may not seek work during the postpartum period (Warner et al., 1996).

Several other researchers have also reported that financial stress and socio-economic status (SES) is a significant risk factor for the development of PPD (Beck, 2001; Dearing, Taylor, & McCartney, 2004; Jardri et al., 2006; O'Hara and Swain, 1996; Patel, Rodriquez, & DeSouza, 2002; Rich-Edwards et al., 2006; Rubertsson, Waldenstrom, Wickberg, Radestad, & Hildingsson, 2005; Rubertsson, Wickberg, Gustavsson, & Radestad, 2005; Segre, O'Hara, Arndt, & Stuart, 2007; Sherman-Slate, 2005). Beck (2001) states that women at risk for PPD may experience a number of stressors that often include financial difficulty related to their demographic status – a stressor exacerbated after childbirth due to the costs of childcare. Single mothers, in particular, with a low income may have fewer resources at their disposal to prepare for motherhood.

Logsdon, Birkimer, and Usui, (2000) found a high incidence of depression in their sample of low socio-economic status African American postpartum women. Cooper et al. (1999) found a PPD prevalence rate of 34.7% in Khayelitsha, a very poor peri-urban settlement near Cape Town. This is roughly three times the expected rate internationally. Halbreich and Karkun (2006) revealed a wide range of reported prevalence of PPD that ranges from almost 0% to almost 60%. They state that one of the factors resulting in variability in PPD that is reported may be as a result of differences in socio-economic environments, such as levels of social support, or perceived social support, poverty, stress, and nutrition. Some researchers have found higher rates of depressive disorders in selected ethnic minorities (e.g. Onozawa, Kumar, Adams, Dore, & Glover, 2003) and some have indicated that the interaction between ethnic status and income increases the risk for depression (Belle, 1990; Golding & Lipton, 1990). Assessing a woman's psychosocial history and presence of stressful life events in early pregnancy, her employment status, as well as her psychiatric history may help the practitioner determine her risk for recurrent or sustained antenatal and PPD.

2.4.7.17 Childcare stress.

Beck's (2001) meta-analysis of PPD to determine the magnitude of the association between PPD and various risk factors revealed that one of the 13 strongest predictors of PPD was childcare stress. Leung (2002) studied stress, social support, and PPD in the context of Chinese culture. Results show that antenatal depression, social support factors and stress factors – including global stress level and specific childcare stress level – were all significant association factors and predictors of PPD. Major themes that emerged in reported stress and support related to postpartum adjustment. They included, amongst others, childcare competence, adjustment to the new roles, baby related problems, childcare arrangement, and support and stress from helpers and from health care professionals.

In a subsequent study, Leung et al. (2005) identified correlations between demographic variables and PPD, and psychosocial variables and antenatal depression in Hong Kong Chinese women. One of the major three predictors in this group was childcare stress. Honey, Bennett and Morgan (2003) found that screening tools for PPD that included maternal reports of childcare stress, assisted in considerably increasing the predictive performance of the screening measures.

2.4.7.18 High stress levels and adverse life events.

Stress associated with life events such as marriage, family structure, housing, occupation, and geographic mobility, have long since been correlated with PPD (Heitler, O'Hara et al., Paykel, Sosa et al., Telles, as cited in Kruckman and Smith, 2006, section 3, paragraph 3). O'Hara and Swain (1996) examined the relationship between adverse life events and PPD. The findings from their meta-analysis indicate that stressful life events during pregnancy is a significant risk factor for later PPD. Subsequent studies have also identified stressful life events as a significant risk factor in PPD (e.g. Grazioli & Terry, 2000; Robertson et al., 2004; Seguin et al., 1999; Srisaeng, 2004).

Stressful life events were identified as a significant risk factor for PPD in various cultural groups. In a study by Leung et al. (2005) on a sample of Hong Kong Chinese postpartum women, one of the major three predictors of PPD was postnatal perceived stress. Ghubash and Abou-Saleh (1997) studied postpartum psychiatric illness in an Arab culture. They found that women with past psychological problems, previous and ongoing marital difficulties and other stresses, and who show early postpartum psychological symptoms, are highly vulnerable to PPD.

Kim and Buist (2005) report on lack of social support as a key risk factor in the development of PPD. Their findings reveal that the isolation experienced by Korean immigrants in Australia is likely to be a significant stress for new mothers from a Korean cultural background. Results by Dennis et al. (2004) indicate that immigration within the last 5 years was amongst the factors that predicted depressive symptomatology at one week after delivery. They recommend that recent immigrant status as a risk factor needs further examining.

According to Mason, Rice, and Records (2005), various life experiences, such as physical or sexual abuse, may impact on how a woman subjectively perceives the normal developmental processes of labour, delivery, and postpartum recovery. According to Records and Rice (2002), a woman's experience of labour is significantly affected by a history of abuse and has been shown to contribute to the development of PPD. The participants in these studies related how their prior abuse influenced their thoughts and views of their labour, delivery, and postpartum experiences and felt that their PPD originated from the combined recall of trauma events and the labour and delivery experience. As labour progressed, they developed a cognitive frame of reference in response to their former abuse experiences. In situations like these, the woman's perception of her labour and delivery experience serves as a trigger that kindles a posttraumatic stress response (Seng & Hassinger, 1998). Records and Rice (2002) state that this response placed them at risk for PPD. Mason et al. (2005) state that the perceived context of abuse combined with feeling overwhelmed and a sense of inadequacy and helplessness, may contribute to the emergence as well as the severity of PPD.

Faisal-Cury, Tedesco, Kahhale, Menezes, and Zugaib (2004) examined PPD and its relationship with life events and patterns for coping. They found no association between PPD and life events. They did, however, find that depressed puerperal women resort to inadequate coping strategies, such as distancing. They indicate that this pattern of coping may be an etiological factor of the PPD as well as a reaction to their difficult life environment.

Knowledge of the risk factors for PPD suggested in this study and in other investigations is important for appropriate assessment by health providers. Ideally, brief questionnaires ought to be used routinely in a variety of settings (e.g., clinic or home) to assess stressors, intimate relationship quality, levels of support, self-esteem, and depressive symptoms as indicators of risk for adverse mental health outcomes. Early identification of mothers with compromised mental health and prompt intervention are essential for the well-being of both mothers and infants.

It is critical that health practitioners recognize the significance of postpartum depressive symptoms and the potential negative ramifications for the mother, her children, and family as a whole. Certain activities that promote mental health in both the antenatal and postpartum periods should be encouraged for positive mental health outcomes. These activities should aim to help reduce chronic stress, boost self-esteem, and strengthen the relationship women have with their partners. Furthermore, continued investigation into the prevalence and degree of postpartum depressive symptomatology is essential (Affonso et al., 2000).

2.4.8 Consequences of postpartum depression.

A depressed antenatal or postpartum woman is often plagued with guilt and anxiety. Her appetite and sleep are affected and she may feel that she is not able to care for her baby adequately. PPD robs a mother of the joy of new motherhood. The insidious aspect of postpartum psychiatric illness is that it will – eventually if not immediately – encompass the entire family.

PPD has been associated with poor maternal functional outcomes, such as substance abuse, loss of employment, suicidal behaviour and death by suicide. Adverse effect has also been reported in terms of low self-esteem, marital relationship and partner's mood state. Apart from the usual symptoms of depression, the mother may have obsessive thoughts about harm that could come to her child. She may also have intrusive thoughts concerning hurting herself or her child.

Posptartum depression has been associated with varied aspects of child outcome, even when current adverse circumstances were taken into account. These included the child's physical health, cognitive development, the mother-infant relationship, emotional development and social competence.

The postpartum period is, normally, a time of readjustment in a marriage with renegotiation of roles. A strong supportive marital relationship tends to "survive" the baby and come out stronger. A couple who experienced problems in their marriage prior to the birth of their baby are, however, at a greater risk of marital breakdown (Roan, 1997). Increased financial pressure, a marital partner feeling left out by the spouse's infatuation with the baby, and especially failure to renegotiate household and parenting responsibility with resulting unequal sharing of tasks are amongst the most likely causes of marital conflict (Roan, 1997; Larsen & O'Hara, 2002).

Studies show that marital stress is a major consequence of PPD (Burke, 2003; Larsen & O'Hara, 2002). Pregnancy and the postpartum period often coincides with the onset or increase in marital discord or domestic violence, all of which can have a deleterious effect on children (Burke, 2003).

Higher rates of depressive disorder have been found in men whose wives or partners have PPD (Bielawska-Batorowicz & Kossakowska-Petrycka, 2006; Burke, 2003; Cox, 2005; Davey, Dziurawiec, & O'Brien-Malone, 2006; Goodman, 2005; Pinheiro et al., 2006; Roberts, Bushnell, Collings, & Purdie, 2006; Schumacher, Zubaran, & White, 2008). Depressed mood in men during the postpartum period was correlated with poor family economic situation, low social support, poor marital relationship and antenatal expectations about what life with an infant would be like (Bielawska-Batorowicz & Kossakowska-Petrycka, 2006).

Some researchers postulate that mental illness in a mother may lead to a more active parenting role of the father in order to buffer the deficit in the mother-infant relationship (e.g. Albertsson-Karlgren, Graff, & Nettelbladt, 2001). Goodman (2005), however, offers evidence that fathers do not provide a buffering effect when a mother is depressed. Instead, it is suggested that depression in the mother has a negative outcome on father-infant interaction, which may increase possible risk to child development. When a new mother is severely depressed there is a much greater likelihood that her partner will develop depression too, which may have emotional and behavioural implications for the child (Schumacher et al., 2008). In light of the above, a family-focused approach to the assessment and treatment of PPD is needed, which includes the assessment of fathers for mood disorders in the postpartum, especially when their partner is depressed.

Responsive maternal contact as well as a healthy environment is important for the infant's normal and healthy development (Pound, 2006). The postpartum period is, however, a sensitive time due to the presence and demands of the developing infant. The care provided by a mother to her infant during this period may be compromised if she is suffering from postnatal depression or postpartum psychosis. Recent literature has found a link between maternal PPD and decreased parental participation in activities that promote the infant's development (McLearn, Minkovitz, Strobino, Marks, & Hou, 2006). Furthermore, findings from research done in developing countries suggest that poor physical health and malnutrition in infants is related to poor maternal mental health (Rahman, Iqbal, & Harrington, 2003).

Some mothers with PPD have obsessive or intrusive thoughts about harming their babies or themselves (Barr, 2003; Jennings et al., 1999; Kruckman & Smith, 2006). Chandra, Vankatasubramanian and Thomas (2002) report that the presence of depression with psychotic symptoms predicted infanticidal ideas and behaviour, especially where the psychotic ideas were directed towards the infant.

To a large extent mothers establish their infants' social environment and mediate their experiences of the external world. A mother's role is largely to provide a secure base from which a young infant or child can begin to explore the outside world. Compared to nondepressed women, depressed mothers' interactions were found to be impaired, and they expressed behaviour that had a negative impact on their children (Hart, Field, & Nearing, 1998; Weinberg & Tronick, 1998; Wolf, De Andraca, & Lozoff, 2002). Interaction was both less contingent and less affectively attuned to the infant's behaviour (Milgrom & Westley, 2003; Stanley, Murray, & Stein, 2004).

The unresponsive or rejecting care associated with PPD may have an acutely negative impact on young infants who are especially vulnerable during this critical imprinting period, and who are particularly dependent on their caregivers (Campbell & Cohn, 1991). An infant or young child whose needs have been rebuffed or neglected by a depressed and withdrawn mother will be generally less willing and able to interact with the environment (Jacobsen, 1999; Leiferman, 2002).

Depression in mothers was first in the list entitled, "Most significant mental health issues impeding children's readiness for school" set out by the Mental Health Policy Panel for the Department of Health Services in 2002 (as cited in Bennett and Indman, 2003, Consequences of untreated mood disorders, para. 1). Disruption in the early mother-infant interaction has a significant impact in later cognitive and behavioural problems in children of depressed mothers (Beck, 1995; Cornish et al., 2005; Grace, Evindar, & Stewart, 2003; Kurstjens & Wolke, 2001; Milgrom & Westley, 2003; Pound, 2006). Galler et al. (2004) confirmed these findings even when background variables such as less maternal education, young maternal age at the time of her first pregnancy, fewer home conveniences and more children in the home were controlled for. Murray (2001) found that depression had no long-lasting damaging effects on the child's cognitive development where the child was in a non-deprived family environment. Murray (2001) reports good environmental circumstances may help buffer any negative impact.

The first few months postpartum are a highly sensitive period for the development of a relationship between the mother and her infant. There is a significant risk of insecure attachment by the infant if the mother has experienced depression during this time (Moehler, Brunner, Wiebel, Reck, & Resch, 2006; Murray, 2001). Bonding disorders in the mother-infant interaction include irritability, aggressive and hostile impulses, lack of maternal emotion, pathological thoughts, and outright rejection. Impaired bonding is not uncommon in mothers who are referred for psychiatric help, and is present in 29% of mothers diagnosed with PPD (Brockington et al., 2001). Sagami, Kayama and Senoo (2004) found that aggressive parenting behaviour was strongly related to PPD. In severe instances, it can lead to child abuse or neglect.

Maternal depressive symptoms were also found to be related to low social competence and low adaptive functioning in children (Luoma et al., 2001; Murray, Sinclair, Cooper, Ducournau, & Turner, 1999; Zapata, 2005). Milgrom and Westley (2003) and Josefsson (2003) report increased temperamental and behavioural difficulties in children of depressed mothers. Kestler (2006) confirmed these findings and found evidence that maternal depression is related to infant stress regulation with greater increases in cortisol level. Murray (2001) found a correlation between behavioural problems and PPD even when parental conflict and a recent depressive episode in the mother were considered. Beck (1998b) reported a small yet significant effect on children's emotional and cognitive development, which appeared to weaken as the infants grew older.

The particular circumstance in which adverse effects are related to depression has been explored. Factors such as the nature, duration and severity of depression as well as the context in which it occurs with respect to other risk and protective factors (e.g., socio-economic status) have been suggested as moderators of the effects of PPD on infant outcomes (Brennan et al., 2000; Essex, Klein, Miech, & Smider, 2001; Murray, Fiori-Cowley, Hooper, & Cooper, 1996). Murray (2001) points out that depression occurring during the early months resulted in a higher rate of significant delays in mental development. Grace et al. (2003) assert that chronic or recurrent maternal depression, rather than solely PPD, are probably related to child outcomes. Kurstjens and Wolke (2001) also found significant interactions where maternal depression was major and had early-onset with repeated episodes. Zapata (2005), who examined the association between maternal depression during the first three years postpartum and child social competence and problem behaviours at first grade level, found that exposure to non-maternal care by 24 months buffered the negative impact of chronic depression.

Puckering (2005) points out that children of depressed mothers have needs which are often overlooked by mental health services and recommends that steps be taken to protect the development of these children. Children of mothers with long-term or chronic depression should be observed for learning and behaviour problems, as well as affective disorders, especially children who come from a deprived family environment who have only had maternal care.

2.5 Conclusion

This chapter provided an overview of perinatal mood disorders with a more detailed description of PPD in particular. The symptoms, prevalence and clinical course of PPD were discussed and perspectives on the etiology of PPD were addressed. Risk factors for PPD were discussed at length as this study examined the presence of these factors in this South African sample. Early detection and treatment of PPD is crucial considering that numerous women are affected by perinatal mood disorders and suffer from its negative impact on themselves as mothers, their infants and their families. The following chapter provides an overview of screening measures used to screen for PPD with particular focus on the Postpartum Depression Screening Scale.

CHAPTER 3

SCREENING FOR POSTPARTUM DEPRESSION

3.1 Chapter Preview

The relationship between mental illness and childbirth has been illustrated by medical professionals for centuries. Women have a greater risk of developing a severe mood disorder after childbirth. Their risk of being admitted to a psychiatric hospital in the first month after delivery is much greater than at any other time in life (Kendell, Chalmers & Platz, and Paffenbarger as cited in Stein, 2007, p. 637). PPD is a major health issue which affects, on average, 13% of childbearing women world-wide, regardless of their cultural background (O'Hara & Swain, 1996).

Postpartum depression is frequently undetected and under diagnosed by health practitioners. This is particularly true in developing countries where mental health in general is typically ignored (Reichenheim & Harpham, 1991). Studies have indicated that up to 80% of women who developed PPD do not report their symptoms to their physicians and as a result are not diagnosed – despite an increase in the awareness of the impact that depression has on mothers, children, and families (Kelly et al., 2001; Whitton et al., 1996; Yonkers et al., 2001). This is of great concern because the consequences of PPD can have severe implications for the family's welfare as well as the child's psychological development.

Missed diagnosis has been found to be frequent in situations which lack structured methods for evaluating mental health status (Evins et al., 2000; Goldsmith, 2007; Reid et al., 1998). Many general practitioners have come to realise that PPD is a serious, identifiable, and treatable illness, yet, screening for PPD isn't always done, and if it is, the use of a screening tool specifically designed for screening for PPD is uncommon (Seehusen et al., 2005). Many general practitioners simply enquire casually about a new mother's mental status, or are of the opinion that screening takes too much effort, which accounts for one major reason why PPD is under diagnosed (Kumar & Robson, 1984; O'Hara, 1995, Seehusen et al., 2005).

There are also reasons why some mothers who have developed PPD do not disclose their symptoms. One reason is that some mothers harbour guilt about feeling depressed after giving birth when society seems to expect it to be a time of joy (O'Hara, 1995). Another reason is the stigma surrounding mental illness which is still prevalent among some people (Keshen & MacDonald, 2004). Some women are embarrassed to complain to their doctors about certain physiological symptoms, like insomnia, as they expect that it is normal to experience these in the months following childbirth (Epperson, 1999). Apart from PPD going undetected frequently, the percentage of women who refer themselves for assistance with PPD has been found to be quite low (Murray et al., 2003).

Numerous women with PPD do not realise that they have the illness. A study by Whitton et al. (1996) found that, of women who had been diagnosed with PPD, over 90% of the women realized something was wrong, but less than 20% of the women reported their symptoms to a health care provider and only one-third of the women believed they had postpartum depression.

3.2 Screening for Postpartum Depression

The high rate of depression found amongst mothers of young children signifies a compounded public health problem, and highlights the necessity to improve detection, treatment, and prevention. PPD has the potential to severely affect the mother's health, the development and health of her infant, as well as the mother-infant relationship, and is therefore of concern to primary and mental health care professionals (Barr, 2008; Leiferman, 2002; Hobfoll, Ritter, & Lavin, 1995; Wickberg & Hwang, 1997; Wolf et al., 2002).

Mothers may contemplate harming themselves as well as harming their infants. PPD can also have devastating effects on the mother's partner, and influence their plans for future children. A survey done by Peindl, Zolnik, Wisner, and Hanusa (1995) indicated that 32% of the women in their study, who had experienced PPD changed their reproductive plans rather dramatically and made the decision not to have more children. Contributing factors were their fear that this mood disorder may recur, the cost of treatment, and the anguish their families experienced as a result of their depression. Greater marital dissatisfaction is evident in husbands whose wives are depressed in the postpartum period (Zelkowitz & Milet, 1996). Furthermore, the spouse or partner of a depressed mother has a higher rate of psychiatric disorders than the spouse or partner of a mother who is not depressed (Areias, Kumar, Barros, & Figueiredo, 1996).

Barr (2008) found that mothers with PPD experienced a delay in adapting to motherhood and termed the interaction they have with their infants "mechanical infant caring" which describes the manner in which mothers with PPD undertake infant care.

Studies have also shown that depressed mothers have a tendency to express behaviours which cause them to be less sensitively attuned to their babies (Murray, 1992; Cooper et al., 1999) and which have a negative impact on their children. These mothers may be disengaged, withdraw or be overly intrusive in their interaction with their children (Field, 1995; Hart et al., 1998; Weinberg & Tronick, 1998; Wolf et al., 2002). Children born to depressed mothers may have long term developmental problems as well as adverse behavioural, cognitive, and emotional outcomes due to poor mother–child interactions (Beck, 1998b; Cooper et al., 1999; Murray, Fiori-cowley et al., 1996). Poor maternal mental health has been associated with poor physical health and malnutrition in infants in developing countries (Rahman et al., 2003). Sleep problems in children has also been associated with maternal depression (Armstrong, O'Donnell, McCallum, & Dadds, 1998; Armstrong, van Haeringen, Dadds, & Cash, 1998).

The debilitating effect that PPD has on new mothers and the long term negative effects it has on child development may be decreased by the early identification of PPD and intervention during pregnancy and the early postpartum period (Canuso, 2008; Leiferman, 2002; Montgomery, 2001; Cooper et al., 1999). Delayed treatment due to late detection of the disorder may lead to a lengthening in the duration of the postpartum mood episode (England et al., 1994; Goldsmith, 2007).

Women typically have a reasonable amount of contact with health services during their pregnancy, labour, and the postpartum period. This is an ideal opportunity for health practitioners to provide information to mothers about PPD and to identify those mothers who seem to be experiencing symptoms of PPD for early intervention (Austin & Lumley, 2002). According to Walther (1997) "the four-to-six-week postpartum visit may be the ideal time to assess women for depression, and the first well baby appointment should not be a missed opportunity for assessment as well" (p. 107). A recent study by Sheeder, Kabir, and Stafford (2009) to determine the prevalence and incidence of maternal depression in the first 6 months postpartum found that screening mothers at 2 months after childbirth detected most mothers who become depressed during the first 6 postpartum months.

Postpartum depression is treatable, but only when the mothers who suffer from it are identified. It is imperative that new mothers are screened routinely for PPD so that those at high risk for PPD are identified. In primary care setting, training health care professionals to identify those mothers at risk and those who are experiencing symptoms of PPD, and to make appropriate referrals for psychosocial care and intervention may assist in reducing adverse outcomes (Austin, 2003; Austin & Priest, 2005).

The majority of health care providers are educated on postpartum mental illness and discuss the risks of postpartum mental illness with prospective parents. Formal questionnaires or depression scales, however, are not typically used (Goldsmith, 2007; Honikman, 2008). Furthermore, they tend to focus on mild emotional reactions as opposed to major mood and anxiety disorders. Researchers in the field of PPD emphasise the need for improved methods for identifying women who may be at risk of developing postpartum mental illness as well as more effective methods for the prevention, early intervention, and treatment thereof (Austin & Lumley, 2002; Buist et al., 2002; Canuso, 2008). In situations where the clinician's professional attention is typically directed mainly at the physical health of the mother and her infant, a screening questionnaire may be an effective method to detect depression in the mother. Nishizono-Maher et al. (2004)

examined the role of self-report screening questionnaires for PPD and conclude that utilising a questionnaire such as the EPDS has "certainly created a sense of openness about postnatal depression and postnatal psychiatric problems in general in community health centres" (p189).

Screening both high-and low-risk populations of women has been deemed necessary in order to minimize depressive symptoms and impairment associated with postpartum mental illness as well as on enhancing parenting efficacy (e.g. Austin & Priest, 2005; Carter et al., 2001). This may be achieved by implementing widespread screening for maternal depression. Baker and Oswalt (2008), Beck and Gable (2000, 2001c), Canuso (2008), Georgiopoulos, Bryan, Wollan, and Yawn (2001), Hanna et al (2004), and Milgrom et al (2011) are amongst the researchers who recognise the serious nature of PPD and emphasize the need for psychometrically sound postpartum screening instruments in an effort to improve detection of PPD in women during the first year following delivery.

3.3 Screening Measures

Psychosocial screening measures as well as assessment programmes may be grouped into two broad categories, namely a 'symptom-based' approach and a 'risk-based' approach. Certain centres use a combination of methods (Austin, 2003, Murray & Cox, 1990). Methods that are symptom-based methods rely on self-report measures that have been validated as suitable measures for screening for maternal distress symptoms in a variety of settings (Ross, Gilbert Evans, Sellers, & Romach, 2003). Risk-based methods involve asking patients about the presence of risk factors for PPD. This method has also been seen as a valuable strategy as some risk factors serve as strong predictors of a patient's susceptibility to PPD (Czarkowski, 1999; Llewellyn, Stowe, & Nemeroff, 1997; Misri, 2000).

The use of structured assessments and screening measures in postpartum primary care settings has led to an increase in the rate of detection of PPD in comparison to the use of unstructured clinical interviews (Evins et al., 2000; Goldsmith, 2007; Keshen & MacDonald, 2004). Goldsmith (2007) encourages the use of validated screening instruments by nursing practitioners in routine postpartum visits. The use of a validated screening tool as opposed to asking general questions about the mother's mood provides a standardized baseline against which the mother's future responses can be measured. Beck (2003) encourages neonatal care providers to familiarise themselves with the spectrum of postpartum mood disorders as well as reliable screening tools for PPD. Beck (2003) asserts that this "will aid [neonatal care providers] in both the anticipation of and routine universal screening for PPD" (p. 37).

The benefits of using screening measures include being able to identify women in need of mental health services, to detect depression in under-served populations, and to prevent mental health problems in mothers and their children (Boyd et al., 2002; Munoz, Le, & Ippen, 2000). Although self-report instruments are not able to provide a diagnosis for major depressive disorder, they have proven effective in identifying women in need of further evaluation as well as women who have a high risk for developing depression (Mu~noz et al., 2000). These mothers can be referred for appropriate treatment and counselling, and as a result the negative sequelae of PPD can be prevented (e.g., Chabrol et al., 2002).

A number of self-report measures have been used in the postpartum assessment of depressive symptomatology. Boyd, Le, and Somberg (2005) recommend the routine use of psychometrically sound and brief self-report instruments. Many of these are, however, general depression instruments, such as the Beck Depression Inventory (BDI) and the Inventory of Depressive Symptomatology (IDS). General depression measures may identify certain features of normal postpartum adjustment, such as fatigue and sleep disturbance, as pathological. The Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) is a self-report measure which was originally developed to assess for general depression severity and has been used frequently by researchers to screen for postpartum depression.

Three instruments which have been developed specifically to measure PPD symptoms are the Bromley Postnatal Depression Scale (BPDS; Stein & Van den Akker, 1992), the Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987), and the Postpartum Depression Screening Scale (PDSS; Beck & Gable, 2000, 2001b). Most studies of PPD have, however, used either the Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987) or a general depression scale, such as the BDI I and II (BDI; Beck et al., 1961). The PDSS was developed more recently in response to research that supported the need for a new screening instrument specific to postpartum depression. These measures are discussed in more detail below.

3.3.1 The Beck Depression Inventory (BDI and BDI-II).

The Beck Depression Inventory (BDI) and the BDI-II are general depression inventories. Both the BDI and BDI-II consist of 21 items with a 4-point Likert rating scale with scores ranging from 0 to 63 (Beck et al., 1961). The BDI-II is a revision of the BDI. Symptom content in the BDI-II was revised to correspond more closely to the diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 1994) for depressive disorders.

The BDI is a commonly used self-report measure in both research and clinical practice and has demonstrated its value in assisting with the identification of major depressive disorder as well as monitoring treatment for this disorder. The BDI has also been used extensively in PPD research. It does, however, rely on somatic symptoms of depression and as a result has been criticized for use with postpartum women. Da Silva Magalhães, Pinheiro, Horta, Pinheiro, and Da Silva (2008) addressed this issue when they examined the validity of the BDI in the postpartum period by comparing factor scores for both postpartum women and their partners. They found that women did not only have elevated scores on the somatic symptoms factor, but also had higher scores than their partners on the depression severity factor. The results of this study reinforce the validity of the BDI in the postpartum period due to a similar proportion of somatic symptoms and little factor variance between the mothers and their partners.

Jolley and Betrus (2007) caution against the conclusions derived from studies in postpartum samples which relied mostly on the BDI to assess for the presence of depression. An over represented depressive symptom score on the BDI may be due to the inclusion of symptoms such as fatigue and sleep disturbance (Troutman & Cutrona, 1990; Whiffen, 1988). Ugarriza (2000) stated that the BDI did not address symptoms such as anxiety, irritability, guilt, tearfulness, and feelings of being overwhelmed which are symptoms that are typically associated with postpartum depression. A number of researchers have reached the conclusion that the BDI may not be an adequate instrument for studies that screen for postpartum depression (Harris, Huckle, Thomas, Johns, & Fung, 1989; Huffman, Lamour, Bryan, & Pederson, 1990; Ugarriza, 2000; & Whiffen, 1988) and it has limited sensitivity when screening for minor depression (O'Hara et al., 1984).

The BDI has moderate correlations with instruments which screen specifically for anxiety, depression, postpartum depression, and general distress. The BDI seems to be similarly correlated with both depression and anxiety, which suggests that its items capture symptoms of both. Although the performance of the BDI-II with women during the postpartum period has not been extensively researched, the limited data has proven good concurrent validity with measures of postpartum depression, although it has been pointed out that some symptoms the BDI-II assesses can be problematic, such as general sleep disturbances and loss of energy (Beck & Gable, 2001a).

3.3.2 The Inventory of Depressive Symptomatology (IDS) and Quick Inventory of Depressive Symptomatology (QIDS).

The 30 item Inventory of Depressive Symptomatology (IDS; Rush et al., 1986; Rush, Gullion, Basco, Jarrett, & Trivedi, 1996) and the 16 item Quick Inventory of

Depressive Symptomatology (QIDS; Rush et al., 2003) are designed to measure the severity of depressive symptoms, including all the criterion symptom domains designated by the American Psychiatry Association Diagnostic and Statistical Manual of Mental Disorders - 4th edition (DSM-IV; APA, 1994) needed for the diagnosis of a major depressive episode. The QIDS as well as the IDS are available in the self-rated (QIDS-SR16 and IDS-SR30) and clinician versions (QIDS-C16 and IDS-C30). These measures may be used to screen for depression, but have predominantly been used for assessing symptom severity. An advantage of the IDS is that it provides a syndromal diagnosis of minor depressive disorder or major depressive disorder, in addition to assessing the severity of depressive symptoms. The usual time frame for assessing symptom severity is the seven day period prior to assessment. Questions are answered on a 4-point Likert scale.

The QIDS-SR16 is a shorter version of the IDS-SR30 and is more time-efficient for use in daily practice and in clinical research. It focuses only on the nine DSM-IV criterion symptom domains. The QIDS ratings were constructed by selecting only those items from the longer 30 item version that were needed to test for the nine DSM-IV criterion diagnostic symptom domains. The QIDS scoring system converts responses to the 16 separate items into the nine DSM-IV symptom criterion domains. The nine domains consist of 1) sad mood, 2) self criticism, 3) concentration, 4) interest, 5) suicidal ideation, 6) sleep disturbance (initial, middle, and late insomnia or hypersomnia), 7) energy or fatigue, 8) psychomotor agitation or retardation, and 9) decrease or increase in appetite or weight or both. The QIDS total score ranges from 0 to 27. The QIDS-SR16 does not include items which assess melancholic, atypical, or their commonly associated symptoms. The IDS, however, includes all of the QIDS items, as well as distinct mood quality, mood reactivity, diurnal mood variation, anxious mood, irritable mood, sexual interest, capacity for pleasure, bodily aches and pains, phobic or panic symptoms, leaden paralysis, digestive problems, and interpersonal rejection sensitivity (Rush et al., 1996). The IDS as well as the QIDS rate symptoms for the preceding 7 days, regardless whether the symptoms have been recent, chronic, or long-standing).

The IDS and the QIDS are useful for clinical and research purposes as both versions are sensitive to change, with psychotherapy, medications, or somatic treatments. The psychometric properties of the QIDS-SR16 and QIDS-C16, as well as the longer 30-item versions, have been established in various samples (Rush et al., 2003; Rush et al., 2005; Rush et al., 2006; Trivedi, Rush, Crismon, et al., 2004; Trivedi, Rush, Ibrahim, et al., 2004). Furthermore, Trivedi, Rush, Ibrahim et al (2004) reported that the total score of the QIDS-SR16 was highly correlated with the IDS-SR30 total score in 544 adult outpatients with major depressive disorder. When comparing the IDS-C30, IDS-SR30, QIDS-C16, QIDS-SR16, equal sensitivity to symptom change was found, indicating high concurrent validity for all four scales.

Both versions of the IDS have been used in postpartum depression, although the performance of the QIDS in postpartum depression is only recently being investigated. Yonkers et al (2001) demonstrated excellent sensitivity, good specificity and moderate PPV of the IDS in English and Spanish speaking postpartum women. Bernstein et al. (2008) examined the differences in the clinical features between postpartum and non-postpartum women using the QIDS-SR16. The two groups of women, who were matched on the basis of age, all met DSM-IV criteria for non-psychotic major depressive disorder.

The major characteristics of depression in both groups were low energy level and restlessness/agitation. The non-postpartum group reported higher levels of sad mood and reduced interest as well as more suicidal ideation. The postpartum depression group, on the other hand, reported that sad mood was less prominent, while decision-making and concentration were impaired, and psychomotor symptoms (restlessness/agitation) were prominent. The QIDS-SR16, which screens for these symptoms, can be considered a useful measure in the assessment of PPD. Questions that assess agitation and restlessness as well as decision-making and concentration ability should be included in screening measures for PPD due to the symptomatic differences between postpartum depression and other depression.

Yonkers et al. (2001) administered the IDS to Spanish and English speaking women during the postpartum period. Their results indicate that the IDS has good specificity (the proportion of women correctly identified as depressed), excellent sensitivity, and moderate positive predictive values (PPV), even when a 13% prevalence rate is assumed. Preliminary evidence of the IDS demonstrates promise of its validity with postpartum women, however further data is needed to establish its reliability to screen during the postpartum period.

3.3.3 The Bromley Postnatal Depression Scale (BPDS).

The Bromley Postnatal Depression Scale (BPDS; Stein & Van den Akker, 1992) was developed to assess both current and previous episodes of PPD. It is a 10-item questionnaire that includes open-ended and yes or no questions. Unlike other screening measures for PPD, the BPDS makes it possible for women to report their mood and behaviours, for all births, both during the antenatal and postpartum period, in order to explore the longitudinal course of PPD. It includes a chart which indicates when the current episode of postpartum depression started, how long it lasted, as well as when it was the worst. For this reason the BPDS has been considered unique.

According to Boyd et al. (2005) the BPDS does not have a recommended cut-off score. The determination of possible postpartum depression is made by examining the mother's self-report of the duration and severity of symptoms and seeking of assistance. Clinical training is therefore required in order to interpret the responses. Limited data is available on the psychometric properties of the BPDS. When Boyd et al. reviewed postpartum depression screening measures, only one published study of BPDS was found, which utilised a self-report measure to determine a DSM-III major depressive disorder diagnosis in calculating sensitivity and specificity. Self-report measures are not generally considered gold standard diagnostic instruments.

3.3.4 The Edinburgh Postnatal Depression Scale (EPDS).

The Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987) was developed to screen specifically for postpartum depression and is the most widely used screening questionnaire for PPD. The EPDS is a brief 10 item questionnaire which is scored from 0 to 3 according to the severity of the symptom experienced in the previous 7 days. A cut-off score for probable depression has been suggested at 12 or 13, and at 9 or 10 for possible depression (Cox et al., 1987).

The 10 items on the EPDS were derived from instruments that screen emotional well-being in the general population. The EPDS has been validated in a number of countries, including the UK (Cox et al., 1987; Murray & Carothers, 1990), Australia (Boyce, Stubbs, & Todd, 1993), and Canada (Zelkowitz & Milet, 1995). It has also been translated into many languages, including Spanish (Garcia-Esteve, Ascaso, Ojuel, & Navarro, 2003), Dutch (e.g., Berle, Aarre, Mykletun, Dahl, & Holsten, 2003) Chinese, Vietnamese (Barnett, Matthey, & Gyaneshwar, 1999), Italian (Benvenuti, Ferrara, Niccolai, Valoriani, & Cox, 1999), Portuguese, Finnish, Bengali (Fuggle, Glover, Khan, & Haydon, 2002), Swedish, Arabic, and Turkish (Aydin, Inandi, Yigit & Hodoglugil, 2004).

Eberhard-Gran, Eskild, Tambs, Opjordsmoen, and Samuelson (2001) carried out a systematic review of 18 validation studies of the EPDS which were published from 1987 to September 2000. They found the sensitivity estimates of the EPDS to be high in most cases. They also found, however, that a substantial proportion of mothers identified by the EPDS as depressed were false positive cases.

The review of the EPDS was updated by Boyd et al. (2005) who conducted a literature search of publications during October 2000 to December 2004. Their results show that the EPDS demonstrates moderate to good reliability properties across samples from a wide variety of countries and languages, with levels of reliability ranging from 0.73 to 0.87 (Boyd et al., 2005). Test-retest reliabilities fell within the good to moderate range, with the values decreasing as the period between administrations increased. Boyd et al. suggest, however, that different cut-off scores may be warranted for different cultural groups.

Gibson, McKenzie-McHarg, Shakespeare, Price & Gray (2009) suggest that the EPDS performs best when a higher cut-off point is used and for mothers who are comfortably able to express their distress in English. Gibson et al. (2009) performed a systematic review of validation studies of the EPDS to determine whether it compares favourably to a structured clinical interview for the detection of antepartum depression and postpartum depression across a variety of settings as well as in different languages of administration. Unfortunately the degree of heterogeneity amongst the studies did not enable them to perform a meta-analysis and to make statistical comparisons of the EPDS across different settings. They do, however, acknowledge that the utility of the EPDS rests in its free availability, how easily it is administered, and its general acceptability to women when given sympathetically.

The EPDS has been validated against the DSM-IV criteria for depression on a cohort of South African women from a low-income, socially disadvantaged urban community (Lawrie, Hofmeyr, De Jager & Berk, 1998). The sample was small, however, and the study had many limitations. Most participants had undergone a cesarean section and thus were a select group. The wording was changed in several items, although this reportedly did not affect the meaning of the scale. The EPDS was read to participants in this study to accommodate illiterate women as literacy rates among South African women differ considerably. Furthermore, due to the variety of languages spoken by South African women, the EPDS was translated by multilingual nursing sisters if necessary, which imposed certain limitations on the reliability of the data. Lawrie et al. (1998) did, however, conclude that the EPDS, administered verbally, is a valid screening instrument in this type of urban South African community. At a threshold of 11 or 12 the EPDS identified 100% of women with major depression and 70.6% of women minor depression. For major and minor depression combined sensitivity of the EPDS was 80%, specificity 76.6%, and positive predictive value 52.6%.

The EPDS has moderate to good correlations with other depression measures. Although it does not have a subscale for anxiety, the EPDS does screen for the presence of anxiety symptoms as well as depression symptoms (Brouwers, van Baar, & Pop, 2001). Rowe, Fisher and Loh (2008) point out, however, that the EPDS is not able to distinguish these conditions. Their study indicates that EPDS total scores were able to distinguish successfully between the categories of "neither diagnosis" and a diagnosis of "co-morbid major depression and anxiety" or "major depression alone". An "anxiety alone" diagnosis, however, could not be distinguished from "depression alone" nor from "neither diagnosis" on the basis of EPDS scores that were not significantly different from each other.

Pallant, Miller, and Tennant (2006) used Rasch analysis to determine whether the EPDS measures a unidimensional construct of depression or whether it measures two separate aspects – depressive feelings and anxiety – as has been suggested by other researchers (Brouwers et al., 2001, Ross, Gilbert Evans et al., 2003). Pallant et al (2006) did not find evidence to support the alternative structure separating depression items (items one, two, and eight) and anxiety items (items three, four, and five). Two sets of items were identified in the principle component analysis of residuals, but the Rasch logit-based person estimates derived from the subsets did not differ significantly from each other and thus supported a unidimensional construct of depression. Furthermore, results from their study question the viability of the original ten-item EPDS as a undimensional measurement of depression as it was found to “fall short of the rigorous standards of measurement defined by the Rasch model.” (p. 7). They suggest that the EPDS would be a more psychometrically robust scale if items seven and eight were removed.

The EPDS screens for cognitive as well as emotional symptoms of PPD. Apart from one item which measures sleep difficulty (as the postpartum recovery period rather than a mood disorder may impact on this) the EPDS deliberately excludes somatic symptoms of depression. The scale will not detect mothers with personality disorders, phobias, or anxiety neuroses. Muzik et al. (2002) found that new mothers with anxiety disorders scored significantly lower on the EPDS than mothers with a major depressive disorder – by an average of 5 points. They suggest that an alternative screening measure be used to identify mothers with postpartum anxiety symptoms. Beck and Gable (2000) point out that the EPDS does not measure the factor of ‘irritability’ – a factor they consider important in order to screen fully for PPD. Herz (as cited in Beck and Gable, 2000, p. 274) regards irritability to be an important component of PPD, and Beck and Gable (2000) agree that it should be included in a scale screening for this disorder. Furthermore, the EPDS does not contain any items written in the context of a woman’s experience as a new mother, such as ‘loss of control’, ‘loneliness’, ‘obsessive thinking’, and ‘irritability’. This has been another identified limitation of the EPDS – its items do not screen specifically for PPD, but are similar to those of a general depression instrument (Beck & Gable, 2000), and scores may be elevated by concurrent psychiatric illness, general emotional distress, or general medical conditions (Smith, Brunetto, & Yonkers, 2004).

Guedeney, Fermanian, Guelfi, and Kumar (2000) examined three cases of false negatives of major depressive episodes which were not identified as potential cases by the EPDS. Comparisons between the EPDS and two other self-report questionnaires, the GHQ-28 (General Health Questionnaire), and the CES-D (Center for Epidemiological Studies Depression Scale), indicate that the EPDS may be better at identifying depressed postpartum women with anxiety and anhedonic symptomatology rather than women with psychomotor retardation as the main symptom in depression.

Navarro et al. (2007) found that both the 12 item version of the General Health Questionnaire (GHQ-12) and the EPDS were useful brief screening tolls for assessing postpartum psychiatric morbidity. They found that both scales had good specificity and sensitivity when the diagnoses were broadened to include depression, anxiety, and adjustment disorders. Good concurrent validity (0.80) was indicated between both instruments.

A double-test strategy used by Lee et al. (2000), which involves the application of two complementary rating scales of symptoms and functioning, the EPDS and the GHQ-12, indicates that utilising more than one screening measure may assist in correctly identifying depressed women and also improve the overall cost-effectiveness of PPD screening programs. Lussier, David, Saucier, and Borgeat (1996) administered the EPDS and the BDI simultaneously to postpartum mothers and found that despite the two instruments claiming to measure the same phenomenon, were quite differently attuned to various facets of postpartum distress, and not equal in eliciting their expression:

> The two self-report instruments seem to tap into different dimensions centering on the presence or acknowledgment of different items or symptoms, which give a different phenomenological picture. Discrepancy occurs when one facet of depressive symptomatology clearly predominates, with the result that distress is picked up by one scale yet remains undetected by the other. (p. 87)

Lussier et al. (1996) discuss examples where the subject's symptomatology, if skewed in one direction, would result in divergent classifications. A woman feeling miserable and scared, for example, would most probably be identified by the EPDS but could be overlooked by the BDI if she did not feel guilty.

According to Lussier et al. (1996), the EPDS is better at reflecting affective upheavals, while the BDI is better at gauging cognitive and attitudinal dysfunction. From another viewpoint, the EPDS may come across as an acknowledgment of feeling and the BDI as an acknowledgment of incapacitation. The BDI tends to be oblivious of a more labile or anxious expression of distress, but seems more sensitive to a breakdown of coping mechanisms. The EPDS, on the other hand, may fail to adequately report a depressive constellation where the subject is "beyond weeping". They recommend that if detection of a range of disability is sought, that multiple assessment strategy is necessary until one instrument can be proven to achieve thoroughness of screening on its own.

3.3.5 The Postpartum Depression Screening Scale (PDSS).

The Postpartum Depression Screening Scale (PDSS; Beck & Gable, 2000, 2001b, 2002) is a 35-item self-report measure that was developed to assess the overall severity of postpartum depression symptoms. It is used to indicate whether the mother needs to be referred for further diagnostic evaluation, and can be used as a framework in therapy for developing a treatment program that targets the specific areas of distress and dysfunction. (Beck & Gable, 2002).

The PDSS assesses seven dimensions: Anxiety/Insecurity, Sleeping/Eating Disturbances, Cognitive Impairment, Emotional Lability, Guilt/Shame, Loss of Self, and Contemplating Harming Oneself. Each dimension consists of 5 items, giving a total of 35 items – each a statement describing how a mother may be feeling after the birth of her baby. The statements originated from actual quotes from women who had participated in the authors' research on PPD (Beck & Gable, 2000). This has resulted in an important characteristic of the PDSS – that it is able to identify the classic symptoms of PPD, such as irritability and anxiety, which are symptoms that are not typical of depression outside the postpartum period. Furthermore, the PDSS allows for the feeling of being overwhelmed and for fatigue, which are universal after childbirth, but do not necessarily indicate PPD. Women are asked to indicate their degree of disagreement or agreement with each statement according to how they have felt during the past two weeks. They indicate their responses on a Likert-type scale with a response format varying from strongly agree (1) to strongly disagree (5).

The PDSS is not appropriate for use in the first two weeks postpartum as it may yield a false-positive screen for PPD. This early postpartum period is commonly associated with mood swings and symptoms of postpartum blues, which are transitory and are a separate clinical phenomenon from PPD.

All items on the PDSS are negatively worded. Agreement with an item thereby indicates that the mother's mood concurs with the psychologically distressing symptom. A higher score on the PDSS indicates higher levels of PPD symptomatology. Lower PDSS scores indicate that the mother experiences fewer symptoms and suggests that her postpartum adjustment is relatively normal. The PDSS has an Inconsistent Responding (INC) index which provides an indication of response validity.

The PDSS is presently readily available in Spanish and English (Beck & Gable, 2003). In recent years it has also been translated into other languages including Chinese (Li, Liu, Zhang, Wang, & Chen, 2011), Thai (Vittayanont, Liabsuetrakul, & Pitanupong, 2006), and Portuguese (Cantilino et al., 2007).

According to Beck and Gable (2002), the PDSS should be easy to read and comprehend for anyone with at least third-grade reading skills. Information on the validity of the PDSS along with sensitivity values, specificity values, and PPVs when using the major depression cut-off score will be discussed at length during the course of this chapter.

3.4 Conceptual Basis of the PDSS

In 1992, C. T. Beck published a phenomenological study of the lived experience of PPD (Beck, 1992). C. T. Beck had conducted in-depth interviews with 12 women with PPD from a support group which Beck co facilitated. The interview focussed on how the women interact, how they regard their circumstances, and how these processes change.

From the transcribed interviews, C. T. Beck identified 45 significant statements concerning the mothers' experience of PPD and clustered them into 11 themes that described the essence of this experience: obsessive thoughts, contemplation of death, unbearable loneliness, loss of self, suffocating guilt, cognitive impairment, loss of previous interests and goals, loss of control of emotions, uncontrollable anxiety, insecurity, and loss of all positive emotions.

Despite the fact that PPD had received considerable research attention by 1993, little of it was qualitative in nature. That being the case, C. T. Beck believed that some aspects of the experience of PPD remained under explored. As well, because previous studies had never demonstrated an unequivocal link between PPD and the physiological changes associated with pregnancy and childbirth, there were undoubtedly other factors at play (e.g., psychosocial, environmental, etc).

PPD received considerable research attention but little of it was qualitative in nature. C. T. Beck's phenomenological study (Beck, 1992) aimed to explore the experience of PPD in greater depth. Furthermore, C. T. Beck opted for a qualitative approach to the topic because she believed that the Beck Depression Inventory (BDl; Beck et al., 1961) failed to accurately capture the real experiences of PPD that she saw in her clinical practice. Research evidence corroborated C. T. Beck's observations, rendering the content validity of the BDI for PPD questionable and in need for further investigation.

From this work, C. T. Beck (Beck, 1993) then developed a substantive theory of postpartum depression using grounded theory, and called it "Teetering on the Edge." The basic social psychological problem that emerged was loss of control. Mothers with PPD tried to cope with this problem using a four-stage process (Beck, 1993):

1. Encountering terror. This is the first stage of PPD. Mothers experienced relentless obsessive thinking, horrifying anxiety, and enveloping fogginess.
2. Dying of self. Isolation, alarming unreality, and thoughts or attempts at self-harm were experienced during this second stage.
3. Struggling to survive, the third stage of PPD, centred on the mothers' attempts to survive by battling the system, praying for relief, and turning to support groups for comfort and support.
4. Regaining control. In this final stage, regaining control, the mothers experienced unpredictable transitioning, mourned lost time, and went through a process of guarded recovery.

In 1996, Beck published the findings of a phenomenological study (Beck, 1996c) investigating the meaning of experiences which postpartum depressed mothers had when interacting with their infants and older children. In this study nine themes emerged, the essence of which were as follows:

- Postpartum depression overtaking mothers' bodies and minds, depriving them of feelings of joy, and preventing them from reaching out to their infants;
- Feeling overwhelmed by the responsibilities of taking care of their children and terrified of not being able to cope;
- Distancing themselves emotionally from their children to survive;

- Lack of desire to interact with their children, and at times, failing to respond to their infants' cues;
- Irrational thinking and guilt;
- Uncontrollable anger and fear of harming child;
- Perception that postpartum depression was causing their relationship with older children to deteriorate;
- Feelings of loss;
- Putting the needs of their children above their own in an effort to minimize the negative effects of PPD on their children.

Beck's qualitative research program on postpartum depression (Beck, 1992, 1993, 1996c) provided the conceptual basis for the development of the PDSS. The PDSS was designed so that its item content would reflect the phenomenology of new motherhood

3.5 Development of the PDSS

3.5.1 Generation of items.

The pilot form of the PDSS was composed of seven dimensions: anxiety/insecurity, sleeping/eating disturbances, cognitive impairment, emotional lability, guilt/shame, loss of self, and contemplating harming oneself. The 6 to 8 pilot items within each symptom dimension were written to reflect the content from the clinical interviews of C. T. Beck's qualitative research – each item a statement describing how a mother may feel after the birth of her baby. These items were then analysed to determine their content validity.

3.5.2 Item content validity.

The expert judgement method (Gable & Wolf as cited in Beck & Gable, 2000, p. 275) was used to ensure content validity for the pilot form of the PDSS. This method comprised two approaches: Firstly, a panel of five content experts reviewed the PDSS individually. Apart from their professional expertise in postpartum depression, four of the five experts had also personally experienced this mood disorder. Secondly, a focus group of 15 graduate students in nursing reviewed the PDSS. These graduate students' clinical specialties were either psychiatry or obstetrics.

The conceptual as well as the operational definitions of the seven symptom dimensions were assessed to determine the content validity of the PDSS. The content experts and focus group members were given the conceptual and operational definitions for each of the seven PDSS symptom dimensions. They were asked to judge how well each item fit the symptom dimensions to which it was assigned. The rating scale ranged from 1 (strongly disagree) to 5 (strongly agree). The mean ratings of fit for the pilot items ranged from 4.00 to 5.00 for the expert group and from 3.73 to 5.00 for the focus group members, suggesting that the judges found that the pilot items adequately described the symptom content of postpartum depression (Beck & Gable, 2000).

Editorial changes were then made, certain items were deleted, and some new items were added to the PDSS based on the reviews of the qualitative comments made by the expert panel and the focus group members. This process yielded a 56-item pilot version of the PDSS, with seven 8-item subscales representing the symptom dimensions.

This revised pilot version was given to 10 mothers within 8 weeks postpartum to review for further assessment of the clarity and readability of the items. No additional suggestions to improve the items were made. Psychometric testing of the PDSS pilot version then took place (Beck & Gable, 2000).

The reliability of the PDSS was assessed to determine which items could be deleted to create a briefer final version. This sample, the development sample, was also used to determine the reliability and validity of the final 35-item PDSS. The sample comprised 525 women who were between 2 weeks and 6 months postpartum, with a mean number of 6 weeks postpartum (Beck & Gable, 2000).

Subsequent research examined the construct validity of the PDSS along with its sensitivity, specificity, and predictive values (Beck & Gable, 2001c). The sample used in this study, the diagnostic sample, comprised 150 mothers within 12 weeks postpartum. The psychometric properties of the PDSS will be presented in the following section with data analyses from both the development and the diagnostic samples.

3.6 Psychometric Properties of the PDSS

3.6.1 Reliability.

An important aspect of reliability is internal consistency. This refers to the average intercorrelations among items in a test or subscale. Items designed to measure the same construct should be highly intercorrelated on a reliable test. The statistic used to measure internal consistency is Cronbach's coefficient alpha. According to Nunnally and Bernstein (as cited in Beck & Gable, 2002, p. 35), it is generally agreed that a measure of an emotional construct should have a minimum coefficient alpha of 0.70.

Analysis confirmed that the responses to the eight items assigned to each of the dimensions in the pilot version of the PDSS were internally consistent with coefficient alpha exceeding 0.75 for all scales. This made it feasible to delete items from each dimension based on the item content as well as the correlation for the respective items with the remaining items which define the dimension. It was made possible to delete three items from each dimension using this process. This allowed the length of the survey to be reduced to five items per dimension while still maintaining sufficient reliability levels and the targeted content coverage (Beck & Gable, 2000; 2002).

The dimension-level reliabilities range from 0.83 (anxiety/insecurity and sleeping/eating disturbances) to 0.94 (loss of self). For an affective instrument these reliability levels are considered high. All items have comparatively high correlations with their targeted dimensions (Beck & Gable, 2000; 2002).

Readability statistics were computed for the now 35-item final version of the PDSS. The Flesch Reading Ease score was 92.7, indicating that the scale requires a third-grade or better reading ability (Beck & Gable, 2002).

The data from the development sample was used to calculate the internal consistency estimates and item analyses were then calculated for the 35-item final version of the PDSS. These results are presented in Table 2. Excellent internal consistency for the final version is demonstrated, with an alpha coefficient of 0.97 for PDSS total score and coefficient ranging from 0.83 to 0.94 for the seven symptom content scales (Beck & Gable, 2000; 2002).

The reliability of the PDSS was further demonstrated in the diagnostic sample. Alpha estimates and item analyses for this sample appear in the columns on the right of Table 2. An alpha coefficient of 0.96 was computed for the PDSS total score and alphas ranged from 0.80 to 0.91 for the content scales (Beck & Gable, 2002).

Individual items on the final PDSS version have moderate to high correlations with their respective scales. Item 28 correlates only moderately ($r = 0.39$) with the Suicidal Thoughts scale. The reliability of this scale remains high though (alpha = 0.86) when Item 28 is kept in. Furthermore, the content of item 28 was judged by clinical experts to be a good fit with the operational definition of the scale. These considerations justified not deleting Item 28 from the scale, thereby maintaining the five-items per scale structure (Beck & Gable, 2002).

Table 2 Item Analysis and Internal Consistency Estimates by Standardization Sample for 35-Item PDSS

		Development Sample (N=525)			Diagnostic Sample (N=150)		
		Correlation with Content Scale	Content Scale Alpha if Item Deleted	Total Score / Content Scale Alpha	Correlation with Content Scale	Content Scale Alpha if Item Deleted	Total Score / Content Scale Alpha
PDSS Total Score				0.97			0.96
Sleeping/Eating Disturbances (SLP)				**0.83**			**0.85**
1	I had trouble sleeping even when my baby was asleep.	0.64	0.79		0.60	0.84	
8	I lost my appetite.	0.57	0.81		0.64	0.83	
15	I woke up on my own in the middle of the night and had trouble getting back to sleep.	0.61	0.80		0.66	0.82	
22	I tossed and turned for a long time at night trying to fall asleep.	0.67	0.78		0.78	0.79	
29	I knew I should eat but I could not.	0.63	0.79		0.63	0.83	
Anxiety/Insecurity (ANX)				**0.83**			**0.80**
2	I got anxious over even the littlest things that concerned my baby.	0.62	0.80		0.60	0.76	
9	I felt really overwhelmed.	0.61	0.80		0.64	0.75	
16	I felt like I was jumping out of my skin.	0.66	0.79		0.52	0.79	
23	I felt all alone.	0.65	0.79		0.64	0.75	
30	I felt like I had to keep moving or pacing.	0.61	0.80		0.55	0.78	
Emotional Lability (ELB)				**0.89**			**0.86**
3	I felt like my emotions were on a roller coaster.	0.75	0.86		0.68	0.83	
10	I was scared that I would never be happy again.	0.69	0.87		0.67	0.84	
17	I cried a lot for no real reason.	0.74	0.87		0.70	0.83	
24	I have been very irritable.	0.75	0.86		0.74	0.82	
31	I felt full of anger ready to explode.	0.72	0.87		0.64	0.84	
Mental Confusion (MNT)				**0.91**			**0.86**
4	I felt like I was losing my mind.	0.80	0.89		0.68	0.83	

		Development Sample (N=525)			Diagnostic Sample (N=150)		
		Correlation with Content Scale	Content Scale Alpha if Item Deleted	Total Score / Content Scale Alpha	Correlation with Content Scale	Content Scale Alpha if Item Deleted	Total Score / Content Scale Alpha
11	I could not concentrate on anything.	0.77	0.90		0.72	0.82	
18	I thought I was going crazy.	0.77	0.90		0.63	0.84	
25	I had a difficult time making even a simple decision.	0.78	0.90		0.69	0.83	
32	I had difficulty focusing on a task.	0.78	0.89		0.68	0.83	
Loss of Self (LOS)				**0.94**			**0.91**
5	I was afraid that I would never be my normal self again.	0.85	0.93		0.75	0.89	
12	I felt as though I had become a stranger to myself.	0.86	0.92		0.76	0.89	
19	I did not know who I was anymore.	0.81	0.93		0.78	0.88	
26	I felt like I was not normal.	0.85	0.92		0.80	0.88	
33	I did not feel real.	0.82	0.93		0.76	0.88	
Guilt/Shame (GLT)				**0.90**			**0.86**
6	I felt like I was not the mother I wanted to be.	0.79	0.86		0.79	0.81	
13	I felt like so many mothers were better than me.	0.77	0.87		0.74	0.82	
20	I felt guilty because I could not feel as much love for my baby as I should.	0.70	0.88		0.59	0.86	
27	I felt like I had to hide what I was thinking or feeling towards the baby.	0.71	0.88		0.56	0.86	
34	I felt like a failure as a mother.	0.77	0.87		0.76	0.82	
Suicidal Thoughts (SUI)				**0.93**			**0.86**
7	I have thought that death seemed like the only way out of this living nightmare.	0.88	0.90		0.85	0.80	
14	I started thinking that I would be better off dead.	0.82	0.91		0.71	0.82	
21	I wanted to hurt myself.	0.80	0.91		0.73	0.82	
28	I felt that my baby would be better off without me.	0.72	0.93		0.39	0.90	
35	I just wanted to leave this world.	0.85	0.90		0.82	0.79	

(Beck & Gable, 2002, p. 36-37).

3.6.2 Validity.

The validity of a psychological test can be defined as the test's ability to assess accurately those psychological characteristics that it purports to measure. There are several types of validity. Each type of validity has a different explanatory role in demonstrating the usefulness and accuracy of a test (Anastasi, 1988).

Content validity refers to whether the test item content adequately samples the behaviour that is being measured. Expert rater studies were performed where experts in postpartum depression rated the extent to which the PDSS pilot items correctly described the symptom content of postpartum depression (Beck & Gable, 2001b). Item content validity of the PDSS was addressed in more detail earlier in a description of the development of the measure.

Establishing construct validity is important for a measure like the PDSS. Construct validity addresses how well a test performs in measuring a theoretical psychological characteristic. The effectiveness of the PDSS depends on whether it can accurately capture and quantify the inner psychological states that constitute postpartum depression. Construct validity was assessed using confirmatory factor analysis and item response theory.

3.6.2.1 Confirmatory factor analysis.

The examination of construct validity was based empirically on the data obtained from actual respondents by means of confirmatory factor analysis. The results of the confirmatory factor analysis of the PDSS, listing the standardized weights for the five items assigned to each of the seven dimensions, are shown in Table 3. Each of the weights is sufficiently high with a minimum *t* value of 14.79 (Beck & Gable, 2000). This indicates that all of the items fit the hypothesized model. Goodness-of-fit indices were also calculated. The Tucker-Lewis index of 0.87 and the root mean-square residual of 0.05 were considered to be supportive of model fit. This information, as well as the evaluation of the modification indices, suggests that the construct validity of the proposed seven-factor solution could be supported for these data.

3.6.2.2 Item response theory.

Construct validity was also examined using item response theory techniques. Firstly, the adequacy of the definition for each dimension was empirically determined. Secondly, the "model fit" data was examined, concerning how well the 5-point Likert response format worked for these items and the respondents. The Facets program (Linacre as cited in Beck and Gable, 2000, p.276) was used to perform the one-parameter Rasch latent trait analysis. This allowed for further examination of construct validity concerning meaningful score interpretations.

Item response theory technique was deemed important as it addresses the adequacy with which the attitude continuum underlying each construct was assessed by the respective items – thereby contributing meaningful construct validity information. More complete score interpretation are made possible when the items which define the construct are spread across the respective attitude continuum (Beck & Gable, 2000).

Table 3 Confirmatory Factor Analysis: Maximum-Likelihood Dimensions and Loadings in the Development Sample (N = 525)

	Item	I	II	III	IV	V	VI	VII
Sleeping/Eating Disturbances (SLP)								
1	I had trouble sleeping even when my baby was asleep.	0.71						
8	I lost my appetite.	0.62						
15	I woke up on my own in the middle of the night and had trouble getting back to sleep.	0.72						
22	I tossed and turned for a long time at night trying to fall asleep.	0.78						
29	I knew I should eat but I could not.	0.67						
Anxiety/Insecurity (ANX)								
2	I got anxious over even the littlest things that concerned my baby.		0.68					
9	I felt really overwhelmed.		0.69					
16	I felt like I was jumping out of my skin.		0.73					
23	I felt all alone.		0.77					
30	I felt like I had to keep moving or pacing.		0.66					
Emotional Lability (ELB)								
3	I felt like my emotions were on a roller coaster.			0.80				
10	I was scared that I would never be happy again.			0.84				
17	I cried a lot for no real reason.			0.76				
24	I have been very irritable.			0.76				
31	I felt full of anger ready to explode.			0.74				
Mental Confusion (MNT)								
4	I felt like I was losing my mind.				0.84			
11	I could not concentrate on anything.				0.79			
18	I thought I was going crazy.				0.85			
25	I had a difficult time making even a simple decision.				0.83			
32	I had difficulty focusing on a task.				0.81			
Loss of Self (LOS)								

	Item	I	II	III	IV	V	VI	VII
5	I was afraid that I would never be my normal self again.					0.87		
12	I felt as though I had become a stranger to myself.					0.89		
19	I did not know who I was anymore.					0.85		
26	I felt like I was not normal.					0.90		
33	I did not feel real.					0.85		
Guilt/Shame (GLT)								
6	I felt like I was not the mother I wanted to be.						0.87	
13	I felt like so many mothers were better than me.						0.83	
20	I felt guilty because I could not feel as much love for my baby as I should.						0.72	
27	I felt like I had to hide what I was thinking or feeling towards the baby.						0.74	
34	I felt like a failure as a mother.						0.82	
Suicidal Thoughts (SUI)								
7	I have thought that death seemed like the only way out of this living nightmare.							0.92
14	I started thinking that I would be better off dead.							0.85
21	I wanted to hurt myself.							0.83
28	I felt that my baby would be better off without me.							0.75
35	I just wanted to leave this world.							0.91

(Beck & Gable, 2002, p. 40)

Examining the spread of the item scale values across the attitude continuum illustrated the differentiation of each of the seven attitude constructs. The item spread in each dimension was regarded as good for the types of items and participants in the study. Items which defined the anxiety/insecurity dimension were especially well spread across the attitude continuum, making it easier and more meaningful for the researchers to describe a person with both high and low scores on this dimension due to a greater comprehensive understanding of the construct on the basis of the content of the respective items (Beck & Gable, 2000).

The response options for the Likert categories of the PDSS (presented in Table 4 below) were examined to determine whether there was an "ordered attitude continuum" in which higher responses corresponded to higher levels of agreement.

The frequency and percentage of people selecting each option was examined and results show that the responses were spread adequately across all the options even though option 5 (strongly agree) was used less frequently for all dimensions. Results further indicated that higher response options on the 5-point category corresponded to higher levels of agreement with the items and more of the targeted dimension. This finding strongly supports the meaningful assessment of the attitude constructs. The 5-point Likert response categories was shown to contribute to the supportive construct validity findings, and were found to operate properly for these items and for participants.

Table 4 Postpartum Depression Screening Scale: Likert Response Category Fit Statistics

Dimension	Response Option	Frequency	Percent	Fit
Sleeping/eating disturbances	1	641	30	-1.26
	2	556	26	-0.73
	3	245	11	-0.30
	4	504	23	0.13
	5	207	10	0.89
Anxiety/insecurity	1	658	29	-1.73
	2	553	24	-0.82
	3	315	14	-0.14
	4	504	22	0.34
	5	266	12	1.32
Emotional lability	1	565	26	-1.81
	2	527	24	-0.95
	3	309	14	-0.13
	4	479	22	0.59
	5	297	14	1.57
Cognitive impairment	1	456	23	-2.03
	2	614	31	-1.18
	3	339	17	-0.17
	4	390	20	0.59
	5	188	9	1.93
Loss of self	1	426	24	-2.84
	2	591	33	-1.49
	3	290	16	-0.27
	4	317	18	0.97
	5	156	9	2.47
Guilt/shame	1	551	31	-2.10
	2	584	32	-1.14
	3	216	12	-0.27
	4	284	16	0.55
	5	163	9	1.44
Contemplating harming oneself	1	331	31	-2.72
	2	432	41	-1.38
	3	144	14	-0.26
	4	97	9	0.63
	5	56	5	1.23

Note: Fit is defined as the average logit scale score for people selecting the respective option.
(Beck & Gable, 2000, p. 281)

3.7 Comparative Analysis of the Performance of the PDSS with Other Depression Instruments

The PDSS demonstrates correlations in the good range with the BDI-II (r = 0.81) and the EPDS (r = 0.79). This indicates that all three instruments measure similar aspects of depression. A recent systematic review of the evidence suggests that the PDSS and the EPDS appeared to be more sensitive in screening for postpartum depression than the Beck Depression Inventory (Gaynes et al., 2005).

Beck and Gable (2001a) compared the performance of the PDSS with the EPDS and the BDI-II. The results are illustrated in Table 5. The PDSS demonstrated higher levels of sensitivity and specificity in the detection of PPD than the BDI-II or the EPDS. They found that, when using the published recommended cut-off scores, the specificity of the PDSS was 98% and the sensitivity was 94% for major depressive disorder.

When screening for both minor and major depressive disorder, the PDSS yielded the highest combination of specificity (72%) and sensitivity (91%). They also found that the PDSS identified a considerably higher percentage of women (94%) diagnosed with major depressive disorder, compared to the EPDS (78%) and the BDI (56%). When the PDSS screening performance was compared qualitatively to the EPDS, the PDSS appeared more sensitive than the EPDS for symptoms related to anxiety, sleep disturbance, and mental confusion.

Table 5 Sensitivity, Specificity, Positive and Negative Predictive Values of the PDSS, EPDS, and BDI-II

	Major Postpartum Depression			
Instrument/Cut-off Score	**Sensitivity (%)**	**Specificity (%)**	**Positive Predictive Value (%)**	**Negative Predictive Value (%)**
PDSS / 80	94	98	90	99
EPDS / 12	78	99	93	96
BDI-II / 20	56	100	100	93
	Major or Minor Postpartum Depression			
Instrument/Cut-off Score	**Sensitivity (%)**	**Specificity (%)**	**Positive Predictive Value (%)**	**Negative Predictive Value (%)**
PDSS / 60	91	72	59	95
EPDS / 9	59	86	64	82
BDI-II / 14	57	97	90	83

PDSS, Postpartum Depression Screening Scale; EPDS, Edinburgh Postnatal Depression Scale; BDI-II, Beck Depression Inventory-II
(Beck & Gable, 2001a).

The BDI's psychometric properties have established it as a robust instrument. Its use, however, as a preferred measure for postpartum depression is questionable. It's specificity for PPD in particular has criticized. Scores on the BDI may be inflated because normal postpartum somatic symptoms are similar to symptoms of depression, while mild depressive episodes may not be detected at all due to it being a measure of general depression (Affonso et al., 2000; Campbell & Cohn, 1991).

The PDSS, unlike the EPDS and the BDI, was based on the conceptual definition of PPD (Beck & Gable, 2001a):

> PPD is a mood disorder that can begin any time during the first year after delivery. Loss of control of emotions, thought processes, and actions is the basic problem of this experience. Symptoms may include a withdrawal of positive emotions, inability to concentrate, insecurity, loneliness, anxiety, difficulty sleeping and/or eating, guilt, and/or shame, obsessive thinking, emotional roller coaster, and contemplating harming oneself. (p. 243)

The PDSS is the only instrument out of these three depression instruments that contains items measuring all these cardinal symptoms (Beck & Gable, 2001a). When the content validity of the EPDS is compared with the PDSS, there are five symptoms derived from the themes in C. T. Beck's phenomenological study of postpartum depression that are not addressed by the EPDS (Table 6). These are loss of control, loss of self, obsessive thinking, cognitive impairment, and loneliness (Beck, 1992). The EPDS, therefore, does not entirely take into consideration irritability, anxiety, and other symptoms that are prevalent among postpartum women.

The PDSS was able to differentiate cognitive impairment and anxiety where neither the BDI nor the EPDS was able to detect them (Beck & Gable, 2001a; Clemmens, Driscoll, & Beck, 2004). Furthermore, the PDSS was more accurate in differentiating sleep disturbances than the BDI. The EPDS was unable to detect any sleep disturbances (Clemmens et al., 2004).

A shortcoming of the EPDS, according to Yonkers and Sampson (2000), is that it is influenced by concurrent psychiatric illness, general emotional distress, and general medical conditions. The EPDS is, according to Halbreich and Karkun (2006), an excellent measure for the purpose of detecting the dimension of depression for which it was developed. They recommend, however, that more culturally sensitive and flexible instruments are needed for the plausible array of postpartum disorders.

Table 6 Comparison of the Item Content of the PDSS' Seven Dimensions with the BDI-II and the EPDS

PDSS Dimension	BDI-II	EPDS
Sleeping	X	X
Eating disturbances	X	
Anxiety / insecurity		X
Emotional lability	X	X
Cognitive impairment	X	
Loss of self		
Guilt / shame	X	
Contemplating harming oneself	X	X

(Beck & Gable, 2001a)

Beck and Gable (2001a) discuss some possible sources for the lack of agreement among the three instruments used in their study. The time frame covered by each instrument varies. The PDSS specifies "over the past two weeks", the BDI-II states "during the past two weeks, including today", and the EPDS enquires how the respondent has felt "in the past 7 days, not just how you feel today".

Furthermore, the instruments differ in terms of the way the items are stated. The EPDS contains both positive and negative worded items, but the BDI-II and the PDSS do not. Recording the total score of items related to these opposite mood sates is questionable, according to Watson, Clark, & Tellegen (as cited in Beck & Gable, 2001a, p. 248). In a depression instrument, the presence of negative moods may differ from the absence of positive moods, and these mood states should be seen as independent (Condon & Corkindale, 1997). The use of both positive and negative item stems has long since not been viewed by instrument developers as good measurement practice (Gable & Wolf, as cited in Beck & Gable, 2001a, p. 248).

The number of items in a depression instrument also plays a role. If only one or two items are changed on an instrument consisting of only a small number of items, it can significantly alter a person's assignment to either the depressed or nondepressed category. Condon and Corkindale (1997) recommend that an instrument containing a larger number of items be used when screening for postpartum depression.

Depression instruments also typically focus on different components of this mood disorder. A mother may screen positive on one instrument, but negative on another when one component of depression predominates over another. Awareness of the differential sensitivity of the depression instrument and how the targeted depression dimension has been operationally defined is therefore important.

A study by Boyd et al. (2005) suggests that the target sample should also be considered when selecting a screening measure. They reviewed published literature on the psychometric properties of self-report depression instruments which were administered during the postpartum period. The screening measures they reviewed included the five screening measures discussed in this chapter, as well as the The Zung Self-Rating Depression Scale (Zung SDS), The General Health Questionnaire (GHQ), and The Center for Epidemiological Studies Depression Scale (CES-D). They make some recommendations about the use of these self-report instruments for various samples and suggest that the GHQ be considered for comorbid conditions in addition to PPD. The IDS seems promising for use with ethnically diverse and urban samples, the BPDS is useful for an assessment of previous history of PPD, and that the BDI-II or the PDSS may be warranted when screening highly educated, predominantly Caucasian samples. Their review also shows that the EPDS has been the most researched measure with moderate psychometric properties, and that the BDI-II and the PDSS appear to be promising screening measures.

3.8 Conclusion

This chapter provided an overview of the different screening measures that are available that assist in assessing, identifying and treating postpartum women who present with depression. It is crucial to detect and treat women with depression in the early stages of the illness, given that so many women suffer from perinatal mental illness, and also considering the morbidity it causes in the mother as well as in her infant. The PDSS was found to be a reliable screening scale. Internal consistencies for the PDSS are excellent on both the individual and the total dimensions. Validity information was found to be promising. The PDSS demonstrates excellent sensitivity and specificity values. Positive predictive values (PPV) were good when using the major depression cut-off score. The PDSS, which was based on the conceptual definition of PPD, seems better able to identify women who may have major depressive disorder as the PPV rates are superior for major depression when compared with screening for minor and major depression.

CHAPTER 4

CROSS-CULTURAL ASSESSMENT

4.1 Chapter Preview

Frequently cross-cultural research involves the application of instruments in various linguistic groups (Van de Vijver & Tanzer, 1997) and, in this study, an existing postpartum depression screening measure developed for the American culture, was adapted into Afrikaans. Therefore, certain issues regarding the compilation of instruments for use in cross-cultural research and cross-cultural application of tests needs to be addressed. This chapter will look specifically at cross-cultural assessment, factors influencing cross-cultural assessment, methodological considerations for cross-cultural assessment, ethical guidelines for adaptation of cross-cultural assessment measures, and translating assessment measures.

4.2 Cross-Cultural Assessment

Cross-cultural assessment is the evaluation of behaviour and attributes by obtaining measures of these under different cultural conditions and by comparing them in order to establish cross-cultural uniformities and differences (Van Ede, 1996). Irrespective of how large or small the cultural difference, cultural groups often share "a large part of their everyday life-worlds, a country, and also a common humanity" (Retief, 1988, p. 183). Knowledge of these uniformities can be used to develop a pan-human theory of human behaviour, while knowledge of differences makes us aware of variations caused by the influence of different cultural conditions (Van Ede, 1996).

Comparative studies across ethnic groups and cultures attempt to elucidate discrepancies among human beings and thus try to achieve a better understanding of human society and behaviour. Researchers in this field label their research as transcultural, cross-national, cross-cultural, or cross ethnic. A dilemma for these comparative studies is to compile instruments that do not discriminate against individuals. Certain individuals may not have been exposed, in their ethnic, cultural or subcultural group, to the issues required by the instrument. For instance, the Minnesota Multiphasic Personality Inventory contains various implicit references to the American culture and extensive adaptations would be required before it could be used in other languages and cultures (Lucio, Reyes-Lagunes, & Scott, 1994).

The comparison of people from different cultural groups has long since become an important part of behavioural science (Manaster & Havighurst, 1972). Baron and Byrne (1994) agree that efforts to understand social behaviour must take careful account of cultural factors. Attention to the effects of cultural factors is an increasingly important trend in modern social psychology. According to Anastasi and Urbina (1997) the problem associated with assessing people who have highly dissimilar cultural backgrounds was recognized in the United States as early as 1910 when large groups of immigrants had to be assessed. The issue of cross-cultural assessment has received increasing attention since the middle of the 20th century when assessment measures were needed in newly developing nations in Africa and elsewhere to decide on admission to educational facilities and for individual counselling.

4.2.1 Multicultural assessment in South Africa.

4.2.1.1 Instrument development versus translation and adaptation.

There is considerable evidence (e.g., Van Ede, 1996; Van Eeden & Prinsloo, 1997; Van de Vijver, 2002; Van de Vijver & Poortinga, 1997; Van de Vijver & Lonner, 1995) to suggest that interest in international comparative studies of cross-cultural research is growing. With this growth has come the need to adapt (or translate) psychological instruments for use in multiple cultures and languages. This is especially pertinent in a linguistically and culturally diverse country like South Africa.

Most measures available in South Africa were developed in the United States of America or the United Kingdom and tend to be more appropriate for westernized English-speaking people (Foxcroft et al., 2006). It would seem impractical and virtually impossible to develop one measuring instrument which would be appropriate for the entire South African population. There would also be many obstacles to overcome in developing a measure suitable for all South Africans, such as

- The measure would need to exhibit appropriate levels of semantic and conceptual equivalence across cultures and languages;
- The procedures through which it is administered must minimize any problems created by lack of normative equivalence;
- The use of a multicultural team approach is likely to be extremely costly and time consuming;

- It becomes virtually impossible to make cross-national comparisons unless the instrument is translated and adapted for all South Africa's population groups in order to make cross-national comparisons.

It seems more appropriate to maximise the use of available, internationally relevant measures as far as is possible across cultural groups, rather than to embark on a totally new screening measure. Selecting an internationally well-researched measure and adapting and translating it for local conditions is also more time and cost effective. Existing measures have the advantage of being accompanied by the attributes of familiarity, experience, and often a vast body of research data. Existing measures can serve as a baseline for modification of culturally loaded test items in the South African context, and the gradual development of localized norms. Despite the advantages of translating and adapting an existing measure, there are numerous methodological issues that need to be addressed, such as bias and equivalence, and whether the measure is culture fair.

4.2.1.2 Progression of psychological assessment in South Africa.

Psychological assessment in South Africa has followed international trends. Measures were imported from overseas from the early 1900's (Foxcroft as cited in Van de Vijver & Rothmann, 2004, p.2). Claassen points out that, initially, psychological measures were developed separately for the English and the Afrikaans-speaking populations (as cited in Van de Vijver & Rothmann, 2004, p. 2) and were only initiated with the White population (Huysamen, as cited in Van de Vijver & Rothmann, 2004, p. 2) – who were, and still are, a minority of the population group. Abrahams and Mauer argue that this discrimination meant that all population groups in South Africa were not adequately represented in the standardisation samples used to derive norm tables, and that the constructs being measured were different from those which the tests had been designed and standardised for (as cited in Van de Vijver & Rothmann, 2004, p. 2).

Biesheuvel explored the effects of potential bias problems associated with cross-cultural assessment in a South African context. He underlined the importance of schooling, home environment, and nutrition, as well as other factors on the cognitive performance on tests in a multicultural society (as cited in Van de Vijver & Rothmann, 2004, p. 2). The apartheid policy in South Africa resulted in a paucity of research on the bias and equivalence of assessment measures between 1960 and 1984 (Claassen and Owen as cited in Van de Vijver & Rothmann, 2004, p. 2). This changed in the 1980's, however, with a renewed interest in the comparison of cultural groups on various assessment measures in order to address issues of bias and equivalence. Since then concern has been expressed about the effectiveness and relevance of some assessment measures used in South Africa (Sibaya, Hlongwane, & Makunga, as cited in Van de Vijver & Rothmann, 2004, p. 2).

The first democratic elections in 1994 resulted in South Africa being regulated by a new constitution in which quality of individuals and basic human rights are guaranteed. This has also impacted on psychological assessment in South Africa and placed the cultural appropriateness of psychological tests and their usage in the spotlight. This led to

South Africa's new Employment Equity Act 55 of 1998, Section 8 (Government Gazette, 1998, as cited in Van de Vijver & Rothmann, 2004) which stipulates that

> Psychological testing and other similar assessments are prohibited unless the test or assessment being used – (a) has been scientifically shown to be valid and reliable, (b) can be applied fairly to all employees, and (c) is not biased against any employee or group. (p. 1)

The expectations and demands raised by this Act puts a great deal of pressure on psychologists to ensure that tests are fair and unbiased. This would be quite a feat in a country which is as linguistically and culturally diverse as South Africa is. A primary goal for assessment professionals in South Africa is, and ought to be, to bring current practice in line with legal demands. This requires the development of new instruments and the validation of existing instruments for use in multicultural groups. This Act may ultimately "enhance the professional level of psychological practice by putting multicultural assessment on the agenda of the profession and by stimulating the development of new tests and even new testing practices" (Van de Vijver & Rothmann, 2004, p. 1).

Research in South Africa which addresses bias and equivalence of assessment measures has become an increasingly explored topic. Van de Vijver and Rothmann (2004) state, however, that "much more research is needed on the equivalence and bias of assessment tools used in South Africa before psychology as a profession can live up to the demands implied in the Equity Act" (p. 2).

4.3 Culture-Fair Tests

In order to address multiculturalism, attempts were initially made to develop tests that were culture-free (Cattell as cited in Van de Vijver, 2002, p. 546; Foxcroft et al., 2006). Classic culture-free tests were developed to eliminate the influence of parameters such as reading, speed, and language. Previously, researchers believed that measures could be developed which were free from cultural influences and could be applied in all cultures and reflect comparable findings. Anastasi and Urbina (1997) and other writers (e.g., Manaster & Havighurst, 1972) maintain that it is useless to try to devise a test that will not be affected by cultural influences. Anastasi (1988) states

> We now recognize that hereditary and environmental factors interact at all stages in the organism's development and that their effects are inextricable intertwined in the resulting behavior. For man, culture permeates nearly all environmental contacts. Since all behavior is thus affected by the cultural milieu in which the individual is reared and since psychological tests are but samples of behavior, cultural influences will and should be reflected in test performance. It is therefore futile to devise a test that is free from cultural influences. (p. 345)

Researchers soon realized that it was impossible to develop a test that was free from any cultural influences and existing cultural measures should not be seen as interchangeable but rather as assisting in providing different types of cross-cultural comparisons (Foxcroft et al., 2006; Grieve, 2006; Mushquash & Bova, 2007; Plank, 2001). Tseng has expressed concern regarding how appropriate and useful it is to apply a conventional assessment instrument to individuals from diverse cultural backgrounds (as cited in Mushquash & Bova, 2007, p. 57). Further concerns are expressed by Butcher, Nezami, and Exner that, regardless of diverse cultural backgrounds, crucial decisions, and treatment plans are formulated according to the outcomes of clinical assessment tools that were developed for the general population (as cited in Mushquash & Bova, 2007, p. 57). Consequently, test developers focused more on culture-reduced, culture-fair, or culture-common tests in which the aim was to remove as much cultural bias as possible and include only behaviour that was common across cultures (Jenson and Cattell & Cattell as cited in Van de Vijver, 2002, p. 546; Foxcroft et al., 2006; Hogan, 2007).

If an instrument is translated from English, any comparison of groups rests on the assumption that test adaptation was culture-fair (Zeidner, Matthews, & Roberts, 2004). “A culture-fair test is equally appropriate for members of all cultures and comprises items that are equally fair to everyone” (Kitayama & Cohen, 2007, p. 561), in other words a culture-fair test tries to eliminate any social or cultural advantages, or disadvantages, that a person may have due to their upbringing.

According to Manaster and Havighurst (1972), a culture-fair test should have the following characteristics:

- It taps aspects of experience that are common to all people to whom the test will be administered, based on factors such as common family systems, language, objects in every day life, and number systems.

- It is designed to provoke an equal degree of intrinsic interest in subjects from the different cultural groups to whom it will be administered.
- It uses a language that is widely familiar and directions are stated in simple operational terms that are easily understood and have the same meaning for all the subjects to whom the test is administered.

It is highly improbable that any single measure could be designed that would incorporate all these characteristics if it had to be administered cross-culturally or cross ethnically. Culture-fair testing is a contentious issue and some authors believe that culture-fair testing is a myth that perpetuates xenophobic and racist agendas and they contend that adjusting for culture is not simply a matter of new norms or adjusting the interpretation of test scores. They believe that it involves an entire new set of testing skills to understand how the person views the experience and whether they understand what is expected. In addition, the researcher needs to interpret the results in the light of this understanding (Barrett & George, 2004).

Two major problems with culture-fair instruments were described by Anastasi (1988), namely, a lack of sufficient knowledge of the cultures concerned by outside designers of culture-fair instruments, and secondly, comparability becoming a matter of intuitive judgment rather than objective standardization.

For the purpose of cross cultural assessment, instruments may be grouped into three general categories (Van de Vijver, 2002):

1. Instruments with a known reliability and validity in Western groups. To what extent these measures retain their psychometric properties after translation would need to be determined empirically.
2. The development of new instruments that are designed to function in a cross-cultural context. These have been referred to as "culture-free", "culture-fair", and more recently, "culture-reduced".
3. Culture-specific instruments that are developed because existing instruments are considered invalid, unreliable, and do not explore the target construct in other cultural groups. The instrument may be newly developed or based on major or minor adaptations of existing measures.

4.4 Factors Influencing Cross-Cultural Assessment

Shuttleworth-Jordan (1996) advocates that a clear distinction should be drawn between the following factors in the consideration of cross-cultural test influences:

- Racial differences (i.e., ethnic factors); and
- Socio-cultural differences (i.e., factors such as primary language, current language usage, socioeconomic status, preschool socialization experiences, levels of education, and test sophistication) as these are frequently associated with racial differences, and are known to account for significant variations in test performance.

She further stresses the importance of recognizing the complex and evolutionary nature of socio-cultural influences in planning appropriate test procedures. Ardila (1995) also points out that it is important to distinguish between the variable of formal education and the variable of culture which includes factors such as familial socialization, primary language, and meaning ascribed to tests. He further states that language is a variable which needs to be considered as a factor that can have test effects in its own right as it is strongly associated with both cultural background and level of education. Grieve (2006) states that "a test score has no meaning unless it is viewed in context" (p. 229). Hambleton (1994) agrees

> There are many factors which affect cross-cultural/language comparisons which need to be considered whenever two or more groups from different language/cultural backgrounds are compared, especially when an instrument is being developed or adapted, or scores are being interpreted. However, often it is necessary that some of these factors are *not* merely taken into account, but that practical steps be taken to either minimize or eliminate the likely (unwanted) effects of these factors on any cross-cultural/language comparisons that are made. (p. 233-234)

Hambleton (1994) identified some pertinent factors in the social context which include schooling, language, culture, and environmental factors. This section will review some of the factors affecting cross cultural assessment.

4.4.1 Schooling.

The level of schooling attained and, in the South African context, the quality of education received indirectly influences the outcome on intelligence measures (Grieve, 2006; Nell, 2000). Holding et al. (2004) found that there is a strong relation between scores on intelligence measures and scholastic and academic achievement. However, in South Africa this situation is further complicated due to the apartheid regime, where the previously disadvantaged learners received a poorer quality of education than their privileged counterparts (Grieve, 2006; Shuttleworth-Edwards et al., 2004).

4.4.2 Language.

Language is regarded as the most important moderator of performance on assessment measures (Grieve, 2006; Nell, 2000). Poor performance on a measure could be attributed to language difficulties as opposed to ability if the measure was administered in a language other than the test-takers home language. Generally even if you are bilingual, it takes longer to process information in another language. Furthermore, according to the American Educational and Research Association (AERA), the APA, and the National Council on Measurement Education (NCME) an individual who knows two languages may not test well in either of them (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999). We can think and discuss so much better in our own language and hence would more than likely perform better in a test that is written and administered in our home language than in a test in a second language. Thus, language becomes a potential source of bias.

Translating the measure could offer a solution. This, however, could also pose difficulties such as some languages do not have the concepts and expression required by a measure. Further, translating items could affect their level of difficulty. Another complication in South Africa is that some learners are schooled in a language other than their home language, which may compromise both languages, and places test-takers in a doubled disadvantage situation (Grieve, 2006).

Respondents need to understand the language of the assessment measure and need to respond by means of language, as most measures require the use of language. Respondents who are not proficient in the language of the test may introduce construct irrelevant components to the testing process (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999).

Cultural groups may differ in their language spoken. They may also differ in terms of the way in which verbal expressions are formally structured, even if they speak the same language. Some cultural groups, for example upper-middle class North Americans and North Europeans, encourage a highly structured, rational, and orderly use of language, while other cultural groups use language more loosely, with less logical structure and less clear-cut meaning. Furthermore, different cultural groups may assign different meanings to commonly used expressions. Respondents from one cultural or ethnic group will therefore differ to other cultural or ethnic groups in their performance to the extent that they are familiar with the questionnaire's language as well as expressions associated with that language.

Hay (2002, p. 23) considered the cultural diversity of her South African sample and points out that

> it is sweeping and broad to assume cultural differences only between Black and White South Africans. There are probably as many cultural differences between English and Afrikaans Westerners, between English-speakers and speakers of other European languages, between Zulus and Sothos, and so on. It is a mistake to assume that little cultural difference exists where groups speak the same language and believe in the same God.

4.4.3 Culture.

Culture may be defined as the learned attitudes and behaviour that are characteristic of a particular social group or organization which is passed from generation to generation (Ponterotto, Casas, Suzuki, & Alexander, 1995). Our culture influences the way we learn, think, and behave. It is an integral part of our environment and cannot be isolated as a factor on its own. Further, the content of any measure reflects the culture of the test developer and the country in which it is to be used. Therefore test-takers who do not share this culture will be at a disadvantage. As discussed in section 4.2, there are no culture-free measures but practitioners are expected to be sensitive to cultural fairness in assessment and not assume equivalence between cultures. A further problem is that there are variations in acculturation, which refers to the "process by which people become assimilated into a culture" (Grieve, 2006, p. 232).

Researchers have for instance, found cultural differences with respect to child birth, neonatal care, and infant and child rearing practices (Anastasi, 1988; Rebelsky & Daniel, 1976). These differences between various cultural and language groups are a function of not only the different traditions, norms, and values, but of different worldviews and interpretations as well (Hambleton, 1994). It is therefore entirely possible that the same construct is interpreted and understood in completely different ways by two different groups. The concept of intelligence, for example, is known to exist in almost all cultures. Lonner found that in many Western cultures this concept is associated with being able to produce responses very quickly, whereas in Eastern cultures, intelligence is often associated with slow thoughtfulness, reflection, and saying the right thing (as cited in Van de Vijver & Poortinga, 2005, p.39). Researchers in the past have neglected to consider these differences and made fallacious assumptions about individuals belonging to different cultures (Bhamjee, 1991).

4.4.4 Environmental factors.

Environmental factors determine the types of learning experiences and opportunities to which we are exposed, which in turn, affects our ability and the level to which we use that ability. Environmental factors can be grouped into distal factors (e.g., socio-economic status and enriching social environment) and proximal factors (e.g., socialization experiences in the home; Grieve, 2006).

4.4.4.1 The home environment.

Certain child rearing practices have been linked to promoting development of competence and cognitive abilities. These include parental responsiveness and the provision of home stimulation (Grieve, 2006).

4.4.4.2 Socio-economic status.

Socio-economic status refers to the person's social standing. The major indicators of SES are education, occupation, and income. The test-takers SES is important as it determines the type of facilities that are available (e.g., schools, libraries, clinics, and other social services), the opportunities that present themselves, and the attitudes of others (Grieve, 2006).

4.4.4.3 Urbanization.

Urbanization is generally found to influence cognitive scores, with urban children outperforming their rural counterparts (Mwamwenda, 1995). The reasons for this could be attributed to an invigorating urban environment that stimulates cognition, access to education at an early age, higher parental levels of educations, and so forth (Grieve, 2006).

4.5 Methodological Considerations in Cross-Cultural Assessment

Both bias and equivalence are fundamental concepts in cross-cultural assessment, in that they refer to the characteristics of a cross-cultural comparison of an instrument rather than the intrinsic properties. Bias and equivalence are concepts that are closely related. Bias refers to "factors that show a differential impact on scores in cultural populations, while equivalence involves the implications of bias on the scope for comparing scores" (Van de Vijver, 2002, p. 548). The equivalence of a measure (or lack of bias) is a prerequisite for valid comparisons across cultural populations (Van de Vijver & Tanzer, 1997), if bias occurs, the equivalence of the scores is challenged (Van de Vijver, 2002). Throughout the history of psychological research there have been many sweeping generalizations about differences in traits and abilities of cultural populations. When examined more closely, however, these generalizations were often based on inadequate psychometric measures. To avoid such blundering statements it would be advisable to demonstrate the absence of bias (i.e., equivalence) instead of simply making the assumption (Poortinga & Malpass, 1986). It is imperative to determine the cultural appropriateness of an instrument and there should be an empirical investigation into the item bias, differential item functioning, and construct equivalence for the different subgroups (Foxcroft et al., 2006).

4.5.1 Bias.

Bias or specifically, test bias, refers to "whether a measure is differentially valid for different subgroups" (Foxcroft et al., 2006, p. 5). Bias occurs when score differences for a construct are observed which do not correspond to differences in the underlying trait, attitude, or ability across cultural groups (Van de Vijver, 2002; Van de Vijver & Tanzer, 1997). Bias challenges the construct validity of an item or measure (Van de Vijver, 2002). Thus, it is imperative when adapting an instrument, that any unfair advantage or disadvantage to a test-taker, irrespective of their cultural, social, economic, or linguistic background, is eliminated (Foxcroft et al., 2006). There are three types of bias: construct bias, method bias, and item bias (Van de Vijver, 2002; Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Poortinga, 1997). These types of bias will be discussed in more detail below.

4.5.1.1 Construct bias.

Construct bias occurs if the construct measured as a whole (e.g., postpartum depression) is not identical across cultural groups (Ægisdóttir et al., 2008; Van de Vijver & Tanzer, 1997). For example, the appropriateness of the item content differs between the two language versions of the measure (Ægisdóttir et al., 2008).

4.5.1.2 Method bias.

Method bias stems from the characteristics of the measure or from its administration (Ægisdóttir et al., 2008; Van de Vijver, 2001; Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Poortinga, 1997). Three types of method bias exist, namely, sample bias; instrument bias; and administration bias. Sample bias occurs when the samples used differ in a variety of relevant characteristics other than the intended construct. Administration bias includes all sources of bias that are caused by administering the instrument (e.g., interviewee is not fluent in the language of the test). Instrument bias refers to biases that occur due to the characteristics or design, or both, of the instrument itself (Van de Vijver & Tanzer, 1997).

4.5.1.3 Item bias

Item bias or differential item functioning refers to distortions at the item level (Ægisdóttir et al., 2008). Biased items have a different psychological meaning across cultures. This has an impact on the comparison of total test scores across cultures. Thus individuals from different groups who have the same ability have a different probability of getting the item right.

Item bias may occur as a result of various factors although common causes are poor translation, poor item formulation (e.g., complex wording; ambiguity in the original item which may elicit different interpretations), low familiarity with the item content in specific cultures or inappropriateness of the item content for certain cultures, or in other words, not be equally relevant or appropriate for the cultural groups being compared (Ægisdóttir et al., 2008; Malpass & Poortinga, 1986; Van de Vijver & Poortinga, 1997; Van de Vijver & Tanzer, 2004), or the influence of some things which are considered culture-specific, for example connotations associated with the item wording or nuisance factors (Van de Vijver & Tanzer, 2004). Van Haaften and Van de Vijver provide an example of item bias which was caused by inappropriate item content. The item "watched more television than usual" had to be removed from a Western coping questionnaire when it was applied to Sahel dwellers who did not have electricity in their homes (as cited in Van de Vijver & Tanzer, 2004). Most studies of bias focus on exploring and testing for item bias (Van de Vijver & Tanzer, 1997).

Van de Vijver and Tanzer (2004) discuss strategies to identify and reduce the three types of bias mentioned above. These strategies are summarised in Table 7 below.

Table 7 Strategies for Identifying and Dealing with Bias in Cross-cultural Assessment

Type of Bias	Strategies
Construct bias	• Decentering (i.e., simultaneously developing the same instrument in several cultures) • Convergence approach (i.e., independent within culture development of instruments and subsequent cross-cultural administration of all instruments)
Construct bias and/or method bias	• Use of informants with expertise in local culture and language • Use samples of bilingual subjects • Use of local surveys (e.g., content analyses of free response questions) • Non-standard instrument administration (e.g., “thinking aloud”) • Cross-cultural comparison of nomological networks (e.g., convergent/discriminant validity studies, monotrait–multimethod studies, connotation of key phrases)
Method bias	• Extensive training of administrators (e.g., increasing cultural sensitivity) • Detailed manual/protocol for administration, scoring, and interpretation • Detailed instructions (e.g., with sufficient number of examples and/or exercises) • Use of subject and context variables (e.g., educational background) • Use of collateral information (e.g., test-taking behavior or test attitudes) • Assessment of response styles • Use of test–retest, training, and/or intervention studies
Item bias	• Judgmental methods of item bias detection (e.g., linguistic and psychological analysis) • Psychometric methods of item bias detection (e.g., differential item functioning analysis) • Error or distracter analysis • Documentation of “spare items” in the test manual which are equally good measures of the construct as actually used test items

(Van de Vijver & Tanzer, 2004, p. 128).

4.5.2 Equivalence.

The attainment of equivalent measures is perhaps the central issue in cross-cultural comparative research (Van de Vijver, 2001; Van de Vijver & Leung, 1997b). Equivalence refers to the level at which the item or test scores can be compared across cultural or language groups (Van de Vijver, 2001; Van de Vijver, 2002). For measures to be equivalent, individuals with the same or similar ability on a construct should obtain the same or similar scores on the different language version (e.g., translation equivalence) of that instrument otherwise the instrument is considered biased and the two versions of the instrument are non-equivalent. For example, a score of 10 on an unbiased scale for depression has the same psychological meaning in all cultural or language groups studied (Van de Vijver, 2002). Instruments need to be equivalent if meaningful comparisons are to be made between the two subgroups (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999; Kanjee, 2006). Without demonstrated equivalence, numerous rival hypotheses (e.g., poor translation) may account for observed cross-cultural differences (Ægisdóttir et al., 2008).

There are three types of equivalence, construct equivalence, measurement unit equivalence, and scalar equivalence (Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Poortinga, 1997). In addition, linguistic differences can easily invalidate the results of a study. If a psychometric measure is poorly translated, it doesn’t matter how sound your methodology is (Onkvisit & Shaw, 2004). Therefore, it is equally important to discuss linguistic equivalence. These types of equivalence will be discussed in more detail below.

4.5.2.1 Construct equivalence.

Construct equivalence (also referred to as functional equivalence and structural equivalence) means that the same underlying psychological construct is measured across all cultural groups in spite of whether or not the measurement of the construct is based on identical instruments across all cultures (Van de Vijver & Tanzer, 2004).

4.5.2.2 Measurement unit equivalence.

Measurement unit equivalence refers to the level of equivalence that can be obtained when two metric measures have the same measurement unit but have different origins across groups (Ægisdóttir et al., 2008; Van de Vijver & Tanzer, 2004). For example, the two language version of a measure may appear the same, but equivalence is threatened if the two groups vary in their familiarity with Lickert-type answer format (method bias). Similarly, if the two groups vary in response style (acquiescence), a score of 4 on a 5-point scale may not have the same meaning for the two groups (Ægisdóttir et al., 2008).

4.5.2.3 Scalar equivalence.

The highest level of equivalence is scalar equivalence, which can be obtained when two metric measures have the same measurement unit and the same origin. This type of equivalence assumes completely bias-free measurement (Van de Vijver & Tanzer, 2004).

4.5.2.4 Linguistic equivalence.

Linguistic equivalence must be ensured when cross-cultural studies are conducted in different languages. Linguistic equivalence requires the research to pay particular attention to potential translation problems. It is therefore recommended that translators pay attention to idiomatic vocabulary, grammatical, and syntactical differences in language, as well as the experiential differences in cultures as expressed in language (Onkvisit & Shaw, 2004). The various translation techniques that can be employed to enhance equivalence will be discussed in section 4.6.

4.6 Ethical Guidelines for Adaptation of Cross-Cultural Assessment Measures

With increased globalisation and the substantial costs of test development, the use of westernized psychological tests in South Africa is widespread, therefore, having measures that are reliable and valid, and can be used for our diverse languages and cultures is crucial. However, the guidelines or standards for the translation and cultural adaptation of instruments that have been established in psychology to date are limited (Van Widenfelt, Treffers, De Beurs, Siebelink, & Koudijs, 2005).

Anastasi (1988) suggested three approaches to the development of tests for different cultures or subcultures. The basic approach is to compile an instrument that taps aspects of cultural experience which are common to many cultures, and validating the resulting measure against local criteria in the cultures where it will be administered.

Without the necessary precautions it cannot be assumed that a test is relatively free form culturally restricting elements, yet this repeated validation in various cultures has often either been neglected or poorly controlled. However, it is unlikely that a measure would fully meet these requirements across a wide range of cultures or ethnic groups.

A second approach is to compile a measure within one culture and then administering it to individuals from different cultural groups. This procedure is typically followed when the object of assessment is the prediction of a local criterion within a specific culture. The criterion itself is usually culturally loaded, therefore the test validity may drop if the cultural loading of the test is reduced. Care should be taken not to regard a measure constructed within a single culture as a universal yardstick for measurement. This approach enables the researcher to determine the cultural distance between groups as well as the individual's degree of acculturation.

The third approach is to construct different measures within each culture and to validate them against local criteria only. In this type of approach an individual's result are compared to local norms and no cross-cultural comparisons are attempted. This seems to negate the purpose of cross-cultural and cross-ethnic research.

The International Test Commission (ITC) was formally established in 1978 is an "Association of national psychological associations, test commission, publishers and other organizations committed to promoting effective testing and assessment policies and to the proper development, evaluation and uses of educational and psychological instruments." (International Test Commission, 2001). The following definition of an instrument adaptation guideline was proposed by the ITC (Hambleton, 1994)

> An instrument adaptation guideline is a practice that is judged as important for conducting and evaluating the adaptation or parallel development of psychological and educational instruments for use in different populations. (p. 233)

Hambleton (1994) predicted that substantially more adaptations might be expected in the future as:

- international exchanges of tests and instruments become more common;
- credentialing exams are adapted for use in multiple languages; and
- interest in cross-cultural research increases.

In 1992 the ITC began a project to prepare guidelines for translating and adapting psychological instruments and other tests, as well as establishing score equivalence across different language or cultural groups. In 2000 the ITC released Guidelines for Adapting Educational and Psychological Tests (Hambleton, 2001; International Test Commission, 2001). These guidelines have become the benchmark for cross-cultural test adaptation around the world (Foxcroft et al., 2006). The ITC subsequently developed the "International Test Commission Guidelines for Translating and Adapting Tests – Version 2010" (International Test Commission, 2010) which further addresses issues of fairness and bias in test use and sets standards for the professional practice of assessment. These guidelines emphasize the importance of cross-cultural validity of measures as well as their constructs across different populations and cultures. The 22 guidelines for adapting psychological and educational tests or instruments are organized into four categories (International Test Commission, 2010):

1. Context, which addresses concerns about construct equivalence in the language groups of interest;
2. Test development and adaptation, which includes the guidelines which arise in the process of adapting an instrument, from selecting translators to statistical methods for analyzing empirical data to investigate score equivalence;
3. Administration, which addresses guidelines having to do with the ways in which instruments are administered in multiple language groups. This includes everything from selecting administrators, to the choice of item formats, to establishing time limits; and
4. Documentation/score interpretations.

The ITC uses the term "adaptation" rather than "translation". This is because the former term is broader and instrument adaptation guidelines seemed to more accurately reflect the process of preparing a test or instrument for use in a second language or culture. Translation is always part of the adaptation process, but is only one of a number of steps that must be carefully carried out to produce a test or instrument that is equally valid in two or more languages and cultures.

Several professional bodies have since provided clear standards and guidelines that need to be adhered to when using psychological tests, these include the Standards for Educational and psychological Testing (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) and the Guidelines for Computer-based Tests and Interpretations (APA, 1986).

In an attempt to address issues of fairness and bias in test use, the need arose to develop standards for the professional practice of testing and assessment. Led by Bartram from the United Kingdom, the ITC developed "International Guidelines on Test Use - Version 2000" (International Test Commission, 2000), which like in many other countries was adopted by South Africa.

The international guidelines for test use (International Test Commission, 2000, p. 12) states that when individuals from diverse groups (e.g., groups differing in terms of age, gender, education, cultural background, or ethnic origin) are assessed all reasonable efforts should be made to ensure that the following best practice guidelines are followed:

- The tests are unbiased and appropriate for the various groups that will be tested.
- The constructs being assessed are meaningful in each of the groups represented.
- Evidence is available on possible group differences in performance on the test.
- Evidence relating to differential item functioning (DIF) is available, where relevant.
- There is validity evidence to support the intended use of the test in the various groups.
- Effects of group differences not relevant to the main purpose (e.g., differences in motivation to answer, or reading ability) are minimised.

- In all cases, Guidelines relating to the fair use of tests are interpreted in the context of local policy and legislation.

Ethical guidelines have also been issued (APA, 1993; International Test Commission, 2000). It has, however, been reported that practitioners, test developers, and test publishers generally do not adhere to these guidelines. Furthermore, many measures which have been translated have not been re-normed and re-validated, and there are no appropriate precautions provided in the test manuals (Barrett & George, 2004).

The ITC guidelines (International Test Commission, 2000, p18) for multilingual instruments require the following:

- Each language or dialect version has been developed using a rigorous methodology meeting the requirements of best practice.
- The developers have been sensitive to issues of content, culture, and language.
- The test administrators can communicate clearly in the language in which the test is to be administered.
- The test-taker's level of proficiency in the language in which the test will be administered is determined systematically and the appropriate language version is administered or bilingual assessment is performed, if appropriate.

The standards for educational and psychological testing (Standard 9.7; American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) state

> When a test is translated from one language to another, the methods used in establishing the adequacy of the translation should be described, and empirical and logical evidence should be provided for score reliability and the validity of the translated test's score inferences for the uses intended in the linguistic groups to be tested. (p. 99).

Standard 9.9 states that if multiple language versions are intended to be comparable, then empirical evidence of test comparability should be reported. The comprehensive set of 22 guidelines provided by the ITC for improving the translation and adaptation of educational and psychological instruments are presented in Appendix A. Hambleton (2004, pp. 65-70) summarized these guidelines and notes the following nine key steps that should be addressed when adapting or translating any assessment instrument:

1. Explore the construct equivalence in the languages and cultures of interest.
2. Decide whether test adaptation or translation is the best option.
3. Choose well qualified translators.
4. Translate or adapt the instrument using the appropriate design.
5. Review the adapted version and make the necessary changes.
6. Conduct a small pilot with the adapted test.

7. Conduct a validation investigation.
8. Place the scores of both the translated and original instruments on a common scale.
9. Document the process and prepare the manual for test users.

4.7 Translating Assessment Measures

In South Africa, many of the psychological measures are in English and translating and adapting these measures would alleviate some of the biases associated with psychological tests. This section looks at the current methods researchers employ in multilingual studies.

4.7.1 Techniques in translating instruments.

Test translation refers to the process of converting a measure from one language to one or more other languages (e.g., from English to Afrikaans), while still retaining the original meaning (Foxcroft et al., 2006). Translating a psychological instrument is more complex than simply rewriting the text into a different language (Bracken & Barona, 1991; Brislin, 1980, 1986; Geisinger, 1994; Hambleton, 1994), it needs to take into consideration the original context of the source instrument as well as reflect the cultural understanding of the target language (Bracken & Barona, 1991). Therefore, an appropriate translation requires a balance between psychological, linguistic, and cultural considerations (Hambleton, 1994; Van de Vijver & Hambleton, 1996). Employing a proper translation methodology is critical as it affects the equivalence of the multilingual versions and the measures' cross-cultural validity. Further, researchers should also bear in mind that test instructions need to undergo the same translation method as the items (Ægisdóttir et al., 2008).

Numerous techniques have been developed for translating, adapting, and re-norming psychological instruments for cultures and languages other than their initial target population (Ferraro, 2002; Fletcher-Janzen et al., 2000; Nell, 2000). According to Bracken and Barona (1991) the successful translation of tests is dependent on following a comprehensive multistep translation and validation process. The translation techniques and processes they describe are (a) source to target language translation, (b) blind back-translation, (c) translation – back-translation repetition, (d) committee review, (e) pilot testing, (f) field-testing, and (g) norm development. Brislin (1980) proposed translation methods, such as back-translation, bilingual, committee, decentering, and pretests. Onkvisit and Shaw (2004) refer to the following translation techniques: back translation, parallel-blind translation, committee approach, random probe, and decentering. However, Van de Vijver and Tanzer (1997) refer to only two translation procedures, namely, the translation-back-translation procedure and the committee approach. Bracken and Barona (1991) advocate that the most commonly applied technique is the back-translation technique while Kanjee (2006) states that the common designs used are forward-translation and back-translation. This section will discuss some of the proposed translation techniques.

4.7.1.1 One way or bilingual translation.

This involves the translation from the original to the target language by a translator who is truly bilingual and also sufficiently educated to be familiar with the concepts of the test and subject matter (Bracken & Barona, 1991). The test translator must also be knowledgeable with the target culture, the construct being assessed, and the principles of assessment (Hambleton & de Jong, 2003; Van de Vijver & Hambleton, 1996). This will assist in minimizing item biases that may result from literal translations. This technique is akin to first step in Bracken and Barona's (1991) multistep translation process mentioned earlier. It is considered to be an uncomplicated translation method. In some instances the researcher may opt to have a few translators conduct a one-way translation of the instrument. This method is less time consuming and less expensive than other methods. Limitations of this method are, however, that no comparison of the final survey version is made and information may be lost through literal translation (McGorry, 2000).

4.7.1.2 Forward-translation.

A forward translation or simple direct translation simply implies translating the instrument into the language chosen (e.g., from English into Afrikaans), by a single translator or a group of translators (Kanjee, 2006; McGorry, 2000). Although this method is more cost effective, there may be a loss of information through literal interpretation (McGorry, 2000). This first version would then be given to a pilot group of test-takers to answer and then the test-takers would be questioned by judges as to the meaning of their responses. The judges would then decide if the responses given reflect a reasonable representation of the test items in terms of cultural and linguistic understanding. If a high percentage of the test-takers present a reasonable representation of an item, the item is regarded as being equivalent. A valuable advantage is that the functioning of any item is provided directly by the test-takers (Kanjee, 2006). The disadvantage, however, is that there are many confounding variables (e.g., personal, cultural, linguistic) that may affect this process and hinder the results. Another disadvantage is that this technique is very labour intensive and time-consuming (Kanjee, 2006).

4.7.1.3 Modified direct translation.

Geisinger (1994) proposed that some steps be taken to overcome some concerns with the forward translation. He also suggested that a panel or committee of experts do independent checks on the original translation as follows: “(a) review the items and react in writing, (b) share their comments with one another, and (c) meet to consider the points made by each other and to reconcile any differences of opinion” (p. 306). He further recommends that the original translator meet with the panel on two occasions, first so that the panel members can explain their concerns, and then again so that the translator can give explanations and clarify why the measure was drafted as it was.

Modified direct translation is more informative than a simple forward translation due to opportunity for discussion among committee members as well as discussions between the translator and the committee. This procedure increases the security of the translation, provided the translator and committee members are competent and are able to reach consensus in the meetings. This technique is, however, likely to consume substantial amounts of effort, time, and money. As a result, practicality is somewhat low (Behling & Law, 2000). This approach also does not involve enough collaborative efforts that are needed to produce a well-translated instrument (Pan & De La Puente, 2005).

4.7.1.4 Parallel blind translation.

In a parallel-blind translation, the measure is translated by several translators independently and then the translators meet to compare their versions and resolve any differences. Once the differences are resolved they jointly present the translated instrument to the researcher (Behling & Law, 2000; Onkvisit & Shaw, 2004). According to Guthery and Lowe (1992) the parallel blind technique has two advantages: speed and researcher control. The process is faster than conventional back-translation because the translators work in parallel as opposed to in sequence. However, this technique lacks source language transparency, in that unless the researcher is bilingual his or her role in the process is limited. For instance, the researcher would not be able to identify if the translators share certain misconceptions or if what appears to be an agreement between them is actually due to their unwillingness to criticize one another's translations (Behling & Law, 2000).

4.7.1.5 Committee approach.

A committee (or cross-translation) approach is where a group of experts (such as cultural, linguistic, and psychological) prepare a translation (Nasser, 2005; Van de Vijver

& Tanzer, 2004). The committee approach differs from the parallel-blind technique due to the fact that the former allows committee members to discuss the research questions with each other during the translation (Onkvisit & Shaw, 2004). If all the translations are the same, then the translation is considered valid (Nasser, 2005). Often researchers combine the committee approach with the back translation technique (Van de Vijver & Leung, 1997b). Major strengths in using this approach is that this collaborative effort improves the quality of the translation, reduces bias, and reduces misconceptions that a single person may bring (Ægisdóttir et al., 2008) This is especially true, if the members have complimentary areas of expertise (Van de Vijver & Tanzer, 2004). A disadvantage of this approach is the absence of an independent evaluation of the adequacy of the translation. If the researcher is not fluent in the translated language, then additional evidence will be needed to evaluate the quality of the committee's work (Van de Vijver & Leung, 1997b). In addition, translators may be reluctant to criticize one another or may miss information relevant to the intended group due to similar cultural backgrounds and education (McGorry, 2000).

4.7.1.6 Pilot-testing or pretest.

Once a translated version of an instrument has been agreed upon and approved by the bilingual review committee, it can be very beneficial to administer it to a small group of people representative of the target population for pilot testing and minor adjustment. The results can be compared to the results obtained from the original language sample (Hambleton & Patsula, 1999; Van Widenfelt et al., 2005). A trained examiner who is fluent in the target language should administer the measure. Pilot testing is a helpful step in instrument adaptation, however, a review by Guillemin, Bombardier, and Beaton (1993) indicates that a pretest or pilot test is done by comparatively few studies on translated measures. An interview or discussion with the pilot participants is also a valuable step as it allows the examiner to determine the pilot participants' reactions to the test instructions, response categories, and items. The examiner should also take note of verbal and non verbal expressions, such as looks of puzzlement, confusion, giggles, or other responses to items that may indicate that the item is perceived as confusing, bizarre or inappropriate, and hence suggest possible translation failure (Bracken & Barona, 1991). These should be discussed by the review committee to explore possible reasons for the inappropriate examinee responses.

Pilot participants that are selected should vary in social and economic background, geographic region, gender, and age. Regional differences in expression should also be considered when using the feedback of pilot participants (Bracken & Barona, 1991; Van Widenfelt et al., 2005). After the initial pilot testing data has been obtained it is essential to meet again as a team to consider any necessary changes. Further adaptations can be made by the team of translators based on the outcome of the pilot testing (Geisinger, 1994). If uncertainty remains about an item, two alternatives for that item can temporarily be included in the version for further testing (Van Widenfelt et al., 2005).

4.7.1.7 Field-testing.

Field testing typically follows the pilot testing. This procedure is essentially the same as pilot testing but differs in magnitude. Examiners should be attentive of any consistent problems related to specific items or test directions. Formal item analyses can be done on the results obtained from the field testing if the sample is large enough. (Bracken & Barona, 1991).

4.7.1.8 Random probe.

A random probe entails placing probes at random locations in both the original and translated measures during pretesting in order to ensure that the respondents understand the items in the same way (Onkvisit & Shaw, 2004). The respondents are then asked to explain why they responded as they did to certain items (Behling & Law, 2000). According to Guthery and Lowe (1992), “if the respondent’s justification to the original answer is strange, then the intent of the question is not being conveyed” (p. 10). This technique is cheap, simple, and quick. However, researchers may need to supplement this technique with a more rigorous procedure, as it provides limited information and is not innately source language transparent (Behling & Law, 2000).

4.7.1.9 Decentering.

Decentering is defined by Eremenco, Cella, & Arnold (2005) as:

> A process in which the source instrument and its translation are open to modification in an iterative process, so that the meaning is equivalent between them. The opposite is to have a translation process in which the source is unchangeable, thereby requiring all adjustments to be made to the translation only. (p. 228)

This translation technique is termed "decentering" because the researcher does not center in either the original language or the target language. Decentering modifies the back-translation technique by considering the original and translated versions as equally important and both are open to modifications (Beck et al., 2003; Ægisdóttir et al., 2008; Geller, Vinokurov, & Martin, 2004; Flores, 2006). If problems are identified in the original measure (e.g., words in the original language that have no equivalence in the translated language), then it should be modified in order to be more easily translatable (Ægisdóttir et al., 2008; Nasser, 2005; Onkvisit & Shaw, 2004). There is constant comparison between the two measures and the original measure is retrospectively modified in order to enhance its translatability. Thus in this process the original measure becomes a draft and is revised to fit the new research situation (McGorry, 2000). Typical modifications made are to words or concepts that are difficult to translate or are culture specific (Van de Vijver & Leung, 1997b). Marin and Marin acknowledge that the use of decentering may lengthen the translation process but it does help to achieve a fully equivalent language version (as cited in Beck et al., 2003, p. 68).

Decentering has a number of advantages. Like back-translation, it is both informative and source language transparent. Further, it provides the researcher with the opportunity to check the reasoning of one translator against another and allows for better equivalence because the source and target versions are equally subject to change, thereby bringing both closer in meaning to the desired concept. The decentering technique, according to Van de Vijver and Leung (1997b), echoes "the goals of the culture-free and culture-fair test movement" (p. 39). This technique is not, however, very practical as it is labour intensive requiring a substantial number of translators that are multicultural, multilingual, and have expertise in the construct under study (Behling & Law, 2000; McGorry, 2000; Van de Vijver & Leung, 1997b). Furthermore, in instances where an instrument has already been validated and is widely used, the instrument developer is likely to be averse to implementing changes to the measure in light of translations. This may be the case regardless of the possibility that decentering would improve the equivalence of the translations with the established version (Eremenco et al., 2005).

4.7.1.10 Back-translation.

The back-translation method, also known as the double translation method, involves the translation of items from the original into the target language by one or more bilingual translators. This material is re-translated back into the original language by another bilingual translator or team of bilingual translators, yielding the back-translated version. Richard Brislin, a cross-cultural psychologist, was the first to write extensively about back translation as a method to ensure a quality translation of a test or measuring instrument (Brislin, 1970; 1976; 1986; Brislin, Lonner, & Thorndike, 1973; Stansfield & Bowles, 2007).

In order to judge the quality and determine the equivalence of the measures, the researcher checks for errors between the original and back-translated versions of the measure and consults with the translators about the possible reasons for any inconsistencies, mistranslations, lost words, and changes in meaning. Once these issues have been addressed further revisions to the translated version may occur. It is then back-translated yet again and compared to the source document. This process of creating a back translation and comparing it with the original version is repeated until the two versions agree. Although the researcher may go through several rounds of revisions, the original-language version of the measure is considered the standard against which the translated version is compared (Beck et al., 2003; Brislin, 1970, 1986; Geller et al., 2004; Kanjee, 2006; McGorry, 2000; Nasser, 2005; Onkvisit & Shaw, 2004). Marin and Marin (as cited in Beck et al., 2003, p. 68) point out that this is what differentiates back-translation and decentering. Decentering regards the original language version as well as the translated version as equally important.

This method has been widely applied and can identify various kinds of errors. Various researchers have stated that the back-translation method is particularly useful for checking the semantic equivalence of the translations of measures in different languages used in cross-cultural research (Beck & Gable, 2003; Bracken & Barona, 1991; Prieto, 1992). The back-translation technique has been used successfully to translate from English to Afrikaans (e.g., Shillington, 1988), as well as in developing a Spanish version of the PDSS (Beck & Gable, 2003). An advantage of using the back-translation technique is that it enhances the reliability and accuracy of the translated instrument as it offers opportunity for revisions through the translation process (Bracken & Barona, 1991; Geisinger, 1994; Van de Vijver & Hambleton, 1996; Van de Vijver & Leung, 2001).

Some researchers have raised concerns regarding translating and adapting an instrument from one language to another (Geisinger, 1994; Hui & Triandis, 1985; Van de Vijver & Hambleton, 1996). Some problems with the back-translation method have been identified. Lack of agreement between the source document and the back translation may be due to problems with the back translation as opposed to problems with the initial forward translation. The back translation is just as likely to contain translation errors (omissions, mistranslations, insertions) as is the forward translation (Stansfield & Bowles, 2007).

When the translator is aware that the forward translation will be validated by a back translation, it may influence the translator's approach to forward translation. The translator may produce a very literal forward translation to help ensure that the back translation will produce a document that is very similar to the original version. This type of translation is, according to Stansfield and Bowles (2007, p. 2) "likely to produce stilted rather than natural expression and result in a test that is difficult to read, and hence less accessible to the examinee population."

If the back-translation version seems to lack equivalence in meaning to the original version, it is not easy to determine whether the differences are as a result of poor translation, or cultural and linguistic differences inherent in cross-cultural research. Furthermore, when the translated version is similar to the original version, it remains uncertain about the nuances of meaning across languages and cultures (Geller et al., 2004). Back-translation may lack equivalence in meaning and still demonstrate spurious lexical equivalence, thus giving the researcher a false sense of security (Birbili, 2000; Brislin, 1970, 1976).

A translated instrument, while linguistically correct, may have poor quality from a psychological perspective. An example provided by Hambleton (1994, p. 235) that illustrates this point, is the test item "Where is a bird with webbed feet most likely to live?" The Swedish translation of the English "bird with webbed feet" into "bird with swimming feet" provides a much stronger clue and thus a direct translation would have given the Swedish test-takers an unfair advantage. Further, if the original language version does not have an equivalent term for the translated version, psychometric properties or constructs could be lost in the translation (Nasser, 2005).

Brislin, Lonner, and Thorndike recommend that a multiple translation method be used to ensure semantic equivalence (as cited in Beck et al., 2003). The AERA, APA, and NCME (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) have also subsequently recommended that back-translation should not be a stand alone procedure as it may provide an artificial similarity of meaning across languages but not be the best version of the new language. They recommend that a more iterative process akin to test development and validation be considered in order to ensure that similar constructs are measured across versions. A process involving successive iterations of forward translations and revisions thereof would work equally well, according to Stansfield and Bowles (2007), and the translation equivalence of the two versions can be more cost effective and can take less time.

Bracken and Barona (1991) recommend that researchers use a bilingual committee of judges to compare the original or back-translated version with the translated version of the measure to ensure that the translation is appropriate for the test-takers. Van de Vijver and Hambleton (1996) suggest that the team of translators should have combined expertise in psychology and linguistics. Beck et al. (2003) used multiple translation methods to help ensure the semantic equivalence in translating the Postpartum Depression Screening Scale (PDSS) into Spanish. The multiple methods they employed included back-translation, the committee approach, pretest techniques, and alternate forms equivalence.

The primary concern of translating any instrument is to produce a version that is both linguistically and culturally equivalent to the original. Having reviewed most of the translation techniques, it should be noted that there will always be concepts that cannot be translated into certain languages or that cannot be asked in a meaningful way in certain cultures (Onkvisit & Shaw, 2004).

4.7.2 Translation procedure.

The translation procedure chosen will depend on whether a new instrument is being developed or whether an existing instrument is being translated for a multilingual context. The former is known as simultaneous development while the latter is referred to as successive development. From a methodological perspective, the first option is easier as difficult items such as local idioms which are often difficult to translate can often be avoided (Van de Vijver & Tanzer, 2004). However, in developing countries such as

South Africa the cost of test development necessitates the use of existing instruments. Three options are available to researchers in the successive development method, namely, application, adaptation, and assembly (Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Tanzer, 2004). These are outlined below.

4.7.2.1 Application.

This option entails the literal translation of an instrument into a target language and it implicitly assumes that the underlying construct is appropriate for each cultural group. The literal translation is commonly used in test translations. (Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Tanzer, 2004).

4.7.2.2 Adaptation.

This option entails the literal translation of part of the items, changes in other items, or the creation of new items – or any combination of these. For some instruments, it is unrealistic to assume that simple translation would yield construct equivalence for the target cultural group. For example, a measure of anxiety may contain items that need to be reworded to ensure culturally idiosyncratic expressions of the construct are included. Adaptations are based on the premise that using the literal translation would yield a biased instrument. The State-Trait Anxiety Inventory has been adapted into 40 languages, and this approach was used to ensure that the underlying constructs – state and trait anxiety – were equivalent across language groups (Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Tanzer, 2004).

4.7.2.3 Assembly.

This option entails the adaptation of an instrument to such an extent that it is practically a new instrument. This option is used when construct bias, caused by differential appropriateness of the item content for the majority of the items, threatens a direct comparison. Another reason for using this option would be an incomplete overlap of the construct definition across cultures. For example, aspects of the construct that are salient for some cultures are not covered by the instrument. Researchers contend that Western personality instruments do not cover the personality constructs of other cultures, such as the Filipino and Chinese cultures (e.g. Cheung et al. and Church as cited in Van de Vijver & Tanzer, 2004, p. 123).

4.8 Conclusion

This chapter provided an overview of cross-cultural assessment and the factors that influence cross-cultural assessment. Both bias and equivalence are pivotal concepts that need to be considered in cross-cultural assessment as they help to determine the cultural appropriateness of an instrument. The types of bias and equivalence and how they impact on test scores across different cultural groups as well as how they influence the translation and adaptation of measures were outlined. Guidelines have been provided by the ITC which have become the benchmark for cross-cultural test adaptation. These were summarized along with the steps that should be addressed when adapting or translating any assessment instrument. A number of techniques have been developed for translating, adapting, and re-norming psychological instruments for cultures and languages other than their initial target population. The techniques that researchers and test developers use in multilingual studies were addressed. The following chapter outlines some cultural approaches to the understanding of childbirth and related mental disorders as well as how these impact on adapting a postpartum depression screening measure cross-culturally.

CHAPTER 5

A CULTURAL APPROACH TO PERINATAL MOOD DISORDERS

5.1 Chapter Preview

Comparative and cross-cultural researchers are faced with deceptively simple questions that need to be considered. Do given syndromes exist in all cultures? If so, then are the syndromes as common in every setting? Are the clinical features the same? Are there any major differences in etiology, course, and outcome? And is there any difference in how they are managed and treated?

Childbirth is a universally similar physiological event for women. It does, however, occur in a socio-cultural context causing the experience to be filtered, mediated, and directed, at an individual as well as a social level, by culturally constituted frameworks. The transition to motherhood is therefore experienced and conceptualized according to a person's specific beliefs, values, and attitudes.

Most of the research on PPD has considered the biological and psycho-social etiologies such as hormonal changes, maternal age, psychiatric history, and level of support. Although these are important contributing factors, the impact of socio-environmental factors, the cultural patterning of childbirth, and the post-partum period as an etiology in PPD also need consideration. This chapter will also discuss some cultural approaches to the understanding of childbirth and related mental disorders as well as how these impact on adapting a postpartum depression screening measure cross-culturally.

5.2 Paradigms of Mental Illness

Biological, psychological, cultural, and sociological theories have all sought to explain the onset of mental illness. The Universalist approach regards mental illness as a disease that has the same set of symptoms, the same diagnosis and treatment, and the same prognosis across the world. This approach is based on the medical model which regards factors like organic brain disease due to either genetic, biochemical or physiological causes as the contributor/s in the onset of mental illness, with pharmacotherapy as a main treatment option. The medical model is a leading approach in psychiatry in Western societies. This paradigm, which regards mental illness as a biological or disease model, is the way in which mental illnesses in the United States of America is categorized and classified (Goldbort, 2006).

The paradigm that regards culture or environment or society as the core contributor/s in the onset of mental illness is the sociological or environmental model. This model views mental illness as violations of, or deviations from certain norms in society. Treatment for an individual from this perspective would require changing societal issues contributing to the individual's stress, such as poverty or sexism.

Mental illness should, however, be viewed as a multifaceted illness that requires the philosophical underpinning of both these paradigms. The exclusion of either one in seeking to understand the impact of mental illness does a disservice to improving the outcome for individuals with mental illness.

A number of studies have indicated that the etiology of PPD lies in multiple factors – psychological, familial, hormonal-biological, social, and cultural (Beck, 2001; Clay & Seehusen, 2004; Halbreich, 2005; Kruckman & Smith, 2006; Leung, 2002; O'Hara & Swain, 1996). Kirmayer, and Lazarus and Folkman, emphasized that cultural factors have a significant impact on one's emotional state (as cited in Bina, 2008, p. 569). Some cultural practices and beliefs can significantly influence PPD, either positively or negatively (Bina, 2008). Cultural factors, along with social, psychological, and biological perspectives, must be taken into account in order to fully comprehend PPD (Bina, 2008; Cox, 1999; Harkness, 1987; Leung, 2002). It is also important to consider all the correlates of PPD across different populations to determine whether PPD is a universal experience or a condition that is specific to Western cultures, and to determine how the illness is expressed in other cultures.

5.3 Prevalence of PPD Across Different Cultures

Research about PPD has mostly been carried out in Western cultures (Affonso et al., 2000). Furthermore, the reported prevalence of PPD in non-Western cultures is variable, with prevalence rates varying from 0% to 40%. The reason for the discrepancy in PPD prevalence is uncertain, but researchers believe that it may be due to any of the following factors: that PPD manifests differently in different cultures, low prevalence rates in some cultures may be due to cultural protective factors, the diagnosis of PPD may be more unacceptable in some cultures or not used at all, or that the clinical criteria documented in the DSM-IV-TR is not sufficient to incorporate cross-cultural diagnostic standards (APA, 2000; Fitch, 2002; Kleinman, 2004; Miller, 2002; Posmontier & Horowitz, 2004; Yoshida, Yamashita, Ueda, & Tashiro, 2001). Attempts to investigate the relationship of postpartum traditional practices with PPD amongst non-Western and Western cultures are made more difficult in light of the factors suggested above.

Stern and Krukman's review (1983) of the cultural aspects of PPD advocates that the phenomenon "postpartum depression" is a culture-bound Western syndrome that is not likely to occur in a non-Western society. They maintain that a significant contributing factor to the onset of PPD in Western societies is a lack of organized social support. More recent publications (e.g. Affonso et al., 2000) indicate that PPD does, however, cross cultural boundaries and is not a culture-bound illness. Posmontier and Horowitz (2004) comment that Stern and Kruckman (1983) failed to address the possibility that expressions of PPD may vary according to culture.

The birth of a child, especially the first child, is arguably a significant life event for any women – or man – regardless of their cultural background. A new baby in a household also impacts on the family's financial budget and on the work load of the mother. Hormonal changes are dramatic during pregnancy and shortly after delivery, and their contribution to depressive symptoms postpartum has been indicated (Ahokas et al., 1999; Altemus et al., 2004; Bloch et al., 2000; Bloch et al., 2003; Epperson et al., 2006).

Halbreich and Karkun (2006) state that "if a comprehensive bio-psychosocio-economic model is applied to the postpartum period (as it should), then it is difficult to explain how such a significant life event would not result in distress in at least some mothers in any culture." (p109). Despite evidence of the psychological, socio-ecomonic, and hormonal impact of childbirth on women, the question whether PPD is specific to certain cultural contexts and whether it is influenced by cultural factors has often been raised.

In an attempt to answer this question, numerous researchers have sought to determine the prevalence of postpartum psychiatric illness in various cultures and countries and explored the socio-cultural factors associated with childbirth. Epidemiological studies and survey results from a variety of different cultures across the world report increasingly high rates of PPD (Rahman et al., 2003). Some examples include studies from India (Patel et al., 2002), Turkey (Inandi et al., 2002), United Arab Emirates (Ghubash & Abou-Saleh, 1997), China (Wang, Jiang, Jan, & Chen, 2003), Hong Kong (Chan & Levy, 2004; Lee, Alexander, Yip, Leung, & Chung, 2004), and Latina and African American women (Yonkers et al., 2001). For the most part, these studies show no substantial difference in the rates of PPD and that the risk factors for PPD are similar.

Researchers agree that PPD is a universal experience, even though it may be referred to by a different name by various cultures. Cox (1999), for instance, maintains that PPD is not limited to certain cultures and states that PPD is readily identified in traditional African cultures too. Cox (1999) points out, however, that there is a paucity of research and literature on postpartum mental disorders in African countries. This may be due to the lack of resources and also possibly due to an attitude that these disorders are of no serious consequence and occur infrequently.

Halbreich and Karkun's (2006) review of the literature on the prevalence of PPD and depressive symptoms in a variety of countries found that PPD was prevalent in 40 countries – although in some countries there were very few reports while other countries, including South Africa, had high prevalence rates. They attribute the variability in reported PPD due to a multitude of cross-cultural, socio-economic, and environmental variables, along with biological vulnerability factors.

The Transcultural Study of Postnatal Depression (TCS-PND) was done across several cultures simultaneously to determine the universality of the concept of postpartum depression. This study also examined and compared the correlates of PPD, its prevalence, the psychosocial origins, as well as the consequences of PPD (Asten, Marks, & Oates, 2004). In the initial phase of the study, Oates et al. (2004) explored the understandings, views, and beliefs regarding what constituted happiness or unhappiness antenatally and during the postpartum period. A common theme emerged across all centres in the 11 countries that participated in the study which revealed that a morbid state of unhappiness occurred after delivery with similar characteristics and attributes. Not all centres, however, recognized it as a specific illness with a definite name – like postpartum depression. Participants described characteristics that met the criteria for diagnosing PPD and attributed the unhappiness to family and marital problems, as well as practical and emotional support. Oates et al. (2004) concluded that new mothers from non-Western societies may be protected from becoming depressed due to the role of social support present in their communities.

Gorman et al. (2004), in keeping with the goal of the TCS-PND study (to develop, translate, and validate PPD research measures for use in different countries and cultures), used and adapted the SCID (Structured Clinical Interview for DSM-IV Disorders) and the EPDS to determine whether the rates of PPD vary across different cultures. They concluded that the overall estimated rate of major depression in the postpartum period – 12.3% - was almost identical to the rate reported by O'Hara and Swain (1996) – 12% - in a meta-analysis of 59 PPD studies conducted across several European, Western, and non-Western countries over the preceding 20 years.

Goldbort's (2006) literature review on the transcultural analysis of PPD also examined women from various cultures to determine whether PPD is a universal experience. This review demonstrates that, although it may be labeled another way in different cultures, PPD is a universal experience. Non-Western cultures tend to use the term 'unhappiness' for PPD. An Ethiopian study found that postpartum mental distress was explained in terms of social adversity and was not considered to constitute a specific mental health illness afflicting postpartum women despite recognizing depressive symptoms (Hanlon, Whitley, Wondimagegn, Alem, & Prince, 2009).

The risk factors for PPD were similar cross-culturally, and included factors like a history of depression or mood disorders, an unplanned or unwanted pregnancy, significant stress in the previous year, child care stress, low social support, marital problems, and fatigue. One notable exception found which impacted on PPD was the sex of the infant. Indian, Turkish, and Chinese cultures favoured a male infant over a female infant. Goldbort (2006) further reports that non-Western cultures do not typically attribute the cause of PPD to biological or medical reasons, but rather to social and environmental reasons, such as lack of support, financial concerns, poverty, and lack of support – factors that have also been found to contribute to PPD in Western societies (e.g. Horowitz & Goodman, 2005; Logsdon, Birkimer, Simpson, & Looney, 2005).

The majority of studies reviewed by Goldbort (2006) utilized the EPDS to screen for PPD, and the prevalence rates found corresponded to PPD prevalence rates in Western cultures. Halbreich and Karkun (2006) point out that the prevalence estimates have been reported to be greater in studies where self-report measures were used compared to interview-based studies (Ghubash & Abou-Saleh, 1997; Gotlib, Whiffen, Mount, Milne, & Cordy, 1989; Wickberg & Hwang, 1997). Samples and sampling methods also often differ across cultures and studies. This may contribute to the variation in prevalence rates (Eberhard-Gran, Eskild, Tambs, Samuelsen, & Opjordsmoen, 2002). Researchers have also investigated whether women from different cultures respond to self-report questionnaires in a different manner as reporting biases may impact on prevalence rates (Halbreich & Karkun, 2006; Yoshida et al., 1997). It was thought that women's cultural context, perceptions, and beliefs, and also whether there is a stigma associated with mental health in their culture may cause women to overestimate or underestimate their responses to self-report questionnaires (Dankner, Goldberg, Fisch, & Crum, 2000; Stuchbery, Matthey, & Barnett, 1998).

5.4 Environmental and Cultural Influence on PPD Prevalence

Differences across cultural and environmental norms may explain some of the variance in PPD prevalence rates (Halbreich & Karkun, 2006). The range of psychosocial experiences that are involved in childbirth is not likely to be the same in different countries and cultures. Any number of socio-economic and environmental factors that are subject to culture-specific standards may impact on reporting styles across groups, and marked differences have been found across cultural groups (Dankner et al., 2000; Kumar, 1994). These include, amongst others, antenatal and postpartum access to healthcare, procedural differences, nutrition, religious customs, gender roles, organization of family structure, variations in the nature of marriage, quality of care available, actual or perceived levels of social support, social responses to a new birth, stress and adverse life circumstances like poverty or the perception of poverty, attitudes concerning pregnancy, motherhood, and mental illness, childrearing practices, and biological vulnerability factors. It is therefore possible to assume that there will be substantial differences between cultures in the incidence of depression, how soon after childbirth depressive symptoms occur, and other factors associated with childbirth.

Edge, Baker, and Rogers (2004) found no differences in the levels of depressive symptoms between White British women and Black Caribbean women. They did, however find clear indications that the psychological and social correlates of depression differed between these ethnic groups. This has implications for the theoretical models concerning the causes of perinatal depression, as these were predominantly constructed from studies of White women.

PPD was found to be less prevalent within certain traditional cultural settings (Halbreich & Karkun, 2006). Many non-Western societies have certain rituals and proscriptions that accompany the transition to motherhood and offers guidance and support as the mother adapts to her new role and responsibilities. Some researchers believe that this reinforces the maternal role transition and assists in relieving the new mother of psychological and physical burden which may protect her from depression (e.g. (Cox, 1996, 1999; Dankner et al., 2000; Seel as cited in Oates et al., 2004).

Stern and Kruckman (1983) regard the lack of organized social support in Westernised cultures as a significant contributing factor to the onset of PPD. Bina's review (2008) of the impact of cultural factors on PPD concludes that not having cultural traditions may lead to an increase in PPD and that cultural rituals and traditions may lessen the impact of PPD. Furthermore, cultural rituals may potentially have a negative effect on a mother's postpartum mood if she does not perceive the rituals as helpful to her.

In certain eastern cultures, for example, a postpartum woman rests in bed for the first 3–6 weeks after childbirth while her mother or mother-in-law takes care of the infant and household chores (Huang & Mathers, 2001). Pillsbury suggests that this emotional and material support may boost a mother's self-esteem and help protect her from the stressful and demanding period of early motherhood (as cited in Lee et al., 1998, p. 436). Halbreich and Karkun (2006) warn, however, that there may be a delay in the onset of depression to later during the postpartum period, despite these supportive practices. Depression at around 2 or 3 months postpartum may therefore be a reaction to receiving very little postpartum support relative to the early postpartum period and being confronted with the harsh realities and demands of motherhood.

Certain Eastern and West-Asian cultures, and the Japanese culture in particular, differ considerably from Western cultures regarding attitudes towards childbearing, marriage, and social support for new mothers (Halbreich & Karkun, 2006). In some

Asian cultures a depressed mood is regarded as self-indulgent. The interest of one's family has a higher priority than individual interests. Self-definition is defined in terms of relationships and social roles, and a person's self-esteem relies on properly fulfilling these roles rather than cultivating individual potential. Therefore, an Asian woman who fulfils her role in her family and society is typically regarded as healthy. This perception of identity is contrary to the westernized concept of encouraging individualism, introspection, self-actualisation, and other self-notions (Furnham & Malik, 1994).

According to Morsbach, Sawaragi, Riddell, and Carswell (as cited in Halbreich and Karkun, 2006, p. 108) a Japanese woman's status is increased when she delivers a healthy baby and she may endure more psychological and physical discomfort for the sake of her infant's well being. This, coupled with the prohibition on crying in the first month after delivery, may result in mothers restricting emotional expression and underreporting symptoms of PPD (Halbreich & Karkun, 2006). Personal difficulties and emotions – which are considered a weakness – are encouraged to be suppressed as there is a heavy stigma attached to the diagnosis of a mental disorder. It has been suggested that the Japanese people's reluctance to express emotion and their reputed stoicism may account for the low prevalence rate of PPD in Japan (Kumar, 1994; Hau & Levy, 2003; Yoshida et al., 2001).

The stigma attached to mental illness is not limited to the Japanese culture. In other cultures females with symptoms of depression also did not seek support from healthcare services. Chandran, Tharyan, Muliyil, and Abraham (2002) argued that this may not only be due to the stigma associated with mental illness, but also the belief that a mother's symptoms of depression are a “normal” experience associated with childbirth, or that they reflect a temporary period of maladjustment that will subside.

African cultures, like some Asian cultures, are known to place an emphasis on collective values and interests of the group as well as extended community support. This is in contrast to the Western societies’ focus on promoting individual well-being and interest (Fouche et al., 1998).

The Transcultural Study of Postnatal Depression (TCS-PND) also revealed that, in the 11 countries that participated in the study, treatment by healthcare professionals for “morbid unhappiness” in the postpartum period was not necessitated (Oates et al., 2004). Widespread difference in the availability and utilisation of services for postpartum mothers and their infants has been reported (Chisholm et al., 2004; Huang & Mathers, 2001). This is another significant cultural factor that may influence reports of PPD prevalence rates as it puts mothers at increased risk for developing PPD (Halbreich & Karkun, 2006). The availability of health care professionals like psychiatric nurses, psychologists, social workers, and others who may provide care and support for women with PPD varies between countries – and even within countries. A study in a poor and over-crowded per-urban settlement in South African led researchers to conclude that there is a need for interventions aimed at preventing or ameliorating PPD and the associated consequences of PPD in the relationship between mothers and their infants (Cooper et al., 1999).

Patients from some cultures are also likely to seek assistance from traditional healers first before consulting someone from the medical profession. African[1] women who live predominantly in rural communities have a high regard for traditional beliefs and customs as they tend to have limited contact with Westernised medicine and methods of health care. These traditional cultural practices have a strong supportive function in these communities (Fouche et al., 1998). According to Rahim and al-Sabiae, mothers who have a long and difficult labour who do not have medical attention may also be at increased risk for PPD (as cited in Halbreich & Karkun, 2006, p. 109).

Cox (1999) describes some facets of perinatal psychiatry that require a specific socio-cultural perspective which are based on the reviews by Kumar (1994), O'Hara (1994), and Cox (1996) on cross-cultural issues within this field. In addition to the particular attitudes, knowledge, and skills that practitioners require when working in the field of perinatal psychiatry, the following facets also need consideration when working with people from different cultures (Cox, 1999, p. 105):

- Perinatal rituals, for example, the postpartum check-up, routinely taking iron tablets, socially sanctioned `lying in period'.
- Rites of passage including the separation, liminal, and reincorporation phases.
- Changing family structures: impact of, and reasons for, increase in single parenting, divorce, and separation.

[1] The term 'African' as used herein, refers to those people of the African continent who share a philosophy of life termed 'African', as opposed to 'Western' or 'Eastern'.

- Kinship systems: the family and grandparents acquiring new roles.
- Naming and other religious ceremonies, for example, baptism, churching, other traditional ceremonies to declare legitimacy.
- Civil and religious understandings of the commitment implied in a long-term relationship – such as a marriage or cohabiting.
- The status of child bearing in society – dubious in the West, highly regarded in Africa and Asia.
- The structure of the family and in particular its support systems and kinship networks, like the availability of co-wives, peer support, and grandparents – especially the availability of the mothers' mother.
- Folk or popular names for perinatal mental health problems, such as blues, PPD, and psychosis.
- Choice of presenting symptoms of a perinatal mood disorder, for example, a headache, palpitations, feeding problems, not coping, and fatigue.
- Choice of healer (obstetrician for hormones; psychiatrist for antidepressants; general practitioners or health visitors for advice about baby, feeding, and sexual problems).

5.5 Symptom Definition and Expression Across Cultures

Symptom definition and symptom expression accounts for one of the foremost problems in studying PPD across different cultures (Reichenheim & Harpham, 1991; Wolf et al., 2002). In order to fully comprehend postpartum mood disorders certain cultural factors must be taken into account together with social, biomedical, and psychological perspectives (Cox, 1999, p.103). Kruckman and Smith (2006) point out that the way a woman experiences non-psychotic PPD may be both cushioned and exacerbated by a number of socio-cultural factors. In different social worlds the manner in which a woman's depression is confronted, discussed, and managed varies. Furthermore, the course of the depression is influenced by cultural meanings and practices.

Several writers have indicated that culture determines what constitutes an illness as well as the appropriate response to that illness (e.g. Furnham and Kuyken, and Prince as cited in Furnham and Malik, 1994, p.107). Therefore, a person's cultural background, with its taboos and expectations, influences the way in which psychological factors and biological changes are perceived and acted upon. Culture influences the manner in which symptoms are experienced as well as the idioms used to describe them. This in turn has an impact on how that person responds to it, how the illness is described to a health practitioner, the decisions about treatment and the likelihood of certain outcomes like suicide (Furnham and Bochner, and Rack as cited in Furnham and Malik, 1994, p. 107; Kleinman, 2004).

Littlewood, a cross-cultural psychiatrist states that "current evidence suggests that the somatic symptoms of endogenous depression do seem to be universal" (as cited in Furnham and Malik, p. 107). Bashiri and Spielvogel (1999) argue to the contrary and claims that dysphoria and depressive illness manifest and are interpreted differently in non-Western and Western societies. Cultural attitudes, beliefs, ways of thinking, and cultural norms for behaviour and emotional responses have an impact on how an individual experiences depression and seeks help. Furthermore, the languages of some cultures do not have as many words to describe depressive experiences as others.

It seems clear that PPD is not a culture-bound Western syndrome. It should therefore not be assumed that the method for evaluating it is culture-free. If broad or unstandardised diagnostic categories are used it creates uncertainty about the boundaries for a syndrome or illness and may also lead to observer error.

Understanding postpartum experiences, how depressive symptoms are expressed, and how it is assessed across different cultural groups are important considerations when screening for PPD as these may vary across cultural groups. Some cultures have their own indigenous definitions of PPD along with explanations of what causes PPD that are not outlined within the Western DSM-IV classification system (Bashiri & Spielvogel, 1999). Using standardized Western diagnostic classification systems and methods may be culturally insensitive as it increases the risk that some signs or symptoms which are prevalent in non-Western cultures will be missed (Okano et al., 1998). This may even be the case when the examiner is a local, but is more Westernised than the individual being assessed (Ghubash & Abou-Saleh, 1997).

People from Western societies tend to describe their distress in symptoms of depression whereas in non Western societies, it is expressed in somatic complaints. Asian, African and Hispanic cultures are more likely to express depression through somatisation (Bashiri & Spielvogel,1999; Park & Dimigen, 1995). Chang found that the difference in depression ratings across different cultures was mainly attributed to somatisation (as cited in Furnham and Malik, 1994). The Black classification group in his study was characterized by a mixture of affective and somatic complaints, the White classification group by cognitive and existential concerns, and the Chinese group by somatic complaints. Chinese people do not report feeling sad, but rather complain that their hearts are being squeezed and that they feel weighed down and exhausted (Kleinman and Good, as cited in Bashiri and Spielvogel, 1999, p.82) or they express boredom, discomfort, and symptoms of dizziness, pain, and fatigue (Kleinman, 2004). Lee, Yip, Chiu, Leung, and Chung (2001) add that Chinese women tend to mention physical symptoms of depression like "wind illness", "wind inside the head", or head numbness. Japanese women are not inclined to express their depressed feelings, but rather express emotional complaints by referring to concerns about childcare or physical problems and symptoms (Yoshida et al., 1997; Yoshida et al., 2001).

Somatisation and hypochondriasis are typical of how depression is expressed in African cultures (O'Hara as cited in Bashiri and Spielvogel 1999, pp. 82-83). Nigerians typically describe symptoms of depression by referring to nausea or vomiting and feeling "hot in the head" (Jinadu and Daramola as cited in Halbreich and Karkun, 2006, p. 107). Nigerians suffering from depression may also describe their symptoms as ants that keep creeping in parts of their brain (Kleinman and Good as cited in Bashiri and Spielvogel, 1999, p. 82). North American and Europeans are more likely to emphasize affective symptoms (Park & Dimigen, 1995). In Western research "Have you ever felt that life isn't worth living?" is a common screening question but one which had no meaning for mothers from Bengali who could not conceive of such a possibility (Watson and Evans as cited in Halbreich and Karkun, 2006, p. 107).

The cultural variation of depressive symptomatology can be found in the frequency of appearance of certain symptoms. Jablensky, Sartorius, Gulbiant and Ernberg (as cited in Bashiri & Spielvogel, 1999, p. 83) found that guilt feelings were more prevalent in a Swiss sample (68%) than in an Iranian sample (32%), who had more somatic symptoms (57%) with only 27% of the Canadian sample reporting somatisation. Suicidal ideation was more prevalent in a Canadian sample (70%) than in a Japanese sample (40%).

5.6 Cultural Factors, Beliefs, and Rituals Associated With Pregnancy and Childbirth in South Africa

Numerous studies across different countries have indicated that PPD is a universal experience. It is expected that there will be very little difference, if any, between the White population of English and Afrikaans-speaking South Africans in their beliefs about, and rituals associated with childbirth as they have essentially experienced the same social knowledge due to being socialized in the same culture. The same may be said for the Coloured population with a westernized upbringing. Some of the Black participants in this study come from areas of adverse circumstances in urban townships where traditional African customs and upbringing may be more prevalent, but not likely as prevalent as in rural areas.

Poverty, unemployment, unwanted pregnancy – often due to rape – and AIDS remains a problem amongst many South Africans – particularly the Black population. Private health care is expensive and free medical care is not always easily accessible even though it is available. These factors are an additional burden to these mothers. South

Africa is also affected by extreme and violent crime. Antenatal exposure to extreme societal stressors, like attempted murder or witnessing a violent crime, is indicated as one of the strongest predictors of PPD in an urban South African cohort (Ramchandani et al., 2009).

Antenatal rituals for White South African women, both English and Afrikaans speaking, are similar to those of North American women and other Western countries. Rituals include baby showers with gift giving to celebrate the imminent arrival of a new baby, regular visits to a general practitioner or obstetrician for antenatal check-ups, antenatal classes in preparation for childbirth, and a 6 week postpartum visit to an obstetric practice. Childbirth most often takes place in a hospital and the mother typically remains in hospital for 3 to 4 days after delivery while nursing staff assist her with recovery and with her baby. The mother's return home from the hospital seems to be a time when she is most vulnerable in the role transition of becoming a mother. Many mothers find themselves feeling isolated and lack support from family members which is common practice in some other cultures. It is customarily regarded that a mother is ready to resume full domestic and marital responsibilities at 6 weeks postpartum. Financial pressure forces more and more families to rely on a double income and working mothers are expected to return to work after 3 months of maternity leave, or sooner if no maternity leave is granted.

Collective responsibility and interdependence are fundamental beliefs of African cultures. Grandmothers play an important role, but generally the entire family and even those who are not biological relatives may all participate in a number of child-rearing functions (Wile & Arechigo, 1999). Hence the African proverb: It takes an entire village to raise a child.

A culturally specific action which is adhered to by some African parents when their unmarried daughter becomes pregnant, is to demand both “umgezo” (cleansing of ritual impurity and bad luck thought to be caused by premarital pregnancy) and “inhlawulo” (the payment of damages) from the family of the man responsible for impregnating their daughter. This requires the payment of money, cattle or goat, as recognition of responsibility and of good faith on the man’s behalf (Preston-Whyte & Zondi, 1989).

In traditional Zulu culture, men have not been allowed to be present at childbirth as it has always been the concern of women alone. The midwives are older women of the “umuzi” past child-bearing age. Mothers and children are isolated until the baby’s umbilical cord falls off – usually for a period of 5 to 10 days. During this time the mother is considered unclean and potentially harmful to their husband’s ancestors in the homestead. The mother may only eat food prepared by the midwife using a special spoon and dish, and may not touch ordinary utensils. After the period of isolation the mother is purified through a sprinkling of “intelezi” and can then resume her normal life and the father may see his child for the first time (Klopper, 1998).

Literature about the beliefs and rituals associated with pregnancy and childbirth in other African cultures of South Africa is scarce. The author could also not find any relevant research regarding how PPD manifests in the different South African cultures. It has, however, been indicated that the traditional African cultural value system is fading in the more cosmopolitan areas of Africa compared to the rural areas (e.g. Owoeye, Aina, & Morakinyo, 2006).

5.7 Use of PPD Screening Measures Across Different Cultures

Comparative cross-cultural research that uses a measuring instrument developed in one culture and subsequently translated to a different language for use in another culture should never assume that the measuring instrument is tapping the same construct(s) in exactly the same manner for each cultural group. Byrne and Campbell (1999, p. 571) argue that in this type of research, the principal psychometric issue should focus on:

- The extent to which the conceptual notion of the construct being measured (e.g., depression) is portable across the cultures of interest; and
- The extent to which the operationalisation of the construct, as measured by the items of the selected instrument (e.g., the BDI), is portable across cultures.

Byrne and Campbell (1999) recommend that researchers and practitioners not only look beneath the surface of item scores, but also always question how appropriate the conceptual and philosophical aspects of the assessment measure is when utilised in a different culture.

The instrument that has been used most frequently in international research into PPD is the Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987). The EPDS has been translated into numerous languages and has been used to screen for PPD in many countries. Many studies where the EPDS was used to screen for PPD indicate a significant variation in the level of depressive symptoms within and between countries, and recommend that different cut-off scores are warranted for different cultural groups.

Halbreich and Karkun's (2006) review of the literature on the prevalence of PPD and depressive symptoms in a variety of countries found that cut-off scores ranged from 9 to 13. The EPDS developers recommended culturally sensitive cut-off points with a range of 9–10 to 13–14 for different populations. The Western standard cut-off score is 12 or 13. Researchers have determined optimal EPDS cut-off scores for various cultures in order to improve the instrument's specificity or sensitivity, or both (e.g. Affonso et al., 2000; Cryan et al., 2001; Yoshida et al., 1997; Yoshida et al., 2001). The inconsistency in estimated EPDS specificity and sensitivity may explain the variances in the prevalence of PPD (Eberhard-Gran et al., 2002). Barnett et al (1999), for instance, concluded that a higher cut-off score (14 or 15) on the EPDS was recommended to identify women with PPD for a Vietnamese-speaking sample in Australia compared to English and Arabic-speaking samples, for whom a cut off score 9 or 10 was indicated (as cited in Boyd et al., 2005, p. 147). Yoshida et al. (2001) found that the EPDS was useful for Japanese women but recommend a much lower cut off score of 8 or 9 due to their reluctance to express depressed mood on self-report questionnaires.

Halbreich and Karkun (2006) regard the EPDS to be an excellent instrument for detecting the dimension of depression for which it was developed. It has, however, been recommended that more culturally sensitive and flexible measures are needed for the range of postpartum mental disorders as the EPDS has not proven to be a valid screening tool across different cultural groups (Bashiri & Spielvogel, 1999; Gibson et al., 2009; Halbreich & Karkun, 2006). Goldbort (2006) suggests that the PDSS be used in proposed PPD multicultural studies in the United States.

Gibson et al. (2009) offer various explanations for the wide range of values for sensitivity and specificity of the EPDS at all cut-off points across samples drawn from various countries with different cultures and socio-environmental conditions in the studies that were reviewed. The methods utilized as well as the populations varied significantly between the studies. The samples were drawn from urban and rural areas, from both poor and affluent women, and from countries with diverse cultural attitudes to the expression of feelings and distress. Screening was performed at different times in the peripartum period, in different clinical settings and countries, and was administered in different languages.

Further important factors to consider that would contribute to the heterogeneity of results is that the diagnostic interviews and criteria used were different. In addition, the screening and diagnostic instruments used in the studies were developed to detect depression according to how the condition is understood and expressed in Western societies and do not accurately screen for the presence of significant distress in other cultural settings (Evans et al., 2001; Gibson et al., 2009). A number of researchers have doubted the validity of applying a Western-based diagnostic system to a world population composed of around 80% non-Western people (Bashiri & Spielvogel, 1999). The close-to zero incidence of PPD reported in some cultures may therefore only be a reflection on a westernised concept of PPD or its EPDS representation. A culturally specific reporting style should also be considered (Halbreich & Karkun, 2006).

Applying the EPDS to cross-cultural samples has resulted in some difficulties. In some contexts difficulties have been reported in the way items are understood by respondents. For example, difficulty in understanding the meaning of items which, according to the researchers, required introspection, were reported in a study conducted in India with a Hindi translation of the EPDS (Banerjee, Banerjee, Kriplani and Saxena as cited in De Bruin, Swartz, Tomlinson, Cooper, & Molteno, 2004). Furthermore, languages differ in their range and differentiation of words to denote mood symptoms. Icelandic researchers recognized this problem when they had to revise 2 of the 10 items of the EPDS as Icelandic women had difficulty understanding the differences between four EPDS items (O'Hara, 1994).

Parry (1996) describes two significant threats to the validity of psychiatric instruments when applied in Africa. Firstly, psychopathological states and culturally distinctive behaviour must be differentiated to avoid confusion. Some behaviour which may be deemed acceptable and is tolerated in one culture may be unusual or unacceptable in another (Sartorius as cited in Parry, 1996, p. 178). Gillis, Elk, Ben-Arie, and Teggin argue that in some African cultures, for instance, "delusions" and "hallucinations" are not unusual occurrences among normal people and are thought to result from encounters with ancestors (as cited in Parry, 1996, p. 178). Secondly, the content of a psychiatric instrument may impact on responses culturally. Items on the instrument which refer to actual life experiences or particular objects, such as watching television or riding a rollercoaster, may lead to biased responses as respondents may not be familiar with or have access to the objects referred to (Buntting and Wessels, as cited in Parry 1996, p. 178). Thirdly, the format of the psychiatric instrument may impact on responses. Gillis et al. also point out that most standardised instruments have an interrogative style which may be foreign to the practice of many Africans, especially in sub-Saharan Africa (as cited in Parry 1996, p. 178).

Strategies have been proposed to deal with these types of problematic issues. Both Kirmayer and Kleinman suggest that an anthropological study may be conducted prior to undertaking an epidemiological study to explore how the population under study understands mental illness as well as investigate their cultural forms of expression and classification of illness (as cited in Parry 1996, p. 178). Based on the above, instruments may be supplemented by the addition of questions that may be analysed separately which enquire about specific cultural phenomena or the somatic expressions of mental illness (Swartz, Ben-Arie, & Teggin as cited in Parry 1996, p. 178). Interviewers or interpreters familiar with the subtle cultural nuances may be used to ask respondents to explain responses (Kortmann, as cited in Parry, 1996, p. 178-179). In some cultures culturally-sensitive interviews are essential when self report instruments do not elicit positive responses. Halbreich and Karkun (2006) recommend that 'such interviews should also include inquiries on complaints and symptoms pertinent to the local culture even if they initially do not fit current westernized molds, are time consuming, and thus more expensive' (pg 109).

In order to develop and harmonise PPD research instruments for use in various countries and cultures, it is necessary to understand the various cultures' beliefs, attitudes, and customs about pregnancy and childbirth. To enable direct comparisons of the incidence of postpartum disorders and possible manifestations thereof that are unique to some cultures, the same procedures should be utilized across samples and sample sizes should be adequate. Only then can meaningful comparisons be made (Halbreich & Karkun, 2006).

Understanding the nature of postpartum disorders across different cultures around the world will help to clarify the underlying mechanisms and differences between culture specific and universal aspects of postpartum disorders. Understanding these will assist in the identification of specific risk factors for certain cultures and thereby help to identify women who are at risk (Halbreich & Karkun, 2006). Furthermore, it will facilitate the development of culture specific preventative and treatment interventions.

5.8 Conclusion

Childbirth takes place in a socio-cultural context. It is therefore important to consider how cultural factors, beliefs, taboos, and rituals contribute to the understanding of childbirth and perinatal mental illness. It must also be given due consideration in the adaptation of screening measures for cross-cultural use. It is clear the PPD is a universal concept, even though it may have different names and manifestations in different cultures. The main purpose of this study was not to discuss the anthropological nature of PPD in different cultures. This chapter has, however, provided an overview of the cultural patterning of childbirth, which, considering South Africa's cultural diversity is likely to influence the assessment of PPD.

CHAPTER 6

AFRIKAANS-SPEAKING SOUTH AFRICANS

6.1 Chapter Preview

The naming of the diverse peoples who have populated South Africa in the past and present is often a difficult and delicate matter. The term "Afrikaner" has come into use with the passage of time and the development of a separate identity since the arrival of the first European settlers in South Africa. This chapter focuses on the history of the Afrikaans-speaking people, the development of the Afrikaans language, and demographic features of the Afrikaans population in South Africa today.

6.2 Definition of Terms

6.2.1 Afrikaner.

The term Afrikaner has often been used interchangeably with that of "Boer", which literally meant farmer, but then came to characterize a particular species of the genus Afrikaner. English South Africans often refer to the Bantu population as Africans. However, the translation of African is Afrikaner, a word which Afrikaners were not prepared to use generically. More recently, the favoured term to describe the Bantu-speaking population has been "Blacks" or "Africans" (Le May, 1995; Giliomee, 2003).

Le May (1995) states that anyone who is rash enough to attempt to interpret the Afrikaner people is perplexed at once by difficulties of definition. The definitive Afrikaans dictionary published in 1950 defines an Afrikaner as “One who is Afrikaans by descent or birth; one who belongs to the Afrikaans-speaking population group”. Defining the Afrikaner by language alone is too broad as it would include, for example, the Cape Coloured people. The Shorter Oxford English Dictionary defines an Afrikaner as a White native of South Africa. Giliomee (2003) states that the term Afrikaners for Whites was first used in the early eighteenth century, but it had to vie with designations like burgher, Christian, Dutchmen, and Boer. It was not until the mid-twentieth century that the term was reserved only for White Afrikaans-speakers. From the 1980’s the term started to become racially inclusive.

There are people who are classified as belonging to one of the Black groups who speak Afrikaans as a first language. However, they comprise a very small fraction of the Afrikaans-speaking population.

The term Afrikaner has been used to discuss the history of the Afrikaans-speaking people, but has otherwise deliberately been avoided in this thesis. This is because the term has many emotional and political connotations, and has been used by White Afrikaans-speakers as an exclusive term to distinguish themselves, not only from White English-speakers, but also from Coloured Afrikaans-speakers.

For the purposes of this study, the Afrikaans-speaking population will be those who have Afrikaans as their first home language and consider themselves to be Afrikaans-speakers, regardless of which population group they belong to. It is expected however, that they will typically be members of the Coloured or White classification groups.

6.2.2 Culture.

Human beings are social creatures (Baron & Byrne, 1994). They generally live out their span as members of groups. Countless studies in social psychology have shown that the groups which individuals belong to greatly affect their attitudes, values, perceptions of the world, and ultimately the person's very identity of who they are. Cultural, racial, and ethnic groups are social definitions that may be used to categorise people. Science Daily (n.d.) points out that ethnic groups are defined substantially by distinctive cultural attributes, behavioural, linguistic, or religious practices. Members of an ethnic group typically maintain a strong cultural continuity over time.

Culture consists of interrelated components of material artifacts, social, and behavioural patterns, and mental products. Cushner and Brislin (1996) refer to culture as

> A set of human-made objective and subjective elements that in the past have (a) increased the probability of survival, (b) resulted in satisfaction for the participation in an ecological niche, and thus (c) become shared among those who communicate with each other because they had a common language and lived in the same time-place. (p. 10)

6.2.3 Cultural group.

The term "cultural group" refers to the common philosophical tenets, which underlie the intellectual collective functioning of the group and includes such things as religious beliefs, traditions, historical folktales, language, and rituals. White (1959) describes culture as

> An extrasomatic, temporal continuum of things and events dependent on symboling. Specifically and concretely, culture consists of tools, implements, utensils, clothing, ornaments, customs, institutions, beliefs, rituals, games, works of art, language, etc. All peoples in all times and places have possessed culture; no other species has or has had culture. (p. 3)

6.2.4 Ethnic group.

Ethnic group refers to perceived cultural characteristics, specific to a particular group. These characteristics commonly include nationality, religion, and dress. Ethnic groups are usually subgroups of racial groups rather than vice versa (Kinloch, 1974). Pogge (1997) states that to constitute an "ethnic group", a set of persons must satisfy three conditions, namely: commonality of descent, commonality of continuous culture, and closure. Pogge (1997) elaborates that

> Members of the set must understand themselves as descendants of members of an historical society (in a broad sense, including tribes, principalities, and the like, as well as systems of interacting tribes or principalities). They must share a common

culture, or partial culture, which they take to be connected, through a continuous history, with the culture of their ancestors (however different from the latter it may have become in the process). And the group must contain all, or nearly all, of the persons who, within the relevant state, are taken to share the descent and culture definitive of the group. (p. 193-194)

Pogge (1997) points out that the first condition is necessary to distinguish ethnic groups from mainly religious and from mainly linguistic groups. The second condition is necessary to distinguish ethnic groups from mainly racial groups, and the third is necessary to distinguish ethnic groups from subgroups.

6.2.5 Racial group.

Shillington (1988) refers to the term "racial group" as perceived physical characteristics, specific to a particular group. She points out that pigmentation differences are the most commonly utilised, and that the consequences of such a social definition include awareness of subordinate group differences by the race group itself and their utilisation by the elite to rationalise prejudice and discrimination.

The Coloured/White dichotomy can be described to some extent by any of the above terms. However, none of these terms refers to the legal distinction that is made between the terms "Coloured" and "White" in South Africa.

6.2.6 Classification group.

According to Omond (1985) the term “classification group” refers to “a racial group defined by law” (p. 21). The term “classification group” is often preferred over “ethnic group”, “racial group” or “cultural group” in South Africa. The reason for this is that membership of the White as opposed to the Coloured group is not determined only or necessarily by membership of an ethnic, racial or cultural group. It is determined by present day law and as such is uniquely South African (Shillington, 1988).

6.3 Historical Overview

In 1488 Bartholomeu Dias, a Portuguese sailor, was the first recorded European to traverse the South African coast in his desperate search for a sea-route from Europe to the riches of the East. A permanent settlement was soon established on the southern tip of the continent by the Dutch while many hundreds of ships – Dutch, French, British, Portuguese – called on this coastline for fresh supplies of water, wood, and food en route to the East (Rissik, 1994).

The South African history of the Afrikaans-speaking people began in 1652 when Jan van Riebeek, a member of the Dutch East India Company (DEIC), arrived at the Cape of Good Hope with some ninety men to establish a permanent base, a fort, and a foothold on the southern tip of Africa. Most of these early settlers were immigrants from Western Europe most of whom were Dutch but also included French, German, Swedish, Danish, and Belgian immigrants. They were sent to the Cape of Good Hope to establish a refreshment station from which they could supply Dutch sailors with fresh vegetables to prevent scurvy.

Most of the European immigrants came from the lower rungs of society and many were illiterate or semi-literate peasants, artisans or laborers employed by the Company as sailors or soldiers. For the first three decades most of the immigrants were single Dutch males. In 1688 a party of fewer than two hundred French Huguenots arrived to join the DEIC settlers (Le May, 1995). They had fled from religious persecution in France and were composed mostly of families. Religion was, in fact, a binding force among the early white-skinned settlers and placed them in contrast with the heathen, dark-skinned indigenous people (Giliomee, 2003).

The Dutch-speaking settlers made an effort to prevent the French immigrants from speaking French and from forming a cohesive group. They were forced to speak Dutch in public places such as schools and churches "so that they could learn our language and morals, and be integrated with the Dutch nation" (Böeseken, as cited in Giliomee, 2003, p. 11). Some authorities took a more lenient stance toward the French immigrants and permitted them to form a church congregation but a tougher policy was imposed in 1701. This policy instructed that the necessary measures be taken to ensure that the French language would gradually become extinct and disappear. This policy of forced cultural assimilation was largely successful and by 1750 no one under the age of forty could still speak French.

Apart from the French women, the female European immigrants were Dutch. During the eighteenth century the German language also made an appearance on the Cape scene with the arrival of single male Germans immigrants. A typical German immigrant of these times had been driven to Holland in search of employment through poverty and the absence of other means of help and waited to be recruited as a soldier or a sailor by the VOC. The Germans were largely single males, spoke diverse dialects, and married Dutch of French women. The language of their children was Dutch, or what the German traveler Henry Lichtenstein, early in the nineteenth century, called "an abbreviated forcible Afrikaans Dutch" (Trapido, as cited in Giliomee, 2003, p. 12). No effort was made this time to accommodate the immigrants' religious sensibilities. Permission for a Lutheran Church was not granted until 1780 by which time the principle of one language and one church for the European community had become well established (Giliomee, 2003).

Colonization was never the Company's policy, yet colonization was made necessary by the exercise of strict economy. The colony therefore expanded largely in spite of, rather than because of, the policy of the DEIC who began allocating land to settlers and permanent farms were being established. The settlers made the new land their own and cut most of their family and community ties with Europe. Their numbers started increasing through immigration, through starting their own families, and some mixed with the local Khoisan people (also known as Hottentots). Mixing continued and was diversified by the arrival of slaves from Madagascar, Mozambique, West Africa, Angola, Malaysia, Indonesia, and Java as shortage of labour proved to be a major problem.

During the first seventy-five years of Company rule there was no rigid racial division. Fenwick and Rosenhain (1991) report that in the early days of the colony several White men married Black women and that Coloured or mixed race slaves were born within the company, often to a free White father and a slave mother. People of mixed racial origins were prominent both as burghers and free Blacks, and did not appear to suffer any racial discrimination. The frequency of racial mixing was due in the first place to the huge gender imbalance in the White population. By 1700 there were twice as many men as women in the adult burgher population in the Cape district. In the interior, the ratio was three to one. Marriages between White men and fair-skinned non-White women were common during the first seventy-five years. Sexual liaisons outside of wedlock and casual sex were common, especially in the slave lodge where local European men as well as sailors and soldiers satisfied their sexual urges. Sailors from various Western European countries were allowed ashore to "relax" and, according to an early Dutch writer quoted by Venter (1974), "Female slaves are always ready to offer their bodies for a trifle, and towards evening one can see a string of soldiers and sailors entering the company's slave lodge where they misspend their time until the clock strikes nine" (p. 20).

In 1685, High Commissioner Hendrik Van Reede of the DEIC visited the Cape Colony and noted that there were approximately 57 mixed race children in the colony (Fenwick & Rosenhain, 1991). He prohibited marriages between Europeans and "heelslag" or full-blooded slave women (of pure Asian or African origin). He permitted marriages with "halfslag" (meaning that the father was White) women with the intention of assimilating such half-castes into the European population. The ban was, however, never enforced. These children were brought up with a knowledge of the Dutch language and Dutch customs, which made it easier for the colonists to train them as servants.

Giliomee (2003) writes that it was through the relationships with these slaves and semi-free servants that the Dutch language was turned into Afrikaans.

By the middle of the 18th century, liaisons between the settlers and other racial groups were strongly frowned upon by the White public who were concerned about the mixing of races. The predominant language was Dutch until the British took over the administration of the troubled Cape Colony in 1795 following the French capture of Holland during the Napoleonic Wars. This gesture was made to keep open the strategic sea-route to Britain’s vast, valuable Indian territories. It was returned to the Dutch government in 1803, but Britain recaptured it in 1806 and administered it in various geographic shapes and political forms until union in 1910.

Giliomee (2003) commented that, at the time of the British conquest of the Cape, all the ingredients for the development of a new group were present. These ingredients were: a specific spoken language, a particular religious doctrine, identification with Pan-Dutch traditions and an awareness of the “differences” between people of different races. These ingredients differentiated the earlier settlers not only from the indigenous people and slaves, but also from the British settlers. The Afrikaners – the name now more common than in the eighteenth century – became a colonized people in a different sense. They were now British subjects, enjoying the rights that went with the status but ruled by a foreign nation.

According to Le May (1995) by 1806, the year of permanent British occupation, the White population was estimated at 18 000, of which the majority were Dutch. In 1820 nearly 5000 British immigrants landed at the Cape Colony where Port Elizabeth is today, having been promised portions of land to farm. Rissik (1994) reports that they endured years of poverty and hardship as their farms were on the frontier and they were effectively the buffer between the Colony and the Xhosa tribes. The battle for land between the two groups led to a number of wars and skirmishes – and large doses of ill will.

Despite their difficulties the 1820 Settlers, as they were called, made their mark as craftsmen, traders, and farmers. Their cultural contributions soon became firmly embedded in the nature of South Africa. They also played a major role in the administration of the Cape as the British style of governing changed from autocratic colonial power to an ever more representative system in the 1850s. The British influence was strong, not only in government, law, and administration, but also in the broader social and cultural sense (Rissik, 1994).

A move that was introduced by the British, which finally led to the official abolition of slavery in the Cape in 1834, caused great discontent among the Dutch-speaking settlers. They objected and moved away in small, separate groups in what was to be known collectively as “The Great Trek”.

The Great Trek, the first mass migration of immigrant South Africans, began in 1835 and only ended in 1848. This was a deliberate and premeditated exodus from British rule. Those who took part in it became known as the Voortrekkers, the pioneers. At that time they also referred to themselves as emigrant farmers or “trek Boers”.

They eventually formed communities in what was known the Transvaal, Orange Free State, Northern Cape, and Natal. The communities were still somewhat discrete units and it was only after the discovery of gold and diamonds and the concomitant influx of "uitlanders" or foreigners, that a real sense of nationalism was felt. The early Voortrekkers tended to band together to prevent too much contact with the non-farming newcomers. The Anglo-Boer War further strengthened the feelings of cohesiveness among the Boers, as they have become known. During the early part of the 20th century these feelings developed into a pride in the past, to the formation of a specific culture based on religious teachings and to the birth of a new African language, Afrikaans.

6.4 The Development of Afrikaans

6.4.1 The history of the Afrikaans language.

The ground for the beginnings of a new language, Afrikaans, were set in 1652 when the DEIC established a halfway house at the Cape. Those first Dutch settlers came into contact with the languages spoken by the indigenous Khoi people and those of the later settlers. High Dutch may have been the official language, but as the settlement grew and the settlers dispersed a new language developed. High Dutch became mingled with loan words from French, German, English, Malay, and Portuguese-Creole, and was constantly influenced by the dialects of the indigenous inhabitants (Le May, 1995; Rissik, 1994).

The transformation of Dutch at the Cape seems to have been quite rapid, although not as rapid as those who previously sought to explain Afrikaans as a Creole language would have had us believe. At the beginning of the twentieth century it was argued that the process was completed in a period of about thirty years after the initial settlement. Not many people agree with that school of thought anymore, as documentary evidence has been found which proves that, although all the salient features of Afrikaans which demarcate it from Dutch were present by the middle of the eighteenth century, many of them continued to compete with the original Dutch structures until the late nineteenth century (Mesthrie, 1995).

Linguistics believed that by 1850 Afrikaans had developed in most part into the language it is today (Rissik, 1994). What is now Afrikaans was, according to Mesthrie (1995), in the 1860s an unstandardized language of hearth and home, with various designations. Le May (1995) states that in the 1870s serious attempts were made to transform this new language into a literary language. Furthermore, he adds that it took another half-century before Afrikaans replaced High Dutch as an official language in South Africa (Mesthrie, 1995).

Afrikaans struggled against the English and Dutch languages – the early colonial powers – for recognition as a medium of cultural expression. General Hertzog remarked that the Afrikaners had to wage a language struggle in an attempt to stop considering themselves as “agterryers” (standing in the back line). Dutch and English speakers would look down upon Afrikaans speakers as Afrikaans was merely considered a dialect and language of the poor Whites. Consequently the Afrikaners developed feelings of inferiority and persecution in the early days of their culture and language development. General Hertzog insisted that a sound sense of White nationhood in South Africa would have to be based on the recognition of both English and Afrikaans cultures. He also encouraged the Afrikaans-speaking community to establish a separate identity to overcome the relative social, cultural, and economic backwardness they experienced. The outcome of political battles in South Africa succeeded in shaping a more exclusive

Afrikaner identity (Giliomee, 2003). By the beginning of the 20th century Afrikaans was generally recognized as a cultural language and vernacular (Rissik, 1994). Furthermore, there was a strong identification with Afrikaans as a public symbol of nationality with South Africa – its only home – and with indigenous or local forms of cultural expression – such as adherence to the Reformed faith (Giliomee, 2003).

Towards the end of the 19th century and the beginning of the 20th century there was much debate about whether Afrikaans was a language in its own right. D.F. Malherbe, who had studied linguistics, maintained that Afrikaans, with its simple and regular structure, was not a dialect, but indeed a language unto its own. D.F. Malan was also a key figure in the promotion of Afrikaans. He 1904 he remarked that Afrikaners would only become strong if they were united. He further stated that Afrikaners needed to realize they had their own heritage, based on their nationality, character, religion, and language (Booyens, as cited in Giliomee, 2003, p. 366). In 1908 Malan took the first step in his public career when he issued this ringing call: "Raise the Afrikaans language to a written language, let it become the vehicle for our culture, our history, our national ideals and you will also raise the people who speak it" (Pienaar, as cited in Giliomee, 2003, p. 366).

By 1907 a number of language associations had been established in Bloemfontein, Cape Town, and Pretoria to promote Afrikaans. The language battle was not over yet and Afrikaans was still opposed. A loyalist section in the English press, "The Star", questioned, with reference to call for English-speakers to become bilingual, whether Dutch or Afrikaans was meant – "Any man who knows the real Dutch language is painfully aware of what a truly stupid patois this South African "taal" is, and it must be a source of surprise and astonishment to the serious inquirer why such a degenerated branch of an originally sound language is so stubbornly maintained in its provoking ugliness" (The Star, 30 September 1910, as cited in Giliomee, 2003, p. 367). The *Cape Times* frequently published letters from Readers in which Afrikaans was denounced as "kitchen", "degenerate", "hotch-potch", "decaying", and "a mongrel" language which was only fit for "peasants and up-country kraals" (*Cape Times*, 4 May, 1901, and Zietsman, as cited in Giliomee, 2003, p. 367).

It was at around this time that Cornelis Jacob Langenhoven made a great effort to win the argument that Afrikaners should use Afrikaans for all purposes. He maintained that the fight to maintain Dutch was futile because of its complexity and the simple grammatical structure of Afrikaans, as opposed to Dutch, offered a better alternative to English. He challenged those who argued that it should not be taught at schools and universities and cried

> It [Afrikaans] is our highest honor, our greatest possession, the one and only White man's language which was made in South Africa and which had not come ready made from overseas … [it is] the bond which joins us as a nation together, the expressed soul of our volk. (Kannemeyer, as cited in Giliomee, 2003, p. 369)

He blundered, however, by calling it a White man's language instead of recognizing its multiracial origins also spoken by Coloured people.

Eugene Marais, editor of "Land en Volk" in Pretoria, advocated the use of Afrikaans as a written language and used Afrikaans and Dutch in the paper. Marais's path later crossed with Gustav Preller at the Pretoria-based De Volkstem, a newspaper that was started by the Boer Generals. Preller had a passion for Afrikaans. As a language activist he fought against the view that depicted Afrikaans as a low-class tongue. He insisted that a distinction be made between the language of the street and servants, and the language of civilized Afrikaans. He called for a new identity for Afrikaners as modern, increasingly urbanized people and strongly supported the use of Afrikaans as a written language, which also served to develop a distinctive nationality among Afrikaners in South Africa (Giliomee, 2003). The aim was nothing less than "to build a nation from words." (Hofmeyr as cited in Giliomee, 2003, p. 372)

In 1914 Langenhoven successfully proposed that Afrikaans be used as an alternative to Dutch for instruction in primary schools (Giliomee, 2003). In the mid-1920s a re-created Afrikaans had become a fully standardized national language and it was generally recognized as a cultural language and vernacular (Mesthrie, 1995).

The Bible was translated into Afrikaans in 1933. This was a major step towards standardizing the language and also served to enhance its credibility among the many Boers, Coloureds, and others who spoke it.

6.4.2 The influence of other languages.

The majority of the Afrikaans vocabulary, according to Rissik (1994), is derived from Dutch, but changed quite substantially, especially in pronunciation. Although there are strong grammatical similarities between Afrikaans and Dutch, it has a far less complex structure, making it a fairly easy language to learn. Mesthrie (1995) reports that it is estimated that about 90-95 per cent of the present-day Afrikaans vocabulary originated from 17th century colloquial Dutch as opposed to contemporary or even from 19th centuryDutch. Words of English, German, French, Portuguese, and Malay origin are also liberally sprinkled throughout Afrikaans. The Coloured population had considerable influence in shaping the Afrikaans vocabulary as it is used today due to their mixed backgrounds (Rissik, 1994). The Afrikaans language borrowed words from almost all of South Africa's diverse cultures as a result of the mixed racial origins of the Afrikaner. Numerous Afrikaans words were coined – particularly for local plants and animals. There are words from African languages like "mampara", which means "an untrained or stupid person" and most often is used as a form of gentle rebuke, or "babelas" from the Nguni language which means "hangover" (Rissik, 1994).

Mesthrie (1995) comments English has had a great influence on Afrikaans and that Afrikaans has, in many ways, developed in a similar direction to English in its degree of analysis, for example, the loss of gender. This grammatical change had, however, started long before the arrival of the English at the Cape in 1795. The two languages also share many structures and vocabulary as both English and Dutch are closely related Germanic dialects.

The absence of Cape Dutch written texts prior to 1830, when the British were in possession of the Cape Colony for about 30 years, makes studying the origins of Afrikaans a difficult task. Examining the effect that English was having on Afrikaans in the 19th and early 20th centuries was an equally arduous – if not impossible task. At that stage, Afrikaans speakers were not in the position they are in today to provide written evidence of the inroads that English was making into their language (Mesthrie, 1995). Only two nineteenth-century works exist which acknowledge that English had an effect on Dutch at the Cape. One was written by A.E. Changuion, which dates from 1844, and the other by N. Mansvelt, which dates from 1884. Both authors were schoolteachers from Holland. When they arrived at the Cape they were appalled at the state of their language as spoken in the Colony and set about trying to purify the Dutch language of their colonial brethren.

It has been stated that Afrikaans had come into being as a new tongue by the late eighteenth century (Raidt, 1989) and all that occurred thereafter was a settling of the dust on this new reality. Mesthrie (1995), however, argues to the contrary. He asserts that the linguistic transformation that would take place after the British occupation of the Cape in 1795 was to be as great as – if not eventually greater than – all the changes that had previously taken place. He describes the English influence on Afrikaans as a story without end, as follows:

> It is taking place now to a degree that is perhaps without precedent in the history of European languages. Such an argument may not be regarded favorably by many in South Africa because, I contend, Afrikaners refuse to see the many inroads that English had made and is making into their language, in terms of language change (or language change in progress), and persist in regarding them as mere interference phenomena which can and should be removed by education. (p. 223)

According to Mesthrie (1995) the mutual influence of the two languages is inevitable. However, the influence was greater in one direction than the other and occurred to such and extent that it eventually passed from the realm of interference into that of true language change, producing a hybrid that is a unique product of South African society. Afrikaans as it is now spoken is a true reflection of the reality of present-day South Africa. It is both the overwhelming influence of English on Afrikaans, and the traditional differences between Afrikaans and Dutch, that serve to demarcate Afrikaans from Dutch and enhance its character as a separate language. The Netherlands, the Dutch people and their language have become a foreign country, people, and language to Afrikaans-speaking South Africans. Their English-speaking compatriots have inadvertently assimilated Afrikaners, culturally and to an ever-increasing degree linguistically.

6.4.3 Landmarks in the extension of the functions of Afrikaans.

Afrikaans had achieved certain landmarks, which include its adoption as a language of instruction in schools from 1914. English had become the official language in 1910, but this made the Boers so unhappy that by 1925 they had seen to it that Afrikaans had become the second and equal official language. Further landmarks include the publication of the first complete Bible in Afrikaans (1933). A remarkable proliferation of governmental vocabulary began when virtually all state publications had to appear both in English and in Afrikaans. Somewhat later language activists fought to establish Afrikaans as a language of technology and specialized disciplines. By 1985, at least 250 technical dictionaries covering a wide range of fields had been produced.

Today numerous South Africans use either English or Afrikaans, or both, as a means of cross-cultural communication. Afrikaans is also used extensively on radio and television, and has become a language of religion, education, and science. There are Afrikaans language newspapers across a broad political spectrum, as well as many famous Afrikaans authors of all races.

6.5 Linguistic Diversity in South Africa

South Africa is certainly a land of linguistic and cultural diversity and the nation's people are often talking across a language and cultural barrier. South Africa's language situation is characterized not only by the number and variety of African, Asian, and European languages that coexist, but also by alternative varieties of these languages – including the Afrikaans of the Coloured population (Mesthrie, 1995). South Africa has eleven official languages, namely, Afrikaans, English, Ndebele, Northern-Sotho, Southern-Sotho, Swazi, Tsonga, Tswana, Venda, Xhosa, and Zulu.

Zulu is the most common home-language and is spoken by 23.8% of the population. Xhosa follows and is spoken as a home-language by 17.6% of the population. Afrikaans is spoken by 13.4%, Sepedi by 9.4%, and English and Swazi are each spoken by 8.2% of the population (Lehohla, 2009).

Some of South Africa's linguistic characteristics are similar to those of other developed nations despite the high degree of linguistic diversity in the country. English is South Africa's language of wider communication and is widely spoken throughout the country – by members of virtually all the different ethno-linguistic groups, and is also taught in schools. Furthermore, there is a high level and degree of bilingualism and even multilingualism. This reflects the extensive intergroup contact that continues, in spite of the legacy of apartheid, to characterize South African society. The literacy rate in South Africa is considered impressive by third world standards. It is still low, though, and is skewed disproportionately toward certain groups at the expense of others Mesthrie (1995) further adds that "the notion of South Africa as a fourth world society (i.e., one in which elements of both the first and third worlds coexist) clearly makes a great deal of sense from the perspective of the country's linguistic situation." (p. 321)

The future of South Africa's language situation is likely to remain essentially unchanged, according to Mesthrie (1995). Linguistic changes that do occur will fall into one of four well-documented linguistic processes: language change, language spread, language emergence, and language death. Mesthrie (1995) further emphasizes that regardless of the nature of political change in South Africa, it is virtually certain that linguistic diversity will remain a feature of social life for generations to come, and that bilingualism and multilingualism will remain commonplace for many, if not most, South Africans well into the future.

As indicated earlier, Afrikaans represents the third largest language group in the South Africa. Contrary to many foreigners' beliefs, a large number of Afrikaans speakers are not White. A large percentage of the Coloured population speaks Afrikaans as a first language. Although it is accepted that the Afrikaans language is common to many Coloured and White people, there is some controversy as to whether the people belonging to these classifications groups are similar.

6.6 Afrikaans-Speaking People: The Coloured – White Dichotomy

The Coloured person has been defined by The Population Registration Act number 30 of 1950 (Omond, 1985) as someone who is not a Bantu and also not a White person. Similarly, according to the Population Registration Amendment Act, Number 64 of 1967 (Omond, 1985), a White person has been defined as someone who

> In appearance obviously is a White person and is not generally accepted as a Coloured person, or is generally accepted as a White person and is not in appearance obviously not a White person. [Furthermore]… his habits, education, and speech and [his] deportment and demeanor shall be taken into account. (p. 22)

This description still does not give much clarity concerning the differences between Coloured and White people (if, in fact, there are differences) or what their identities are (if they have them). According to Venter (1974) points of view vary from the assertion that Coloured and White people have largely the same genetic base to an absolute refusal to believe that Coloured people have any White origins. The latter group presumes that Coloured people originate solely from Hottentots, Khoisan people, and slaves. Mason states that it became the trend in the Western Cape to refer to people of mixed descent as 'Coloureds' or 'Cape Coloureds' (as cited in Giliomee, 2003, p. 110).

The Afrikaans language is common to many Coloured and White people. Their membership to these subgroups will be addressed in more detail below.

6.6.1 Classification and identification of Coloured and White Afrikaans-speakers.

The myth of the Coloured identity was explored by Van der Ross (1979):

> It is claimed by some that there is a special identity, peculiar to the Coloured people. They have, according to this claim, an identity, which they share with no other population group, and this sets them apart in a very special sense. (p. 2)

He refutes the assumptions that Coloured people all have the same origin, are necessarily easily recognizable, and that they have their own specific culture. He also asserts that the mixed nature of the Coloured person's composition precludes him from having a separate identity.

The heterogeneity of the Coloured people is further emphasized by the *Theron Report* (1976). This report discusses the common bonds that may hold Coloured people together:

> The most important positive binding element between Coloureds is probably their being South African. The most negative binding element is probably the biological typing of Coloureds in terms of biological characteristics, for example variations in skin pigmentation, hair texture, and facial features in so far as these are perceptible and are used by other groups as criteria for exclusion from their own ranks. (par. 21.4)

White Afrikaans-speaking people may be classified according to four broad identification patterns (Giliomee, 1975):

1. The first viewpoint is held by those who see Afrikaans-speaking Whites "as a distinct White volk, membership of which is clearly defined" (Giliomee, 1975, p. 32).
2. A second viewpoint is that the Afrikaans-speaking White belongs to a larger White population group into which he is increasingly being assimilated. The identity of Afrikaans-speaking Whites is seen to be tied up with the identity of the White population group as a whole.
3. A third viewpoint attempts to redefine the Afrikaans-speaking White person in terms of cultural attributes Afrikaans-speakers are seen as a cultural group which seeks to express its cultural heritage through its language. Furthermore, this cultural group is seen as a political entity existing in a plural society, with its members believing in values which transcend communal interests.
4. The last major viewpoint states that the Afrikaans-speaker should see himself as belonging to a number of different groups within the broader community, one of which is cultural. Membership of this classification group is not seen as being linked to a particular political position.

The concern with "identity" appears to be particularly marked in some sections of the White Afrikaans-speaking community, which is a contrast to the Coloured people who appear to de-emphasize or even disregard the notion of identity.

The Afrikaans-speaking Whites' identity becomes a reality in so far as it can be distinguished from their identities of other groups. The diversity of opinion presented above does not enable a definitive statement to be made in terms of the identification of Coloured or White people. It does, however, indicate the ambiguity of the situation and therefore it would seem to be appropriate to explore possible implications of classification for Afrikaans-speaking people.

6.6.2 Implications of classification.

The position the Coloured people have been in has perhaps been the most difficult of all the South African cultural groups. Yet, Coloured people do not see themselves as having a separate identity. Most Coloured people subscribe to the Western culture, although many align themselves with Black people and even consider themselves to be Black. This ambivalent position is described by Mann (2007) as being 'marginal'. A marginal personality is characterized by feelings of insecurity, hypersensitivity, and self pity which develop due to someone who desperately wants to be accepted by a privileged group but who is excluded from finding membership within that group (Mann, 2007).

The ambivalent position of the modern Coloured person is to some extent due to attitudes on the part of White and Coloured people to themselves and each other. However, the attitudes of Coloured people to Black people are also complex. Generally, relationships with Black people have been cordial. Despite this Coloured people do not completely identify with Black people.

A few decades ago, Lison (1977) studied Afrikaans- and English-speaking students – both Coloured and White – and found that the development of certain personality characteristics stems from the ambiguous position of the Coloured people:

> The Coloured students are by no means at home within the society. Their personality strongly reflects a person (sic) whose position within society is uncertain, and their severity of social maladjustment is greater. The Coloured female is untrusting of others, is introverted, socially insecure, and has difficulty in establishing close, meaningful relationships. Her male colleague too, is struggling. He is a person with feelings of inferiority and inadequacy, low self-esteem, and a disorganization of thought processes. (p. v)

Opinions on the position of the White Afrikaans-speaking population are also complex. In the 1980's, De Klerk believed that the Afrikaans-speaking White population was moving towards a greater feeling of cohesiveness within the South African community as a whole (De Klerk, 1984). Archibald (1969) described them as an emergent minority. Historically, Afrikaans-speaking Whites have largely lagged behind their English-speaking counterparts in terms of social status and economic dominance. This has changed, but it has to a large extent shaped the Afrikaans-speaking White population of today. According to Archibald (1969), the minority status experience of the Afrikaans-speaking White has been morally disabling and has had a profoundly negative effect on their personality development.

Schlemmer (1974) disputed Archibald's (1969) view. Schlemmer (1974) points out that the Afrikaans-speaking White group has largely emerged from the minority position and from traditional ties and has instead become "a bureaucratically organized White elite" (p. 204).

It is believed that the still-conservative orientation of the young Afrikaans-speaking White population has had an effect on the personality development of the Afrikaans-speaking White youth (Archibald, 1969; Orpen, 1970; Viljoen & Grobler, 1972). Orpen (1970) makes the point that Afrikaans-speaking Whites have, to a great extent, internalized the authoritarian norms that prevail in South African society, and have accepted them with little question.

In order to conclude the discussion on Afrikaans-speaking people, pertinent statistical data relating to demographic variables is presented in the next section.

6.7 Demographic Features

Presenting an accurate demographic picture of the Afrikaans-speaking people of South Africa is difficult, as the statistics available from the Central Statistical Services are presented in terms of classification group rather than language group.

6.7.1 Geographical region.

According to mid-year population estimates in 2009, there were 4 433 100 Coloured people in South Africa The largest concentration of Coloured people was found in the Western Cape province (61%), followed by the Eastern Cape Province (12%), and the Northern Cape (10%; Lehohla, 2009). Whites in South Africa as a whole was estimated at 4 472 100 people (Lehohla, 2009; Stats SA, 2009), with the majority living in the Gauteng province (41%) followed by the Western Cape (19.4%) and Kwazulu Natal (11%).

6.7.2 Language.

The distribution of home languages in South Africa, as recorded by the last major census in 2001, indicates that 13.35% of the total population spoke Afrikaans, of which 53% are Coloured and 42.4% are White. Afrikaans is the third most predominant language in South Africa, preceded by 23.8% who spoke Zulu and 17.6% who spoke Xhosa (Stats SA, Population census, 2001, as cited in Lehohla, 2009, section 2.18).

6.8 Conclusion

This chapter has addressed the Afrikaans-speaking people by providing a historical overview of the Afrikaner in South Africa. The development of Afrikaans was discussed which focused on the history of the Afrikaans language and the influence other languages had on the development of Afrikaans. Linguistic diversity in South Africa was addressed as well as the Afrikaans-speaking classification groups. This led to a discussion of the Coloured and White classification groups in South Africa as well as a brief outline of their geographical and language distribution. It is hoped that this chapter has provided the reader with a more comprehensive understanding of the Afrikaans-speaking population. The next chapter details the methods and procedures that were utilized in this study.

CHAPTER 7

RESEARCH DESIGN AND METHODOLOGY

7.1 Introduction

This chapter addresses the research design and describes the research methodology employed in this study. The sample and sampling procedure is discussed, the measures used are briefly described, and the translation and administration procedures are outlined. Finally the methods used to analyze the data are described.

7.2 Primary Objective of the Research

Postpartum depression (PPD) is not uncommon – with up to 20 percent of all mothers, in all circumstances suffering from this type of depression. PPD is not always easy to identify without screening measures and may develop slowly any time during the first year of the baby's life. Every mother is different and may have a different combination of symptoms. Some may be more anxious or irritable than sad. It may be mild or severe. Some mothers have been depressed ever since the pregnancy, and sometimes "The Blues" just don't seem to go away. Some mothers manage well initially and then their mood becomes darker and darker. If untreated, it can adversely affect a mother's functioning as well as her infant's development. Screening all mothers after birth is therefore very important to ensure that they get the necessary help and support they need. With this in mind, the primary objectives of this research were to:

- Address the problem of the unavailability of suitable PPD screening measures for the majority of Afrikaans-speakers by providing an Afrikaans version of an existing PPD screening measure – the PDSS - that was developed for use with an American culture and that has not been standardized on a South African population;
- Determine the validity and the reliability of the PDSS and the Afrikaans PDSS for English and Afrikaans speaking South African mothers based on the Rasch measurement model procedures;
- Determine how well the PDSS, EPDS and QIDS correlate when used with a South African sample;
- Determine the magnitude of the relationship between a positive screen for PPD in a South African sample and known risk factors for PPD.

7.3 Research Methods and Designs Used in the Study

Both qualitative and quantitative methodologies were used. Qualitative analysis was performed using two translation techniques, namely Brislin's back-translation method advocated by Brislin (1970) and the committee approach. It is, however, considered unlikely that any one result can provide unequivocal evidence for such linguistic equivalence of a test (Kline, 1993). Rather, a whole series of results can build up a composite picture, which overall could demonstrate the equivalence of a test. With this in mind, various quantitative methods from the Rasch rating scale measurement model were also used to examine the validity and the reliability of the Afrikaans PDSS (Linacre, 2009).

7.3.1 Multiple translation method: Brislin's back-translation method and the committee approach.

Brislin's back-translation method together with the committee approach was used in this study to qualitatively explore the linguistic equivalence of the PDSS and the Afrikaans PDSS. The multiple translation method was also used to translate the QIDS into Afrikaans. Brislin, Lonner, and Thorndike recommend that a multiple translation method be used to ensure semantic equivalence (as cited in Beck et al., 2003). The back-translation method involves the translation of items from the original into the target language by a translator. This material is then translated back into the original language by another translator. The original version and the back-translation are compared to determine the efficacy of the translation.

The back-translation method has been demonstrated to be especially useful in cross-cultural research for checking the equivalence of the translations of measures in different languages (Bracken & Barona, 1991; Prieto, 1992). The back-translation technique has been used successfully to translate from English to Afrikaans (e.g., Shillington, 1988).

A committee (or cross-translation) approach involves a group of experts (such as cultural, linguistic, and psychological) in preparing a translation (Nasser, 2005; Van de Vijver & Tanzer, 1997). The committee members discuss the instrument's questions with each other during a collaborative effort to improve the quality of the translation, minimise bias, and reduce misconceptions (Ægisdóttir et al., 2008; Onkvisit & Shaw, 2004). Researchers often combine the committee approach with the back translation technique (Van de Vijver & Leung, 1997b).

7.3.2 Item response theory and the Rasch measurement model.

An item response theory (IRT) model, specifically the Rating scale model, a formulation of an extended Rasch model, was employed in this study as implemented by Winsteps (Linacre, 2009). IRT, also referred to as latent trait theory, is a paradigm for the design, the analysis, and the scoring of questionnaires and other instruments. A fundamental purpose of IRT is that it provides a theoretical framework which enables researchers to evaluate how well tests and measuring instruments work, and more specifically, how well the individual items on these measures work (Hambleton, Swaminathan, & Rogers, 1991). In a multidimensional instrument IRT allows researchers to determine how adequately the attitude continuum which underlies each dimension is assessed by the respective items in the instrument (Beck & Gable, 2001d). IRT is frequently used for developing and refining measuring instruments (Hambleton et al., 1991) and assessing performance across groups using conformable items where all the respondents did not need to respond to all the items (Andrich, 2004).

IRT models are based on the assumption that the items that are being analysed are unidimensional, in other words, a single construct, or single dominant affect or attitude is measured (Chou & Wang, 2010; Harvey & Hammer, 1999). Most IRT models assume unidimensionality, in other words, all the items measure the same latent trait or underlying construct (Chou & Wang, 2010). The latent trait is the "unobserved characteristic that is presumed to be responsible for the observed responses that are made to the test's items" (Harvey & Hammer, 1999, p. 356). The latent trait is therefore that which is intended to be measured and "is defined by the items or agents of measurement used to elicit its manifestations or responses." (Linacre, 2009, p. 429).

Another assumption of IRT is local independence, meaning that a person's responses to one item are statistically independent to responses on any other items (Beck & Gable, 2001d; Linacre, 2009, p. 392). In local independence the latent trait measured, in this case PPD, is the only factor affecting the response to an item. This means that once the contribution of the latent trait to the data is removed, only random and normally distributed noise is left (Chou & Wang, 2010, p. 719; Linacre, 2009, p. 392). The local independence of items therefore implies that no residual associations are left in the data after the latent trait has been removed (Pallant et al., 2006). This means that all covariance among the items occurs as a result of the association between the items and the underlying construct being measured (Edelen, Thissen, Teresi, Kleinman, & Ocepek-Welikson, 2006). Local independence is associated with unidimensionality because a data set can only be unidimensional when item responses are locally independent based on a single latent variable (Embretson and Reise as cited in Chou and Wang, 2010, p. 719).

Another fundamental premise of IRT is that an item characteristic curve (ICC) or function can describe the relationship between a respondent's item performance and the group of traits that underlie the item performance (Hambleton, Swaminathan, & Rogers, 1991). Accordingly, as the level of the latent trait increases, so too does the probability that the respondent will endorse items and/or select categories that signify higher levels of agreement with the items (Beck & Gable, 2001d).

A number of different models have been developed within IRT. The different IRT models are named by the number of parameters that are used to describe the items of a questionnaire. Three popular IRT models are the one, two, and three parameter logistic models.

According to Yu (2010), IRT and Rasch measurement models are similar to each other in terms of computation, but their philosophical foundations differ immensely. Whereas IRT models may use up to three parameters, the Rasch model utilises only one parameter. The Rasch measurement model is often regarded to be a one-parameter IRT model. (Furr & Bacharach, 2007). The Rasch measurement model is, however, different to the one-parameter IRT model as it offers a completely different approach to conceptualizing the relationship between data and the theory (Andrich, 2004b; Royal, 2010). IRT attempts to fit a model to the observed data whereas the Rasch measurement model specifies that the data fit the model (Andrich, 2004b). The IRT approach would therefore adjust the model parameters to reflect the patterns that are observed in the data. A Rasch approach, on the other hand, specifies what the requirements are for fundamental measurement and emphasizes fitting the data to the model before any claims concerning the presence of a latent trait in the test or measuring instrument may be considered valid (Andrich, 2004b). Although IRT and the Rasch measurement model have diverse views regarding model-data fit, they are similar in that they take both person and item attributes into consideration in assessment methods, in contrast to classical test theory.

The Rasch measurement model is based on the assumption of a unidimensional measurement model (Bond & Fox, 2001). The Rasch measurement model assumes that if a person responds to a unidimensional construct then he or she ought to respond as expected according to his or her ability level (also referred to as trait levels) and according to the item difficulty level (Smith, Conrad, Chang, & Piazza, 2002).

Unidimensionality means the questions measure a single dominant affect or attitude (Smith et al., 2002) which, in this study, is the typical emotions or symptoms for the degree of depression experienced. Unidimensionality can always be determined on a particular level of reduction. Depression, for example, is unidimensional on a higher level but is multidimensional on a more basic level (Biondi, Picardi, Pasquini, Gaetano, & Pancheri, 2005). Smith et al (2002) state that "if an instrument is composed of multiple subscales, then unidimensionality refers to the set of items for each subscale" (p. 191). All items from the same subscale should therefore load on the construct measured by that subscale, and not on any other subscale.

In the Rasch measurement model, the item difficulty is the single item characteristic that is assumed to influence a respondent's performance (Rasch as cited in Beck & Gable, 2001d, p. 7; Smith, 2004). Or stated differently, the probability of a specific response by a specific person on a specific question is a function of the person's "ability" (level of depression) or theta (θ), and the "difficulty level" of the item (or d). In this sense "difficulty level" refers to the difficulty of endorsing an item (yes or no in a dichotomous case, or more or less as in the traditional 5-point Likert scale). The "ability" therefore indicates the degree of a latent variable, such as depression, that the item is meant to measure.

A distinguishing attribute of the Rasch measurement model is that the item difficulty parameter and the person ability parameter can be estimated separately from each other (Schumacker, 2004). As a result it yields a test free person ability calibration because the person's ability may be estimated independently of the sampling distribution of the test items. The Rasch measurement model also makes sample free item difficulty calibration possible where item difficulty is estimated independently of person abilities (Schumacker, 2004).

A number of different Rasch measurement models have transpired to address the vast number of psychometric needs across various disciplines (Schumacker, 2004). The various measurement models provide the means for constructing interval measures from raw data. The family of Rasch measurement models include the dichotomous model – the simplest measurement model, the partial credit model, the rating scale model, binomial trials, Poisson counts, ranks, many-faceted, and multidimensional models (Wright & Mok, 2004). The different Rasch measurement models are defined by either the way in which a respondent is able to record a response to an item or the different response formats, by the number of dimensions in a multidimensional model, by the number of facets in the data collection design, or a combination of these factors (Schumacker, 2004). The Rasch rating scale model employed in this study describes the probability that a person will endorse a particular rating scale category on a specific item of a rating scale. In rating scale analysis the number of thresholds refers to the number of response categories. There is only one threshold in the dichotomous Rasch model.

The Rasch model makes it possible to construct linear additive measures from "counts of qualitatively-ordered observations, provided that the structure of quantity is present in the data" (Linacre and Wright as cited in Salzberger, 2010, p. 1275). In an instrument where the test is unidimensional, or where the subscales are unidimensional, Rasch analysis is able to order the items of the scale or subscale from least to most difficult on a continuum. Rasch analysis is also able to calibrate person affect measures and place the respondents on the continuum – a linear scale – according to their item agreements. The person and item calibrations have equal interval units of measures on the continuum and are termed "logits". Logits are calibrated data with a mean of 0 and a SD of 1. A negative logit represents an item that is easy to endorse, whereas a positive logit represents an item that is hard to endorse (Smith et al., 2003; Wright & Stone, 1999). Logits less than -2 or greater than +2 are very easy or very hard to endorse respectively (Maree, 2004).

Since logits are linear metric units they may be used to compute item difficulty, trait ability, and item-fit indices to analyse the psychometric properties of an instrument for a certain population. To determine whether an instrument is appropriate for the sample, the overlap between item difficulty and trait ability distributions on the logit scale are examined (Hong & Wong, 2005).

Rasch analysis provides indicators of how well every item fits within the underlying construct using linear metric units, providing the researcher with insight regarding the relative difficulty of items and therefore allows for examining the construct validity of an instrument (Overston as cited in Kyriakides, Kaloyirou, & Lindsay, 2006, p. 785; Wu & Chang, 2008). Construct validity can only be achieved if the structure of the variable is supported by the item calibrations and if person characteristics can be substantiated by their placement on the variable (Wright & Stone, 1999). Construct validity may therefore be determined by comparing both person ability levels and item difficulty levels on the logit scale. The difficulty indices allow for the examination of how well the items span the continuum giving an indication of how well the items measure what they are intended to measure. Better construct validity is achieved if the items are well differentiated or spread out on the logit continuum. This allows for a more complete score interpretation for both high and low scoring respondents because the content of the respective items provide a more adequate definition of the construct (Beck & Gable, 2000; 2001d; 2003; Bond & Fox, 2001; Bond, 2003; Smith, 2004).

Unidimensionality is analysed by principal components factorial analysis of the residuals as well as by analysis of fit statistics or indices (mean-square infit and mean-square outfit) – a necessary quality control technique to determine the validity of test items and person responses. Fit analysis is an important part of using latent trait models, like the Rasch model, if the interpretation of the calibration of results is to be meaningful. When the parameters of a Rasch model are estimated, they may be used to calculate the expected response pattern for every item and person. Comparison of observed and expected response patterns yields the fit statistics for the persons and items (Linacre, 2009, p. 428; Wright & Stone, 1999, p. 47). Fit statistics therefore enable researchers to determine how well the data cooperates with the construction of measurement and if and where misfit occurs – in other words, person and item response patterns that do not meet the requirements of the model and do not contribute to useful measurement. Confidence may be placed in person measurement and item calibration when the fit statistics fall within an acceptable range for the study (Smith, 2004; Wright & Stone, 1999). Person fit to the Rasch model is an indication of whether the participants responded consistently to the items, or whether their responses were erratic. Item fit to the Rasch model is an indication the items performed logically. Item misfit may occur when the item is confusing, too complex, or it measures a construct other than what it was intended to measure.

Two aspects of fit are reported, namely infit and outfit. Non-standardized person fit and item fit are reported for infit and outfit as a mean-square statistic (MNSQ) and as a standardized value (*t*-statistic). MNSQ statistics are reported in this study. The infit statistic gives more weight to person performance closer to the item value. In other words, persons whose ability is close to the item's difficulty level should provide more sensitive insight regarding that item's performance. Outfit statistics are not weighted. They are more sensitive to the influence of outlying scores. Aberrant infit scores are normally a greater reason of concern than aberrant outfit scores. More attention is therefore paid to infit scores than to outfit scores by those who use the Rasch model (Bond & Fox, 2001).

MNSQ statistics of 1.00 are ideal values by Rasch specifications. Linacre (as cited in Chiang, Green, & Cox, 2009, p. 266) states that the values for differentiating fit and misfit should be sufficiently flexible to allow for researcher judgment. However, MNSQ statistics between 0.50 and 1.50 are considered to be productive for measurement (Linacre, 2002).

According to Smith, Schumacker, and Bush (1998, p. 78) the MNSQ statistic is dependent on sample size and relying on a single critical value for the MNSQ can result in an under detection of misfit. Wright (as cited in Smith et al., 1998, pp. 78-79) provides a formula for calculating the critical value for mean squares which takes the sample size into account:

$$Critical\ value\ MNSQ(infit) = 1 + \frac{2}{\sqrt{x}}$$

$$Critical\ value\ MNSQ(outfit) = 1 + \frac{6}{\sqrt{x}}$$

Where x = the sample size. If this formula is applied to the samples in this study, the critical value for the MNSQ infit would be 1.15 for both the English and Afrikaans samples. The critical value for the MNSQ outfit for the Afrikaans sample (n = 178) would be 1.45, and 1.44 for the English sample (n = 187). The MNSQ infit value in this calculation is particularly stringent, and more in line with values for high stakes questionnaires, for which a range of 0.80 to 1.20 is recommended (Wright & Linacre, 1994). Values between 0.60 and 1.40 are more suitable for rating scales (Wright & Linacre, 1994). Bond and Fox (2007) also recommend that MNSQ infit and outfit values for persons and items be in the range of 0.60 to 1.40 for a Likert scale. Based on these recommendations a range of 0.60 to 1.40 was selected for differentiating between fit and misfit items and persons.

MNSQ statistics that are greater than 1.40 may suggest a lack of construct homogeneity with other items in the scale (Doble & Fisher, and Green, as cited in Hong & Wong, 2005, p.132). Items with MNSQ statistics which are smaller than 0.60 may suggest the presence of item redundancy.

7.4 Advantages of Item Response Theory and the Rasch Measurement Model over Classical Test Theory

Classical test theory (CTT) is a methodological approach which employs conventional techniques to analyse data. CTT is popular and has its purposes but it also has shortcomings which causes it to mask vast amounts of important information (Royal, 2010). “Modern test theory”, or item response theory (IRT), is a measurement theory that was developed to address some shortcomings of CTT (Lord as cited in Beck & Gable, 2001d, p. 5).

IRT methods are gaining popularity in wide variety of psychological assessments and are not only limited to the traditional measures of aptitude and ability and measures with dichotomously scored items. This may be partly attributed to an increase in computer availability and advances in computer software in recent decades which has favoured the more computationally demanding techniques of IRT relative to those based on CTT (Harvey & Hammer, 1999). The use of IRT models capable of analysing items which are rated by means of ordered-category scales, such as the Likert scale, or unordered, nominal scales have gained increasing attention. Harvey and Hammer (1999) state that “the addition of these polytomous models renders the IRT approach applicable to virtually any type of standardized psychological assessment instrument” (p. 354). Furthermore, it has been predicted that IRT-based methods will, to a large degree, replace CTT-based methods in future years (Harvey & Hammer, 1999).

Two different approaches within IRT developed as a result of different viewpoints of the most prominent IRT pioneers, one articulated by Rasch (1960), and the other by

Lord and Novick (1968) and Birnbaum (1968). The primary difference between their approaches was concerned with the how the relationship between the data and the model was conceptualised (Andrich, 2004a). These two approaches are referred to as the traditional paradigm and the Rasch paradigm by Andrich (2004a). The traditional paradigm contends that measurement models must fit the data whereas the Rasch paradigm contends that the data must conform to the measurement model.

Although these traditional paradigm and the Rasch paradigm have have diverse views on model-data fit, they both offer several advantages over the classical test theory (CTT) in terms of test development and evaluation as well as the scoring process. The term "IRT" is used in the next section to refer broadly to these two paradigms, both which take person and item attributes into account. Where the Rasch measurement model, as an extension of IRT, has a unique advantage over the traditional paradigm in IRT, it will be indicated. Some of the advantages of IRT-based methods are highlighted below:

7.4.1 Focus on item-level.

IRT provides a more holistic, integrative view of item performance as it focuses more on the item-level than CTT, which places a greater focus on test-level indices of performance such as the overall reliability coefficient of an instrument (Harvey & Hammer, 1999). The sample dependency of item and test indices (like reliability indices, p-values, and item-total correlations) and the item dependency of person ability is a major limitation of CTT. Furthermore, although CTT is able to quantify the total sample difficulty or item discrimination, it is unable to effectively combine and present this information simultaneously in a convenient format (Smith et al., 2002).

7.4.2 Better construct interpretation.

IRT models were established on the assumption that the items being analysed are essentially unidimensional (Bond & Fox, 2001; Smith et al., 2002). This does not, however, restrict the model only being applied to instruments that measure a single variable. Instruments which are composed of multiple subtests or dimension can also be analysed using a unidimensional IRT model as each subtest or dimension is analysed separately. The assessment of unidimensionality is important as it provides evidence of the construct validity of a measure (Van der Ven & Ellis, 2000). IRT analysis enables researchers to examine the construct validity of instruments more thoroughly and can result in finer construct interpretation. This allows for a more thorough description of high- and low-scoring respondents (Beck & Gable, 2001d).

7.4.3 Better measurement precision across the continuum of the variable.

In IRT-based methods, items with higher discrimination result in higher levels of information which indicates better measurement precision, and therefore lower undesirable errors of measurement. This is because information in IRT is inversely related to the standard error of measurement (SEM). In CTT-based methods it is the concept of reliability which indicates better measurement precision, and it is reliability that is inversely related to the SEM (Harvey & Hammer, 1999). In CTT, however, only one reliability estimate is given and, because the SEM is calculated using the reliability estimate, CTT provides only one SEM which is applied to all the scores – despite the knowledge that extreme scores are less precise. CTT lacks procedures that would make it possible to determine how measurement error varies across the different ability levels (Smith et al., 2002). CTT assumes that the test is equally precise across the possible test scores and a single number, like the internal-consistency reliability coefficient, for example, is used to quantify the measurement precision of a test (Harvey & Hammer, 1999). IRT-based methods provide a SEM for each person and each item. This enables the researcher to determine the accuracy of item location or person ability estimates which cannot be accomplished with CTT-based methods.

7.4.4 Test development.

IRT-based methods have advantages of CTT-based methods when items need to be selected for test development. By using IRT-based methods the test developer is more easily able to determine the effect of deleting or adding a certain item or set of items by exploring the test's combined information function for the items being examined (TIF) and the test standard error (TSE) function for an item pool. Investigating the change in graphic curvature of the TIF or TSE functions after deleting or adding items and comparing this to the desired performance curve allows the test developer to tailor the test closely to desired specifications. Test developers using CTT-based methods need to rely on far less sensitive measures like the test's global coefficient alpha or standard error of measurement (SEM; Harvey & Hammer, 1999).

7.4.5 Information on category functioning.

Another advantage of using Rasch analysis for the validation of latent trait measures is that it provides additional information on the category functioning of an instrument. Linacre states that this may serve to increase the measurement accuracy of the instrument (as cited in Wu and Chang, 2008). Winsteps (a Rasch analysis software program) provides rating scale category counts, rating scale category fit statistics, average measures, and step measures to assist researchers in determining whether there are potential problems with the functioning of the rating scale (Linacre, 2009). Determining whether the average measures or step calibrations advance monotonically across the rating scale categories, for example, enables researchers to optimize the effectiveness of the scale categories (Linacre, 2004).

7.4.6 Scoring methods.

IRT based methods offer substantial advantages over the scoring methods typically used in CTT-based tests. More sources of information can be considered simultaneously using IRT to estimate a respondent's score, including the specific items that were answered correctly or incorrectly, and the item's properties, like difficulty level and discrimination. This makes it possible to assess the degree to which the IRT model being used provides a good fit to the individual's response pattern, to produce better estimates of the latent trait scores, and to produce quantitative estimates of the "quality" or likelihood of an observed response profile (Harvey & Hammer, 1999). A limitation of CTT-based methods is that it is not possible to determine how a person may respond to a certain item. As different metrics are used for persons and items it is not possible to predict the outcome of the interaction between item difficulty and person ability.

In Rasch measurement the scores are linear and are mapped onto a common metric with the same calibrations and steps. Co-calibration, a process in Rasch measurement, makes it possible for a number of instruments, which purport to measure the same construct, to measure in the same unit. This is possible even when the separate instruments have a different number of items, a different number of rating scale points, and rating scale points with different labels (Smith, 2004). Rasch measurement therefore assists in determining convergent validity between instruments that purport to measure the same construct (Smith, 2004).

7.4.7 Differential item functioning.

One of the major assumptions of the Rasch measurement model is the generalisability aspect of construct validity when the data fits the model. Parameter invariance is one major characteristic of the Rasch model (Smith, 2004, p. 109). Invariance of parameters (item invariance and person invariance) is an important distinction of the Rasch measurement model from other latent trait models and CTT (Bond, 2003; Smith, 2004, p. 109). This refers to the extent to which inferences regarding item calibrations or person measures are invariant, within measurement error, across different groups, contexts, tasks, or time frames. Only the Rasch measurement model has sufficient statistics for estimating item and person parameters (Smith, 2004, p. 109). Examining item and person invariance "places the boundaries and context to which the frame of reference for interpretations can be extended or constrained." (Smith 2004, p. 110).

If parameter (or measurement) invariance is established then it means that there will be an equal probability that two individuals from different cultural or demographic groups will respond in the same way to an item, given that both individuals are at the same level of the latent trait being measured. Once parameter invariance is established, the differences on an instrument's scores accurately reflect the differences on the latent characteristics assessed by the latent trait or construct. Parameter invariance is determined through analysis of differential item functioning (DIF), which refers to distortions at the item level (Ægisdóttir et al., 2008; Küçükdeveci, Sahin, Ataman, Griffiths, & Tennant, 2004).

Psychological assessments should be free from DIF as items with DIF differ in psychological meaning across cultures or groups and have the potential to impact on the comparison of total test scores across cultures or groups. Therefore individuals from different cultures groups who have the same ability have a different probability of getting the item right (Hambleton et al., 1991).

Harvey and Hammer (1999) regard CTT-based methods of assessing bias as being limited as they do not allow for distinguishing between a scenario where the subgroups have different means and the test is biased and a scenario where the means are different, yet the test is not biased. IRT techniques on the other hand offer a powerful alternative for examining DIF (Harvey & Hammer, 1999). Assessing bias, or analysis of invariance, can be conducted by using CTT-based methods by examining differences in item means by group or time, but such analyses are greatly simplified via use of the Rasch measurement model (Andrich, 2004; Chiang et al., 2009) due to its assumption of parameter invariance.

7.4.8 Administrative efficiency and item banking.

Administrative efficiency and item banking, where items may be selected from a calibrated item pool for every individual being assessed, differ significantly between IRT-based and CTT-based testing. Using a CTT-based approach, it is difficult to compare the performance of persons taking different forms of an assessment. It is also not possible to compare scores obtained from the same set of items unless the entire data set is available or a certain type of imputation method is used (Smith et al., 2002). The assumption in CTT-based testing is that the whole item pool will be administered to each individual whereas IRT-based testing allows for selecting different items for different individuals, or selecting a different number of items for different individuals. This is because the IRT model results in instruments that are sample free and test free (Schumacker, 2010). A sample-free measure means that the item characteristics do not vary with the sample being researched, in other words, the instrument transcends the group measured. When an instrument is test free, several items at a variety of difficulty levels may be omitted from the scale without influencing the respondent's score. Furthermore, in a test-free instrument, it is not necessary for every person to complete the entire scale (Wright as cited in Beck and Gable, 2001d). The family of Rasch measurement models in IRT is rather robust when data is missing, and comparative estimates can still be made even for those individuals who did not respond to all items, as well as for items that only some individuals responded to (Wright & Mok, 2004).

In the context of Rasch measurement, the process of co-calibration of instruments that purport to measure the same construct, makes it possible to select a different mix of items depending on the desired precision at different locations on the variable (Smith, 2004). This has the advantage of tailoring the selection of test items and reducing testing time by limiting the number of items, or by administering more tests in the same amount of time while still producing a test with its highest degree of measurement precision for a specified latent trait (Harvey & Hammer, 1999). Typical scoring methods used in CTT-based approaches are highly dependent on each individual having the same list and number of items.

7.4.9 Additivity.

Additivity is an advantage that the Rasch measurement model has over other IRT models and over CTT. Additivity refers to the properties of the measurement units which are called logits (logarithm of odds). IRT techniques like Rasch analysis are capable of constructing linear measures from counts of qualitatively-ordered observations, provided the data fit the Rasch model (Linacre and Wright as cited in Salzberger, 2010, p.1275).

The ordering of items on a continuum (item difficulty calibration) and calibrating person affect measures by means of linear metric units (logits), which maintain the same size over the entire continuum, allows for computing multivariate parametric statistical techniques (Smith et al, 2002). Of all the IRT-based models, only Rasch analysis strives to provide invariance in scientific measurement with respect to estimates of item difficulty and person ability. The Rasch family of measurement models is the only model that produces linear measures, gives estimates of precision, and is able to separate the parameters of the measuring instrument and the object being measured (Wright & Mok, 2004).The use of an interval level of measurement for person and item estimates, as opposed to an ordinal level measurement like in CTT-based techniques, means that invariance of item and person estimate values always remain relative (Bond & Fox, 2007, p. 71). Although CTT-based statistical models typically assume an interval scale of measurement in order to allow parametric statistical techniques, they are based on raw scores that are mostly from ordinal scales of measurement that do not support the mathematical operations needed to compute means and standard deviations. When logits as opposed to raw scores are used, researchers are better able to calculate means and variances and this allows for a more accurate determination of reliability (Schumacker, 2004; Smith, 2004). This will be discussed in more detail in the section that follows.

7.4.10 Superior reliability estimates.

Reliability, according to Schumacker (2004), is typically defined as ‘the consistency of responses to a set of items or the consistency of scores from the same instrument or parallel-forms instrument. Reliability is also defined as the degree to which scores are free from measurement error.' (p. 243)

In CTT, five different types of reliability coefficients are generally used, depending on the test situation. These are: 1) test-retest reliability, 2) rater consistency, 3) alternate forms reliability, 4) split-half, and 5) internal consistency. Rasch measurement models can also be used to compute reliability for various test administration designs. Rasch-based methods allow for identifying measurement error in the same type of testing situations, provides reliability estimates and individual SE's, and is able to yield more diagnostic information regarding individual person and rater performance. Rasch-based methods have been described as more advantageous to traditional methods in CTT (e.g. Harvey & Hammer, 1999; Schumacker, 2004; Smith, 2004) as they allow the researcher to pinpoint those individuals who exhibit consistent, declining, or improved performance on different forms of the same test, or on retesting. Those individuals who show declining or improved performance may then, for diagnostic purposes, be identified and the reasons explored. Furthermore, the rater reliability design provides more extensive information regarding individual raters, like rater consistency, severity levels, and potential bias (Schumacker & Smith, 2007).

Rasch analysis for calculating internal consistency also has distinctive advantages over traditional measures of internal consistency. In CTT, Cronbach alpha is the traditional measure of internal consistency, or reliability coefficient, indicating the extent to which items measure a single construct. It examines the average inter-item correlation of the items in a questionnaire (Cortina, 1993). If all items in a questionnaire are measuring the same construct (without any error), then Cronbach alpha will be equal to one. If there is no shared variance in the items, then only measurement error is reflected which results in Cronbach alpha being equal to zero (Hinton as cited in Spiliotopoulou, 2009). A Cronbach alpha value of one does, however not necessarily imply unidimensionality of the questionnaire (Helms, Henze, Sass, & Mifsud, 2006). The presence of more than one construct may be determined by factor analysis. When there is a one factor solution, the Cronbach alpha is likely to be high, which indicates that the items are measuring the same latent construct. A Cronbach alpha value equal or greater than 0.70 is conventionally regarded as an acceptable level of internal consistency (Bland and Altman as cited in Spiliotopoulou, 2009). Caution should be taken, however, when judging estimates of internal consistency as a low coefficient alpha value might not always indicate problems with the construction of the tool and a high value does not always suggest adequate reliability. Spiliotopoulou (2009) indicates that these reports might rather be a reflection of the data characteristics of the construct and suggests that researchers, reviewers, and practitioners should consider several guidelines for interpreting internal consistency estimates. These guidelines may include consideration of the variability of the data, whether the statistical tool is appropriate for the level of measurement of data, whether the data are normally distributed and linear, the scale's length and width, and the sample size.

Cronbach alpha utilises nonlinear raw scores in calculating the sample variance, and like other traditional estimates of reliability, normally include extreme scores (i.e. zero scores and perfect scores) that do not have any error variance (Schumacker & Smith, 2007). Including these scores therefore decreases the average error variance which results in an increase in the reported reliability (Schumacker & Smith, 2007; Smith, 2004). The sample variance is therefore potentially misleading (Smith, 2004). Rasch analysis, on the other hand, typically excludes extreme scores due to their SEMs being infinitely large and that they provide little information regarding the person's location on the underlying variable (Linacre as cited Smith, 2004, p. 99).

Concern has also been raised about the use of raw scores in the SEM (Schumacker & Smith, 2007). The classical method of estimating the SEM uses the reliability coefficient and the score standard deviation (SD):

$$SEM = SD_x(1 - R)^{1/2}$$

Where SD_x represents the observed spread of the sample raw scores and R represents the reliability estimate. The average error variance for the test, and hence the confidence intervals around the scores are represented by the SEM. When determining the precision of every score on the scale, this method may be misleading due to extreme scores being less precise than central scores (Smith, 2004).

Smith (2004), Schumacker (2004), and Schumacker and Smith (2007) address these concerns within the context of Rasch measurement, where each person's ability and each item's difficulty are indicated on a linear scale as logits, as opposed to raw scores on an ordinal scale – provided the data fit the model. These estimates, due to them being on a linear scale, are more appropriate for calculating means and variances (Smith, 2004). Schumacker (2004, p244.) concurs that 'reliability determination in the Rasch model is more directly interpretable because logits (linear measures) rather than raw scores (ordinal measures) are used. Logits rather than raw scores are used in Rasch analysis because logits satisfy the following criteria for measurement: logical ordering, linear scales, and objective comparisons. The calibration of items and persons on a common linear scale provides information on criterion-referenced and norm-referenced information for person measures and item calibrations.

A further advantage of Rasch-based methods is that it yields a direct estimate of the modelled error variance for each estimate of a person's ability and item's difficulty rather than sample dependent averages used in CTT (Schumacker, 2004; Wright in Smith 2004, p. 96). CTT lacks procedures for determining how measurement error varies across person ability levels (Smith et al., 2002). The SEs in Rasch models provide a quantification of the precision of every person measure and item difficulty. They can also be used to describe the confidence intervals in which each item's 'true' difficulty, or person's 'true' ability lies. The individual SE may be more useful than a sample or test average which overestimates the error variance of persons with extreme scores. Should a group estimate of reliability be required, the individual SE may be squared and summed to yield a correct average error variance for the sample (as opposed to the error variance for an 'average' person sampled) which is then used to calculate formulas for internal consistency. A superior estimate of internal consistency is produced due to numerical values being linear (provided the data fit the Rasch model), and due to the actual average error variance of the sample being used as opposed to the error variance of an 'average' person. The result is a person variance that is adjusted for measurement error, which represents the 'true' variance in the person measures. Furthermore, in Rasch measurement, person separation reliability (person reliability estimate) is calculated as the ratio of the adjusted true variance to the observed variance. This represents the proportion of variance that is not due to error.

Correlation-based reliability estimates (including KR-20 and Rasch person reliability) are unfortunately nonlinear and suffer from ceiling effects as their estimates are restricted in range from zero to one. The Rasch measurement model addresses these shortcomings by yielding a person separation index and an item separation index which have a range from zero to infinity.

"Separation" in Rasch analysis is the measure of the spread of both items and persons in standard error units. The separation index should exceed 1.00 for the instrument to be minimally useful. Higher separation values represent a better spread of persons and items along a continuum. Lower separation values indicate redundancy in the items and less variability of persons on the trait. The separation value provides evidence of reliability, with higher values yielding higher reliability.

The item separation index allows the researcher to determine whether the items discriminate different levels of person performance and therefore provides evidence of "test" reliability (Linacre, 2009). Conventionally, only a person reliability estimate is reported, which also provides an indication of test reliability (Linacre, 2010). Larger item separation indices demonstrate better confidence in the spread of items across the targeted continuum (Beck & Gable, 2001d; Bond & Fox, 2001).

Rasch analysis also produces a person separation index. The person separation index enables the researcher to determine whether persons are able to discriminate differences in item calibration (Linacre, 2009). The person separation index is on a ratio scale and is able to compare the true distribution of person measures (in logits) with their measurement error, which results in an indication of the spread of person measures in SE units (Fisher as cited in Smith, 2004). The higher the person separation index, the more spread out the persons are on the variable being measured and the better the reliability. According to Linacre (2009) a separation index of 2 signifies that high measures are statistically different from low measures.

It is useful to examine the person separation index across several analyses of the same data, as an increase in person separation index signifies an increase in reliability even when Rasch person reliability remains unchanged due to it's maximum value of one (Smith, 2004; Schumacker & Smith, 2007).

In Rasch measurement, the person reliability estimate (person separation reliability) provides evidence for internal consistency reliability. It is calculated as the ratio of the adjusted true variance to the observed variance:

Person Reliability Estimate = True Variance / Observed Variance.

This represents the proportion of variance that is not due to error. The Rasch person reliability estimate is conceptually equivalent to Cronbach's alpha, but is computed without extreme scores making its value lower than that for Cronbach's alpha. Winsteps provides a person reliability estimate as well as an item reliability. CTT does not typically compute an estimate of item reliability (Linacre, 2009).

7.5 Participants and Sampling Procedures

For the purpose of this study three different categories of participants are needed: 1) participants for the translation process; 2) participants for the administration of the PDSS, the EPDS, and the QIDS; and 3) participants for the administration of the Afrikaans version of the PDSS, the Afrikaans version of EPDS, and the Afrikaans version of the QIDS. Two different sampling procedures were employed in this study. The first pertains to the translation process, and the second to the administration of the screening questionnaires. These are outlined below.

7.5.1 Participants for the translating process.

The main purpose of this study was to provide an Afrikaans version of an existing PPD screening measure – the PDSS – and to determine the reliability and validity of the Afrikaans PDSS and the English PDSS on respective South African mothers. The QIDS-SR and the EPDS are two additional screening questionnaires that were selected as convergent instruments to provide additional data on the construct validity and the equivalence of constructs across the translations. At the time this study was undertaken, an Afrikaans version of the EPDS was available from Postnatal Depression Support Association South Africa (PNDSA), but not an Afrikaans version of the QIDS-SR. Therefore accredited translators and persons with a thorough knowledge of the subject matter were required for the translation of the PDSS as well as the QIDS.

A non-probability sampling technique – purposive sampling – was used to select the translators, back-translators, as well as experts to review the translated versions. The reason for this is that the researcher wanted to ensure that individuals who translated the PDSS met certain requirements: a) they had to be bilingual in English and Afrikaans; b) they must have had experience in translating in these languages; and c) they must have some knowledge about the subject matter they will be translating. Knowledge about the subject matter being translated is important to avoid literal translations being made, which could cause misunderstanding in the target population (Hambleton, 1994). In order to meet these requirements, translators registered with the South African Translators Institute, who were accredited in English and Afrikaans translating, and who had translated subject matter in the field of psychology were selected for the back-translation process. The purpose of the PDSS or QIDS was briefly explained to the translators responsible for translating the respective questionnaires. The experts who were selected to evaluate and review the translations were all psychologists with a clinical background and experience in the assessment of depression and were bilingual in English and Afrikaans.

Six people were involved in the back-translation process and refining of the Afrikaans version of the PDSS. Two bilingual professionals and translators registered with the South African Translators Institute (SATI) translated and back-translated the PDSS. One psychologist from PNDSA and two psychologists, both senior lecturers in the field with experience in the validation of psychological measures, evaluated and reviewed the translations. The author of the PDSS, Professor Cheryl Beck, was also involved in the final revision.

The translation of the QIDS into Afrikaans was performed by four individuals. Two translators from SATI (different translators to those mentioned above) translated and back translated the QIDS into Afrikaans. The two psychologists/senior lecturers who evaluated the translations of the PDSS also evaluated and reviewed the Afrikaans translation and back-translation of the QIDS. The author of the QIDS was also consulted for advice and permission on the metric conversion of items 8 and 9.

7.5.2 Participants for the English PPD screening process.

A total of 187 English-speaking postpartum mothers of mixed parity were selected through convenience sampling for screening. Participants were therefore selected on the basis of availability. An advantage of this technique lies in the relative ease with which the sample can be made available. However, a disadvantage of this technique is that the sampling method may be seen as arbitrary and not a true representation of the population which limits the generalisability of the results.

Participants were eligible if they met the following criteria:

- Mothers between 4 and 16 weeks postpartum;
- A South African citizen, residing in South Africa;
- Able to speak and read English or Afrikaans fluently; and
- Gave birth to a healthy baby without a disability.

7.5.3 Participants for the Afrikaans PPD screening process.

The same inclusion criteria as listed above applied to this sample. This sample comprised 178 Afrikaans-speaking postpartum mothers of mixed parity who were also selected through convenience sampling. A total of 365 mothers (187 English and 178 Afrikaans) were therefore screened for PPD in this study.

7.6 Measures

Data were collected with a demographic questionnaire, comprising socio-demographic and obstetric data, the Postpartum Depression Screening Scale (PDSS), the Afrikaans version of the PDSS, the Edinburgh Postpartum Depression Scale (EPDS), and the Quick Inventory for Depressive Symptomatology – Self Report (QIDS-SR16).

7.6.1 Demographic questionnaire.

The demographic questionnaire collected data about the mother's home language, language proficiency in either English or Afrikaans, ethnic group, marital status, education level, employment status, mother's age, obstetric history, perception of level of care during labour and delivery, perception of level of support after childbirth, psychiatric history, baby's health, baby's current age and gestational age at birth, and the baby's sex. Questions relating to known risk factors for PPD were also included.

7.6.2 The Postpartum Depression Screening Scale (PDSS).

The PDSS (Beck & Gable, 2000) is a self-report, 35-item Likert response scale consisting of seven dimensions, each containing five items. The dimensions include Sleeping/Eating Disturbances, Anxiety/Insecurity, Emotional Lability, Cognitive Impairment, Loss of Self, Guilt/Shame, and Contemplating Harming Oneself. Each item describes how a woman may feel after the birth of her child. The mother is asked to indicate her degree of agreement or disagreement on a five-point scale from (1) strongly disagree to (5) strongly agree. The woman is asked to circle her answer which best describes how she has felt over the past 2 weeks. After completing the PDSS, or its Afrikaans translation, the mothers in this study were asked to indicate if there were any items that they found difficult to understand.

The PDSS can be completed by the mother in 5 to 10 minutes. The scale may be administered by any health practitioner the postpartum woman comes into contact with. The conceptual basis of the PDSS is based on Beck's series of qualitative studies on PPD (Beck, 1992, 1993, 1996c).

The PDSS is intended to provide an overall score for PPD, but also considers the multidimensionality of PPD and gives seven subscale scores. The summative scoring results in a total score range from 35 to 175. The total score may be sorted into one of three categories: 1) normal adjustment (total score of <59), 2) significant symptoms of PPD (total score of 60 to 79), and 3) positive screening for PPD (total score of ≥80). The psychometric properties of the PDSS are presented in chapter 3.

7.6.3 The Edinburgh Postnatal Depression Scale (EPDS).

The Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987) designed to screen for the risk of PPD in women by measuring emotional and cognitive symptoms of PPD and sleep difficulty. The EPDS excludes somatic symptoms of depression as this may be affected by normal postpartum recovery rather than signify a mood disorder. The EPDS does contain items which pertain to anxiety specifically, but opinions are divided on whether the EPDS screens for the presence of anxiety as well as depression (Brouwers et al., 2001; Pallant et al., 2006).

The EPDS is a 10-item self report measure with a 4-point Likert scale. Each of the 10 questions has 4 answer choices that are scored between 0 and 3. The EPDS total score ranges from 0 to 30. The total score is obtained by adding the scores for each item. The cut-off point of the EPDS was calculated to be 12 or 13 for probable depression, and at 9 or 10 for possible depression (Cox et al., 1987). Boyd et al. (2005) have suggested, however, that different cut-off scores may be warranted for different cultural groups. Cultural groups may vary in the manner that depressive symptoms and postpartum experiences are expressed. This has an impact on the assessment of PPD as the symptom presentation may differ across cultural groups, and therefore also the optimal scores for PPD screening (Affonso et al., 2000; Barnett et al., 1999; Bashiri and Spielvogel, 1999).

The EPDS is a widely used screening scale for PPD and demonstrates moderate to good reliability properties across samples from a wide variety of countries and languages (e.g., Barnett et al., 1999; Benvenuti et al., 1999; Berle et al., 2003; Garcia-Esteve et al., 2003). The EPDS has moderate to good correlations with other depression measures (e.g.

Flynn, Sexton, Ratliff, Porter, & Zivin, 2011) and has been found to be a valid screening instrument, when administered verbally, in an urban South African community in a study by Lawrie et al. (1998). The above mentioned factors, and that the EPDS is available freely as a screening tool for PPD, and is brief and easily administered made it a desirable instrument to include in this study.

7.6.4 The Quick Inventory for Depressive Symptomatology – Self Report (QIDS-SR16).

The Quick Inventory of Depressive Symptomatology (QIDS; Rush et al., 2003) is derived from the 30-item Inventory of Depressive Symptomatology (IDS). The 16-item QIDS is a shorter, more time-efficient version of the IDS and is used in daily practice and in clinical research. The 16 items were identified as needed to rate the nine criterion domains of major depression: sleep disturbance, psychomotor disturbance (agitation and retardation); appetite or weight disturbance or both (appetite increase or decrease and weight increase or decrease), depressed mood, decreased interest, decreased energy, worthlessness or guilt, concentration or decision making, and suicidal ideation.

Just as there are two versions of the IDS with identical items: a clinician rating (IDS-C30) and a self-report (IDS-SR30), there are also two versions of the QIDS: Quick Inventory of Depressive Symptomatology – Clinician Rating (QIDS-C16) and the Quick Inventory of Depressive Symptomatology – Self Report (QIDS-SR16; Rush et al., 2003).

The (QIDS-SR16) was selected for use in this study due to it being a self report measure of depressive symptom severity; it provides a specific assessment of al the core criterion DSM-IV symptoms of MDD; its brevity was considered ideal for the population being screened; it has demonstrated highly acceptable psychometric properties and it has proven useful as a brief rating scale of depressive symptom severity in both research and clinical settings (Rush et al., 2003). The IDS – and the QIDS – were designed to assess depression for a patient population, but the IDS has thus far proven to have excellent sensitivity, good specificity and moderate PPV when administered to women during the postpartum period (Yonkers et al., 2001). Research has shown that the QIDS-SR (16) correlates well with the IDS-SR (30) (0.96) and the Ham-D (24) (0.86), and that the QIDS-SR (16) is as sensitive to symptom change as the IDS-SR (30) and HAM-D (24), signifying high concurrent validity for all three scales (Rush et al., 2003).

The QIDS-SR16 total scores range from 0 to 27. The total scores were obtained by adding the scores for each of the nine symptom domains of the DSM-IV MDD criteria. To score domains which consist of more than one item, the highest score of the item relevant for each domain is taken. The QIDS-SR16 takes approximately 5-7 minutes to complete. Table 8 presents the thresholds that are recommended for major depression screening with the two versions of the QIDS.

Table 8 Severity Thresholds for the QIDS-C16/QIDS-SR16

	QIDS-C16	QIDS-SR16
No depression	≤ 5	≤ 5
Mild	6-10	6-10
Moderate	11-15	11-15
Severe	16-20	16-20
Very Severe	≥ 21	≥ 21

7.7 Procedure

7.7.1 Procedure for the translation of the PDSS.

A multiple method translation incorporating the back-translation method together with a committee (or cross-translation) approach was used in this study. Permission was sought from Western Psychological Services (WPS) for the translation of the PDSS into Afrikaans. Once approval was given, an accredited translator registered with SATI translated the PDSS into Afrikaans. This material was then back-translated into the original language (English) by another accredited translator registered with SATI. The promoter of the study examined, and commented on, the quality of the Afrikaans translation. The original version was compared to the Afrikaans translation, and finally the original version to the back-translation to determine whether there were significant discrepancies. Suggestions were made to improve 8 of the 35 items. A bilingual psychologist, who is also a senior lecturer in the field, with experience in adapting and translating psychological questionnaires, compared the original version to the back translation. The original version was then compared to the Afrikaans version and discrepancies in linguistic equivalence were pointed out and better alternatives to 14 items were suggested to improve the Afrikaans translation.

The translation was evaluated further by a board member from PNDSA, a bilingual psychologist with extensive experience in the assessment and treatment of PPD. It was suggested that she review all three versions – the original, the back-translation, and the Afrikaans version – along with the comments and suggestions for further improvement made by the study promoter and the bilingual psychologist. The PNDSA psychologist then made recommendations for further improvement to the Afrikaans translation and suggested alternatives to 19 items that would keep the translated version as close as possible to the original version while keeping the language simple and easy to understand.

All the evaluators' comments and recommendations for improvement were then incorporated. The promoter of the study and the researcher evaluated these and, together with the recommendations made for improved quality of the Afrikaans translations, selected the most suitable Afrikaans translations. However, discrepancies with items 16 (I felt like I was jumping out of my skin.) and item 33 (I did not feel real.) remained. The author of the PDSS, Cheryl Beck, was contacted to provide insight into the real meaning of these two items. With her clarification the most suitable Afrikaans translation was selected. The standard translation, back-translation, adjustment sequence has been utilised in many studies requiring the translation of instruments into African languages (Parry, 1996).

7.7.2 Procedure for the translation of the QIDS-SR.

Permission was sought from the author of the QIDS-SR for the translation of the screening scale into Afrikaans so that it could be used as an additional screening scale for the purposes of this study. Once approval was obtained, an accredited translator registered with SATI translated the QIDS-SR into Afrikaans. The Afrikaans version was then back-translated into English by another accredited translator registered with SATI. The researcher and the promoter of the study examined, and commented on, the quality of the Afrikaans translation. The promoter of the study compared the original version to the Afrikaans translation, and the original version to the back-translation to determine whether there were significant discrepancies. The bilingual psychologist/senior lecturer who assisted with the translation of the PDSS also compared the original QIDS-SR to the back translation and the original version to the Afrikaans version. Discrepancies in linguistic equivalence were pointed out and better alternatives were suggested to improve the Afrikaans translation. The researcher and promoter of the study evaluated these comments and recommendations and selected the most suitable Afrikaans translation.

7.7.3 Procedure for the screening process.

A variety of professionals who have contact with postpartum women were approached and informed about the research. These included nursing staff at antenatal, postnatal, and immunisation clinics, obstetricians, general practitioners, staff at maternity hospitals, psychologists, and individuals offering antenatal and postnatal exercise classes. They were asked to assist by identifying participants for inclusion in the study.

Pamphlets were distributed to those professionals who agreed to assist with the study by referring mothers for participation, regardless of whether or not they presented with symptoms of depression or anxiety. Referring professionals were also given information about the research and recruiting lists on which mothers could complete their contact details if they wished to participate or be contacted with more information about the research. Mothers who were interested in participating could opt to contact the researcher in person or could leave her contact details on the recruiting list provided. Regular contact was maintained with referring professionals in order to obtain this information and to encourage the referral of additional mothers.

Mothers who expressed an interest in participating in the research were contacted by the researcher to determine whether they met the following criteria:

- A South African citizen, resident in South Africa;
- Able to speak and read English or Afrikaans fluently;
- Were between 4 and 16 weeks postpartum;
- Gave birth to a healthy baby without a disability.

Referring professionals were also encouraged to refer antenatal mothers for participation in the study. The researcher made contact with these mothers to discuss the research and to make arrangements for participation once their babies were delivered. The researcher then followed up with these mothers between 4 and 16 weeks after their expected due date.

Where possible, the researcher assessed mothers in person by paper/pencil administration. Mothers with internet access could opt to complete the questionnaires confidentially online via a secure, password-protected website. Participants who wished to participate online obtained this information from the researcher. Online assessments allowed the researcher to assess mothers from across South Africa. This method of assessment also meant that mothers could be assessed with minimal disruption in the postpartum period as they were able to complete the questionnaires at home, online, and in their own time. Individuals who suffer from disorders that affect their ability to complete self-report measures reliably and validly were not asked to volunteer for this study. During the paper/pencil administration the participation criteria were discussed with the mothers. The researcher was able to determine from interaction with the mothers whether they were coherent and understood the administration procedure. Those mothers who opted to participate online were provided with information regarding the research and the participant criteria prior to completing the research questionnaires. The researcher was of the opinion that those mothers who could communicate their intention to participate via email or telephonically with the researcher and who could subsequently navigate successfully through the website pages during participation did not suffer from disorders that would negatively impact their ability to participate.

The researcher found that about one third of the mothers who had indicated their desire to participate online refrained from doing so. The researcher assumed that this was due to the time-consuming and demanding role of the early postpartum period. The researcher followed up with these mothers by sending a written reminder about the research and added that participation is especially valued considering the demands of early motherhood. Mothers who still refrained from participating were not pursued.

Anastasi (1988) states individuals who are assessed for research purposes should be assessed by a suitably qualified person and the participants should receive feedback from the assessment. In accordance with this recommendation, individuals were informed prior to completing the screening questionnaires that they would receive feedback in the form of brief reports of the results of their screening. It was assumed that this would serve as an incentive for participants to complete the questionnaires accurately according to how their mood has been, thereby allowing for more reliable results.

Mothers who screened positively for symptoms of PPD were referred for counselling, to a PPD support group – if one was available in their vicinity – and for further assessment by their doctor if required. Mothers were also given the contact details for PNDSA for additional support and information.

7.8 Ethical Considerations

Good psychological research can only be made possible with mutual respect between the participant and the researcher. The participant should also have confidence in the researcher. Therefore a number of ethical guidelines must be considered when conducting research with human participants (British Psychological Society, 2009). The following ethical principles were followed to ensure that the guidelines as stipulated by the British Psychological Society (BPS, 2006) were adhered to:

- Required ethical approval was obtained from the Western Psychological Services (WPS), who holds copyright for the PDSS, to translate the PDSS to Afrikaans for use in local populations, and to adapt the PDSS for on-line administration, in English and Afrikaans, via a secure on-line environment for administration and scoring.
- Permission was obtained from the author of the QIDS-SR for the translation of the screening scale into Afrikaans so that it could be used as an additional screening scale for the purposes of this study. The author of the QIDS was also consulted for advice and permission on the metric conversion of items 8 and 9.
- Approval was obtained by The Royal College of Psychiatrists for using the EPDS as an additional screening scale for the purposes of this study and for online administration on a password-protected website.
- Ethical clearance was obtained from the University of Pretoria.
- The researcher provided mothers with information regarding the objectives of the study and obtained their informed consent prior to participation. For online participation a procedure was followed, recommended by Kraut et al., (2004), whereby mothers clicked a button on an online form to indicate that they have read and understood the consent form before they could complete the research questionnaires.
- Being involved in a research experience may be a safe and anonymous means for participants to explore thoughts and feelings that they may not want to confide to family and friends (Cooper, Turpin, Bucks, & Kent, 2005). The researcher may also find evidence of psychological problems of which a participant may be

unaware (BPS, 2007). With these factors in mind, and the knowledge that many women with PPD are reluctant to reveal their postpartum distress, the researcher recognised that mothers may use this study as an opportunity to explore symptoms, thoughts, and feelings that they may be experiencing. It was therefore regarded as important to provide the mothers with the results of the screening. In particular, mothers who screened positively for PPD, mothers who indicated the presence of suicidal thoughts or thoughts of harming their babies, and mothers who were unaware that their symptoms would result in a positive screen for PPD, needed to be followed up with information about PPD, information about where to seek support or treatment, and prompt referral to their doctor if required.

- A secure password-protected website was established for participation online. Participants had to contact the researcher in order to obtain a username and password for online participation.
- The participants' biographical questionnaires as well as their screening questionnaires were anonymised and scored by the researcher prior to being sent to the University of Pretoria for statistical analysis to ensure anonymity and confidentiality of the results.
- Individuals who suffer from disorders that affect their ability to complete self-report measures reliably and validly were asked not to volunteer for this study.
- Mothers were informed that participation in the study was voluntary.
- Screening for symptoms of PPD was done at no financial cost to the participants, nor were the participants financially reimbursed for their participation.

- The participants' information was treated with utmost confidentiality and a mother's data was destroyed if she decided to withdraw. No participants wanted to withdraw from the study after completing the research questionnaires.
- Referring health practitioners did not have access to the results of the screening.

7.9 Data Analysis

7.9.1 Descriptive statistics for the PDSS.

The participants' demographic questionnaire data were collated and charted. The data collected from mothers included demographic information as well as known obstetric and psychosocial risk factors for PPD. The descriptive statistics for the English and Afrikaans samples in this study were examined to determine if there were significant differences between them. Furthermore, the demographic and obstetric characteristics of the participants and their PDSS screening results across three screening outcome categories were investigated.

7.9.2 Qualitative data analysis.

Qualitative analysis of the screening questionnaire items was done to arrive at a satisfactory Afrikaans translation of the PDSS. This was achieved by familiarising the translators with the subject matter of the inventory, and then comparing the items on the PDSS to the items on the back-translated version. Face validity was used to determine whether the items in both instruments appeared to measure similar concepts.

7.9.3 Quantitative data analysis.

7.9.3.1 Rasch analysis.

Rasch analysis was conducted as implemented by Winsteps software (Linacre, 2009). The specific measurement model employed was the Rating Scale Model, which is a formulation of an extended Rasch model based on IRT.

The main objective of this study was to analyse the PDSS and the Afrikaans PDSS in South African mothers within the Rasch framework. This would allow for determining the validity and reliability of these screening scales in a South African sample.

Given that the Rasch model distributes items along a level of difficulty, it was possible to determine whether some individual items on the PDSS and Afrikaans PDSS, or in turn, on each of the PDSS and Afrikaans PDSS dimensions, were harder to endorse than others. The psychometric properties of the PDSS and an Afrikaans translation of the PDSS were examined within the Rasch framework to determine how well the items defined the underlying construct of PPD in a South African sample. The PDSS was, however, developed as a multidimensional construct of PPD, incorporating seven individual dimensions. Rasch analysis was also performed to determine how adequately the attitude continuum which underlies each PDSS dimension (or construct) was assessed by the five items which constitute the dimension. These additional analyses of the dimensions were considered essential due to the fact that PPD is a multi-faceted phenomenon.

Fit statistics were computed to show how well the raw data fit the Rasch model. The Rasch model assumes that the items assess a unidimensional or single construct. The PDSS was, however, developed to assess the multidimensional construct of PPD and therefore incorporates seven individual dimensions. Rasch analysis was therefore performed on the PDSS and Afrikaans PDSS as a whole, as well as on each separate dimension.

The hypothesis of unidimensionality is that the items of the same factor should ideally load only on that factor. The assessment of unidimensionality is important as it provides evidence of the construct validity of a measure (Bond & Fox, 2001; Van der Ven & Ellis, 2000). Unidimensionality is equally important to the subtests or, in this case dimensions, that comprise a measure. Assessing the unidimensionality of each dimension of the PDSS and Afrikaans PDSS is therefore an important requirement for unidimensionality of the overall measure. Dimensionality was assessed by examining Rasch principal components analysis of residuals (PCA) and by examining item fit statistics.

PCA residuals were analysed to determine if secondary dimensions were present (Linacre, 2009). The residuals are the difference between the observed and the predicted scores. Using raw data in an analysis leads to non-linearity present in the data being accumulated in the PCA. This analysis was therefore performed using calibrated data (logits).

The item fit statistics – the global infit and outfit mean-square statistics – were also examined to assess the overall fit of the data to a unidimensional structure.

Unidimensionality of the PDSS and Afrikaans PDSS dimensions were determined by individual item fit. This was performed by examining the individual item infit and outfit mean-square statistics. These statistics also provide an indication of how well the data fit a unidimensional Rasch model and if any items misfit was present. Items with mean-square fit values above 1.5 contribute little value to the measure (Linacre, 2009). In this study a range of 0.60 to 1.40, as recommended by Bond and Fox (2007) and Wright and Linacre (1994) for rating scales, was selected to differentiate between fit and misfit persons and items.

When the dimensional structure of the PDSS and Afrikaans PDSS were confirmed, an item analysis using item-total correlations was performed. The Pearson item-total correlation (r_{it}) allows for identifying item misfit thereby providing an indication of the construct validity and whether there are coding problems present. This analysis is similar to the discrimination or item-total correlation in CTT. It does, however, differ in that extreme values are omitted (Maree, 2004). The Pearson item-total correlation (r_{it}) was compared to the expected score (EXP) to determine if discrepancies were evident which could indicate that an item did not fit the dimension well.

Indices of reliability of the PDSS and Afrikaans as well as individual dimensions were determined by Rasch analysis through item and person separation coefficients. Internal consistency reliability was determined by the person reliability estimate. Classic reliability coefficients were also calculated. The item-person map and item and person separation reliabilities were investigated to determine the appropriateness of item difficulty.

The data was also examined to evaluate the effectiveness of the Likert response categories as this impacts on how well the response data defines the dimension. The following six criteria were applied, as suggested by Linacre (2004), to evaluate the appropriateness of the Likert response categories for the PDSS and Afrikaans PDSS:

1. There should be at least ten observations in each category as low frequencies in the category can lead to unstable or imprecise estimates in the step calibrations.
2. There should be reasonably regular observation distribution for each category.
3. The average measures should increase monotonically with each category. People with higher abilities are thus expected to endorse higher categories and people with lower abilities are expected to endorse lower categories.
4. The outfit mean-square statistic for each category should be less than 2. Values greater than 2 suggest the presence of more unexplained variance than explained randomness as anticipated from the Rasch model, therefore indicating that some data did not support the definition of the underlying variable.
5. The step calibrations should advance orderly from easy to hard. An essential conceptual feature of a rating scale design is that a greater amount of the underlying variable in a respondent corresponds to a greater probability that the respondent will be observed in a higher category of the rating scale. When items have disordered categories it causes concern about the appropriateness of the item for measuring the underlying latent variable.
6. Step difficulties should advance by at least 1.40 logits, but by less than 5 logits. If the threshold distance were too wide, a "dead zone" is created in the middle of the category which means that the scale will not be precise in targeting

respondents between two successive categories. A five category rating scale should ideally advance by at least 1 logit (Linacre, 2004). If the advance is less than 1, the categories may need redefining to have wider substantive meaning, or categories should be combined.

When investigating the quality of a new measure, it is important to establish invariance before instruments may be deemed to be equivalent in a measurement sense (Küçükdeveci et al., 2004). Only then do the differences on the screening scales' scores accurately reflect the differences on the latent characteristics assessed by the construct. Invariance is determined through analysis of DIF which is a powerful means of checking for item bias in Rasch analysis (Bond, 2003). The foundation of DIF is to determine whether items have shifted in meaning across different groups or across different time points. Inconsistency in an item's difficulty estimate location across samples, with variation greater than the modelled error is a clear indication that DIF exists and indicates that the item has significantly different meanings for the different groups (Bond, 2003; Bond & Fox, 2007, p.92). Invariance analyses can be conducted via CTT by examining differences in item means by group or time, but such analyses are greatly simplified via use of Rasch model software (Chiang et al., 2009).

In a study by Allalouf and Sireci (1998), a panel of translators and researchers reviewed each DIF item to determine the possible sources of DIF and to formulate general conclusions about the sources of DIF in translated verbal items. The following were found to be the four main causes for DIF:

1. Changes in the difficulty of sentences or words despite an accurate translation;
2. Changes in the meaning of the item or the item content during the translation, thereby creating a different item. This may happen as a result of an incorrect translation causing a change in the meaning of the item, or a word that has a single meaning in the source language is translated into in a word that has more than one meaning in the target language;
3. Changes in the format of the item, for example, longer or shorter sentences in the target language;
4. Items remain the same, but differ in terms of their cultural relevance in the source and target language. The content of an item may, for example, be more familiar to one culture than to another.

Gierl and Khaliq (as cited in Allalouf, 2003, p. 56) identified the following sources of DIF in achievement tests: the addition or omission of phrases or words that affect the meaning of the item; differences in the words or expression either inherent or not inherent to the target language or culture; and format changes of the items.

The translation method used is important to consider in an attempt to reduce the likelihood of measurement bias. Translating an instrument for use in another culture can have a significant impact on the instrument's psychometric properties (Ramirez, Teresi, Holmes, Gurland, & Lantigua, 2006). Back-translation is a method widely used in cross cultural research used for addressing semantic equivalence. Ramirez et al recommend, though, that it be used in addition to other qualitative methods like cognitive interviews and random probes in order to address aspects of item equivalence and conceptual adequacy both within and across populations from diverse ethnic or cultural backgrounds and from different language groups.

Allalouf and Sireci (1998) formulated a basic flow chart which depicts the process involved in identifying the sources of DIF in an item. They recommend that translators use this flow chart, presented in Figure 2, to identify the causes for DIF. This chart was used as a guidline to examine the sources of DIF in this study.

DIF is generally not anticipated prior to administering tests or screening scales. Researchers must therefore rely on post-hoc explanations to determine the presumed causes of DIF. One recommendation in this regard, is to focus on items in a test that did not display DIF in order to help determine how it is different from items that did display DIF (Allalouf & Sireci, 1998).

Analysis of variance was conducted for each item of the PDSS and the Afrikaans PDSS. This analysis allowed the researcher to determine if DIF was present and if items have significantly different meanings across the two samples.

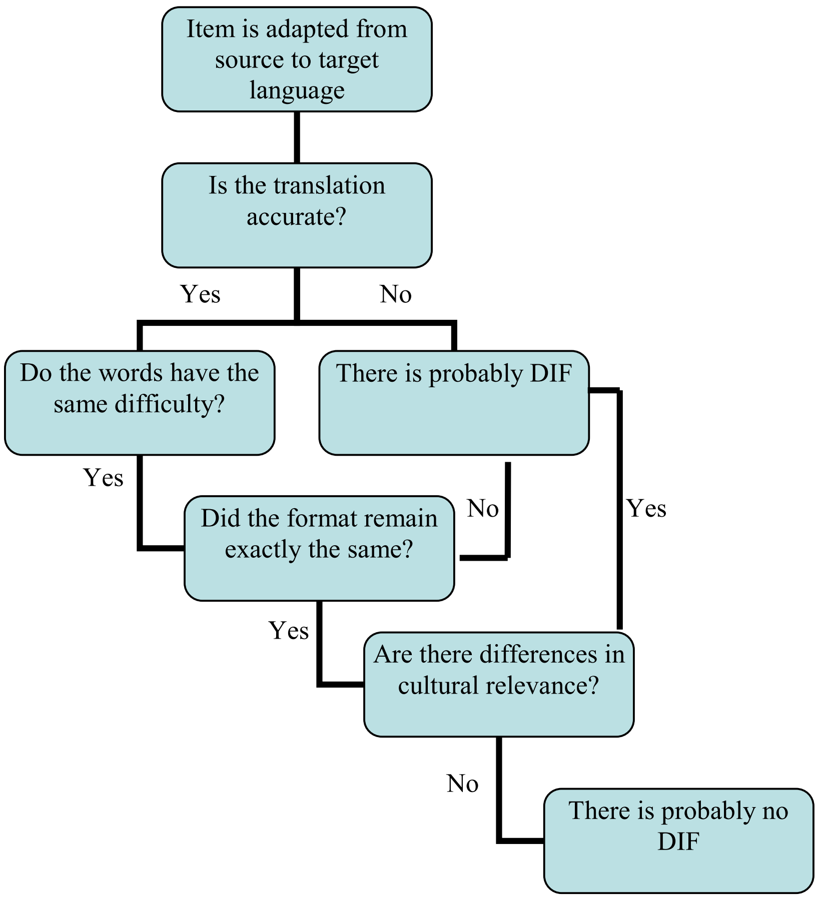

Figure 2 Flow chart for examining the sources for differential item functioning. (Adapted from Allalouf & Sireci, 1998, p. 19).

Convergent validity is a subtype of, and also an important aspect of construct validity that is determined by examining an expected overlap between measures that theoretically measure the same construct. Convergent validity therefore refers to the degree to which a measure is similar to (or correlated with) other measures that it is theoretically predicted to be similar to (or correlate with). High correlations provide evidence of convergent validity (Trochim, 2006). Convergent validity was examined to ascertain whether the PDSS and the Afrikaans PDSS correlate positively with other self-report screening scales for depression, namely the EPDS and the QIDS-SR16 and their respective Afrikaans translations.

7.9.3.2 Multiple regression analysis.

Various factors have been reported to be associated with the development of postpartum mood disorders. The relationship between known risk factors for PPD and scores on the PDSS amongst women in South Africa was investigated through multiple regression analysis. This statistical method was used to analyse the dataset because it is able to depict the relationship between several independent variables (predictor variables) and the dependent variable (the outcome or response) on a continuous scale, such as the severity of PPD.

Multiple regression is able to determine the relative influence of several independent variables (hereafter referred to as "predictor variables" for those variables that may be useful in predicting scores) when they are used to predict or explain a dependent variable (also referred to as the outcome; Field, 2005). The outcome is therefore explained by the predictor variables and how much influence they have. The fact that multiple regression makes it possible to determine how important the predictor variables are and takes into account how important the associations between the predictor variables are, has made it an extremely popular method of data analysis in the past couple of decades (Cramer, 2003; Foster, Barkus, & Yavorsky, 2006).

Based on the literature of risk factors for PPD, predictor variables were selected that were likely to correlate with the dependent variable. The dependent variable in this study (the PDSS score) is described by the following predictor variables, namely a history of psychiatric illness, antenatal depression in recent pregnancy, postpartum blues, feeling negative or ambivalent about expecting this baby, fearful of childbirth, lack of support from the baby's father, lack of support from friends, infant temperament, concern about health related issues regarding the infant, like colic, sleeping and feeding problems, and allergies, difficulty conceiving, and life stress.

The stepwise selection method was used in the multiple regression analysis. This method relies on computer software to select the order in which predictor variables and is based purely on mathematical criteria (Field, 2005). The predictor variables are entered in sequence and the software selects the predictor that has the highest simple correlation with the outcome. The predictor variable is retained if its addition contributes to the model (i.e. the theory that the predictor variable is likely to indicate a high PDSS score). The remaining predictor variables are, however, subjected to re-testing to determine if they are still making a contribution to the success of the model. These remaining predictor variables are removed if the re-testing indicates that they are no longer contributing significantly (Field, 2005). Employing the stepwise method therefore ensures that the minimum possible number of predictor variables is included to predict the outcome variable, in this case, the total PDSS score.

According to Foster et al (2003, p. 30), multiple regression is used to answer three types of question, namely:

1. What is the relative importance of the predictor variables included in the analysis?
2. Does a particular variable add to the accuracy of the prediction?
3. Given two alternative sets of predictors, can it be determined which is more effective? For example, can PPD be predicted better by the mother's demographic characteristics or by obstetric factors?

In simple regression the degree of the relationship between two continuous variables is expressed as a correlation coefficient which may vary from -1.00 to +1.00. When two variables are correlated, then predicting the score on one variable is possible if you know what the score on the other variable is. The stronger the correlation, the closer the scores will be to forming a straight line, and the more accurate the prediction will be (Foster et al., 2003). The scattergram for simple regression, which depicts the relationship between only two variables, is a visual representation of the following regression equation:

$$y = c + m(x)$$

In this equation, *y* (plotted on the vertical axis) is the predicted score on the dependent variable, *x* (plotted on the horizontal axis) is the score on the independent variable, *c* is the constant, or the intercept or point at which the line crosses the *y* axis, and *m* is the regression coefficient or weight, which indicates by how much *x* must be multiplied to obtain the predicted value of *y*. Put in another way, the regression coefficient (*m*) is the amount of change in the dependent variable (*y*) resulting from a one-unit change in an independent (*x*) variable when all the other predictor variables are held constant. The difference between the predicted *y* score and the actual score is known as the residual (Foster et al, 2006).

In this study *y* is the PDSS score. The predictor variables (*x*) are categorical, for example, the presence of a history of depression is coded 1, and no history of depression is coded 0. When the predictor variable (*x*) has a value of 0, the variable disappears and the leaves only the constant value (*c*). All participants therefore start off with the constant value. The presence of any predictor variables (*x*) are therefore expected to add to the overall PDSS score.

Multiple regression is simply an extension of this correlation principle. In multiple regression a prediction to one variable is based on several other variables as opposed to just one, as in simple regression. For every predictor variable that is added, a coefficient is added, so that each predictor variable has its own coefficient (Field, 2005). This enables not only the prediction of the dependent variable, but also determining the relative influence of each of the predictor variables on the outcome – the PDSS total score – and gives an indication of the combined ability of the predictor variables in predicting or explaining the variation in the outcome variable (y). The multiple regression equation is therefore slightly more complex:

$$y = c + m_1(x_1) + m_2(x_2) + m_3(x_3)...m_\kappa(x_\kappa)$$

The aim of multiple regression is to find the regression coefficient (the weight, i.e. m_1, m_2, etc.) for each of the predictor variables (x_1, x_2, etc.) which will produce the values of y which are closest to the actual values (Foster et al., 2003). The regression coefficients therefore maximise the correlation between the predicted y values and the combination of the predictor (x) variables. In this study the relative influence of a number of predictor variables (the known risk factors for PPD) as they relate to the outcome variable (the PDSS total score) was investigated.

When two or more predictor variables correlate strongly with each other, known as collinearity, then making assumptions about the relative contribution of each predictor variable is difficult. SPSS is able to determine if collinearity in the data was present.

SPSS 19 was used in this study to examine the following requirements for multivariate analyses, namely sample size, independence of residuals (Durbin-Watson test), presence of multicollinearity (the variance inflation factor or VIF, the tolerance statistic, and collinearity diagnostics), the influence of outliers (casewise diagnostics using Cook's distance, and Mahalanobis Distance), homoscedasticity and non-linearity (plots of the standardised predicted values against the standardised residuals), and normality of residuals (Field, 2005).

7.9.3.3 Correlation of PDSS, EPDS, and QIDS-SR16 total scores.

Statistical analyses were performed on the total sample (N = 365) to determine the comparison of the participants scores across the three screening scales. Descriptive statistics for the three screening scales were calculated and the frequencies were determined according to the participants screening results at the published cut-off thresholds recommended. Chi-square tests were performed on the categorical depression screening status to compare participants who scored positive for symptoms of PPD on the three measures.

The Pearson correlation was used to investigate the relationship between the PDSS and the EPDS, and the PDSS and the QIDS-SR16. The Pearson correlation measures the strength of the correlation (linear dependence) between two variables. It is sometimes referred to as "Pearson's *r*" and is denoted by *r*. It yields a value that may range from +1 to -1. The stronger the association between two variables are, the closer the Pearson correlation coefficient will be to either +1 or -1, depending on whether the association is positive or negative. A value of 0 signifies that there is no association or linear correlation between the two variables. A general guideline is that a positive coefficient *r* of 0.50 to 1.00 indicates a strong positive association between the variables, and a negative coefficient of -0.50 to -1.00 indicates as a strong negative association between the variables (Cohen, 1988).

CHAPTER 8

RESULTS AND DISCUSSION

8.1 Introduction

This chapter presents the descriptive statistics for the English and Afrikaans samples. Frequency distributions and means are reported along with Pearson Chi-square statistics where significant differences were present between the English and Afrikaans samples.

This is followed by the results of the Rasch analysis. The performance of the PDSS and Afrikaans PDSS was analysed using Rasch analysis to evaluate how well the items contributed to the underlying construct of PPD. The same analysis was also performed with the scales' dimensions. Dimensionality was examined using item fit statistics and principal component analysis (PCA) of standardized residuals. Reliability of the PDSS, the Afrikaans PDSS, and their dimensions were determined by the person reliability estimates and Cronbach alpha. The appropriateness of item difficulty was determined by examining the item-person map and person reliability estimates. The category functioning was also evaluated to determine the effectiveness of the Likert response categories of the PDSS and Afrikaans PDSS. Finally, differential item functioning (DIF) was examined to compare the estimates across the English and Afrikaans samples to determine whether the items have significantly different meanings for the two groups.

The results of the multiple regression analysis, using the stepwise selection method, are presented next. This statistical method was used to analyse the relationship between known risk factors for PPD and scores on the PDSS.

Finally the results of the Pearson correlation are presented. This analysis was performed to determine the relationship between participants' scores on the PDSS, the EPDS, and the QIDS-SR16.

8.2 Descriptive Statistics

Frequency distributions were used to summarise the data and means were calculated where appropriate. Pearson Chi-square statistics were used to determine if significant differences were present between the English and Afrikaans samples. All the p-values were two-tailed and p-values <0.05 were considered statistically significant.

All participants in the study were South African citizens. One participant, although a South African citizen, completed the research questionnaire from abroad. It was determined that she had only lived overseas for a short while and was therefore not excluded from participation. All other participants were resident in South Africa at the time.

Participants' home language is indicated in Table 9. The majority of participants (96.1%) who completed the questionnaires in Afrikaans and the majority of participants who completed the questionnaires in English (92.5%) indicated that they were completing the questionnaires in their home language. A small number of participants (4.7%) indicated that their home language was neither English nor Afrikaans.

All participants who completed the English PDSS had English as a subject at school. One hundred and sixty four participants (87.7%) had English as a first language, and 23 participants (12.3%) had English as a second language. The participants were asked whether they considered themselves fluent in English. Fluency in the language of test administration was a requirement for participation in this study. One participant indicated that she did not consider herself fluent in English. She did however complete grade 12 with English as a first language at high school. As the researcher had also conversed with her successfully in English, it is believed that she judged her English language ability harshly and she was not excluded from participating in the study.

All the participants who completed the Afrikaans PDSS had Afrikaans as a language taught at high school – 167 participants (93.8%) had Afrikaans as a first language, and 11 participants (6.2%) had Afrikaans as a second language. As with the English-speaking participants, the Afrikaans-speaking participants were requested to indicate on the participant information form whether they considered themselves fluent in Afrikaans. All the participants who completed the Afrikaans PDSS considered themselves fluent in Afrikaans.

The demographic characteristics of the mothers are shown in Table 9. Most mothers were White (84.9%), followed by Black (5.2%), Asian (4.9%) and Coloured (4.7%) mothers. The imbalance in the race/ethnic group of the mothers may be attributed to the nature of the study – i.e. the sampling requirement that mothers should be fluent in

English or Afrikaans, the fact that many mothers were recruited from clinics in urban areas and from magazine articles, and that participation could be done online requiring internet access.

As can be seen in Table 9, most of the sample was married (88.8%) or in a de facto relationship (4.1%). All the participants were below 45 years of age. The majority of participants were between the ages of 26 and 35 (78.8%). The mean age of the participants was 30.11 with a standard deviation of 4.17. No significant differences in marital status and age were noted between the English and Afrikaans mothers.

The education level and employment status of the participants are presented in Table 10. Close to a quarter (23.6%) completed grade 12, just over two thirds of the participants (67.4%) either had a degree or a diploma, and 4.4% a trade certificate. No significant differences were noted between the English and Afrikaans samples. Almost half of the participants worked full-time (49.3%), 27.1% were unemployed, followed by 13.2% who were self-employed, and 10.4% who were employed part-time.

Table 9 Demographic Characteristics Stratified by Questionnaire Language: Home Language, Race/Ethnic Group, Marital Status and Age

Demographic Characteristics	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Home language									
English	177	48.5	4	2.2	173	92.5			
Afrikaans	171	46.8	171	96.1	0	0			
Xhosa	7	1.9	2	1.1	5	2.7			
Zulu	5	1.4	1	0.6	4	2.1			
Northern Sotho	2	0.5	0	0	2	1.1			
Southern Sotho	1	0.3	0	0	1	0.5			
Chinese	1	0.3	0	0	1	0.5			
Other	1	0.3	0	0	1	0.5			
Race/ethnic group									
White	310	84.9	160	89.9	150	80.2			
Black	19	5.2	5	2.8	14	7.5			
Asian	18	4.9	0	0	18	9.6			
Coloured	17	4.7	13	7.3	4	2.1			
Other	1	0.3	0	0	1	.5			
Marital status							3.06	3	0.383
Married	324	88.8	163	91.6	161	86.1			
Unmarried	24	6.6	8	4.5	16	8.6			
De Facto Relationship	15	4.1	6	3.4	9	4.8			
Divorced	2	0.5	1	0.6	1	0.5			
Age (in years)							18.07	24	0.800
18-20	6	1.7	2	1.2	4	2.1			
21-25	38	10.5	17	9.5	21	11.2			
26-30	151	41.3	75	42.2	76	40.8			
31-35	137	37.5	70	39.3	67	35.8			
36-40	28	7.6	13	7.3	15	8.1			
40-44	4	1.1	1	0.6	3	1.6			
Missing data	1	0.3			1	0.5			
M	30.11 years		30.21 years		30.01 years				
SD	4.17		4.384		3.943				

Table 10 Demographic Characteristics Stratified by Questionnaire Language: Education Level and Employment Status

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Education level							5.75	7	0.569
Degree or Diploma	246	67.4	117	65.7	129	69.0			
Trade Certificate	16	4.4	10	5.6	6	3.2			
Grade 12	86	23.6	45	25.3	41	21.9			
Grade 11	5	1.4	3	1.7	2	1.1			
Grade 10	6	1.6	1	0.6	5	2.7			
Grade 9	2	0.5	1	0.6	1	0.5			
Grade 8	3	0.8	1	0.6	2	1.1			
Grade 7	1	0.3	0	0	1	0.5			
Employment status							3.62	3	0.305
Full-time	180	49.3	79	44.4	101	54.0			
Unemployed	99	27.1	54	30.3	45	24.1			
Self-employed	48	13.2	26	14.6	22	11.8			
Part-time	38	10.4	19	10.7	19	10.2			

Table 11 presents the number of weeks since birth, the infant's sex and gestational age at birth, and the infant feeding method the mother opted for. Most participants were between 5 and 7 weeks postpartum (32.1%) or 16 weeks postpartum (11.5%). The mean age postpartum was 5.3 weeks (standard deviation 3.768). The mean number of weeks since birth was 5.68 weeks (SD 4.043) for the English participants and 4.9 weeks (SD 3.421) for Afrikaans participants. A significant difference was noted between the English and Afrikaans participants in the number of weeks since birth ($x2$ = 27.07, df = 12, p = 0.008). More English mothers participated at 16 weeks postpartum than expected and substantially less Afrikaans mothers participated at 16 weeks than expected. Furthermore, more Afrikaans mothers participated at 5 weeks postpartum than expected and substantially less English mothers participated at 5 weeks than expected. There was no significant difference in the number of male and female babies born to the English and Afrikaans participants.

In both the Afrikaans and English samples, the majority of infants were born between 38 and 40 weeks postpartum (55.6% and 63.1% respectively). More mothers from the Afrikaans sample gave birth pre-term (25.9%) than mothers from the English sample (18.2%). These results were, however, not statistically significant.

The majority of mothers from both samples opted to breastfeed their babies from birth (Afrikaans: 46.1%; English: 48.7%), followed by mothers who breastfed initially but now bottle feed with formula only (Afrikaans: 21.9%; English: 20.9%). The feeding method of choice did not differ significantly between the English and Afrikaans mothers.

Table 11 Demographic Characteristics Stratified by Questionnaire Language: Number of Weeks Since Birth, Infant's Sex, Gestational Age at Birth, and Feeding Method

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
No. of weeks since birth							27.07	12	0.008**
4 weeks	25	6.8	8	4.5	17	9.1			
5 weeks	43	11.8	29	16.3	14	7.5			
6 weeks	36	9.9	12	6.7	24	12.8			
7 weeks	38	10.4	25	14.0	13	7.0			
8 weeks	35	9.6	17	9.6	18	9.6			
9 weeks	34	9.3	19	10.7	15	8.0			
10 weeks	30	8.2	16	9.0	14	7.5			
11 weeks	22	6.0	12	6.7	10	5.3			
12 weeks	22	6.0	11	6.2	11	5.9			
13 weeks	10	2.7	5	2.8	5	2.7			
14 weeks	15	4.1	6	3.4	9	4.8			
15 weeks	13	3.6	7	3.9	6	3.2			
16 weeks	42	11.5	11	6.2	31	16.6			
M	5.3 weeks		4.9 weeks		5.68 weeks				
SD	3.768		3.421		4.043				
Infant's sex							0.36	1	0.549
Male	174	47.7	82	46.1	92	49.2			
Female	191	52.3	96	53.9	95	50.8			
Gestational age of infant at birth							6.68	4	0.154
≤ 28 weeks	7	1.9	3	1.7	4	2.1			
29 - 33 weeks	11	3.0	9	5.1	2	1.1			
34 - 37 weeks	62	17.0	34	19.1	28	15.0			
38 - 40 weeks	217	59.5	99	55.6	118	63.1			
> 40 weeks	68	18.6	33	18.5	35	18.7			
Feeding method							2.49	3	0.476
Breast fed – from birth	173	47.4	82	46.1	91	48.7			
Initially breastfed but now bottle fed only	78	21.4	39	21.9	39	20.9			
Bottle fed - from birth [a]	58	15.9	33	18.5	25	13.4			

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Combination of breast and bottle	56	15.3	24	13.5	32	17.1			

* $p \leq 0.05$
** $p \leq 0.01$
*** $p \leq 0.001$
[a] bottle fed implies formula milk

Table 12 presents the perceived level of support obtained by the mothers in the postpartum period. More mothers from the English sample indicated that they received sufficient help and support from the baby's father (77%) than mothers from the Afrikaans sample (63.5%). Less mothers than expected from the English sample indicated that they received some help and support from the baby's father, while more Afrikaans mothers than expected indicated that they received some help and support (Table 68a and Table 68b in Appendix F). Overall the amount of help and support mothers received from the baby's father differed significantly between the two samples ($x2$ = 10.09, df = 2, p = 0.006). This is due to a larger percentage of English mothers indicating that they received sufficient help compared to Afrikaans mothers, while a smaller percentage indicated that they received some help and support. If the percentage of mothers who indicated that they received either sufficient help and support or some help and support from the baby's father were combined, then the distribution between the Afrikaans and English samples are strikingly similar at 92.7% for the Afrikaans sample and 92.5% for the English sample. The percentage of mothers who indicated that they received no help and support is similar in both samples (Afrikaans: 7.3%; English: 7.5%).

A similar pattern is seen for help and support obtained from family. The amount of help and support mothers received from extended family differed significantly between the two samples ($x2$ = 10.05, df = 2, p = 0.007). This may be attributed to the differences in expected rates of both sufficient help and support, and some help and support received from the two samples (Table 69a and Table 69b in Appendix F).

If the percentage of mothers who indicated that they received either sufficient help and support or some help and support from extended family were combined, then the distribution between the Afrikaans and English samples are strikingly similar with 87.2% of English mothers and 87.1% of Afrikaans mothers indicating that they received either sufficient or some help and support.

Slightly more English mothers received sufficient help or some help and support from friends (58.8%) compared to mother from the Afrikaans sample (52.8%). The majority of mothers from both samples do not receive additional support from other sources (Afrikaans: 78.7%; English 78.6%).

Table 12 Perceived Level of Support Obtained by Mothers, Stratified by Questionnaire Language

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*Df*	*P*
Support from father							10.09	2	0.006**
No	27	7.4	13	7.3	14	7.5			
Yes	257	70.4	113	63.5	144	77.0			
Some	81	22.2	52	29.2	29	15.5			
Support from family							10.05	2	0.007**
No	47	12.9	23	12.9	24	12.8			
Yes	231	63.3	100	56.2	131	70.1			
Some	87	23.8	55	30.9	32	17.1			
Support from friends							2.34	2	0.311
No	161	44.1	84	47.2	77	41.2			
Yes	129	35.3	56	31.5	73	39.0			
Some	75	20.5	38	21.3	37	19.8			
Support from others							4.24	2	0.120
No	287	78.6	140	78.7	147	78.6			
Yes	50	13.7	20	11.2	30	16.0			
Some	28	7.7	18	10.1	10	5.3			

* $p \leq 0.05$
** $p \leq 0.01$
*** $p \leq 0.001$

The obstetric profile of mothers is presented in Table 13. A total of 38.6% of mothers gave birth by elective caesarean. This was the most common method of delivery in both samples. This was followed by a normal vaginal delivery (27.1%), emergency caesarean (20.3%), and then traumatic vaginal delivery (13.7%). No significant differences were found in the method of delivery or in the rating of care during labour and delivery between the English and Afrikaans mothers. Most mothers rated their care during labour and delivery as being excellent (58.9%), with a further 29.3% rating it as good. Six percent of mothers perceived their care as being poor.

Most participants had only had 1 pregnancy (57.8%), followed by mothers who had two pregnancies (25.5%). Less mothers had 3 pregnancies (11%) and only 5.5% of mothers had a fourth, fifth or sixth pregnancy. Mean gravidity was 1.66 with a standard deviation of 0.939. No significant differences were found in gravidity between the two samples. The majority of mothers who participated in this study (60%) only had 1 child, 27.9% had two children, and 10.1% had three children. Few mothers had more than three children (1.7%). The mean number of children respondents had was 1.54 (*SD* 0.754). No significant differences were found between the English and Afrikaans mothers in the number of children they had.

Participants were asked to indicate whether a health practitioner had diagnosed them with either antenatal depression during, and/or PPD after their recent pregnancy at their postnatal follow-up appointment with their caregiver. If this was the case, they were asked to indicate whether they are currently receiving counselling or psychotherapy. This data is presented in Table 14. Close to a quarter of mothers (23.3%) had not yet had a postpartum follow-up appointment with their caregiver. Nearly half of the participants

(48.5%) indicated that their caregiver did not enquire about the presence of depressive symptoms at their postpartum follow-up, while 28.2% of mothers indicated that their caregivers did.

A small number of participants (5.5%) were diagnosed with PPD after their recent pregnancy and even less (3%) were diagnosed with antenatal depression. These figures were fairly similar across the samples and the differences were not statistically significant. Only 2.2% of these mothers were receiving counselling or psychotherapy for PPD at the time of assessment while 10.4% of mothers were using medication for depression or anxiety. No significant differences were found between the English and Afrikaans samples concerning counselling or psychotherapy and use of medication for depression or anxiety.

Table 13 Obstetric Profile of Mothers Stratified by Questionnaire Language

	Frequency Total (n=365) [b]	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Type of delivery [b]							4.66	3	0.198
Elective caesarean	141	38.6	67	37.6	39.6	74			
Normal vaginal	99	27.1	42	23.6	30.5	57			
Emergency caesarean	74	20.3	39	21.9	18.7	35			
Traumatic vaginal	50	13.7	30	16.9	10.7	20			
Perception of care during labour and delivery							5.25	3	0.154
Excellent	215	58.9	99	55.6	62.0	116			
Good	107	29.3	52	29.2	29.4	55			
Unremarkable	21	5.8	15	8.4	3.2	6			
Poor	22	6.0	12	6.7	5.3	10			
Gravidity [b]							3.00	5	0.700
1st pregnancy	211	57.8	100	56.2	59.4	111			
2nd pregnancy	93	25.5	45	25.3	25.7	48			
3rd pregnancy	40	11.0	20	11.2	10.7	20			
4th pregnancy	15	4.1	9	5.1	3.2	6			
5th pregnancy	4	1.1	3	1.7	0.5	1			
6th pregnancy	1	0.3			0.5	1			
M	1.66		1.7			1.61			
SD	0.939		0.974			0.905			
Number of biological children [b]							3.38	4	0.497
1 child	219	60.0	103	57.9	62.0	116			
2 children	102	27.9	48	27.0	28.9	54			
3 children	37	10.1	22	12.4	8.0	15			
4 children	5	1.4	3	1.7	1.1	2			
5 children	1	0.3	1	0.6					
M	1.54		1.59			1.48			
SD	0.754		0.814			0.691			

[b] Data is missing where totals do not add up to N = 365

Table 14 Current PPD and Antenatal Depression Assessment and/or Treatment of Mothers, Stratified by Questionnaire Language

	Frequency Total (N = 365) [b]	Total (%)	Frequency Afrikaans (n = 178)	Afrikaans (%)	Frequency English (N = 187)	English (%)	X^2	*df*	*P*
PPD diagnosis[a]	20	5.5	10	5.6	10	5.3	0.01	1	0.910
Antenatal depression diagnosis	11	3.0	5	2.8	6	3.2	0.05	1	0.823
Caregiver enquired about symptoms of depression at postnatal follow up							0.56	2	0.755
No	177	48.5	83	46.6	94	50.3			
Yes	103	28.2	53	29.8	50	26.7			
Not been for follow up	85	23.3	42	23.6	43	23.0			
Currently receiving counseling or psychotherapy for PPD [b]							1.30	2	0.523
No	18	4.9	7	3.9	11	5.9			
Yes	8	2.2	5	2.8	3	1.6			
N/A	337	92.3	164	92.1	173	92.5			
Currently using medication for depression or anxiety [b]							2.53	1	0.112
No	323	88.5	163	91.6	160	85.6			
Yes	38	10.4	14	7.9	24	12.8			

[a] Related to recent pregnancy
[b] Data is missing where totals do not add up to N = 365

The psychiatric history of the mothers is presented in Table 15. Most mothers (65.8%) had no history of the psychiatric illnesses listed in Table 15. Almost a quarter of mothers (23.8%) did, however, have a history of depression, while 8.2% had a history of an anxiety disorder, 6.6% had a history of PPD after a previous pregnancy, 3.3% of mothers had had an eating disorder, only 2 mothers (0.5%) had antenatal depression during a previous a pregnancy, and 1 mother (0.3%) indicated that she had a history of obsessive compulsive disorder.

Mothers were asked to indicate whether they think they had PPD (11.5%), some symptoms of PPD (22.2%), or no PPD (41.9%). Mothers could also opt to indicate that they were uncertain about whether or not they had PPD (20.5%), or that they did not really know what PPD was (3.8%). This data is presented in Table 16. A statistically significant difference was found to responses made by mothers from the two samples ($x2 = 10.90$, $df = 4$, $p = 0.028$). Significantly less English mothers than expected indicated that they thought they may have some symptoms of PPD, while significantly more Afrikaans mothers than expected thought they may have some symptoms of PPD. Furthermore, significantly more English mothers than expected thought they did not have PPD, and significantly less Afrikaans mothers than expected thought they did not have PPD.

Table 15 Psychiatric History of Mothers Stratified by Questionnaire Language

Psychiatric History	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)
Depression	87	23.8	46	25.8	41	21.9
Anxiety	30	8.2	8	4.5	22	11.8
PPD after a previous pregnancy	24	6.6	14	7.9	10	5.3
Anorexia	7	1.9	5	2.8	2	1.1
Bulimia	5	1.4	2	1.1	3	1.6
Antenatal depression during a previous pregnancy	2	0.5	1	0.6	1	0.5
Obsessive compulsive disorder	1	0.3	0	0	1	0.5

Table 16 Self Evaluation PPD Statements Chosen by Mothers, Stratified by Questionnaire Language

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Self evaluation							10.90	4	0.028*
I think I may have some symptoms of postpartum depression	81	22.2	52	29.2	29	15.5			
I think I may have postpartum depression	42	11.5	21	11.8	21	11.2			
I do not really know what postpartum depression is	14	3.8	5	2.8	9	4.8			
I know what postpartum depression is and I do not think I am suffering from it	153	41.9	67	37.6	86	46.0			
I feel uncertain about whether or not I may have postpartum depression	75	20.5	33	18.5	42	22.5			

* $p \leq 0.05$

Table 17 contains the peripartum and psychological profile of the mothers. Postpartum blues is fairly common after the birth of a baby, and this was evident in this study with 70.1% of mothers indicating that they had postpartum blues. For most mothers (72.3%) this pregnancy was planned. Some mothers (14.2%) indicated that they had difficulty conceiving, while 7.4% had had fertility treatment with their recent pregnancy. Close to a quarter of the mothers (24.1%) indicated that they had had complications in their pregnancy such as pre-eclampsia or a threatened miscarriage. More than a quarter of the mothers indicated that they were intensely anxious or fearful of childbirth, and 44.1% of mothers had a history of premenstrual dysphoric disorder (PMDD), or PMS. Furthermore, according to their own self-evaluation, nearly half of the mothers indicated that they thought they were perfectionistic. No significant differences were found between the English and Afrikaans mothers' peripartum and psychological profile.

The psychosocial characteristics are presented in Table 18. Women were asked about certain major distressing life events in the past two years which are known risk factors for PPD. Most common events included financial concerns (59.2%), moving house (46.6%), house alterations (36.7%), and changing jobs (31.8%). It should be noted, however, that the last mentioned factor also includes mothers who resigned and opted to be a stay-at-home mother. The researcher determined that in some instances this was chosen to ease the pressure of working full time while having young children and as such, for some participants, the change was not experienced as a major distressing life event, but quite the contrary.

Table 17 Peripartum and Psychological Profile of Mothers Stratified by Questionnaire Language

	Frequency Total (n=365) [b]	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Postpartum blues							1.79	1	0.181
No	109	29.9	59	33.1	50	26.7			
Yes	256	70.1	119	66.9	137	73.3			
Planned pregnancy							0.99	1	0.319
No	101	27.7	45	25.3	56	29.9			
Yes	264	72.3	133	74.7	131	70.1			
Difficulty conceiving [b]							0.18	1	0.669
No	312	85.5	154	86.5	158	84.5			
Yes	52	14.2	24	13.5	28	15.0			
Fertility treatment							2.35	1	0.125
No	338	92.6	161	90.4	177	94.7			
Yes	27	7.4	17	9.6	10	5.3			
Complicated pregnancy							1.55	1	0.213
No	277	75.9	130	73.0	147	78.6			
Yes	88	24.1	48	27.0	40	21.4			
Tokophobia or intensely fearful or anxious of childbirth [b]							0.28	1	0.600
No	269	73.7	133	74.7	136	72.7			
Yes	95	26.0	44	24.7	51	27.3			
History of PMS[a] or PMDD[c]							2.51	1	0.113
No	204	55.9	107	60.1	97	51.9			
Yes	161	44.1	71	39.9	90	48.1			
Consider self a perfectionist							2.75	1	0.097
No	195	53.4	103	57.9	92	49.2			
Yes	170	46.6	75	42.1	95	50.8			

[a] PMS = premenstrual syndrome
[b] Data is missing where totals do not add up to N = 365
[c] PMDD = premenstrual dysphoric disorder

Other common distressing life events that participants experienced are the loss of close friends or family, either through relocation or migration (29%), their spouse or partner changing jobs (28.8%), serious illness of a family member (26%), family problems (26%), being victimised by violence or crime (18.4%), marriage (17.8%), bereavement (17.3%), moving to a different town or city, or migration (16.4%), marital discord (14.8%), and another pregnancy and birth (14.2%). Less common stressful events were job loss or retrenchment (9.6%), serious injury, illness, or personal health problems (7.4%), and a spouse or partner's job loss or retrenchment (7.1%).

The responses to six different life stressors (moving house, moving city or migrating, job changes in mothers, job changes in partners, bereavement, and being victimised by violence or crime) varied significantly between the English and Afrikaans mothers.

A profile of how mothers felt about their pregnancies is presented in Table 19. The majority of mothers were positive about their pregnancies (73.7%), some were ambivalent (18.6%), and a small percentage were negative (5.2%) or predominantly anxious (2.5%). No significant differences were found between the two samples regarding how they felt about their pregnancies.

Table 18 Psychosocial Characteristics of Mothers Stratified by Questionnaire Language

Major Life Stresses in the past 2 years	Frequency Total (n=365) [b]	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
House alterations [b]							1.44	1	0.229
No	230	63.0	118	66.3	112	59.9			
Yes	134	36.7	60	33.7	74	39.6			
Moving house							8.52	1	0.004**
No	195	53.4	109	61.2	86	46.0			
Yes	170	46.6	69	38.8	101	54.0			
Moving city / immigrate [b]							5.34	1	0.021*
No	304	83.3	156	87.6	148	79.1			
Yes	60	16.4	21	11.8	39	20.9			
Job changes: self							6.77	1	0.009**
No	249	68.2	133	74.7	116	62.0			
Yes	116	31.8	45	25.3	71	38.0			
Job changes: partner							9.38	2	0.009**
No	252	69.0	136	76.4	116	62.0			
Yes	105	28.8	38	21.3	67	35.8			
N/A	8	2.2	4	2.2	4	2.1			
Job loss / retrenchment: self [b]							1.91	1	0.167
No	329	90.1	157	88.2	172	92.0			
Yes	35	9.6	21	11.8	14	7.5			
Job loss / retrenchment: partner							0.47	2	0.791
No	331	90.7	163	91.6	168	89.8			
Yes	26	7.1	11	6.2	15	8.0			
N/A	8	2.2	4	2.2	4	2.1			
Financial concerns							0.08	1	0.776
No	149	40.8	74	41.6	75	40.1			
Yes	216	59.2	104	58.4	112	59.9			
Bereavement [b]							5.73	1	0.017*

Major Life Stresses in the past 2 years	Frequency Total (n=365) [b]	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
No	301	82.5	155	87.1	146	78.1			
Yes	63	17.3	22	12.4	41	21.9			
Loss of close friends / family relocating, emigrating, etc. [b]							0.11	1	0.737
No	258	70.7	124	69.7	134	71.7			
Yes	106	29.0	53	29.8	53	28.3			
Serious illness of a family member							2.28	1	0.131
No	270	74.0	138	77.5	132	70.6			
Yes	95	26.0	40	22.5	55	29.4			
Another pregnancy and birth							1.93	1	0.164
No	313	85.8	148	83.1	165	88.2			
Yes	52	14.2	30	16.9	22	11.8			
Marriage [b]							0.51	1	0.475
No	299	81.9	148	83.1	151	80.7			
Yes	65	17.8	29	16.3	36	19.3			
Marital problems							0.16	1	0.694
No	311	85.2	153	86.0	158	84.5			
Yes	54	14.8	25	14.0	29	15.5			
Family problems [b]							0.02	1	0.897
No	269	73.7	131	73.6	138	73.8			
Yes	95	26.0	47	26.4	48	25.7			
Victimised by violence or crime							5.07	1	0.024*
No	298	81.6	137	77.0	161	86.1			
Yes	67	18.4	41	23.0	26	13.9			
Serious injury, illness, or personal health problems [b]							0.01	1	0.935
No	337	92.3	165	92.7	172	92.0			
Yes	27	7.4	13	7.3	14	7.5			

* $p \leq 0.05$
** $p \leq 0.01$
*** $p \leq 0.001$
[b] Data is missing where totals do not add up to N = 365

Table 19 Profile of How Mothers Felt About Their Pregnancies, Stratified by Questionnaire Language

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
How mother felt about expecting a baby							4.38	3	0.223
Positive	269	73.7	125	70.2	144	77.0			
Ambivalent	68	18.6	38	21.3	30	16.0			
Negative	19	5.2	12	6.7	7	3.7			
Other:	9	2.5	3	1.7	6	3.2	3.75	5	0.586
• Anxious	3	0.8	2	1.1	1	0.5			
• Anxious – overwhelmed	1	0.3			1	0.5			
• Anxious – losing baby	2	0.5	1	0.6	1	0.5			
• Anxious – pregnancy	1	0.3			1	0.5			
• Anxious – responsibility	1	0.3			1	0.5			
• Anxious – motherhood and weight gain	1	0.3			1	0.5			

Infant attributes as factors in maternal depression has been discussed in chapter 2. Infants who are temperamentally difficult or irritable are strongly predictive of maternal depression. Furthermore, a mother's psychological distress influences how she experiences her infant's behavioural characteristics. Table 20 outlines the mothers' perceptions of their infants' temperament and specific concerns they have had about their infants. Two thirds of the mothers in the study indicated that they experience their infants' temperament as being good. Infants described with demanding temperament accounted for 22.5% of mothers. Remaining mothers reported infant temperament as fussy (5.5%), difficult (4.1%), or a combination of all these characteristics (1.6%).

The majority of mothers from both samples reported no specific concerns regarding their infants. Of the concerns that were reported, infant colic (26.6%), infant sleeping (25.5%), and infant feeding (22.2%) issues were greater issues for the total sample. Significant differences were found between the English and Afrikaans mothers regarding infant feeding concerns ($x2 = 4.03$, $df = 1$, $p = 0.045$) and concerns regarding infant prematurity ($x2 = 13.21$, $df = 1$, $p < 0.001$). The amount of Afrikaans mothers who were concerned about their infants' prematurity was significantly higher than expected, whereas significantly fewer English mothers were concerned about their infants' prematurity than expected[1].

[1] Some mothers were referred from postpartum support groups. Although the researcher can only speculate, it is possible that a group of mothers were referred from a predominantly Afrikaans-speaking support group for mothers with premature babies.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table 20 Infant Temperament and Concerns Regarding Infant, Stratified by Questionnaire Language

	Frequency Total (n=365)	Total (%)	Frequency Afrikaans (n=178)	Afrikaans (%)	Frequency English (n=187)	English (%)	X^2	*df*	*P*
Infant's temperament according to mother									
Good	242	66.3	116	65.2	126	67.4			
Demanding	82	22.5	42	23.6	40	21.4			
Fussy	20	5.5	6	3.4	14	7.5			
Difficult	15	4.1	10	5.6	5	2.7			
Combination of the Above	6	1.6	4	2.2	2	1.1			
Specific concerns regarding infant									
No problems	159	43.6	81	45.5	78	41.7			
Health concerns	16	4.4	8	4.5	8	4.3			
Colic	97	26.6	42	23.6	55	29.4			
Infant's sleep	93	25.5	44	24.7	49	26.2			
Feeding	81	22.2	48	27.0	33	17.6	4.03	1	0.045*
Allergies	15	4.1	4	2.2	11	5.9			
Premature	39	10.7	30	16.9	9	4.8	13.21	1	0.000***
Other:	22	6.0	11	6.2	11	5.9			
• Postnasal drip	1	0.3	0	0	1	0.5			
• Reflux	9	2.5	3	1.7	6	3.2			
• Difficulty winding	1	0.3	0	0	1	0.5			
• Occasional vomiting	1	0.3	0	0	1	0.5			
• Weight gain issues	2	0.5	0	0	2	1.1			
• Cramps and crying same time each day	1	0.3	1	0.6	0	0			
• Minor disability	1	0.3	1	0.6	0	0			
• Difficulty bonding	1	0.3	1	0.6	0	0			
• Infant dyschezia	1	0.3	1	0.6	0	0			
• Breastfeeding-related problems	4	1.1	4	2.2	0	0			

* $p \leq 0.05$
*** $p \leq 0.001$

8.3 Results of Rasch Analysis of the English PDSS

An IRT model, specifically the Rating Scale Model, a formulation of an extended Rasch model, was employed in this study as implemented by Winsteps (Linacre, 2009). Rasch analysis was performed on the 35-item PDSS and its Afrikaans translation to determine how well the items defined the underlying construct of postpartum depression in a South African sample. The PDSS was, however, developed as a multidimensional construct of postpartum depression, incorporating seven individual dimensions. Rasch analysis was also performed to determine how adequately the attitude continuum which underlies each PDSS dimension (or construct) was assessed by the five items which constitute the dimension. These additional analyses of the dimensions were considered essential due to the fact that PPD is a phenomenon that is composed of multiple components.

The Rasch model assumes that if people respond to a unidimensional construct they ought to respond as expected according to their ability levels and item difficulty levels (Harvey & Thomas as cited in Maree, personal communication, October 8, 2009). Therefore, the probability of a specific response by a specific person on a specific question is a function of the person's ability (level of depression), and the 'difficulty level' of the question (or the degree of depression that the question is meant to measure). Given that the Rasch model allows one to calculate the level of difficulty required to endorse items, it was possible to determine whether some individual items on the PDSS, or in turn, on each of the PDSS dimensions, were harder to endorse than others.

Unidimensionality was evaluated with fit statistics or indices: a mean-square infit and a mean-square outfit. The analysis of fit statistics is a quality control technique that is necessary to determine the validity of person responses and test items. It allows for the monitoring of the responses of persons and items to determine if and where misfit occurs, and how well the data cooperates with the construction of measurement. When fit statistics fall within an acceptable range for the study, confidence may be placed in item calibration and person measurement (Wright & Stone, 1999). Reasonable MNSQ fit values for a rating scale are recommended at 0.60 – 1.40 (Bond & Fox, 2007; Wright & Linacre, 1994).

Item person construct maps were constructed for the PDSS and the Afrikaans PDSS which show the positions of persons and items on a vertical ruler. This map gives an indication of difficulty indices (degree of depression) and how well the items span the attitude continuum, or, in other words, how well the construct has been differentiated.

The data was also examined to evaluate the effectiveness of the Likert response categories as this impacts on how well the response data defines the dimension. The PDSS and the Afrikaans PDSS were then compared to examine differential item functioning to determine if the items have significantly different meanings across the two samples.

8.3.1 Summary of English Rasch analysis: persons and items.

The summary statistics of the non-extreme persons and items[1] for the English PDSS are presented in Table 21a and Table 21b.The average person infit (1.10) and outfit (1.06) is almost 1 indicating that most persons responded according to expectation. The *SD* provides an indication of the variation of in/outfit values (in this case 0.56 and 0.65). One *SD* above and one *SD* below the mean, represents approximately 68 % of the distribution of values (if the distribution is normal), according to the ideal z- or normal distribution graph. The values are slightly higher than the Afrikaans sample, which are 1.07 and 1.03 respectively. The minimum and maximum values for infit (0.28 – 2.94) and outfit (0.16 – 4.38) are extreme (acceptable value 0.60 – 1.40; Bond & Fox, 2007; Wright & Linacre, 1994) indicating that there are some persons that had unexpected responses on the PDSS.

The min of -5.05 logits for the measure is extremely low indicating that one or some women in this sample were not really depressed. The maximum value of 4.19 on the contrary indicates that some were very depressed. The average logit for person ability was -0.80 with a *SD* of 1.63. This is rather wide and means that approximately 68 % of respondent scores fell within -2.43 and 0.83 logits. If this is the case then the minimum and maximum measure values of the PDSS are really extreme.

The PDSS items functioned well with average infit and outfit values (1.02 and 1.05) close to 1 which is the expected Chi-Square value for these indices. The *SD*s were 0.30 and 0.52 respectively. This indicates that there is neither too much nor too little variation and that most of the items fit the Rasch model. The minimum and maximum

[1] Summary statistics for extreme and non-extreme persons for the PDSS are presented in Table 70 in Appendix F.

values for infit (min 0.64; max 2.01) and outfit (min 0.44; max 2.63) indicate that there are some extreme values.

Reliability information for both items and persons on the PDSS is excellent. The person separation index is high at 4.52. The person reliability estimate is .95 with a Cronbach Alpha (KR-20) value of .98 indicating that the items in the PDSS as a whole were sufficiently able to separate the participants in the sample along the continuum. The person reliability estimate is conceptually equivalent to Cronbach's alpha. The formulation differs though and Cronbach's alpha includes extreme scores, whereas Rasch person reliability estimate is computed without extreme scores. The high person reliability (internal consistency) indicates that the items correlate highly with each other, or in other words, that the participants reacted to the various items in a similar manner.

The PDSS demonstrates an item separation index of 6.65. This indicates that the items are well dispersed on the scale and can distinguish between a number of levels of performance.

Table 21a Summary Statistics of 182 Non-Extreme Persons and Items for the English PDSS.

	Raw Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	49.70	35.00	-0.80	0.26	1.10	0.00	1.06	0.00
S.D.	36.50	0.10	1.63	0.17	0.56	1.80	0.65	1.60
Max	138.00	35.00	4.19	1.03	2.94	5.50	4.38	6.40
Min	1.00	34.00	-5.05	0.16	0.28	-4.30	0.16	-3.90

Real RMSE 0.35 True S.D. 1.60 Separation 4.52 Particip Reliability .95
Model RMSE 0.31 True S.D. 1.60 Separation 5.16 Particip Reliability .96
S.E. of participant mean = 0.12

Minimum Extreme Score: 5 Participants

Table 21b English PDSS: Summary of 35 Measured (Non-Extreme) PDSS

	Raw Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	258.30	186.90	0.00	0.09	1.02	0.00	1.05	0.10
S.D.	84.70	0.20	0.63	0.01	0.30	2.40	0.52	2.50
Max	438.00	187.00	1.28	0.11	2.01	7.30	2.63	8.00
Min	105.00	186.00	-1.43	0.08	0.64	-3.60	0.44	-3.40

Real RMSE 0.09 True S.D. 0.62 Separation 6.65 Particip Reliability .98
Model RMSE 0.09 True S.D. 0.62 Separation 7.04 Particip Reliability .98
S.E. of PDSS mean = 0.11

UMEAN = 0.0000 USCALE = 1.0000
PDSS items raw score-to-measure correlation = -1.00
Data points: 6368 Log-likelihood Chi-Square: 12552.26 with 6047 *d.f.* *p*= 0.0000
Global Root-Mean-Square Residual (exluding extreme scores): 0.8580

8.3.2 Rating scale requirements: English PDSS.

This section examines the quantitative functioning of the English PDSS rating scale. Table 22 contains summary statistics for the 5-point Likert response categories used for the PDSS. Summary statistics for the response categories for individual items are discussed later in this chapter.

a) Category observations

The frequency of responses to the categories of the 5-point Likert rating scale can be seen in Table 22. For response category 0 there were 2679 or 41 % of the total responses. The category that had the least responses were 4 (strongly agree) which had only 12 % or 769 responses. In this summary table no category had less than 10 responses.

No category across all items of the PDSS had less than ten observations, although there were individual items which had response categories with less than 10 observations. The overall response pattern indicates that all category frequency counts for the rating scale are sufficiently large. This indicates that locally stable estimates of the rating scale structure may be produced (Linacre, 2004).

Table 22 Summary Statistics for the 5-Point Likert Response Categories Used for the PDSS

Summary of Category Structure (N = 187)

Category	**Label**	**Score**	**Observed Count**	**%**	**Observed Average[a]**	**Sample Expect.**	**MNSQ**		**Structure Calibration**	**Category Measure**
							Infit	**Outfit**		
0	Strongly Disagree	0	2679	41	-2.17	-2.11	0.96	0.99	NONE	(-2.23)
1	Disagree	1	1311	20	-0.86	-1.01	0.98	0.71	-0.86	-0.88
2	Neither Disagree nor Agree	2	698	11	-0.15	-0.30	0.90	0.96	-0.01	-0.09
3	Agree	3	1086	17	0.23	0.37	1.31	1.61	-0.42	0.81
4	Strongly Agree	4	769	12	1.72	1.71	1.01	1.04	1.28	(2.51)

[a] Observed Average is mean of measures in category, not a parameter estimate.

b) Regular observation distribution

All categories were used fairly regularly, although category 0 (strongly disagree) was selected more frequently and has an observed count of 41%. Category 1 (disagree) follows at 20% (interestingly these values are the same in the same in the Afrikaans sample) and category 3 (agree) at 17%. Category 2 (neither disagree nor agree) and 4 (strongly agree) have the least observations (11% and 12% respectively). This indicates that mothers were less likely to choose the middle category (neither disagree nor agree) and the most extreme category (strongly agree) and that redundant categories may exist.

c) Average measures advance monotonically with category

The average measures (expressed as logits) increase from small to large in categories: -2.17, -0.86, -0.15, 0.23 and 1.72. The observed average measures demonstrate values that are fairly close to their expected values.

d) OUTFIT mean-squares less than 2

Outfit mean-squares indicate random noise and unexpected observations in a category. Most categories demonstrate values close to the expected 1.0. Category 3 (agree) had the largest value (1.61) indicating that the category has been used unexpectedly. A value of 1.6 is still considered acceptable for this sample.

e) Step calibrations advance orderly

The step calibrations should advance from easy to hard uniformly. In Table 22 the step calibrations are -0.86, -0.01, -0.42 and 1.28. The pattern is similar to the Afrikaans PDSS with disordered transition between categories 1 and 2 as well as between categories 2 and 3. Figure 3 shows that category 2 does not form a prominent hill on the graph as it should, indicating that it is relatively rarely observed (Linacre, 2004). If either categories 1 and 2 were combined, or categories 2 and 3, it would form a more prominent category and the transition between 1 and 3 will be as expected. Categories 0, 3 and 4 form distinct peaks, while category 1's peak is also somewhat submerged.

f) Step difficulties advance by at least 1 logit

The categories have step difficulties which advance as follows:

Categories 1-2: -.01 – (-0.86) = 0.85

Categories 2-3: -0.42 – (-.01) = -0.41

Categories 3-4: 1.28 – (-0.42) = 1.7

According to Linacre (2004), a five category rating scale should ideally advance by at least 1 logit Linacre (2004, p.274). The width of advances for categories 1 to 2 and categories 2 to 3 is somewhat narrow. This confirms the problematic nature of category 2. The step calibration of categories 2-3 are especially problematic and may indicate substantive problems with the rating scale category definitions when used with this sample. Categories 3-4 demonstrate an adequate step of 1.7 logits.

```
        CATEGORY PROBABILITIES: MODES - Structure measures at
intersections
P      -+---------+---------+---------+---------+---------+---------+-
R  1.0 +                                                             +
O      |                                                             |
B      |00                                                           |
A      |  0000                                                     44|
B   .8 +      00                                                444  +
I      |        000                                           44     |
L      |           00                                       44       |
I      |             0                                    44         |
T   .6 +              00                                44           +
Y      |                00                            44             |
    .5 +                  0                  3333   44               +
O      |                   0             3333    33*3                |
F   .4 +                    00         33        44  333             +
       |                111111*111    3        44       33           |
R      |             111       0  11*3        4           333        |
E      |          111           00 32**2    44               333     |
S   .2 +       111             22**2   1*2**                    333  +
P      |   1111             22233  00   4*1 2222                   33|
O      |111             2222333      ***   111  2222                 |
N      |         22222223333     4444   0000  11111 2222222          |
S   .0 +****************444444444           0000000******************+
E      -+---------+---------+---------+---------+---------+---------+-
       -3        -2        -1         0         1         2         3
      PARTICIPANT [MINUS] PDDS MEASURE
```

Figure 3 Probability curves of observations in each category of the PDSS.

8.3.3 Item person construct map: English PDSS.

Table 23 represents a geographical description of the two facets – participants and PDSS items. In this table the items are shown located at their calibrated measures. This allows for the comparison of both person ability (the presence of depression) and item difficulty (difficult items indicate more depression). A mapped hierarchy of the 35 items is provided along the vertical logit ruler. The items at the bottom of this figure are those that are easier for the participants to agree with. The items at the upper end are those that are more difficult to agree with. PDSS items are positioned according to its measure in logits. Ideally items should be spread out along the vertical logit ruler. This indicates good variable definition and is important for construct validity. From Table 23a it can be seen that in many instances more than two items are positioned on the same logit measure.

It seems as if insufficient items are present at either end of the difficulty level. This may indicate that low-ability (non-depressed) people did not understand the items, were unfamiliar with an expression used, or that the questionnaire is not appropriate for non-clinically depressed people. However, Table 23b shows that the categories in the rating scale cover the spread of person abilities well. The spread of ability (the absence or presence of depression) is much wider than the spread of item difficulty. There is an overrepresentation of items at the mean level and insufficient items at the upper and lower ends of the vertical logit ruler to allow for a proper description of the high and low scoring person and to determine depression accurately. A similar distribution is evident in the Afrikaans sample, but in the English sample the distribution extends more toward the upper end of the vertical logit ruler indicating more English participants who scored higher than the items were able to measure.

From the distribution along the vertical logit ruler, it is evident that a significant proportion of the English sample screened positively for postpartum depression. Another significant proportion of the English sample screened negatively for postpartum depression.

Items 7 and 21 from the SUI dimension were the items that were most difficult to strongly agree with. These are closely followed by the remaining three items from the same dimension. Yet there were still participants who scored higher than the items could measure. This indicates that some measurement precision is lost at the most difficult level.

8.3.4 Item fit: English PDSS.

Item fit for the English PDSS is indicated in Table 24. A range of 0.60 to 1.40 for MNSQ infit and outfit are acceptable limits (Bond & Fox, 2007; Wright & Linacre, 1994). Items 1, 8, 15, 22, and 29 had infit mean-squares greater than 1.40 which indicates that they either do not fit the definition of the constructs they are measuring very well (thus forming another constructs). All these items are from the Sleeping/Eating Disturbances (SLP) dimension and their poorer fit values within the total PDSS may be a reflection that they form a separate dimension. No items were overfitted (i.e. < 0.60).

Table 23a Item-Person Distribution Map for the English PDSS (N = 187)

```
  PARTICIPANT - MAP - PDDS
              <more>|<rare>
  5                 +
                    |
                    |
                    |
                 .  |
  4                 +
                 .  |
                    |
                 .  |
                    |
  3              #  +
                 .  |
                    |
               .# T|
                 .  |
  2              .  +
                    |
                    |
                 #  |
                 #  |T PDSS_21  PDSS_7
  1              .  +  PDSS_14
               .# S|  PDSS_29  PDSS_35
              ###  |S PDSS_16  PDSS_28  PDSS_8
           .#####  |  PDSS_30
           .#####  |  PDSS_15  PDSS_18  PDSS_19  PDSS_20  PDSS_22  PDSS_27
                      PDSS_33
  0       #######  +M PDSS_10  PDSS_11  PDSS_12  PDSS_26  PDSS_32
            .####  |  PDSS_1   PDSS_31  PDSS_34  PDSS_4
           ######  |  PDSS_13  PDSS_17  PDSS_25
            #####  |S PDSS_23  PDSS_5   PDSS_6
           .##### M|  PDSS_2
 -1         .####  +
           .#####  |T PDSS_3   PDSS_9
              .##  |  PDSS_24
             ####  |
              .##  |
 -2           ###  +
                #  |
              ### S|
              .##  |
               ##  |
 -3            .#  +
               .#  |
                   |
               ##  |
                .  |
 -4               T+
                   |
               ##  |
                   |
                   |
 -5            .#  +
                   |
                   |
                   |
                   |
 -6           .##  +
              <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.
```

Table 23b Item Category-Person Distribution Map for the English PDSS (N = 187)

```
  PARTICIPANT - MAP - PDDS - Expected score zones (Rasch-half-point thresholds)
           <more>|  Disagree   Neither D  Agree      Strongly Agree
5                +
                 |
                 |
                 |
              .  |
4                +
              .  |
                 |
              .  |
                 |
3             #  +
              .  |                                   PDSS_2.35
                 |                                   PDSS_1.35
                                                     PDSS_7.35
             .# T|                                   PDSS_1.35
                                                     PDSS_1.35
                                                     PDSS_8.35
              .  |                                   PDSS_2.35
                                                     PDSS_3.35
2             .  +                                   PDSS_2.35
                                                     PDSS_3.35
                 |                                   PDSS_1.35
                                                     PDSS_1.35
                                                     PDSS_1.35
                                                     PDSS_2.35
                                                     PDSS_3.35
                 |                        PDSS_2.25  PDSS_1.35
                                                     PDSS_1.35
                                                     PDSS_1.35
                                                     PDSS_2.35
                                                     PDSS_2.35
                                                     PDSS_3.35
                                                     PDSS_4.35
              #  |                        PDSS_7.25  PDSS_1.35
                                                     PDSS_2.35
                                                     PDSS_2.35
                                                     PDSS_3.35
              #  |T                       PDSS_1.25  PDSS_1.35
                                                     PDSS_3.35
                                                     PDSS_6.35
1             .  +                        PDSS_2.25  PDSS_2.35
                                          PDSS_3.25  PDSS_5.35
                                                     PDSS_9.35
             .# S|             PDSS_2.15  PDSS_1.25  PDSS_2.35
                               PDSS_7.15  PDSS_2.25  PDSS_3.35
                                          PDSS_3.25
                                          PDSS_8.25
            ###  |S            PDSS_1.15  PDSS_2.25  PDSS_2.35
                                          PDSS_2.25
                                          PDSS_3.25
         .#####  |             PDSS_2.15  PDSS_1.25
                               PDSS_2.15  PDSS_1.25
                               PDSS_3.15  PDSS_1.25
                                          PDSS_1.25
                                          PDSS_1.25
                                          PDSS_1.25
                                          PDSS_2.25
                                          PDSS_3.25
         .#####  |                        PDSS_1.25
                                          PDSS_2.25
                                          PDSS_3.25
                                          PDSS_4.25
0       #######  +M PDSS_7.05  PDSS_1.15  PDSS_1.25
```

Table 23b (continued) Item Category-Person Distribution Map for the English PDSS (N = 187)

```
       PARTICIPANT - MAP - PDDS - Expected score zones (Rasch-half-point thresholds)
                <more>|  Disagree    Neither D  Agree       Strongly Agree

    0        #######  +M PDSS_7.05   PDSS_1.15  PDSS_1.25
                                     PDSS_2.15  PDSS_2.25
                                     PDSS_3.15  PDSS_3.25
                                     PDSS_8.15
                .####  |  PDSS_1.05  PDSS_1.15  PDSS_1.25
                          PDSS_2.05  PDSS_1.15  PDSS_2.25
                          PDSS_3.05  PDSS_2.15  PDSS_5.25
                                     PDSS_2.15  PDSS_6.25
                                     PDSS_3.15
               ######  |  PDSS_2.05  PDSS_1.15  PDSS_2.25
                                     PDSS_1.15
                                     PDSS_1.15
                                     PDSS_2.15
                                     PDSS_3.15
                #####  |S PDSS_2.05  PDSS_1.15  PDSS_3.25
                                     PDSS_3.15
                                     PDSS_3.15
                                     PDSS_4.15
               .##### M|             PDSS_1.15  PDSS_2.25
                                     PDSS_1.15  PDSS_9.25
   -1           .####  +  PDSS_2.05  PDSS_1.15
                          PDSS_2.05  PDSS_2.15
                          PDSS_8.05  PDSS_2.15
                                     PDSS_5.15
                                     PDSS_6.15
               .#####  |T PDSS_1.05  PDSS_2.15
                          PDSS_1.05
                          PDSS_1.05
                          PDSS_2.05
                          PDSS_3.05
                  .##  |  PDSS_1.05
                          PDSS_3.05
                          PDSS_3.05
                          PDSS_3.05
                 ####  |  PDSS_1.05  PDSS_3.15
                          PDSS_1.05
                          PDSS_2.05
                  .##  |  PDSS_1.05  PDSS_2.15
                          PDSS_1.05  PDSS_9.15
                          PDSS_4.05
   -2             ###  +  PDSS_3.05
                    #  |  PDSS_1.05
                          PDSS_2.05
                          PDSS_5.05
                          PDSS_6.05
                  ### S|  PDSS_1.05
                          PDSS_2.05
                  .##  |
                   ##  |
   -3              .#  +  PDSS_2.05
                          PDSS_3.05
                   .#  |
                       |  PDSS_9.05
                   ##  |  PDSS_2.05
                    .  |
   -4                 T+
                       |
                   ##  |
                       |
                       |
   -5              .#  +
                       |
                       |
                       |
                       |
   -6             .##  +
                <less>|  Strongly    Disagree   Neither D  Agree
EACH "#" IS 2. EACH "." IS 1.
```

The difficulty level in logits (measure) and the measurement error (model SE) for each item are also indicated in Table 24. The Rasch error estimate, a standard error estimate (referred to as model error or model S.E.) indicates measurement precision (Wright, 1995). Smaller error estimates are better. However, if a respondent (or item) has haphazard responses it will reflect in a larger infit or outfit value. A large SEM means that less confidence can be placed in that respondent's (or item's) estimated score. All measurement error values for the English PDSS are small with values less than 0.12 and a mean of 0.09.

The Pearson item-total correlation (r_{it}) gives an indication of construct validity and whether there may be coding problems. The Pearson item-total correlation (r_{it}) has a range of -1 to +1. Negative or zero values suggest persons or items with response strings that contradict the variable, or no fit. A high negative correlation indicates a reverse coding problem. From Table 24 it can be seen that no negative correlations are evident. Furthermore, similar to the values for the Afrikaans PDSS, all the values are fairly high in spite of some fit problems. Pearson item-total correlation (r_{it}) values range from 0.55 to 0.79 with no negative correlations. Correlations are expected to higher within the PDSS dimensions than in the PDSS total, which, as a whole, may be considered multidimensional.

Table 24 Item Statistics for the English PDSS Total: Misfit Order (N = 187)

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	PDDS
1	281	187	-0.10	.09	2.01	7.3	2.63	8.0	A .55	.72	37.9	49.0	PDSS_1
8	200	187	0.52	.09	1.43	3.4	2.29	4.5	B .60	.68	48.4	52.3	PDSS_8
29	157	187	0.74	.10	1.46	3.1	1.95	3.0	C .58	.65	59.9	59.4	PDSS_29
22	223	187	0.30	.09	1.60	4.6	1.80	3.2	D .61	.69	45.1	51.2	PDSS_22
15	234	186	0.17	.08	1.60	4.6	1.72	3.0	E .62	.70	49.2	50.9	PDSS_15
2	367	187	-0.85	.09	1.36	3.1	1.67	4.6	F .68	.75	41.8	47.3	PDSS_2
31	284	187	-0.15	.08	1.20	1.7	1.61	2.3	G .69	.72	54.9	51.0	PDSS_31
30	212	187	0.39	.09	1.17	1.4	1.40	2.0	H .67	.69	54.9	51.4	PDSS_30
24	438	187	-1.43	.09	1.10	0.9	1.31	2.1	I .74	.77	44.5	46.8	PDSS_24
27	211	187	0.25	.08	0.92	-0.7	1.18	0.9	J .68	.68	58.2	53.4	PDSS_27
13	333	187	-0.48	.08	1.13	1.2	1.05	0.4	K .74	.74	41.8	49.0	PDSS_13
20	217	187	0.21	.08	1.10	0.9	0.93	-0.2	L .67	.67	53.3	53.3	PDSS_20
21	105	187	1.28	.11	1.07	0.5	0.78	-0.5	M .61	.61	68.1	68.5	PDSS_21
23	356	187	-0.68	.08	1.02	0.2	0.99	0.0	N .75	.75	52.2	47.6	PDSS_23
3	409	187	-1.10	.09	1.01	0.1	1.01	0.1	O .77	.77	41.8	47.2	PDSS_3
9	424	187	-1.26	.09	0.99	0.0	0.94	-0.5	P .79	.77	51.1	49.8	PDSS_9
25	314	187	-0.34	.09	0.98	-0.2	0.93	-0.5	Q .76	.74	47.8	48.0	PDSS_25
32	271	187	-0.01	.09	0.96	-0.4	0.87	-0.9	R .74	.72	54.9	48.3	PDSS_32
11	290	187	-0.09	.09	0.96	-0.4	0.96	-0.3	q .74	.73	49.5	47.5	PDSS_11
5	332	187	-0.56	.08	0.87	-1.3	0.94	-0.4	p .76	.74	51.1	47.2	PDSS_5
4	275	187	-0.13	.08	0.92	-0.8	0.94	-0.4	o .74	.72	50.0	48.5	PDSS_4
16	206	187	0.52	.09	0.92	-0.7	0.92	-0.4	n .71	.69	57.7	51.2	PDSS_16
7	106	187	1.16	.11	0.92	-0.4	0.65	-0.8	m .61	.60	70.3	69.7	PDSS_7
17	314	187	-0.40	.08	0.87	-1.2	0.91	-0.5	l .75	.73	52.2	48.5	PDSS_17
6	336	187	-0.55	.08	0.89	-1.0	0.86	-1.0	k .76	.74	46.2	47.8	PDSS_6
28	142	187	0.66	.09	0.88	-0.8	0.59	-1.2	j .65	.62	68.7	65.6	PDSS_28
14	119	187	1.01	.10	0.85	-1.0	0.66	-0.7	i .63	.61	65.9	67.4	PDSS_14
34	266	187	-0.13	.08	0.76	-2.3	0.58	-2.6	h .75	.70	60.4	49.9	PDSS_34
19	242	187	0.10	.08	0.75	-2.4	0.59	-2.4	g .74	.70	59.3	50.5	PDSS_19
35	133	186	0.75	.09	0.71	-2.1	0.44	-1.8	f .65	.61	70.7	66.4	PDSS_35
18	232	187	0.14	.08	0.71	-2.8	0.54	-2.5	e .74	.69	61.0	51.3	PDSS_18
12	255	187	0.03	.09	0.70	-3.0	0.61	-2.7	d .77	.71	53.8	49.4	PDSS_12
26	265	187	-0.10	.08	0.69	-3.1	0.56	-3.1	c .77	.71	61.5	49.2	PDSS_26
10	260	187	-0.02	.08	0.69	-3.1	0.60	-2.6	b .76	.71	56.6	49.1	PDSS_10
33	232	187	0.15	.09	0.64	-3.6	0.51	-3.4	a .76	.70	64.3	51.0	PDSS_33
MEAN	258.3	186.9	0.00	.09	1.02	0.0	1.05	0.1			54.4	52.4	
S.D.	84.7	0.2	0.63	.01	0.30	2.4	0.52	2.5			8.5	6.6	

PARTICIPANT: REAL SEP.: 3.86 REL.: .94
PDDS: REAL SEP.: 6.65 REL.: .98

8.3.5 Dimensionality: English PDSS.

A Rasch principle component analysis (PCA) of residuals was performed. Residuals are the differences between the scores that are predicted by the Rasch model and the actual scores that are observed (Chou & Wang, 2010; Hong & Wong, 2005). The PCA indicates the presence of secondary dimensions (Linacre, 2009) and was performed using calibrated data (logits) as opposed to raw data to avoid non-linearity in data accumulating in the PCA (Maree, personal communication, October 12, 2009). Table 25 indicates the variance explained by the measures and raw unexplained variance. The empirical values match the modelled values perfectly in most instances, which indicate that the measures explain the expected amount of variance in the data.

The variance explained by the measures is 64.60 eigenvalues or 64.9% which means that the measures explains most of the variance and that the English PDSS has a wide spread of items and persons with different abilities, i.e. different degrees of PPD. Raw unexplained variance is 35.1%. Eigenvalues greater than 1.40 are indicative of possible secondary dimensions. The unexplained variance in the first contrast is 3.60 eigenvalues (3.7%), in the second contrast, 3.20 eigenvalues (3.2%), in the third contrast, 3 eigenvalues (3%), in the fourth contrast 2.20 eigenvalues (2.2%), and in the fifth contrast 1.90 eigenvalues (2%). These values indicate that five additional dimensions exist, and that the PDSS is a multidimensional screening scale. The plot in Figure 4 below as well as the loadings of factors in Table 26 also suggest that dimensionality in the PDSS exists.

Table 25 Variance Decomposition of the Observations for the English PDSS Items (N = 187)

	Empirical			Modeled
	Eigenvalue units	%	%	%
Total raw variance in observations	99.60	100.00		100.00
Raw variance explained by measures	64.60	64.90		65.60
Raw variance explained by persons	33.80	33.90		34.30
Raw variance explained by items	30.90	31.00		31.30
Raw unexplained variance (total)	35.00	35.10	100.00	34.40
Unexplained variance in 1st contrast	3.60	3.70	10.40	
Unexplained variance in 2nd contrast	3.20	3.20	9.00	
Unexplained variance in 3rd contrast	3.00	3.00	8.60	
Unexplained variance in 4th contrast	2.20	2.20	6.20	
Unexplained variance in 5th contrast	2.00	2.00	5.60	

```
       STANDARDIZED RESIDUAL CONTRAST 1 PLOT

     -2             -1               0              1               2
     -+---------------+---------------+---------------+---------------+- COUNT
   .7 +                               |                               +
      |                               |    A                          | 1
   .6 +                             B |  C                            + 2
      |                               |                               |
C  .5 +                               |                               +
O     |                               |                               |
N  .4 +                               |                               +
T     |                               |                               |
R  .3 +                 D             |                               + 1
A     |                               F       E                       | 2
S  .2 +                               |     H G                       + 2
T     |             I               K |           J                   | 3
   .1 +                         ML    |                               + 2
1     |        O                    N | P                             | 3
   .0 +-------------------------------|-------------------------------+
L     |                      Q        |                               | 1
O -.1 +                             R |                               + 1
A     |                             q p o                n            | 4
D -.2 +                    m          |                               + 1
I     |                               |                   l           | 1
N -.3 +           k                   | h i           j               + 4
G     |                               |  g                            | 1
  -.4 +                               |           f                   + 1
      |                      c        e          d                    | 3
  -.5 +                       b       |                               + 1
      |                             a |                               | 1
  -.6 +                               |                               +
     -+---------------+---------------+---------------+---------------+-
     -2              -1               0               1               2
                                PDDS MEASURE
 COUNT:        1  1 1   1  1 21 11  6 3 32111 2  12   1  11
```

Figure 4 Standardized residual contrast of English PDSS items.

Table 26 Standardized Residual Loading for the English PDSS (Sorted by Loading)

PDSS Dimension	PDSS Item	Item Content	Loading	Measure	MNSQ Infit	MNSQ Outfit	Entry Number
SLP	1	I had trouble sleeping even when my baby was asleep.	.61	-0.21	1.96	2.74	A
SLP	22	I tossed and turned for a long time at night trying to fall asleep.	.59	0.21	1.65	1.77	B
SLP	15	I woke up on my own in the middle of the night and had trouble getting back to sleep.	.52	0.13	1.66	1.64	C
ANX	2	I got anxious over even the littlest things that concerned my baby.	.33	-0.80	1.28	1.78	D
MNT	32	I had difficulty focusing on a task.	.32	-0.13	0.88	0.82	E
MNT	11	I could not concentrate on anything.	.29	-0.26	0.83	0.95	F
ANX	16	I felt like I was jumping out of my skin.	.25	0.34	0.91	0.91	G
ANX	30	I felt like I had to keep moving or pacing.	.24	0.29	1.16	1.35	H
SLP	8	I lost my appetite.	.24	0.39	1.47	2.00	I
SLP	29	I knew I should eat but I could not.	.20	0.76	1.52	1.76	J
MNT	25	I had a difficult time making even a simple decision.	.19	-0.43	0.90	0.90	K
ELB	3	I felt like my emotions were on a roller coaster.	.19	-1.08	0.94	0.97	L
ELB	24	I have been very irritable.	.10	-1.29	1.03	1.45	M
MNT	4	I felt like I was losing my mind.	.05	-0.17	0.89	0.92	N
ELB	17	I cried a lot for no real reason.	.01	-0.44	0.90	0.88	O
GLT	34	I felt like a failure as a mother.	-.57	-0.10	0.86	0.67	a
SUI	35	I just wanted to leave this world.	-.57	1.01	1.02	0.64	b
SUI	28	I felt that my baby would be better off without me.	-.56	0.91	1.16	0.78	c
SUI	14	I started thinking that I would be better off dead.	-.48	1.18	1.15	0.83	d
ELB	10	I was scared that I would never be happy again.	-.42	-0.05	0.69	0.60	e
SUI	21	I wanted to hurt myself.	-.39	1.36	1.24	0.90	f
SUI	7	I have thought that death seemed like the only way out of this living nightmare.	-.39	1.35	1.21	0.77	g
GLT	6	I felt like I was not the mother I wanted to be.	-.38	-0.58	0.85	0.84	h
GLT	13	I felt like so many mothers were better than me.	-.38	-0.56	1.08	0.99	i
GLT	20	I felt guilty because I could not feel as much love for my baby as I should.	-.36	0.26	1.29	1.00	j
GLT	27	I felt like I had to hide what I was thinking or feeling towards the baby.	-.36	0.31	1.01	1.08	k
LOS	19	I did not know who I was anymore.	-.29	0.08	0.80	0.64	l
ANX	23	I felt all alone.	-.20	-0.71	1.03	0.95	m
LOS	33	I did not feel real.	-.16	0.15	0.63	0.50	n
ELB	31	I felt full of anger ready to explode.	-.16	-0.23	1.28	1.33	o
LOS	12	I felt as though I had become a stranger to myself.	-.12	-0.02	0.67	0.60	p
ANX	9	I felt really overwhelmed.	-.10	-1.19	0.85	0.88	q
LOS	26	I felt like I was not normal.	-.10	-0.09	0.70	0.58	R
LOS	5	I was afraid that I would never be my normal self again.	-.08	-0.54	0.82	0.88	Q
MNT	18	I thought I was going crazy.	-.06	0.14	0.79	0.62	P

8.3.6 Performance of English PDSS dimensions: Rasch analysis of persons and items.

This section presents the results of the Rasch analysis of persons and items for the seven dimensions of the English PDSS. Summary statistics for each dimension is presented in Table 27 and is discussed below. This will be followed by the item fit statistics for the five items that constitute each dimension.

8.3.6.1 Sleeping/Eating Disturbances (SLP) dimension.

Person and item information for the Sleep/Eating dimension is presented in Table 27. Winsteps (Linacre, 2009) eliminated 37 respondents in this dimension with extreme scores who scored all high (4's) or all low (0's) hence the observed count of 150 participants. The average raw score of persons in this dimension is the second lowest of the seven dimensions at 7.20.

Person fit to the Rasch model is an index of whether individuals are responding to items in a consistent manner or whether the responses are erratic or idiosyncratic. The person infit mean-squares statistic = 0.96 with at t-statistic of -0.10, and the outfit mean-square statistic is 0.98 with a t-value of 0.00. These values are near to the Rasch-modeled expectations of 1.00. This indicates that there is neither too much nor too little variation with most participants responding as expected. The *SD* infit and outfit values for this dimension are 0.68 and 0.75 respectively. The minimum and maximum values for infit (0.07 and 3.72) and outfit (0.06 and 4.35) are extreme. This indicates that there are some persons that had unexpected responses to items on the SLP dimension.

Table 27 Summary Statistics for the PDSS Dimensions

Statistic		Sleeping / eating disturbances	Anxiety / insecurity	Emotional lability	Cognitive impairment	Loss of self	Guilt / shame	Contemplating harming oneself
Mean raw score	Items	219.00	313.00	341.00	276.40	265.20	272.60	121.00
	Persons	7.20	8.90	9.10	8.20	7.90	8.40	6.50
Measure (logits)	Items	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Persons	-0.66	-0.34	-0.33	-0.72	-0.89	-0.55	-1.15
Model error	Items	0.09	0.10	0.10	0.11	0.13	0.11	0.15
	Persons	0.57	0.63	0.66	0.69	0.74	0.71	0.67
SD (logits)	Items	0.32	0.90	0.76	0.28	0.60	0.68	0.58
	Persons	1.22	1.69	1.81	2.04	2.28	2.11	1.75
M Infit MNSQ	Items	1.01	1.01	1.02	0.98	1.00	1.02	0.99
	Persons	0.96	1.01	0.98	0.97	0.98	0.97	0.99
M Outfit MNSQ	Items	0.98	1.01	0.97	0.98	0.98	0.95	0.93
	Persons	0.98	1.01	0.97	0.98	0.98	0.95	0.93
Mean Infit (t)	Items	0.00	0.10	0.20	-0.10	0.00	0.00	-0.30
	Persons	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10	0.00
Mean Outfit (t)	Items	-0.20	0.10	-0.30	-0.10	-0.20	-0.60	-0.50
	Persons	0.00	0.00	-0.10	-0.10	-0.10	0.00	0.00
Separation	Items	3.21	8.70	7.25	2.20	4.48	5.51	3.34
	Persons	1.56	2.02	2.10	2.30	2.41	2.38	2.06
Cronbach alpha		.85	.88	.91	.93	.95	.93	.95
Rasch reliability	Persons	.71	.80	.82	.84	.85	.85	.81

MNSQ = mean-square

Reasonable mean-square fit values for a rating scale are recommended at 0.60 – 1.40 (Wright & Linacre (1994). The minimum of -2.70 logits for the items in this dimension is very low indicating that one or some women in this sample did not have symptoms of sleeping/eating disturbances. The maximum value of 2.93, however, indicates that some participants had significant symptoms of sleeping/eating. The average logit for person ability or measure of sleeping/eating disturbances is -0.66 with a model standard error of 0.57 and a *SD* of 1.22. This means that almost 68 % of participants fell within a range of -1.88 and 0.56 logits (assuming that the distribution is approximately normal). Therefore the minimum measure value is (-2.70) is not extreme. The maximum measure value (2.93) is extreme. Extreme values are at least two standard deviations (i.e. 1.22 x 2 = 2.44) from the mean. This is approximately on a 5 % significance level.

The PDSS items on the SLP dimension functioned very well, but this will be confirmed later in the chapter when the items are examined individually in more detail. The average infit and outfit values of 1.01 (t = 0.00) and 0.98 (t = -0.20) respectively are ideal Chi-Square values for these indices. The infit and outfit *SD* values were 0.16 and 0.16. This indicates that there is very little variation and that most of the items in this dimension fit the Rasch model. The minimum and maximum MNSQ statistics for infit (min 0.80; max 1.18) as well as the minimum and maximum MNSQ statistics for outfit (min 0.77; max 1.16) in the SLP dimension are within an adequate range and indicate that the items function well together within this dimension.

Item reliability and item separation indices provide an indication of the measure's ability to define a distinction hierarchy of items along the measured variable. More confidence can be placed in an item's constant placement across different samples when these values are higher (Bond & Fox, 2001). A large item separation index also demonstrates better confidence in the spread of items across the targeted continuum (Beck & Gable, 2001e).

Reliability information for both items and persons on the SLP dimension is also presented in Table 27. The person separation index is 1.56. This value is the lowest of the 7 dimensions and indicates that persons are not as well separated across the SLP dimension as they are on the other PDSS dimensions. The Rasch person reliability estimate is conceptually equivalent to Cronbach alpha (or KR-20 in the dichotomous case), but is computed using logits and does not include extreme scores making its value lower than that for Cronbach alpha (Linacre, 2009). Cronbach alpha is the conventional "test" reliability index which reports an estimated test reliability based on the sample's raw scores and is computed on the complete data set, including extreme scores. The Rasch model's reliability determination, based on logits and excluding extreme scores is the preferred reliability estimate. The Rasch person reliability estimate for the SLP dimension is also lower than other dimensions at .71 along with the Cronbach Alpha (KR-20) value at .85. The SLP dimension therefore demonstrates adequate internal consistency, but it is lower than that of the other PDSS dimensions. Participants are not responding as consistently across the 5 items of this dimension and the PDSS may not be screening the participants' level of sleep and eating disturbances as well as the other facets of PPD. The SLP dimension demonstrates a much lower item separation index of 3.21. The items on the SLP dimension are therefore not as well dispersed on the scale.

8.3.6.2 Anxiety/Insecurity (ANX) dimension.

The person and item information for the ANX dimension is also presented in Table 27. Data for this dimension is presented for 169 participants as Winsteps (Linacre, 2009) eliminated 18 respondents with extreme scores. The average raw score of persons in this dimension is the second highest of the 7 dimensions at 8.90.

The person infit mean-squares statistic is 1.01 with a t-statistic of -0.10, and the outfit mean-square statistic is 1.01 (t = 0.00). These values are near to the Rasch-modeled expectations of 1.00. The *SD* infit and outfit values for this dimension are fairly wide at 0.85 and 0.90 respectively. The minimum and maximum MNSQ statistics for infit (0.03 and 4.75) and outfit (0.03 and 4.83) indicate that there are some persons that had unexpected responses to items on this dimension.

The minimum of -3.81 logits for the items in this dimension is very low indicating that one or some women in this sample did not have symptoms of anxiety/insecurity. The maximum value of 3.79, however, indicates that some participants had significant symptoms of anxiety/insecurity. The average logit for person ability or measure of anxiety/insecurity levels, is -0.34 with a model standard error of 0.63 and a *SD* of 1.69. This means that almost 68 % of participants fell within a range of -2.03 and 1.35 logits (assuming that the distribution is approximately normal). Therefore the minimum and maximum measure values of 3.78 and -3.80 are extreme.

The PDSS items on the Anxiety/Insecurity dimension functioned very well, but this will be confirmed later in the chapter when the items are examined individually in more detail. The average MNSQ indices for both item infit and outfit are ideal at 1.01 (t =

0.10). The infit and outfit *SD* values were 0.11 and 0.21. This indicates that there is very little variation and that most of the items in this dimension fit the Rasch model. The minimum and maximum MNSQ statistics for infit (min 0.89; max 1.20) as well as the minimum and maximum MNSQ statistics for outfit (min 0.79; max 1.40) in the ANX dimension are adequate, although the maximum outfit value is slightly elevated. The values are, however, not as extreme as those for the PDSS as a whole (min infit 0.63; max infit 1.96; min outfit 0.50; max outfit 2.74). This indicates that the items function well together within this dimension.

Reliability information for both items and persons on the ANX dimension shows a person separation index of 2.02. This indicates that persons are sufficiently separated across the ANX dimension. The Rasch person reliability estimate for the Anx dimension is .80 and the Cronbach Alpha value is .88. This indicates that the PDSS items separated the participants well along the continuum. It further demonstrates good internal consistency of responses to items and that the items correlate highly with each other. Participants are therefore responding in a consistent fashion across the 5 items of this dimension.

The PDSS's ANX dimension therefore adequately screens for participants' levels of anxiety. The PDSS ANX dimension demonstrates an item separation index of 8.70. This indicates that the items on the ANX dimension are well dispersed on the scale and can distinguish between a number of levels of performance.

8.3.6.3 Emotional Lability (ELB) dimension.

The person and item information for the ELB dimension can also be found in Table 27. Winsteps (Linacre, 2009) eliminated 21 respondents in this dimension with extreme scores and the data is presented for the remaining 166 participants with non-extreme scores. The average raw score of persons in this dimension is 9.10. This is the highest raw score of the seven dimensions. The infit mean-squares statistic is 0.98 (t = -0.10) and the outfit mean-square statistic is 0.97 (t = -0.10). Both these values are close to the Rasch-modeled expectations of 1.00 with little variation and participants responding as expected with good fit to the Rasch model. The *SD* infit and outfit values for this dimension are rather wide at 0.95 and 1.10 respectively.

The minimum and maximum MNSQ statistics for person infit (0.11 and 5.59) and outfit (0.10 and 9.25) are extreme (acceptable value 0.60 – 1.40). The maximum MNSQ statistic is high for outfit (9.25) – the most extreme of all the dimensions. This indicates that there are persons that had unexpected responses to items on the ELB dimension.

The minimum and maximum values in logits (-3.67 and 3.62 respectively) for the items in this dimension are extreme indicating that one or some women in this sample did not have symptoms of emotional lability and that one or some participants had significant symptoms of emotional lability. The average logit for person ability is -0.33 with a model standard error of 0.66 and a *SD* of 1.81. Therefore, around 68% of participants fell within a range of -2.14 and 1.48 logits. The minimum measure value (-3.65) is not extreme. The maximum measure value (3.62) is extreme.

The items on the ELB dimension appear to have functioned very well. This will be confirmed later in the chapter when they are examined individually. The average infit and outfit values are 1.02 (t = 0.20) and 0.97 (t = -0.3) respectively. The infit and outfit *SD* values are 0.09 and 0.14 indicating that there is very little variation and that most of the items in the ELB dimension fit the Rasch model. The minimum and maximum MNSQ statistics for infit (min 0.89; max 1.15) as well as the minimum and maximum MNSQ statistics for outfit (min 0.73; max 1.16) in this dimension are adequate indicating that the items in the ELB dimension did not have extreme values and function well together.

Reliability information for items and persons on the ELB dimension shows a person separation index of 2.10. This indicates that persons are sufficiently separated across this dimension. The person reliability estimate for the ELB dimension is good at .82 and the Cronbach Alpha value of .91 also indicates good internal consistency of responses to items. This demonstrates consistent responding by participants across the 5 items of this dimension. The PDSS's ELB dimension therefore adequately screens for participants' levels of emotional lability. Items in this dimension are well dispersed on the scale with an item separation of 7.25.

8.3.6.4 Mental Confusion (MNT) dimension.

Person and item information for the MNT dimension is also presented in Table 27. Winsteps (Linacre, 2009) eliminated 36 respondents in this dimension with extreme scores and the data is presented for the remaining 151 participants with non-extreme scores. The average raw score of persons in this dimension is 8.20. The person infit

MNSQ statistic is 0.97 (t = -0.10) and the outfit mean-square statistic is 0.98 (t = -0.10). Both these values are close to the Rasch-modeled expectations of 1.00. Little variation is evident and participants responded as expected with good fit to the Rasch model. The *SD* infit and outfit values for this dimension are fairly wide at 0.81 and 0.86 respectively. The minimum and maximum MNSQ statistics for infit (0.02 and 4.98) and outfit (0.02 and 5.92) are also extreme (acceptable value 0.60 – 1.40) and indicative of some unexpected responses to items on the MNT dimension.

Furthermore, the extreme minimum and maximum values in logits (-4.25 and 4.60 respectively) for the items in this dimension indicate that one or some women in this sample did not have symptoms of mental confusion while one or some participants had significant symptoms. The average logit for person ability is -0.72 with a model standard error of 0.69 and a *SD* of 2.04. Close to 68% of participants fell within a range of -2.76 and 1.32 logits making the maximum measure value (4.59) extreme. The minimum measure value of -4.23 is not extreme.

Although it will be confirmed later in the chapter, the results here suggest that the items on the MNT dimension functioned very well. The average item infit and outfit MNSQ statistics are identical at 0.98 (t = -0.10). The infit and outfit *SD* values are 0.13 and 0.16 indicating that there is very little variation and that most of the items in the MNT dimension fit the Rasch model. The minimum and maximum MNSQ statistics for item infit (min 0.81; max 1.16) as well as the minimum and maximum MNSQ statistics for outfit (min 0.78; max 1.20) in this dimension are adequate. The items function well together within this dimension and did not have extreme values.

Reliability information for both items and persons on the MNT dimension reveals a person separation index of 2.30 indicating that persons are sufficiently separated across the MNT dimension. The person reliability estimate for this dimension is .84 with a Cronbach Alpha value of .93. This shows that responses to items on the MNT dimension demonstrate good internal consistency and that participants are responding in a consistent fashion across the 5 items of this dimension. The items on the MNT dimension therefore adequately screens for mental confusion among the participants. The item separation for this dimension is 2.20. This indicates that the items on the MNT dimension are not very well dispersed on the scale.

8.3.6.5 Loss of Self (LOS) dimension.

The person and item information for the LOS dimension can be found in Table 27. Winsteps (Linacre, 2009) eliminated 47 respondents in this dimension with extreme scores and the data is presented for the remaining 140 participants with non-extreme scores. The average raw score of persons in this dimension is 7.90. Both the person infit and outfit mean-squares statistics are 0.98 (t = -0.10). These values are very close to the Rasch-modeled expectations of 1.00. Little variation is evident with participants responding as expected and indicates good fit to the Rasch model. The *SD* infit and outfit values for this dimension are fairly wide at 0.81 and 0.85 respectively.

The minimum and maximum MNSQ statistics for person infit (0.04 and 4.68) and outfit (0.04 and 4.76) are extreme. This indicates that there are some persons that had unexpected responses to items on the LOS dimension.

The extreme minimum and maximum values in logits (-4.49 and 4.42 respectively) for the items in this dimension indicate that one or more women in this sample did not have symptoms while others had significant symptoms of loss of self. The average logit for person ability is -0.89 with a model standard error of 0.74 and a *SD* of 2.28. Approximately 68% of participants fell within a range of -3.17 and 1.39 logits. The minimum measure value (-4.49) is therefore not extreme. The maximum measure value (4.42) is extreme.

Functioning of the items on the LOS dimension appears to be very good with an average infit and outfit value of 1.00 (t = 0.00) and 0.98 (t = -0.20) respectively. The infit and outfit *SD* values for items are 0.17 and 0.16 respectively indicating little variation in responses and that items in the LOS dimension fit the Rasch model. The minimum and maximum MNSQ statistics for infit are 0.78 and 1.28, while the minimum and maximum MNSQ statistics for outfit are 0.76 and 1.26. While the minimum values are adequate, the maximum values are slightly high in this dimension indicating that some items had extreme values.

On the LOS dimension, the person separation index is 2.41. This indicates that persons are sufficiently separated across this dimension. The person reliability estimate for the LOS dimension is good at .85. The Cronbach Alpha value of .95 also indicates good internal consistency of responses to items. This demonstrates consistent responding by participants across the 5 items of this dimension. The PDSS's LOS dimension therefore adequately screens for participants' feelings of loss of self. Items in this dimension are fairly well dispersed on the scale with an item separation of 4.48.

8.3.6.6 Guilt/Shame (GLT) dimension.

Table 27 also presents the person and item information for the GLT dimension. Winsteps (Linacre, 2009) eliminated 43 respondents in the GLT dimension with extreme scores and the data is presented for the remaining 144 participants with non-extreme scores. The average raw score of persons in this dimension is 8.40. The infit and outfit mean-squares statistics are close to the Rasch-modeled expectation of 1.00 with MNSQ statistics of 0.97 (t = -0.10) for infit and 0.95 for outfit (t = 0.00). Items in this dimension fit the Rasch model with little variation evident and participants responding as expected.

The *SD* infit and outfit values for persons in this dimension are wide at 0.88 and 0.92 respectively. Relative to the other dimensions, the GLT dimension (– along with the ELB dimension) exhibit the most extreme maximum mean-square statistic values of infit and outfit. The maximum MNSQ for person infit is 5.79 (min 0.05) while the maximum for outfit is 6.09 (min 0.04). This indicates the presence of unexpected responses to items on this dimension.

The minimum and maximum values in logits (-4.09 and 3.91 respectively) for items in this dimension is extreme. This indicates that one or more women in this sample did not have symptoms while others had significant symptoms of guilt or shame. The average logit for person ability is -0.55 with a model standard error of 0.71 and a *SD* of 2.11. Almost 68% of participants fell within a range of -2.66 and 1.56 logits. The minimum measure value (-4.09) is therefore not extreme, whereas the maximum measure value (3.91) is extreme.

The performance of items on the GLT dimension is good. Individual item functioning will, however, be examined in more detail later in the chapter. The average infit and outfit values are 1.02 (t = 0.00) and 0.95 (t = -0.60) respectively. The infit and outfit *SD* values are both 0.25 indicating that there is slight variation and that most of the items in this dimension fit the Rasch model. The minimum and maximum MNSQ infit values (0.63 and 1.29 respectively) are adequate. The maximum MNSQ outfit value (1.29) is adequate but the minimum MNSQ outfit value (0.54) is a bit extreme.

Reliability information for the GLT dimension demonstrates a person separation index of 2.38 indicating that persons are sufficiently separated across this dimension. The person reliability estimate for this dimension is .85 with a Cronbach Alpha value of .93. This shows that responses to items on the GLT dimension demonstrate good internal consistency and that participants are responding in a consistent fashion across the items from this dimension. The items on the GLT dimension therefore adequately screens for feelings of guilt or shame among the participants. The item separation for the GLT dimension is very good at 5.51. This indicates that the items on the GLT dimension are well dispersed on the scale.

8.3.6.7 Suicidal Thoughts (SUI) dimension.

The person and item information for the SUI dimension can be found in Table 27. Winsteps (Linacre, 2009) eliminated 112 respondents with extreme scores in this dimension and the data is presented for the remaining 75 participants with non-extreme scores. The average raw score of persons in this dimension is the lowest of the 7 dimensions at 6.50. The person infit mean-squares statistic is 0.99 (t = 0.00) and the outfit mean-square statistic is 0.93 (t = 0.00). These values are near to the Rasch-modeled expectations of 1.00. The *SD* infit and outfit values for persons in this dimension are the narrowest of all 7 dimensions at 0.68 and 0.63 respectively.

The minimum and maximum MNSQ statistics for person infit (0.04 and 3.14) and outfit (0.05 and 3.03) indicate that there are some persons that had unexpected responses to items on the SUI dimension. The maximum infit and outfit values are, however, the lowest of the 7 dimensions.

The minimum and maximum measure values in logits for items in this dimension are extreme at -3.21 (minimum) and 4.38 (maximum) indicating that one or some women in this sample did not have symptoms of suicidal thoughts and that one or some participants had significant symptoms of suicidal thoughts. The average logit for person ability (suicidal thoughts) is -1.15 with a model standard error of 0.67 and a *SD* of 1.75. Therefore, around 68% of participants fell within a range of -2.90 and 0.60 logits.

Item performance on the SUI dimension is good with an average infit and outfit value of 0.99 (t = -0.30) and 0.93 (t = -0.5) respectively. The infit and outfit *SD* values are 0.44 and 0.38 indicating that there is some variation in participant responses. The minimum and maximum MNSQ statistics for item infit are 0.66 and 1.85 respectively, while the minimum and maximum MNSQ statistics for outfit are 0.61 and 1.66. The minimum values are adequate but the maximum values are extreme indicating that some items had extreme values in the SUI dimension.

The person separation index on the SUI dimension is 2.06. Participants are therefore adequately separated across this dimension. The person reliability estimate for the SUI dimension is good at .81. The Cronbach Alpha value of .95 also indicates good internal consistency of responses to items. Participants therefore responded consistently across the 5 items of this dimension indicating that it adequately screens for symptoms of suicidal ideation. Items in this dimension are, however, not as well dispersed on the scale with an item separation of 3.34.

8.3.7 Item Fit Statistics for the PDSS Dimensions.

Item-fit indices (MNSQ) indicate the degree to which individual items define a unidimensional construct (Hong & Wong, 2005). Therefore, to examine the unidimensionality – or in other terms, the construct validity – of a scale, item fit statistics must be computed (Schumacker, 2004). The analysis of fit is an essential part of using latent trait models, like the Rasch model, if the interpretation of the calibration of results is to be meaningful. The parameters of a Rasch model, once estimated, are used to compute the expected, or predicted, response pattern for every item. The comparison of the expected patterns and the observed patterns yields the fit statistics for persons and items. In Rasch measurement, fit statistics are used to assist in identifying and controlling the quality of item and person response patterns that do not meet the requirements of the model and therefore do not contribute to useful measurement. If the data (i.e. items or persons) do fit the model requirements, the estimated ability is believed to correctly represent the respondent's ability, and hence the difficulty parameters are believed to correctly represent the item difficulty (Smith, R. M., 2000; Smith, E. V., 2004). Items or persons that do not fit the requirements of the model will be examined further to determine how they are interfering with the measurement process.

Unstandardised fit estimates (i.e. mean-squares, or MNSQ) are modelled by Rasch analysis to have a mean of 1. Ideally the actual unstandardised item fit statistic would be very close to the expected mean of 1 to indicate that there is little spread from the ideal and that there is a good fit between the item and the Rasch model (Bond & Fox, 2001). Reasonable MNSQ fit values for a self-report rating scale are recommended at 0.60 – 1.40 (Wright & Linacre, 1994).

The Rasch error estimate indicates how precisely the Rasch parameter was estimated. Large error estimates signify haphazard responses to an item.

The Pearson item-total correlation (r_{it}) is the correlation between the total item score and the item. “It is similar to the discrimination or item-total correlation in CTT, although it differs in that extreme values are omitted” (Maree, 2004, p. 7). A negative Pearson item-total correlation (r_{it}) indicates an inverse relationship between the dichotomous item responses and the total raw score, and may indicate the presence of a problem like reverse coding. A general rule is to drop any items with a zero or negative Pearson item-total correlations (r_{it}) correlation (Schumacker, 2004). An item with a low Pearson item-total correlation (r_{it}) value indicates that the item does not fit the construct well and that it may be tapping another dimension. A high positive value suggests good correlation and that the item belongs to a unidimensional construct (Maree, 2004).

Furthermore, when there is a great discrepancy between the observed Pearson item-total correlation (r_{it}) and the expected (EXP) value, it may indicate that the item does not show a good fit with the dimension being measured. When the observed value is much higher than the expected value it may indicate dependency in the data. When the observed value is much lower than expected value, unmodeled noise is possible (Linacre, 2008).

The tables referred to in this section (Table 28 to Table 34) compare the items of the PDSS dimension in terms of their measure order. The items are listed in sequence from most difficult to agree with to easiest to agree with.

8.3.7.1 Sleeping/Eating Disturbances (SLP) dimension.

Table 28 presents the item fit statistics for the items from the SLP dimension. In this dimension the most difficult item to agree with is Item 29 (I knew I should eat but I could not) whereas the easiest to agree with is Item 1 (I had trouble sleeping even when my baby was asleep). Mean-squares for both infit and outfit for items in this dimension are good and all fall within the acceptable range of 0.60 and 1.40. This indicates that little distortion is evident in the measurement system, that the items were well understood by most participants, and that the items appear to fit the definition of the construct well. The SLP items have better fit statistics within the SLP dimension than within the total PDSS, which provides support for the construct validity of this dimension. The Rasch error estimates on this dimension were small and ranged from 0.09 – 0.10, with a mean of 0.09.

The Pearson item-total correlation (r_{it}) values for the SLP dimension support construct validity with values that range from .69 to .78. This also suggests that there are no coding errors in this dimension. The Pearson item-total correlation (r_{it}) values and the expected values of all items in this dimension indicate very little discrepancy. All items in this dimension correlate well and tap into a unidimensional construct of disturbances in sleeping or eating.

Table 28 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Sleeping/Eating Disturbances (SLP) Dimension (n=187)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Sleeping/Eating Disturbances (SLP)						
1	I had trouble sleeping even when my baby was asleep.	-0.49	0.09	1.06	1.05	.78
8	I lost my appetite.	0.21	0.09	1.18	1.16	.71
15	I woke up on my own in the middle of the night and had trouble getting back to sleep.	-0.15	0.09	0.83	0.80	.78
22	I tossed and turned for a long time at night trying to fall asleep.	-0.02	0.09	0.80	0.77	.78
29	I knew I should eat but I could not.	0.45	0.10	1.14	1.09	.69
M		0.00	0.09	1.01	0.98	
SD		0.32	0.00	0.16	0.16	

MNSQ = mean-square

8.3.7.2 Anxiety/Insecurity (ANX) dimension.

The items from the ANX dimension are listed in Table 29 from most difficult to agree with (Item 16: I felt like I was jumping out of my skin) to easiest to agree with

(Item 9: I felt really overwhelmed). No items in this dimension were overfitted i.e. none for infit were smaller than 0.60. Item 2 (I got anxious over even the littlest things that concerned my baby), had an outfit MNSQ statistic that was borderline (1.40). Although infit MNSQ statistics are more likely to indicate problematic fit, this item was monitored for any further discrepancies. The Rasch error estimates on this dimension was small and ranged from 0.09 – 0.10, with a mean of 0.10.

The Pearson item-total correlation (r_{it}) values for the ANX dimension are high and indicate good construct validity and that there are no coding errors. There is not much discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of any items in this dimension. All items in this dimension correlate well and tap into a unidimensional construct of anxiety or insecurity.

Table 29 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Anxiety/Insecurity (ANX) Dimension (n=187)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Anxiety/Insecurity (ANX)						
2	I got anxious over even the littlest things that concerned my baby.	-0.60	0.10	1.20	**1.40**	.78
9	I felt really overwhelmed.	-1.13	0.10	0.94	0.91	.84
16	I felt like I was jumping out of my skin.	1.14	0.10	0.89	0.79	.78
23	I felt all alone.	-0.40	0.09	0.93	0.91	.81
30	I felt like I had to keep moving or pacing.	0.98	0.10	1.07	1.05	.76
M		0.00	0.10	1.01	1.01	
SD		0.90	0.01	0.11	0.21	

Note. Boldface value indicates a high MNSQ statistic that is borderline for problematic fit.
MNSQ = mean-square

8.3.7.3 Emotional Lability (ELB) dimension.

Items from the ELB dimension are listed in Table 30. The most difficult item to agree with is Item 10 (I was scared that I would never be happy again). The easiest item to agree with is Item 24 (I have been very irritable).

All mean-squares for infit and outfit in the ELB dimension are near 1.00 indicating little distortion of the measurement system. Items in this dimension appear to have been well understood by the English participants and the items seem to fit the definition of the construct well. The Rasch error estimates on this dimension was small and ranged from 0.10 – 0.11, with a mean of 0.10.

The Pearson item-total correlation (r_{it}) values for the ELB dimension are all positive high values between .81 and .84, indicating good construct validity. These high values also indicate that there are no coding errors. There is not much discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of any items in this dimension. All items in this dimension correlate well and tap into a unidimensional construct.

Table 30 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Emotional Lability (ELB) Dimension (n=187)

	Dimension / Item	Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Emotional Lability (ELB)						
3	I felt like my emotions were on a roller coaster.	-0.68	0.10	0.96	0.96	.84
10	I was scared that I would never be happy again.	0.88	0.10	0.89	0.73	.84
17	I cried a lot for no real reason.	0.29	0.10	1.10	1.06	.82
24	I have been very irritable.	-1.10	0.11	1.00	0.94	.84
31	I felt full of anger ready to explode.	0.62	0.10	1.15	1.16	.81
M		0.00	0.10	1.02	0.97	
SD		0.76	0.00	0.09	0.14	

MNSQ = mean-square

8.3.7.4 Mental Confusion (MNT) dimension.

Table 31 presents the item fit statistics for the items from the MNT dimension. In this dimension the most difficult item to agree with is Item 18 (I thought I was going crazy), and the easiest was Item 25 (I had a difficult time making even a simple decision). Infit and outfit mean-squares range between 0.78 and 1.20 – all within an acceptable range. Therefore little distortion is evident in the items of this dimension, they were well understood by most participants, and the items appear to fit the definition of the construct well. The Rasch error estimates on this dimension was small and ranged from 0.11 – 0.12, with a mean of 0.11.

In the MNT dimension the high positive Pearson item-total correlation (r_{it}) values indicate good construct validity and no coding errors with values that range from .84 to .87. There is very little discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of the items in this dimension. All items in the MNT dimension correlate very well and tap into a unidimensional construct.

Table 31 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Mental Confusion (MNT) Dimension (n=187)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Mental Confusion (MNT)						
4	I felt like I was losing my mind.	-0.07	0.11	1.09	1.11	.85
11	I could not concentrate on anything.	0.00	0.12	0.97	0.98	.87
18	I thought I was going crazy.	0.39	0.11	0.89	0.78	.84
25	I had a difficult time making even a simple decision.	-0.47	0.12	1.16	1.20	.85
32	I had difficulty focusing on a task.	0.15	0.12	0.81	0.81	.87
M		0.00	0.11	0.98	0.98	
SD		0.28	0.00	0.13	0.16	

MNSQ = mean-square

8.3.7.5 Loss of Self (LOS) dimension.

The items of the LOS dimension are listed in terms of their measure order in Table 32. The most difficult item to agree with is Item 33 (I did not feel real). The item that was the easiest to agree with was Item 5 (I was afraid that I would never be my normal self again).

All infit and outfit MNSQ statistics for items in the LOS dimension are within an acceptable range. The items appear to have been well understood by the English participants and seem to fit the definition of the construct well. The Rasch error estimates on this dimension was small and ranged from 0.12 – 0.13, with a mean of 0.13.

The Pearson item-total correlation (r_{it}) values for the LOS dimension are all positive high values between .88 and .90 indicating good construct validity and no coding errors. The Pearson item-total correlation (r_{it}) values and the expected values of all items in this dimension indicate very little discrepancy. All items in this dimension correlate well and tap into a unidimensional construct.

Table 32 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Loss of Self (LOS) Dimension (n=187)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Loss of Self (LOS)						
5	I was afraid that I would never be my normal self again.	-1.14	0.12	1.28	1.26	.88
12	I felt as though I had become a stranger to myself.	0.25	0.13	1.07	1.02	.89
19	I did not know who I was anymore.	0.39	0.12	0.78	0.93	.89
26	I felt like I was not normal.	-0.05	0.13	0.97	0.92	.89
33	I did not feel real.	0.54	0.13	0.89	0.76	.90
M		0.00	0.13	1.00	0.98	
SD		0.60	0.00	0.17	0.16	

MNSQ = mean-square

8.3.7.6 Guilt/Shame (GLT) dimension.

Item fit statistics from the GLT dimension are listed in Table 33. The most difficult item to agree with is Item 27 (I felt like I had to hide what I was thinking or feeling toward the baby). The easiest item to agree with was Item 6 (I felt like I was not the mother I wanted to be).

Item 34 overfit the model with an outfit MNSQ statistic of 0.54, which is below the acceptable range. The remaining items appear to have been well understood by the English participants and seem to fit the definition of the construct well. The Rasch error estimates on this dimension were small and ranged from 0.11– 0.12, with a mean of 0.11.

The Pearson item-total correlation (r_{it}) values for the GLT dimension indicate good construct validity with high positive values that range from .81 to .90. These high values also indicate that there are no coding errors. There is very little discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of the items in this dimension. All items in this dimension correlate well and tap into a unidimensional construct of feelings of guilt or shame.

Table 33 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Guilt/Shame (GLT) Dimension (n=187)

	Dimension / Item	Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Guilt/Shame (GLT)						
6	I felt like I was not the mother I wanted to be.	-0.84	0.12	0.86	0.82	.90
13	I felt like so many mothers were better than me.	-0.69	0.12	1.05	1.04	.88
20	I felt guilty because I could not feel as much love for my baby as I should.	0.67	0.11	1.29	1.06	.81
27	I felt like I had to hide what I was thinking or feeling towards the baby.	0.83	0.12	1.27	1.28	.81
34	I felt like a failure as a mother.	0.03	0.11	0.63	**0.54**	.88
M		0.00	0.11	1.02	0.95	
SD		0.68	0.00	0.25	0.25	

Note. Boldface values have infit and outfit MNSQ statistics less than 0.60 or greater than 1.40
MNSQ = mean-square

8.3.7.7 Suicidal Thoughts (SUI) dimension.

Table 34 presents item fit statistics for the SUI dimension. The most difficult item in the SUI dimension to agree with was Item 21 (I wanted to hurt myself), and the easiest was Item 28 (I felt that my baby would be better off without me). Item 28 does, however have a high infit mean-square value (1.85) which indicates that responses to this item were unpredictable, possibly due to unmodeled noise or that their data underfit the model. The remaining items from this dimension had infit and outfit mean-squares within an acceptable range that reflect little distortion these items, that they were well understood by most participants, and appear to fit the definition of the construct well. The Rasch error estimates on this dimension were relatively higher than on the previous dimensions, but were still small and ranged from 0.14 – 0.17, with a mean of 0.15.

In the SUI dimension the high positive Pearson item-total correlation (r_{it}) values indicate good construct validity and no coding errors with values that range from .85 to .91. Item 28 shows slight discrepancy between the Pearson item-total correlation (r_{it}) value (.85) and the expected value (.89) and with a slightly elevated infit MNSQ statistic mentioned earlier, also suggests that item 28 may not fit the SUI dimension as well as the other items do.. There is very little discrepancy between these values on the remaining items of this dimension which suggests that they correlate very well and tap into a unidimensional construct.

Table 34 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the English PDSS Suicidal Thoughts (SUI) Dimension (n=187)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Suicidal Thoughts (SUI)						
7	I have thought that death seemed like the only way out of this living nightmare.	0.45	0.16	0.69	0.71	.90
14	I started thinking that I would be better off dead.	0.08	0.15	0.66	0.61	.91
21	I wanted to hurt myself.	0.77	0.17	1.01	0.93	.90
28	I felt that my baby would be better off without me.	-0.76	0.14	**1.85**	**1.66**	.85
35	I just wanted to leave this world.	-0.54	0.15	0.75	0.75	.90
M		0.00	0.15	0.99	0.93	
SD		0.58	0.01	0.44	0.38	

Note. Boldface values have infit and outfit MNSQ statistics less than 0.60 or greater than 1.40
MNSQ = mean-square

8.3.8 Response category statistics: Item option and distractor frequencies for the PDSS dimensions.

The frequency of responses to the 5-point Likert rating scale categories are outlined in the Table 72 to Table 78 in Appendix F and are briefly discussed below. In the English PDSS, the SLP, LOS, GLT and SUI dimensions, category “0” was selected most often in all items. This is particularly evident in the SUI dimension with percentage data counts ranging from 68% (Item 28) to 75% (Item 7). In the ANX, ELB and MNT dimensions, category “0” was selected more often for the majority of items.

Five categories from 4 items of the SUI dimension had less than 10 observations. The remaining dimensions had category observations that ranged from 10 to 140. In general though, the PDSS categories were used fairly regularly across all items.

All items, apart from Item 29, in the PDSS dimensions have average measure values (in logits) which increase gradually with each higher response category. This supports the validity of the 5-point Likert scale for the PDSS with each higher response category corresponding to “more” of the variable being measured. There are, however, a number of categories across all the PDSS dimensions that have outfit MNSQ statistics greater than 1.60 or lower than 0.60. The convergent and discriminant validity of the item categories for the PDSS dimensions is supported by the Pearson item-total correlation (r_{it}) values. In only three items from the PDSS the Pearson item-total correlation (r_{it}) values do not advance steadily. These are items 21, 29, and 35.

8.4 Results of Rasch Analysis of the Afrikaans PDSS

8.4.1 Summary of Afrikaans Rasch analysis: persons and items.

The summary statistics of the non-extreme persons and items[1] for the Afrikaans PDSS are presented in Table 35a and Table 35b. Most persons responded according to expectation with the average person infit (1.07) and outfit (1.03) at almost 1. The *SD* infit and outfit values are 0.52 and 0.61 respectively. According to the ideal z- or normal distribution graph, one *SD* above and below the mean represents approximately 68% of the distribution of values. The minimum and maximum values for person infit (min.0.20; max 3.48) and outfit (min. 0.25; max 4.00) are extreme signifying that some persons had unexpected responses to some items on the screening scale.

The min of -4.63 logits for the measure is extremely low indicating that one or some women in this sample did not have symptoms of PPD. The maximum value of 2.60 logits for the measure indicates that some participants were very depressed – although the maximum is somewhat lower than the maximum for the Eng PDSS (4.32 logits). The average logit for person ability was -0.99 (0.20 lower than the Eng PDSS) with a *SD* of 1.42. This range, although wide, is not as wide as the range for the Eng PDSS. It indicates that approximately 68% of respondent scores fell within -2.41 and 0.43 logits. Therefore the minimum and maximum measure values are extreme.

---

[1] Summary statistics for extreme and non-extreme persons for the Afrikaans PDSS are presented in Table 71 in Appendix F.

Based on these results, the Afrikaans PDSS items functioned well with average infit and outfit values (1.05 and 1.03) close to 1 – only marginally better than the Eng PDSS items. The *SD*s were 0.32 and 0.53, indicating neither too much nor too little variation and that most of the items fit the Rasch model. The minimum and maximum values for item infit (min 0.62; max 2.07) and outfit (min 0.54; max 3.10) indicate the presence of some extreme values.

Reliability information for items and persons on the Afrikaans PDSS is presented in Table 35a and Table 35b. The person separation index is high at 4.28. The person reliability estimate is excellent at .95 with a Cronbach Alpha of .98. This provides evidence of excellent internal consistency of responses to items on the Afrikaans PDSS and indicates that items were able to sufficiently separate the participants along the continuum. The 35 items in the Afrikaans PDSS correlate well with each other and participants are responding in a consistent fashion. The Afrikaans PDSS therefore adequately screens for measured symptoms of PPD.

Reliability is further confirmed with an item separation index of 7.00. This indicates that the Afrikaans PDSS items are well dispersed on the scale and can distinguish between a number of levels of performance.

Table 35a Summary Statistics of 170 Non-Extreme Persons and Items for the Afrikaans PDSS.

	Raw Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	47.80	35.00	-0.99	0.25	1.07	0.00	1.03	0.00
S.D.	34.00	0.10	1.42	0.16	0.52	1.80	0.61	1.80
Max	129.00	35.00	2.60	1.01	3.48	5.20	4.00	5.90
Min	1.00	34.00	-4.63	0.16	0.20	-5.80	0.25	-4.40

Real RMSE	0.32	True S.D.	1.39	Separation	4.28	Particip Reliability	.95
Model RMSE	0.30	True S.D.	1.39	Separation	4.66	Particip Reliability	.96
S.E. of participant mean = 0.11							
Minimum Extreme Score:		8 Participants					

Table 35b Afrikaans PDSS: Summary of 35 Measured (Non-Extreme) PDSS

	Raw Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	232.30	178.00	0.00	0.09	1.05	0.10	1.03	0.00
S.D.	85.70	0.20	0.69	0.01	0.32	2.50	0.53	2.40
Max	417.00	178.00	1.39	0.12	2.07	7.40	3.10	8.50
Min	80.00	177.00	-1.45	0.09	0.62	-3.90	0.54	-3.20

Real RMSE	0.10	True S.D.	0.68	Separation	7.00	Particip Reliability	.98
Model RMSE	0.09	True S.D.	0.68	Separation	7.48	Particip Reliability	.98
S.E. of PDSS mean = 0.12							
UMEAN = 0.0000	USCALE = 1.0000						
PDSS items raw score-to-measure correlation = -1.00[a]							
Data points: 5949		Log-likelihood Chi-Square: 11795.70 with 5743 *d.f.*				*p*= 0.0000	

[a] Approximate due to missing data

8.4.2 Rating scale requirements: Afrikaans PDSS

This section examines the quantitative functioning of the Afrikaans PDSS rating scale. Table 36 contains summary statistics for the 5-point Likert response categories used for the Afrikaans PDSS.

a) Category observations

All the responses for all the Afrikaans PDSS items are collated in Table 36. For response category 0 there were 2453 responses (41% of the total responses). Category 1 had 1174 responses (20% of the total responses). The same percentages were observed for responses to category 0 and category 1 on the English PDSS. The categories that had the least responses were categories 2 (neither agree nor disagree) and 4 (strongly agree) with 9% of the total responses each (observed count of 531 and 520 respectively). All category frequency counts are sufficiently large indicating that locally stable estimates of the rating scale structure may be produced (Linacre, 2004). The response pattern to individual items from the Afrikaans PDSS will be examined in more detail later to determine if there are items with category frequency counts less than 10.

Table 36 Summary Statistics for the 5-Point Likert Response Categories Used for the Afrikaans PDSS

Summary of Category Structure (N = 178)

Category	Label	Score	Observed Count	%	Observed Average[a]	Sample Expect.	MNSQ Infit	MNSQ Outfit	Structure Calibration	Category Measure
0	Strongly Disagree	0	2453	41	-2.29	-2.24	0.96	1.10	NONE	(-2.21)
1	Disagree	1	1174	20	-0.89	-1.03	0.99	0.74	-0.84	-0.92
2	Neither Disagree nor Agree	2	531	9	-0.20	-0.31	0.92	0.86	0.14	-0.17
3	Agree	3	1271	21	0.24	0.31	1.14	1.34	-0.87	0.78
4	Strongly Agree	4	520	9	1.05	1.07	1.13	1.17	1.57	(2.74)
Missing			1	0	-0.87					

[a] Observed Average is mean of measures in category, not a parameter estimate.

b) Regular observation distribution

Category 0 (strongly disagree) was used most frequently (41%), followed by category 3 (agree; 21%) and category 1 (disagree; 20%). Categories 2 (neither disagree nor agree) and 4 (strongly agree) have 50 % less observations indicating that respondents did not endorse the middle category and the most extreme category as expected. These two categories were also used less frequently in the English PDSS which reflects that they may be redundant. Participants may also be more inclined to choose category 3 (agree) than category 4 (strongly agree).

c) Average measures advance monotonically with category

Average measures (expressed as logits) steadily increase from small to large with each category, i.e. -2.29, -0.89, -0.20, 0.24 and 1.05. The observed average measures demonstrate values that are close to their expected values.

d) OUTFIT mean-squares less than 2

Outfit mean-squares indicate random noise with values large than 1.4 indicating unexpected observations in that category (Smith, Wakely, De Kruif, & Swartz, 2003). Most categories demonstrate values close to the expected 1.0. No categories had values over 2. Similar to the English PDSS, category 3 (agree) also had the largest value (1.34) but his value is still acceptable for this sample.

e) Step calibrations advance orderly

Ideally step calibrations should increase uniformly from easy to hard. For the Afrikaans PDSS the step calibration values are: -.84, 0.14, -.87 and 1.57. The transitions between categories 1 and 2, and categories 2 and 3 are problematic. Linacre (2004) suggests that ideally curves should form a series of prominent hills. Figure 5 indicates, however, that the only prominent hills are for categories 0, 3 and 4. The negative value in the table for category 3 (-.87) may be due to the narrowness of categories 2 and 3.

Category 2 does not form a prominent hill meaning that this category is relatively rarely observed. A similar pattern is observed in the categories of the English PDSS.

f) Step difficulties advance by at least 1 logit

The Afrikaans PDSS categories have step difficulties which advance as follows:

Categories 1-2: 0.14 – (-.84) = 0.98

Categories 2-3: -0.87 – 0.14 = -1.01

Categories 3-4: 1.57 – (-0.87) = 2.44

Steps should ideally advance by at least 1 logit when five categories are employed (Linacre, 2004, p.274). The advance from categories 1 to 2 is the lowest (0.98), but is very near to the acceptable value of 1 logit. The advance from categories 2 to 3 is acceptable at 1.01 logits while the advance from category 3 to 4 is adequate at 2.44. The rating scale category definitions appear to function better with the Afrikaans sample than with the English sample.

```
    CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+-
R  1.0 +                                                           +
O      |                                                           |
B      |00                                                         |
A      |  0000                                                     |
B   .8 +      000                                                44+
I      |        00                                             444  |
L      |          00                                         44     |
I      |            00                                     44       |
T   .6 +              0                                  44         +
Y      |               00               333333333    44           |
    .5 +                 0             33        33344             +
O      |                  0          33           44333            |
F   .4 +                   00       3            4     33          +
       |                1111111*11   3            44        33        |
R      |             1111      0 1**           44           333     |
E      |          111          0*  11       44                33   |
S   .2 +       111            33 *222**   44                    333+
P      |   1111             2*222 00   ***22                       |
O      |111             22**3       0*44 11122222                  |
N      |         22222**33      44444 00000 1111122222222          |
S   .0 +***************4444444444          000000******************+
E      -+---------+---------+---------+---------+---------+---------+-
      -3        -2        -1         0         1         2         3
     PARTICIPANT [MINUS] PDSS MEASURE
```

Figure 5 Probability curves of observations in each category.

8.4.3 Item person construct map: Afrikaans PDSS.

Table 37a represents a geographical description of the two facets – participants and PDSS items. More than two items are frequently positioned on the same logit measure. For good variable definition, and construct validity, items should be spread apart – the further, the better. As with the English PDSS, it seems as if there are insufficient items from the Afrikaans PDSS present at either end of the difficulty level. However, Table 37b shows that the rating scale categories cover the spread of person abilities well. Few Afrikaans respondents scored higher than the items were able to measure. The person and item distribution is indicative of some measurement precision lost at the most difficult level.

The distribution of participants indicates that significant proportions of the Afrikaans sample screened either negatively or positively for PPD. As with the English sample, items from the dimension that measures contemplating harming oneself were the items that were most difficult to agree strongly with (Item 7 and Item 21), and are also closely followed by the remaining items from this dimension (items 14, 35, and 28).

Table 37a Item Distribution Map for the Afrikaans PDSS (N=178)

```
PARTICIPANTS - MAP - PDSS
          <more>|<rare>
 3             +
               |
            X  |
               |
               |
            X  |
 2             +
              T|
          XXX  |
            X  |
           XX  |T PDSS_21  PDSS_7
               |  PDSS_14
 1         XX  +  PDSS_35
           XX  |
      XXXXXXX  |S PDSS_15  PDSS_28  PDSS_29
       XXXXXX S|  PDSS_22  PDSS_8
        XXXXX  |  PDSS_18  PDSS_27  PDSS_33
     XXXXXXXX  |  PDSS_1   PDSS_16  PDSS_20
 0  XXXXXXXXXX +M PDSS_11  PDSS_12  PDSS_19  PDSS_25  PDSS_26  PDSS_4
    XXXXXXXXX  |  PDSS_10  PDSS_30  PDSS_32
     XXXXXXXX  |
     XXXXXXXX  |  PDSS_17  PDSS_31  PDSS_34  PDSS_5   PDSS_6
   XXXXXXXXXX  |S PDSS_13  PDSS_2
  XXXXXXXXXXX  |
-1     XXXXXX M+  PDSS_23
         XXXX  |  PDSS_24
    XXXXXXXXX  |T PDSS_3
      XXXXXXX  |  PDSS_9
       XXXXXX  |
           XX  |
-2        XXX  +
          XXX  |
        XXXXX  |
       XXXXXX S|
           XX  |
           XX  |
-3     XXXXXX  +
           XX  |
               |
         XXXX  |
               |
              T|
-4    XXXXXXX  +
               |
               |
               |
           XX  |
               |
-5   XXXXXXXX  +
          <less>|<frequ>
```

Table 37b Item Category-Person Distribution Map for Afrikaans PDSS (N = 178)

```
PARTICIPANT - MAP - PDDS - Expected score zones (Rasch-half-point thresholds)
             <more>|  Disagree   Neither D  Agree      Strongly Agree
    4              +
                   |
                   |
                   |
                   |
                   |                                   PDSS_2.35
                                                       PDSS_7.35
    3              +                                   PDSS_1.35
                   |                                   PDSS_3.35
                X  |
                   |                                   PDSS_1.35
                                                       PDSS_2.35
                                                       PDSS_2.35
                   |                                   PDSS_2.35
                                                       PDSS_8.35
                X  |                                   PDSS_1.35
                                                       PDSS_2.35
    2              +                                   PDSS_1.35
                                                       PDSS_1.35
                                                       PDSS_2.35
                                                       PDSS_3.35
                  T|                                   PDSS_1.35
                                                       PDSS_1.35
                                                       PDSS_2.35
              XXX  |                        PDSS_7.25  PDSS_1.35
                                                       PDSS_1.35
                                                       PDSS_2.35
                                                       PDSS_3.35
                                                       PDSS_4.35
                X  |                        PDSS_1.25  PDSS_3.35
                                            PDSS_2.25
               XX  |T                                  PDSS_1.35
                                                       PDSS_3.35
                                                       PDSS_3.35
                                                       PDSS_5.35
                                                       PDSS_6.35
                   |                        PDSS_3.25  PDSS_1.35
                                                       PDSS_2.35
    1          XX  +                        PDSS_2.25
               XX  |             PDSS_2.15  PDSS_1.25  PDSS_2.35
                                 PDSS_7.15  PDSS_2.25
                                            PDSS_2.25
          XXXXXXX  |S            PDSS_1.15  PDSS_1.25  PDSS_2.35
                                            PDSS_8.25
           XXXXXX S|             PDSS_3.15  PDSS_2.25  PDSS_3.35
                                            PDSS_3.25
            XXXXX  |                        PDSS_1.25  PDSS_9.35
                                            PDSS_1.25
                                            PDSS_1.25
                                            PDSS_2.25
                                            PDSS_2.25
         XXXXXXXX  |             PDSS_1.15  PDSS_1.25
                                 PDSS_2.15  PDSS_1.25
                                 PDSS_2.15  PDSS_2.25
                                            PDSS_3.25
                                            PDSS_4.25
    0  XXXXXXXXXX  +M            PDSS_2.15  PDSS_1.25
```

Table 37b (continued) Item Category-Person Distribution Map for the Afrikaans PDSS (N = 178)

```
PARTICIPANT - MAP - PDDS - Expected score zones (Rasch-half-point thresholds)
             <more>|  Disagree   Neither D  Agree      Strongly Agree
    0  XXXXXXXXXX  +M            PDSS_2.15  PDSS_1.25
                                 PDSS_8.15  PDSS_3.25
        XXXXXXXXX  |  PDSS_2.05  PDSS_1.15  PDSS_3.25
                      PDSS_7.05  PDSS_2.15  PDSS_3.25
                                            PDSS_5.25
         XXXXXXXX  |  PDSS_1.05  PDSS_1.15  PDSS_1.25
                                 PDSS_1.15  PDSS_2.25
                                 PDSS_2.15  PDSS_6.25
                                 PDSS_3.15
         XXXXXXXX  |             PDSS_1.15  PDSS_1.25
                                 PDSS_1.15
                                 PDSS_2.15
       XXXXXXXXXX  |S PDSS_3.05  PDSS_1.15
                                 PDSS_1.15
                                 PDSS_2.15
                                 PDSS_3.15
                                 PDSS_4.15
      XXXXXXXXXXX  |  PDSS_2.05  PDSS_3.15  PDSS_2.25
                      PDSS_2.05
   -1      XXXXXX M+  PDSS_1.05  PDSS_1.15  PDSS_2.25
                      PDSS_2.05  PDSS_3.15  PDSS_3.25
                                 PDSS_3.15
                                 PDSS_5.15
                                 PDSS_6.15
             XXXX  |  PDSS_1.05  PDSS_1.15  PDSS_9.25
                      PDSS_8.05  PDSS_2.15
        XXXXXXXXX  |T PDSS_1.05
                      PDSS_2.05
                      PDSS_2.05
                      PDSS_3.05
          XXXXXXX  |  PDSS_1.05  PDSS_2.15
                      PDSS_1.05
                      PDSS_2.05
           XXXXXX  |  PDSS_1.05  PDSS_2.15
                      PDSS_1.05
                      PDSS_1.05
                      PDSS_2.05
                      PDSS_3.05
                      PDSS_4.05
               XX  |  PDSS_3.05  PDSS_3.15
   -2         XXX  +  PDSS_1.05  PDSS_9.15
                      PDSS_3.05
                      PDSS_3.05
                      PDSS_5.05
              XXX  |  PDSS_1.05
                      PDSS_2.05
                      PDSS_6.05
            XXXXX  |
           XXXXXX S|  PDSS_2.05
               XX  |
               XX  |  PDSS_2.05
                      PDSS_3.05
   -3      XXXXXX  +  PDSS_9.05
               XX  |
                   |
             XXXX  |
                   |
                  T|
   -4     XXXXXXX  +
                   |
                   |
               XX  |
                   |
   -5    XXXXXXXX  +
```

8.4.4 Item fit: Afrikaans PDSS.

Table 38 contains the item fit statistics for the Afrikaans PDSS. A range of 0.60 to 1.40 for infit and outfit MNSQ are acceptable limits. No items had an infit MNSQ less than 0.6. Infit MNSQ statistics were high for items 30, 1, 15, 2, 29, and 8. This means they do not fit the definition of the construct by either forming a secondary construct or dimension. Items 1, 8, 15 and 29 are, in fact, from a separate dimension – the SLP content scale. Misfit in the total Afrikaans PDSS for these items may therefore merely be a reflection that they form a clear construct on their own. A similar trend was seen with items from the SLP content scale in the English PDSS. It is therefore important to place more emphasis on the construct validity of the items within their content scales as opposed to within the total screening scale.

The measure statistic (difficulty level in logits), and Model SE (measurement error) for each item are also presented in Table 38. All measurement error values for the Afrikaans PDSS are small with values less than 0.12 and a mean of 0.9.

Pearson item-total correlation (r_{it}) represents item-total correlation which provides an indication of construct validity and the presence of coding problems. Table 38 shows that there are no zero or negative correlations suggesting that there are no reverse coding problems nor respondents or items with response strings that contradict the variable. All the Pearson item-total correlation (r_{it}) values range are quite high despite some fit problems, and range from .51 to .80.

Table 38 Item Statistics for the Afrikaans PDSS Total: Misfit Order (N = 178)

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD		r_{it}	EXACT MATCH OBS%	EXACT MATCH EXP%	PDSS
30	260	170	-0.23	.09	2.07	7.4	3.10	8.5	A	.52	38.8	47.7	PDSS_30
1	212	170	0.12	.09	1.65	4.9	2.69	6.5	B	.54	39.4	49.6	PDSS_1
15	152	170	0.60	.09	1.82	5.4	1.57	2.3	C	.51	49.4	54.8	PDSS_15
2	313	170	-0.62	.09	1.24	2.0	1.78	4.3	D	.68	35.3	48.8	PDSS_2
29	142	170	0.69	.10	1.50	3.5	1.12	0.6	E	.55	54.1	56.5	PDSS_29
8	170	170	0.45	.09	1.31	2.4	1.02	0.2	F	.60	55.3	52.7	PDSS_8
9	417	170	-1.45	.09	0.98	-.1	1.27	1.8	G	.79	47.1	50.4	PDSS_9
20	203	170	0.19	.09	1.21	1.8	0.86	-0.7	H	.66	50.6	51.1	PDSS_20
22	156	170	0.57	.09	1.19	1.5	0.99	0.0	I	.60	54.1	54.1	PDSS_22
21	85	170	1.32	.12	1.17	1.0	0.72	-1.0	J	.52	71.2	67.6	PDSS_21
7	80	170	1.39	.12	1.14	0.9	0.74	-0.8	K	.51	76.5	69.1	PDSS_7
32	240	170	-0.09	.09	0.96	-0.3	1.14	0.9	L	.69	51.8	48.2	PDSS_32
13	322	170	-0.69	.09	0.98	-0.1	1.13	0.9	M	.76	52.9	48.7	PDSS_13
16	206	170	0.17	.09	1.02	0.2	1.13	0.7	N	.66	53.5	50.6	PDSS_16
24	391	170	-1.23	.09	0.93	-0.5	1.12	0.8	O	.80	55.9	50.1	PDSS_24
35	114	169	0.96	.10	1.11	0.8	0.76	-1.0	P	.57	63.9	61.7	PDSS_35
23	366	170	-1.03	.09	1.10	0.9	1.03	0.2	Q	.77	47.1	48.8	PDSS_23
14	93	170	1.22	.11	1.10	0.7	0.70	-1.1	R	.54	69.4	66.1	PDSS_14
27	189	170	0.30	.09	1.07	0.6	0.80	-1.0	q	.66	51.8	52.0	PDSS_27
17	296	170	-0.50	.09	1.03	0.3	1.06	0.4	p	.73	48.2	48.4	PDSS_17
31	293	170	-0.47	.09	1.04	0.4	0.99	0.0	o	.74	45.9	48.2	PDSS_31
3	394	170	-1.25	.09	0.93	-0.6	0.94	-0.3	n	.80	52.9	50.3	PDSS_3
4	237	170	-0.07	.09	0.94	-0.5	0.87	-0.8	m	.71	49.4	47.8	PDSS_4
5	291	170	-0.46	.09	0.93	-0.6	0.87	-0.8	l	.75	51.2	48.4	PDSS_5
28	146	170	0.66	.10	0.91	-0.7	0.64	-1.7	k	.64	58.2	55.8	PDSS_28
11	237	170	-0.07	.09	0.62	-3.9	0.88	-0.7	j	.75	56.5	47.8	PDSS_11
33	194	170	0.26	.09	0.85	-1.3	0.84	-0.8	i	.68	57.6	51.3	PDSS_33
19	222	170	0.05	.09	0.81	-1.8	0.65	-2.2	h	.72	53.5	48.9	PDSS_19
10	254	170	-0.19	.09	0.80	-1.9	0.77	-1.5	g	.74	53.5	47.6	PDSS_10
6	303	170	-0.55	.09	0.80	-1.8	0.73	-1.9	f	.78	55.3	48.3	PDSS_6
25	238	170	-0.07	.09	0.75	-2.4	0.72	-1.8	e	.74	57.6	47.8	PDSS_25
26	217	170	0.08	.09	0.70	-3.0	0.54	-3.1	d	.73	57.6	49.6	PDSS_26
18	175	170	0.41	.09	0.68	-3.0	0.68	-1.7	c	.69	65.9	52.2	PDSS_18
34	286	170	-0.42	.09	0.67	-3.3	0.59	-3.0	b	.79	56.5	48.4	PDSS_34
12	235	170	-0.05	.09	0.65	-3.5	0.54	-3.2	a	.75	60.6	47.9	PDSS_12
MEAN	232.3	170.0	.00	.09	1.05	.1	1.03	.0			54.2	51.9	
S.D.	85.7	.2	.69	.01	.32	2.5	.53	2.4			8.4	5.7	

PARTICIPANT: REAL SEP.: 4.28 REL.: .95
PDSS: REAL SEP.: 7.00 REL.: .98

8.4.5 Dimensionality: Afrikaans PDSS.

A Rasch principle component analysis (PCA) of residuals (the difference between observed and predicted scores) was performed. The PCA is indicative about the presence of secondary dimensions (Linacre, 2009) and was performed using calibrated data (logits) as opposed to raw data to avoid non-linearity in data accumulating in the PCA. Table 39 indicates the variance explained by the measures and raw unexplained variance. The empirical values match the modelled values reasonably well indicating that the measures explain the expected amount of variance in the data.

The variance explained by the measures is 58.60 eigenvalues or 62.6% which means that the measures explains most of the variance and that the Afrikaans PDSS has a wide spread of items and persons with different abilities, i.e. different degrees of PPD. Raw unexplained variance is 37.4%. Eigenvalues greater than 1.40 are indicative of possible secondary dimensions. The unexplained variance in the first contrast is 4.70 eigenvalues (5%), in the second contrast, 3.00 eigenvalues (3.2%), in the third contrast, 2.50 eigenvalues (2.7%), in the fourth contrast 2.30 eigenvalues (2.5%), and in the fifth contrast 1.80 eigenvalues (1.9%). These values indicate the presence of five additional dimensions, and that the Afrikaans PDSS is a multidimensional screening scale.

The items loading in Table 40 and the plot in Figure 6 below suggests that dimensionality in the Afrikaans PDSS exists.

Table 39 Variance Decomposition of the Observations for the Afrikaans PDSS Items (n = 178)

	Empirical			Modeled
	Eigenvalue units	%	%	%
Total raw variance in observations	93.60	100.00		100.00
Raw variance explained by measures	58.60	62.60		63.90
Raw variance explained by persons	39.40	42.10		43.00
Raw variance explained by items	19.20	20.50		20.90
Raw unexplained variance (total)	35.00	37.40	100.00	36.10
Unexplained variance in 1st contrast	4.70	5.00	13.30	
Unexplained variance in 2nd contrast	3.00	3.20	8.60	
Unexplained variance in 3rd contrast	2.50	2.70	7.20	
Unexplained variance in 4th contrast	2.30	2.50	6.60	
Unexplained variance in 5th contrast	1.80	1.90	5.10	

```
      -2              -1               0               1               2
      -+---------------+---------------+---------------+---------------+- COUNT
   .7 +                                |          C   A   B D           + 4
      |                                |                                |
   .6 +                                |                     E          + 1
C     |                                |  F                             | 1
O  .5 +                         G      |                                + 1
N     |                                |    H                           | 1
T  .4 +                                |                                +
R     |                       I        |                                | 1
A  .3 +                                |                                +
S     |                                |                                |
T  .2 +                             J  |K                               + 2
      |            M                   |   L                            | 2
1  .1 +                O        N      |                                + 2
      |                     R  Q      Pq                                | 4
L  .0 +--------------------------------|p-------------------------------+ 1
O     |                                |                                |
A -.1 +                                |      o                         + 1
D     |                                |                                |
I -.2 +         m                      |  n     l                       + 3
N     |                                | k                              | 1
G -.3 +            j           i      gh                                + 4
      |                      f        c|      d  e                      | 4
  -.4 +                                |          b                     + 1
      |                            a   |                                | 1
      -+---------------+---------------+---------------+---------------+-
      -2              -1               0               1               2
                                 PDSS MEASURE

 COUNT:         1  2   1    11122  11 5 21211 2 112   1   1 11
```

Figure 6 Standardized residual contrast of Afrikaans PDSS items.

Table 40 Standardized Residual Loading for the Afrikaans PDSS (Sorted by Loading)

PDSS Dimension	PDSS Item	Item Content	Loading	Measure	MNSQ Infit	MNSQ Outfit	Entry Number
SUI	35	Ek wou eenvoudig hierdie wêreld agterlaat.	.70	0.96	1.11	0.76	A
SUI	14	Ek het begin dink dat dit beter sou wees as ek dood was.	.69	1.22	1.10	0.70	B
SUI	28	Ek het gevoel dat dit vir my baba beter sou wees sonder my.	.69	0.66	0.91	0.64	C
SUI	21	Ek wou myself seermaak.	.68	1.32	1.17	0.72	D
SUI	7	Ek het gedink die dood sou die enigste uitweg uit hierdie nagmerrie wees.	.60	1.39	1.14	0.74	E
GLT	20	Ek het skuldig gevoel omdat dit vir my gevoel het asof ek nie my baba lief genoeg het nie.	.54	0.19	1.21	0.86	F
GLT	34	Ek het gevoel asof ek as ma misluk.	.48	-0.42	0.67	0.59	G
GLT	27	Dit het gevoel asof ek my ware gevoelens en gedagtes oor my baba moes wegsteek.	.44	0.30	1.07	0.80	H
GLT	6	Ek het gevoel asof ek nie die ma is wat ek wou wees nie.	.35	-0.55	0.80	0.73	I
ELB	10	Ek was bang dat ek nooit weer gelukkig sou wees nie.	.22	-0.19	0.80	0.77	J
LOS	26	Ek het gevoel asof ek nie normaal was nie.	.19	0.08	0.70	0.54	K
LOS	33	Ek het nie eg gevoel nie.	.14	0.26	0.85	0.84	L
ELB	24	Ek was baie geïrriteerd.	.14	-1.23	0.93	1.12	M
LOS	5	Ek was bang dat ek nooit weer my normale self sou wees nie.	.10	-0.46	0.93	0.87	N
ANX	23	Ek het alleen gevoel.	.08	-1.03	1.10	1.03	O
MNT	4	Ek het gevoel of ek van my verstand af raak.	.07	-0.07	0.94	0.87	P
ELB	31	Ek het baie kwaad gevoel en was gereed om te ontplof.	.06	-0.47	1.04	0.99	Q
GLT	13	Ek het gevoel asof baie ander ma's beter as ek was.	.06	-0.69	0.98	1.13	R
LOS	12	Ek het soos 'n vreemde vir myself gevoel.	.03	-0.05	0.65	0.54	q
ANX	30	Ek het gevoel asof ek heeltyd aan die gang moes bly.	-.43	-0.23	2.07	3.10	a
SLP	29	Ek het geweet ek moes eet, maar kon nie.	-.39	0.69	1.50	1.12	b
SLP	8	Ek het my eetlus verloor.	-.37	0.45	1.31	1.02	c
MNT	32	Ek het gesukkel om op 'n taak te konsentreer.	-.37	-0.09	0.96	1.14	d
SLP	15	Ek het in die middel van die nag vanself wakker geskrik en gesukkel om weer aan die slaap te raak.	-.34	0.60	1.82	1.57	e
ANX	2	Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak.	-.34	-0.62	1.24	1.78	f
MNT	11	Ek kon op niks konsentreer nie.	-.32	-0.07	0.62	0.88	g
MNT	25	Ek het dit moeilik gevind om die eenvoudigste besluite te neem.	-.32	-0.07	0.75	0.72	h
ELB	17	Ek het sonder enige rede baie gehuil.	-.30	-0.50	1.03	1.06	i
ELB	3	Ek het gevoel asof my emosies wipplank ry.	-.28	-1.25	0.93	0.94	j
SLP	1	Al het my baba geslaap, het ek gesukkel om te slaap.	-.23	0.12	1.65	2.69	k
SLP	22	Ek het snags lank rondgerol en gesukkel om aan die slaap te raak.	-.22	0.57	1.19	0.99	l
ANX	9	Ek het heeltemal oorweldig gevoel.	-.21	-1.45	0.98	1.27	m
ANX	16	Ek was so angstig ek het gevoel asof ek uit my vel wou spring.	-.21	0.17	1.02	1.13	n
MNT	18	Ek het gedink ek raak gek.	-.09	0.41	0.68	0.68	o
LOS	19	Ek het myself nie meer geken nie.	-.02	0.05	0.81	0.65	p

8.4.6 Performance of Afrikaans PDSS dimensions: Rasch analysis of persons and items.

The results of the Rasch analysis of persons and items for the seven dimensions of the Afrikaans PDSS are presented in this section. The summary statistics for each Afrikaans PDSS dimension as a whole is presented in Table 41 and is discussed below. A discussion of the dimensions' individual item fit statistics will be presented in the section that follows.

8.4.6.1 Afrikaans Sleeping/Eating Disturbances (SLP) dimension.

Table 41 summarizes the person and item information for the Afrikaans SLP dimension. Data for 58 participants with extreme minimum scores were excluded. Data for the remaining 120 participants demonstrate an average raw score of 6.90. The person mean-squares statistics are near to the Rasch-modeled expectations of 1.00. The infit MNSQ is 0.96 (t = 0.00) and the outfit MNSQ is 0.93 (t = 0.00). The *SD* infit and outfit values for this dimension are both 0.65. The minimum and maximum MNSQ statistics for infit (0.02 and 3.48) and outfit (0.02 and 3.32) indicate that there are some persons that had unexpected responses to items on this dimension.

Table 41 Summary Statistics for the Afrikaans PDSS Dimensions

Statistic		Sleeping / eating disturbances	Anxiety / insecurity	Emotional lability	Cognitive impairment	Loss of self	Guilt / shame	Contemplating harming oneself
Mean raw score	Items	166.40	312.40	325.60	225.40	231.80	260.60	103.60
	Persons	6.90	9.40	10.10	7.40	8.30	9.30	6.50
Measure (logits)	Items	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Persons	-0.81	-0.27	-0.06	-1.18	-0.66	-0.28	-1.33
Model error	Items	0.10	0.09	0.10	0.12	0.12	0.12	0.17
	Persons	0.58	0.58	0.63	0.70	0.69	0.65	0.73
SD (logits)	Items	0.25	0.67	0.62	0.29	0.49	0.72	0.67
	Persons	1.14	1.36	1.67	1.87	2.00	1.77	1.99
M Infit MNSQ	Items	1.01	1.02	1.00	0.99	0.98	0.98	0.98
	Persons	0.96	1.00	0.99	0.99	0.94	0.96	0.95
M Outfit MNSQ	Items	0.93	0.98	0.99	0.97	0.96	1.00	0.99
	Persons	0.93	0.98	0.99	0.97	0.96	1.00	0.99
Mean Infit (t)	Items	0.10	0.00	0.00	-0.10	-0.20	-0.20	-0.10
	Persons	0.00	0.00	-0.10	-0.10	-0.10	-0.20	-0.10
Mean Outfit (t)	Items	-0.30	-0.30	-0.10	-0.20	-0.30	0.00	0.00
	Persons	0.00	0.00	-0.10	-0.10	-0.10	-0.10	0.00
Separation	Items	2.16	6.98	5.78	2.11	3.90	5.85	3.52
	Persons	1.34	1.69	2.04	2.04	2.31	2.11	2.12
Cronbach alpha		.87	.84	.90	.91	.93	.93	.94
Rasch reliability	Persons	.64	.74	.81	.81	.84	.82	.82

MNSQ = mean-square

As with the English SLP dimension, the Afrikaans SLP dimension demonstrates minimum and maximum values in logits (-2.69 and 2.23 respectively) that are least extreme of the seven dimensions. Afrikaans participants were therefore also more likely to report the presence of some slight or moderate disturbance in sleeping or eating. The average logit for person ability is -0.81 with a model standard error of 0.58 and a *SD* of 1.14. Approximately 68% of participants therefore fell within a range of -1.95 and 0.33 logits.

On average, the items on the SLP dimension functioned very well. Individual item performance will, however, be discussed later in the chapter. The average item infit and outfit values are 1.01 (t = 0.10) and 0.93 (t = -0.3) respectively. The infit and outfit *SD* values are 0.22 and 0.23 respectively, indicating that there is little variation and that most of the items in the SLP dimension fit the Rasch model. The minimum and maximum MNSQ statistics for item infit (min 0.64; max 1.27) as well as the minimum and maximum MNSQ statistics for item outfit (min 0.65; max 1.25) in this dimension are within an acceptable range. The items in the Afrikaans SLP dimension do not have extreme values and function well together in this dimension.

On the SLP dimension, the person separation index is 1.34. As with the English SLP dimension, the person separation index for the Afrikaans SLP dimension is the lowest of the 7 dimensions and indicates that persons are not as well separated across this dimension as they are on the other dimensions. The person reliability estimate is also lower than other dimensions at .64. The Cronbach Alpha is higher at .87. Internal consistency for the SLP dimension is adequate, although it is lower than that of the other PDSS dimensions. As with the English sample, the participants in the Afrikaans sample are not responding as consistently across the 5 items of this dimension. The PDSS may not be screening the participants' level of sleep and eating disturbances as well as the other facets of PPD. An item separation index of 2.16 for this dimension indicates that the items on the SLP dimension are not as well dispersed on the scale.

8.4.6.2 Afrikaans Anxiety/Insecurity (ANX) dimension.

Person and item information for the Afrikaans PDSS Anxiety/Insecurity dimension is also summarized in Table 41. Winsteps (Linacre, 2009) eliminated 11 participants in this dimension who had extreme scores, hence the observed count of 167 participants. The average raw score of persons in the Afrikaans ANX dimension is 9.4 – the second highest raw score of the seven dimensions. The person infit mean-squares statistic = 1 with a t-statistic of 0.00, and the outfit mean-square statistic = 0.98, also with a t-statistic of 0.00. This demonstrates good fit the Rasch model in this dimension with neither too much nor too little variation and most participants responding as expected. The *SD* infit and outfit values for this dimension are 0.74 and 0.71 respectively.

The minimum and maximum values for person infit (0.03 and 3.68) and outfit (0.03 and 3.28) are fairly extreme. This indicates that there are some persons that had unexpected responses to items on the Afrikaans Anxiety/Insecurity dimension.

The minimum of -3.09 logits for items in the ANX dimension is low. This suggests that one or more women in this sample did not have symptoms of anxiety/insecurity. The maximum of 3.49 logits does, however, indicate that some participants had significant symptoms of anxiety/insecurity. The average logit for person ability or measure of anxiety/insecurity levels, is -0.27 with a model standard error of 0.58 and a *SD* of 1.36. If the distribution were approximately normal, almost 68 % of participants fell within a range of -1.63 and 1.09 logits. The minimum and maximum measure values of 3.49 and -3.09 are therefore extreme.

The PDSS items on the Anxiety/Insecurity dimension functioned very well. Individual item functioning will, however, be examined in more detail later. The average item infit and outfit values of 1.00 (t = 0.00) and 1.01 (t = 0.10) respectively are ideal Chi-Square values for these indices. Infit and outfit *SD* values were both 0.29. This indicates little variation and that most of the items in this dimension fit the Rasch model. The minimum MNSQ infit value for items is adequate at 0.81 while the maximum MNSQ infit value is elevated at 1.59. Outfit MNSQ shows an acceptable minimum of 0.74 but an elevated maximum of 1.54. Although the maximum MNSQ statistics are elevated, they remain lower than those for the Afrikaans PDSS as a whole (max infit 2.07; max outfit 3.10). This indicates that the items function well together within this dimension.

Reliability information for both items and persons on the ANX dimension is also shown in Table 41. The person separation index is moderate at 1.69. The person reliability estimate for this dimension is .74 with a Cronbach Alpha of .84 indicating that the items in this dimension were able to sufficiently separate the participants along the continuum. It further demonstrates good internal consistency of responses to items and that the items in the ANX dimension correlate well with each other. Participants are responding in a consistent fashion across the 5 items of this dimension. The Afrikaans PDSS's ANX dimension therefore adequately screens for participants' levels of anxiety.

The Afrikaans ANX dimension demonstrates an item separation index of 6.98. This indicates that the items on the ANX dimension are well dispersed on the scale and can distinguish between a number of levels of performance.

8.4.6.3 Afrikaans Emotional Lability (ELB) dimension.

The person and item information for the ELB dimension can also be found in Table 41. Winsteps (Linacre, 2009) eliminated 25 respondents in this dimension with extreme scores and the data is presented for the remaining 153 participants with non-extreme scores. As with the English PDSS ELB dimension, the Afrikaans PDSS ELB dimension also demonstrates the highest average raw score (10.10) of persons across the seven dimensions. Both the infit and outfit mean-squares statistics for persons are 0.99 (t = -0.10). These values are close to the Rasch-modeled expectations of 1.00. Little variation is present with participants responding as expected in this dimensions showing good fit to the Rasch model. The *SD* infit and outfit values for this dimension are wide at 0.79 and 0.97 respectively.

The minimum and maximum MNSQ statistics for person infit (0.02 and 4.76) and outfit (0.02 and 6.75) are extreme and are an indication that there are persons that had unexpected responses to items on the ELB dimension.

The minimum and maximum values in logits (-3.53 and 3.72 respectively) for the items in this dimension suggest that one or some women in this sample did not have symptoms of emotional lability while one or more participants had significant symptoms of emotional lability. The average logit for person ability is -0.06 with a model standard error of 0.63 and a *SD* of 1.67. Approximately 68% of participants therefore fell within a range of -1.73 and 1.61 logits. The minimum and maximum measure values of -3.53 and 3.72 are therefore extreme.

On average, the items on the ELB dimension functioned very well within this dimension. The average item infit and outfit values are 1.00 (t = 0.00) and 0.99 (t = -0.10) respectively. The infit and outfit *SD* values are 0.09 and 0.14 suggest that very little variation is present and that most of the items in the ELB dimension fit the Rasch model. The minimum and maximum MNSQ statistics for infit (min 0.89; max 1.12) as well as the minimum and maximum MNSQ statistics for outfit (min 0.82; max 1.15) in this dimension are adequate indicating that the items in the ELB dimension did not have extreme values and function well together.

Reliability information for items and persons on the Afrikaans ELB dimension shows a person separation index of 2.04 which indicates that persons are sufficiently separated across this dimension. The person reliability estimate for the ELB dimension is good at .81 and the Cronbach Alpha of .90 also indicates good internal consistency of responses to items. This demonstrates consistent responding by participants across the 5 items of this dimension. The Afrikaans PDSS's ELB dimension therefore adequately screens for participants' levels of emotional lability. Items in this dimension are well dispersed on the scale with an item separation of 5.78.

8.4.6.4 Afrikaans Mental Confusion (MNT) dimension.

The person and item information for the Afrikaans PDSS MNT dimension is presented in Table 41. Winsteps (Linacre, 2009) eliminated 34 respondents with extreme scores and the data is presented for the remaining 144 participants with non-extreme scores. The average raw score of persons in this dimension is 7.40. The infit mean-squares statistic is 0.99 (t = -0.10) and the outfit mean-square statistic is 0.97 (t = -0.10). Both these values are close to the Rasch-modeled expectations of 1.00. Little variation is evident and participants responded as expected. This indicates that the items in the Afrikaans MNT dimension fit the Rasch model. The *SD* infit and outfit values for this dimension are fairly wide at 0.92 and 0.90 respectively. The minimum and maximum MNSQ statistics for person infit (0.04 and 6.65) and outfit (0.04 and 5.71) are also extreme and are indicative of some unexpected responses to items on the Afrikaans MNT dimension.

The extreme minimum and maximum values in logits (-4.48 and 4.23 respectively) for the Afrikaans MNT items suggest that one or more women in this sample did not have symptoms of mental confusion while one or more had significant symptoms. The average logit for person ability is -1.18 with a model standard error of 0.70 and a *SD* of 1.87. Close to 68% of participants fell within a range of -3.05 and 0.69 logits. The maximum score of 4.23 logits is therefore very high.

Overall, the Afrikaans MNT items functioned very well within this dimension. Individual item performance is, however, examined in more detail in the next section. The average item infit and outfit values are 0.99 (t = -0.10) and 0.97 (t = -0.20) respectively. The infit and outfit *SD* values are 0.21 and 0.24 indicating that there is only some variation and that most of the items in the MNT dimension fit the Rasch model. The minimum and maximum MNSQ statistics for infit (min 0.75; max 1.32) as well as the minimum and maximum MNSQ statistics for outfit (min 0.76; max 1.37) in this dimension are adequate. The Afrikaans MNT items therefore function well together within this dimension and did not have extreme values.

Reliability information for both items and persons on the MNT dimension, as shown on Table 41, indicates a person separation index of 2.04 indicating that persons are sufficiently separated across the MNT dimension. The person reliability estimate for this dimension is .81 with a Cronbach Alpha of .91. The items on the Afrikaans MNT dimension demonstrate good internal consistency and participants responded in a consistent fashion across the 5 items of this dimension. The items on the Afrikaans MNT dimension therefore adequately screen for mental confusion among the participants. An item separation index of 2.17 suggests that the items on the MNT dimension not very well dispersed on the scale.

8.4.6.5 Afrikaans Loss of Self (LOS) dimension.

Table 41 also presents the person and item information for the Afrikaans PDSS LOS dimension. Winsteps (Linacre, 2009) eliminated 46 respondents with extreme scores in this dimension and the data is presented for the remaining 132 participants with non-extreme scores. The average raw score of persons in this dimension is 8.30. The person infit mean-squares statistic is 0.94 (t = -0.10) and the outfit mean-square statistic is 0.96 (t

= -0.10). These values are close to the Rasch-modeled expectations of 1.00. Little variation is therefore evident with participants responding as expected in this dimension and demonstrating good fit to the Rasch model. The *SD* infit and outfit values for this dimension are rather wide at 0.78 and 0.86 respectively.

The minimum and maximum MNSQ statistics for person infit (0.04 and 4.68) and outfit (0.04 and 5.71) are extreme. This suggests that there are some persons that had unexpected responses to items on the Afrikaans PDSS LOS dimension.

The extreme minimum and maximum values in logits (-3.95 and 4.59 respectively) for the items in this dimension indicate that one or more women in this sample did not have symptoms while others had significant symptoms of loss of self. The average logit for person ability is -0.66 with a model standard error of 0.69 and a *SD* of 2.00. Therefore, approximately 68% of participants fell within a range of -2.66 and 1.34 logits. The minimum and maximum measure values of -3.95 and 4.59 are therefore extreme.

The items on the Afrikaans LOS dimension appear to function well, on average, with an average infit and outfit value of 0.98 (t = -0.20) and 0.96 (t = -0.30) respectively. The infit and outfit *SD* values are 0.11 and 0.15 respectively indicating little variation in responses and that these items fit the Rasch model. Both the minimum and maximum MNSQ statistics for item infit (min 0.80; max 1.13) and outfit (min 0.75; max 1.18) are adequate.

On the LOS dimension, the person separation index is 2.31. This indicates that persons are sufficiently separated across this dimension. The person reliability estimate for the LOS dimension is good at .84. The Cronbach Alpha of .93 also indicates good internal consistency of responses to items. This demonstrates consistent responding by participants across the 5 items of this dimension. The Afrikaans PDSS's LOS dimension therefore adequately screen for participants' feelings of loss of self. Items in this dimension are moderately well dispersed on the scale with an item separation of 3.90.

8.4.6.6 Afrikaans Guilt/Shame (GLT) dimension.

The person and item information for the GLT dimension can be found in Table 41. Winsteps (Linacre, 2009) eliminated 53 respondents with extreme scores in this dimension and the data is presented for the remaining 125 participants with non-extreme scores. The average raw score of persons in this dimension is 9.30. The person infit and outfit mean-squares statistics are close to the Rasch-modeled expectation of 1.00 with MNSQ statistics of 0.96 (t = -0.20) for infit and 1.00 for outfit (t = -0.10). Items in this dimension fit the Rasch model with little variation evident and participants responding as expected.

The *SD* infit and outfit values for this dimension are wide at 1.01 and 1.23 respectively. The Afrikaans GLT dimension exhibits the most extreme maximum mean-square statistic infit and outfit values. The maximum MNSQ for person infit is 8.04 (min 0.07) while the maximum for outfit is 9.01 (min 0.06). This indicates the presence of unexpected responses to items on this dimension.

The minimum and maximum values in logits (-3.83 and 4.11 respectively) for items in this dimension is extreme. This indicates that one or more women in this sample did not have symptoms while others had significant symptoms of guilt or shame. The average logit for person ability is -0.28 with a model standard error of 0.65 and a *SD* of 1.77. Approximately 68% of participants therefore fell within a range of -2.05 and 1.49 logits. The minimum and maximum measure values of -4.08 and 3.90 are therefore extreme.

In general, the items in the Afrikaans GLT dimension performed well, although this will be confirmed later when the items of the GLT dimension are examined individually. The average item infit and outfit MNSQ statistics are 0.98 (t = -0.20) and 1.00 (t = 0.00) respectively. The infit and outfit *SD* values are 0.14 and 0.18 indicating that there is slight variation and that most of the items in this dimension fit the Rasch model. The minimum and maximum MNSQ infit values are adequate (min 0.83; max 1.17). The maximum MNSQ outfit value (1.24) is slightly high while the minimum is adequate at 0.83. Some items in this dimension therefore had slightly extreme values.

Reliability information for this dimension demonstrates a person separation index of 2.11 indicating that persons are sufficiently separated across this dimension. The person reliability estimate for this dimension is good at .82 with a Cronbach Alpha of .93. This shows that responses to items on the Afrikaans GLT dimension demonstrate good internal consistency and that participants are responding in a consistent fashion across the items from this dimension. The items on the GLT dimension therefore adequately screens for feelings of guilt or shame among the participants. Items on the Afrikaans GLT dimension are well dispersed on the scale with an item separation index of 5.85.

8.4.6.7 Afrikaans Suicidal Thoughts (SUI) dimension.

The person and item information for the SUI dimension is presented in Table 41. Winsteps (Linacre, 2009) eliminated 105 respondents in this dimension with extreme scores and the data is presented for the remaining 73 participants with non-extreme scores. Similar to the English PDSS SUI dimension, the average raw score of persons in the Afrikaans SUI dimension is the lowest of the 7 dimensions at 6.50. The person infit mean-squares statistic is 0.95 (t = -0.10) and the outfit mean-square statistic is 0.99 (t = 0.00). These values are near to the Rasch-modeled expectations of 1.00. The *SD* infit and outfit values for this dimension are 0.77 and 0.94 respectively.

The minimum and maximum MNSQ statistics for person infit (0.07 and 3.86) and outfit (0.07 and 5.06) indicate that there are persons that had unexpected responses to items on the SUI dimension.

The minimum and maximum measure values, in logits, for items in this dimension are extreme at -4.06 (minimum) and 3.49 (maximum). This suggests that one or more women in this sample did not have symptoms of suicidal thoughts and that one or more participants had significant symptoms of suicidal thoughts. The average logit for person ability (suicidal thoughts) is -1.33 with a model standard error of 0.73 and a *SD* of 1.99. Therefore, around 68% of participants fell within a range of -3.32 and 0.66 logits.

Item performance on the SUI dimension is, in general, good with an average infit and outfit value of 0.98 (t = -0.10) and 0.99 (t = 0.00) respectively. The infit and outfit *SD* values are 0.28 and 0.26 indicating that there is slight variation in participant responses. The minimum and maximum MNSQ statistics for item infit are 0.68 and 1.48

respectively, while the minimum and maximum MNSQ statistics for item outfit are 0.68 and 1.45. The minimum values are adequate but the maximum values are extreme indicating that some items had extreme values in the SUI dimension.

The person separation index on the SUI dimension is 2.12. Participants are therefore adequately separated across this dimension. The person reliability estimate for the SUI dimension is good at .82. The Cronbach Alpha of .94 also indicates very good internal consistency of responses to items. Participants therefore responded consistently across the 5 items of this dimension indicating that it adequately screens for symptoms of suicidal ideation. Items in this dimension are, however, only moderately well dispersed on the scale, with an item separation index of 3.34.

8.4.7 Item fit statistics for the Afrikaans PDSS dimensions.

Tables 42 to 48 compare the items of the Afrikaans PDSS dimensions in terms of their measure order. The items are listed in sequence from most difficult to agree with to easiest to agree with.

8.4.7.1 Afrikaans Sleeping/Eating Disturbances (SLP) dimension.

The items from the Afrikaans SLP dimension are listed in Table 42. The most difficult item to agree with is Item 29 (Ek het geweet ek moes eet, maar kon nie) and the easiest to agree with is Item 1 (Al het my baba geslaap, het ek gesukkel om te slaap) – similar to the English SLP dimension. The infit MNSQ statistics for items from this content scale indicate that the items perform better within the Afrikaans SLP content scale than within the total Afrikaans PDSS. This indicates good construct validity for items from this content scale. The Rasch error estimates for items on this dimension were small at 0.10 for all items.

The Pearson item-total correlation (r_{it}) values for the Afrikaans SLP dimension indicate good construct validity with positive values that range from .71 to .78. These high values also indicate that there are no coding errors. Item 1 has a slight discrepancy between the Pearson item-total correlation (r_{it}) value (.78) and the expected value (.82) and with a slightly elevated infit MNSQ statistic mentioned earlier, also suggests that item 1 may not fit the SLP dimension as well as the other items do. There is not much discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of the other items in this dimension which correlate well and tap into a unidimensional construct of disturbances in sleeping or eating.

Table 42 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Sleeping/Eating Disturbances (SLP) Dimension (n=178)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Sleeping/Eating Disturbances (SLP)						
1	Al het my baba geslaap, het ek gesukkel om te slaap.	-0.46	0.10	1.27	1.25	.78
8	Ek het my eetlus verloor.	0.02	0.10	1.10	1.06	.76
15	Ek het in die middel van die nag vanself wakker geskrik en gesukkel om weer aan die slaap te raak.	0.06	0.10	0.92	0.71	.74
22	Ek het snags lank rondgerol en gesukkel om aan die slaap te raak.	0.08	0.10	0.64	0.65	.77
29	Ek het geweet ek moes eet, maar kon nie.	0.30	0.10	1.14	0.96	.71
M		0.00	0.10	1.01	0.93	
SD		0.25	0.00	0.22	0.23	

MNSQ = mean-square

8.4.7.2 Afrikaans Anxiety/Insecurity (ANX) dimension.

Table 43 lists the items from the Afrikaans PDSS ANX dimension. The most difficult item to agree with is Item 16 (Ek was so angstig ek het gevoel asof ek uit my vel wou spring) and the item that was the easiest to agree with is Item 9 (Ek het heeltemal oorweldig gevoel). These items were also indicated as the most difficult and the easiest to agree with in the English ANX dimension. Item 30 ('Ek het gevoel asof ek heeltyd aan die gang moes bly.') does not fit well with an infit MNSQ statistic of 1.59. This item may be poorly constructed, ambiguous, or does not relate closely to the overall construct. The remaining items demonstrate good fit with values that range from 0.81 to 1.01. (acceptable = 0.60 – 1.40; Wright & Linacre, 1994). The Rasch error estimates on this dimension was small and ranged from 0.08 – 0.10, with a mean of 0.09.

The Pearson item-total correlation (r_{it}) values for the Afrikaans ANX dimension are generally good (.64 to .80) suggesting that coding errors were unlikely. Item 30 (Ek het gevoel asof ek heeltyd aan die gang moes bly) does, however, have the lowest Pearson item-total correlation (r_{it}) value of all items in the Afrikaans PDSS. Furthermore, relative to other items in the Afrikaans PDSS, it differs more significantly from the expected value (.73) for this item which suggests the presence of unmodeled noise. Coupled with a high infit MNSQ statistic (1.59), Item 30 does not fit the ANX dimension as well as the other items do. The items in this dimension correlate well and tap into a unidimensional construct of anxiety or insecurity suggesting good construct validity.

Table 43 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Anxiety/Insecurity (ANX) Dimension (n=178)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Anxiety/Insecurity (ANX)						
2	Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak.	-0.05	0.09	0.83	0.82	.79
9	Ek het heeltemal oorweldig gevoel.	-1.01	0.10	1.01	0.94	.80
16	Ek was so angstig ek het gevoel asof ek uit my vel wou spring.	0.95	0.09	0.81	0.74	.73
23	Ek het alleen gevoel.	-0.36	0.08	0.86	0.87	.79
30	Ek het gevoel asof ek heeltyd aan die gang moes bly.	0.47	0.09	**1.59**	**1.54**	.64
M		0.00	0.09	1.02	0.98	
SD		0.67	0.00	0.29	0.29	

Note. Boldface values have infit and outfit MNSQ statistics less than 0.60 or greater than 1.40
MNSQ = mean-square

8.4.7.3 Afrikaans Emotional Lability (ELB) dimension.

Table 44 lists the items from the Afrikaans ELB dimension The most difficult item to agree with was Item 10 (Ek was bang dat ek nooit weer gelukkig sou wees nie). The easiest item to agree with is Item 24 (Ek was baie geïrriteerd). The same items were noted as the most difficult and easiest to agree to in the English ELB dimension. All mean-squares for infit and outfit in the ELB dimension are near 1.00 and fall within an acceptable range. This suggests little distortion of the measurement system for this dimension. Items in this dimension appear to have been well understood by the Afrikaans participants and seem to fit the definition of the construct – emotional lability – well. The Rasch error estimates on this dimension was small and ranged from 0.10 – 0.11, with a mean of 0.10.

The Pearson item-total correlation (r_{it}) values for the ELB dimension are all high positive values between .77 and .87 which indicate that there are no coding errors. There is very little discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of the items in this dimension. This indicates good construct validity and that all the items in this dimension correlate well and tap into a unidimensional construct.

Table 44 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Emotional Lability (ELB) Dimension (n=178)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Emotional Lability (ELB)						
3	Ek het gevoel asof my emosies wipplank ry.	-0.72	0.11	1.12	1.15	.85
10	Ek was bang dat ek nooit weer gelukkig sou wees nie.	0.78	0.10	1.07	1.01	.77
17	Ek het sonder enige rede baie gehuil.	0.35	0.10	1.04	1.12	.80
24	Ek was baie geïrriteerd.	-0.76	0.11	0.89	0.85	.87
31	Ek het baie kwaad gevoel en was gereed om te ontplof.	0.35	0.10	0.90	0.82	.83
M		0.00	0.10	1.00	0.99	
SD		0.62	0.01	0.09	0.14	

MNSQ = mean-square

8.4.7.4 Afrikaans Mental Confusion (MNT) dimension.

Table 45 presents the item fit statistics for the items for the Afrikaans MNT dimension. The most difficult item to agree with is Item 18 (Ek het gedink ek raak gek). This item was also the most difficult to agree with in the English MNT dimension. The easiest was Item 4 (Ek het gevoel of ek van my verstand af raak). Items in this content scale had infit MNSQ statistics within an acceptable range. Item 4 demonstrates the poorest fit (infit MNSQ = 1.32; outfit MNSQ = 1.37), but its fit statistic still falls within an acceptable range. The items in this dimension were well understood by most participants and appear to fit the definition of the construct well. The Rasch error estimates on this dimension was small and ranged from 0.11 – 0.13, with a mean of 0.12.

High positive Pearson item-total correlation (r_{it}) values in the Afrikaans MNT dimension indicate that there are no coding errors and support good construct validity. The values range from .79 to .86. There is not much discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of any items in this dimension indicating that they correlate very well and tap into a unidimensional construct.

Table 45 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Mental Confusion (MNT) Dimension (n=178)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Mental Confusion (MNT)						
4	Ek het gevoel of ek van my verstand af raak.	-0.34	0.11	1.32	1.37	.79
11	Ek kon op niks konsentreer nie.	0.24	0.13	0.84	0.80	.86
18	Ek het gedink ek raak gek.	0.43	0.12	0.75	0.76	.82
25	Ek het dit moeilik gevind om die eenvoudigste besluite te neem.	-0.13	0.12	0.92	0.80	.85
32	Ek het gesukkel om op 'n taak te konsentreer.	-0.20	0.12	1.14	1.14	.82
M		0.00	0.12	0.99	0.97	
SD		0.29	0.00	0.21	0.24	

MNSQ = mean-square

8.4.7.5 Afrikaans Loss of Self (LOS) dimension.

The items of the Afrikaans LOS dimension are listed in terms of their measure order in Table 46. The most difficult item to agree with is Item 33 (Ek het nie eg gevoel nie). The item that was the easiest to agree with was Item 5 (Ek was bang dat ek nooit weer my normale self sou wees nie). Both these items also ranked as the most difficult and the easiest to agree with in the English LOS dimension. The infit MNSQ statistics for items in the Afrikaans LOS dimension are all within an acceptable range. The items suggest undimensionality and appear to have been well understood by the Afrikaans participants. The Rasch error estimates for items on this dimension were small at 0.12 for all items.

Pearson item-total correlation (r_{it}) values for the Afrikaans LOS dimension are all positive high values between .83 and .87 indicating good construct validity and that there are no coding errors. The Pearson item-total correlation (r_{it}) values and the expected values of all items in this dimension indicate very little discrepancy. All items in the Afrikaans LOS dimension correlate well and tap into a unidimensional construct.

Table 46 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Loss of Self (LOS) Dimension (n=178)

	Dimension / Item	Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Loss of Self (LOS)						
5	Ek was bang dat ek nooit weer my normale self sou wees nie.	-0.94	0.12	1.13	1.06	.87
12	Ek het soos 'n vreemde vir myself gevoel.	0.08	0.12	0.80	0.75	.88
19	Ek het myself nie meer geken nie.	0.10	0.12	0.98	0.92	.86
26	Ek het gevoel asof ek nie normaal was nie.	0.27	0.12	0.94	0.85	.86
33	Ek het nie eg gevoel nie.	0.48	0.12	1.04	1.18	.83
M		0.00	0.12	0.98	0.96	
SD		0.49	0.00	0.11	0.15	

MNSQ = mean-square

8.4.7.6 Afrikaans Guilt/Shame (GLT) dimension.

Table 47 lists the item fit statistics for the Afrikaans GLT dimension. Similar to the English GLT dimension, the most difficult item to agree with here is also Item 27 (Dit het gevoel asof ek my ware gevoelens en gedagtes oor my baba moes wegsteek). The easiest item to agree with was, however, Item 13 (Ek het gevoel asof baie ander ma's beter as ek was).

All items in this dimension had infit MNSQ statistics within an acceptable range. They appear to have been well understood by the English participants and seem to fit the definition of the construct well. The Rasch error estimates on this dimension was small and ranged from 0.11 – 0.12, with a mean of 0.12.

The Pearson item-total correlation (r_{it}) values for the Afrikaans GLT dimension indicate good construct validity with high positive values that range from .80 to .90.

These high values also indicate that there are no coding errors. There is very little discrepancy between the Pearson item-total correlation (r_{it}) values and the expected values (EXP) of the items in this dimension suggesting that they correlate well and tap into a unidimensional construct of feelings of guilt or shame.

Table 47 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Guilt/Shame (GLT) Dimension (n=178)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Guilt/Shame (GLT)						
6	Ek het gevoel asof ek nie die ma is wat ek wou wees nie.	-0.52	0.12	0.85	0.83	.90
13	Ek het gevoel asof baie ander ma's beter as ek was.	-0.85	0.12	1.17	1.24	.89
20	Ek het skuldig gevoel omdat dit vir my gevoel het asof ek nie my baba lief genoeg het nie.	0.79	0.11	0.90	0.87	.82
27	Dit het gevoel asof ek my ware gevoelens en gedagtes oor my baba moes wegsteek.	0.94	0.12	1.12	1.21	.80
34	Ek het gevoel asof ek as ma misluk.	-0.36	0.12	0.83	0.86	.89
M		0.00	0.12	0.98	1.00	
SD		0.72	0.00	0.14	0.18	

MNSQ = mean-square

8.4.7.7 Afrikaans Suicidal Thoughts (SUI) dimension.

Table 48 shows that the most difficult item in the Afrikaans SUI dimension to agree with was Item 7 (Ek het gedink die dood sou die enigste uitweg uit hierdie nagmerrie wees), and, like in the English SUI dimension, the easiest was Item 28 (Ek het gevoel dat dit vir my baba beter sou wees sonder my). The Afrikaans version of Item 28 does, however, also has a high infit mean-square value (1.48) which indicates that this item did not fit the model well or that responses to this item were unpredictable, possibly due to unmodeled noise. The remaining items from this dimension had infit and outfit mean-squares within an acceptable range that reflect little distortion in these items, that they were well understood by most participants, and appear to fit the definition of the construct well. The Rasch error estimates on this dimension was slightly higher than on other dimensions in the Afrikaans PDSS and ranged from 0.16 – 0.19, with a mean of 0.17.

The high positive Pearson item-total correlation (r_{it}) values that range from .85 to .90 support good construct validity for items in this dimension and that there are no coding errors. The Pearson item-total correlation (r_{it}) values and the expected values of all items in this dimension indicate very little discrepancy. All items in the Afrikaans SUI dimension correlate well and tap into a unidimensional construct.

Table 48 Item Difficulty, Fit Statistics, and Pearson Item-Total Correlations (r_{it}) for the Afrikaans PDSS Suicidal Thoughts (SUI) Dimension (n=178)

Dimension / Item		Item difficulty (logits)	*SE*	Infit MNSQ	Outfit MNSQ	r_{it}
Suicidal Thoughts (SUI)						
7	Ek het gedink die dood sou die enigste uitweg uit hierdie nagmerrie wees.	0.74	0.19	1.06	1.06	.85
14	Ek het begin dink dat dit beter sou wees as ek dood was.	0.33	0.18	0.68	0.68	.89
21	Ek wou myself seermaak.	0.43	0.18	0.87	0.90	.85
28	Ek het gevoel dat dit vir my baba beter sou wees sonder my.	-1.09	0.16	**1.48**	**1.45**	.90
35	Ek wou eenvoudig hierdie wêreld agterlaat.	-0.41	0.17	0.80	0.86	.90
M		0.00	0.17	0.98	0.99	
SD		0.67	0.01	0.28	0.26	

Note. Boldface values have infit and outfit MNSQ statistics less than 0.60 or greater than 1.40
MNSQ = mean-square

8.4.8 Response category statistics: Item option and distractor frequencies for the Afrikaans PDSS dimensions.

The frequency of responses to the 5-point Likert rating scale categories of the Afrikaans PDSS are briefly discussed below and are outlined in Table 79 to Table 85 in Appendix F. In the Afrikaans PDSS the SLP, LOS, GLT and SUI dimensions, category "0" was selected most often in all items. This trend was also noted in the same dimensions of the English PDSS. This was particularly evident in the SLP dimension and even more so in the SUI dimension with frequency counts ranging from 63% (item 28) to 78% (item 21). Similar to the English PDSS, category "0" was selected more often for the majority of items in the Afrikaans ANX, ELB and MNT dimensions.

The Afrikaans PDSS items had more categories with less than 10 observations than the English PDSS items had. This was particularly noticeable in the Afrikaans SLP dimension (7 categories) and the Afrikaans SUI dimension (9 categories), but also in the Afrikaans MNT dimension (4 categories) and the Afrikaans ANX dimension (1 category). The remaining dimensions had category observations that ranged from 10 to 138. The remaining categories for the Afrikaans PDSS were used fairly regularly.

All items in the Afrikaans PDSS dimensions have average measure values (in logits) which increase gradually with each higher response category. This supports the validity of the 5-point Likert scale for the Afrikaans PDSS with each higher response category corresponding to "more" of the variable being measured. Similar to the English PDSS, however, there are a number of categories across all the Afrikaans PDSS dimensions that have outfit MNSQ statistics greater than 1.40 or lower than 0.60. The Pearson item-total correlation (r_{it}) values provide support for the convergent and discriminant validity of the item categories for the Afrikaans PDSS dimensions. Some items, however, have values that do not advance steadily. These are items 8, 11, 12, 14, 15, 16, 21, 22, 26, 29, 30 and 35.

When there is a great discrepancy between the observed Pearson item-total correlation (r_{it}) and the expected (EXP) value, it may indicate that the item does not show a good fit with the dimension being measure. When the observed value is much higher than the expected value it may indicate dependency in the data. When the observed value is much lower than expected value, unmodeled noise is possible (Linacre, 2008).

8.5 Items Marked as Difficult to Understand

After completing the PDSS, or its Afrikaans translation, the participants were asked to indicate if there were any items that they found difficult to fully understand. It is important that respondents understand the language of the assessment measure used. Respondents who are not proficient in the language of the measure may introduce construct irrelevant components to the assessment process (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999). To effectively identify women with PPD from different cultures and language groups, there should be no language barrier in the screening process.

Cultural groups may differ in their language spoken. They may also differ in terms of the way in which verbal expressions are formally structured, even if they speak the same language. Furthermore, different cultural groups may assign different meanings to commonly used expressions. Respondents from one cultural or ethnic group will therefore differ to other cultural or ethnic groups in their performance to the extent that they are familiar with the questionnaire's language as well as expressions associated with that language. For this reason participants were asked to mark items they did not fully understand. These items are presented in Table 49 below.

Twelve English participants and eight Afrikaans participants marked items as difficult to understand. Several participants had difficulty understanding a number of items. Item 16 (I felt like I was jumping out of my skin; Ek was so angstig ek het gevoel asof ek uit my vel wou spring) was marked most frequently as difficult to understand, and was also the most frequently marked Afrikaans item (three participants). Item 16 was marked by five English participants and, together with Item 2 (I got anxious over even the littlest things that concerned my baby) were the most frequently marked English items. Other frequently marked items were Item 3 (I felt like my emotions were on a roller coaster; Ek het gevoel asof my emosies wipplank ry), Item 9 (I felt really overwhelmed; Ek het heeltemal oorweldig gevoel), Item 30 (I felt like I had to keep moving or pacing; Ek het gevoel asof ek heeltyd aan die gang moes bly), and Item 33 (I did not feel real; Ek het nie eg gevoel nie).

Table 49 Items Marked by Participants as Difficult to Understand after Completing English PDSS or Afrikaans PDSS

Participants	No. of items marked	Items marked as difficult to understand															
		Item[a]	Item[b]	Item	Item	Item[d]	Item[d]	Item[a]	Item[d]	Item	Item[a]	Item[d]	Item[bd]	Item[ad]	Item	Item[abc]	Item
English																	
E 17	5		2	3				9					24			30	
E 39[e]	1									16							
E 52	6			3	4					16	18					30	33
E 110	3		2	3						16							
E 113	1																33
E 114	1									16							
E 130	1					5											
E 136[e]	5		2				8	9		16						30	
E 152	1							9									
E 154	1																33
E 178	1		2														
E 183	1		2														
Afrikaans																	
A 4	1													26			
A 33	1									16							
A 72	1	1															
A 85[e]	5		2	3				9					24			30	
A 88	7		2						12	16		19			27	30	33
A 116[e]	1			3													
A 149	1																33
A 178	2									16							
Total times item was marked		1	7	5	1	1	1	4	1	8	1	1	2	1	1	5	5

a items with DIF in total PDSS or total Afrikaans PDSS
b items with DIF in content scale
c item had fit problems in Rasch analysis of Afrikaans PDSS Anxiety/Insecurity content scale
d item contributes to INC index
e participant has an INC score of 4 or more.

8.6 Invariance and Differential Item Functioning

Demonstrating reliability and validity are important steps in the cross-cultural adaptation and validation of instruments. Although necessary, these are, however, not sufficient conditions for evaluating cross-cultural validity when the aim is to compare persons across diverse cultures or countries by means of adapted versions of the same instrument. An increasing awareness of the cultural, gender, developmental, and socio-economic influences on psychological constructs has resulted in greater recognition of the need to demonstrate measurement invariance before assuming that measures are equivalent across groups (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999).

Invariance is therefore also a requirement of cross-cultural validation. In simple terms, invariance means that if two respondents from different racial, ethnic, gender (or other) subgroups are at the same level of the trait or construct being measure, then the probability of a respondent from one subgroup affirming an item (in the dichotomous case) will be the same as the probability of a respondent from another subgroup affirming the item (Küçükdeveci et al., 2004 Get another ref here). Bond and Fox (2007) define invariance as a variable which maintains its identity from one occasion to the next. Invariance may encompass stability over time or stability over samples in the order of item logit positions as well as stability of item positions on the logit scale across time or across samples.

Construct-irrelevant variance and construct under-representation are two major threats to validity (Downing & Haladyna, 2004). The construct-irrelevant components of a measure refer to those “variables that systematically (rather than randomly) interfere with the ability to meaningfully interpret scores or ratings …" (Downing & Haladyna, 2004, p.327). These variables do not form part of the construct that is being measured and may include, for example, items that have not been statistically proven to be valid and reliable, items written at an inappropriate reading level, or the use of inappropriate jargon (Downing & Haladyna, 2004). If the responses to a questionnaire (and hence the outcome or results to that questionnaire) are dependent on language proficiency, that dependency may be responsible for construct-irrelevant variance.

Measurement invariance must be established before instruments may be deemed to be equivalent in a measurement sense (Küçükdeveci et al., 2004). If measurement invariance is established, then the differences on the screening scales' scores accurately reflect the differences on the latent characteristics assessed by the construct.

Invariance is determined through analysis of item bias or differential item functioning (DIF). When an item's difficulty estimate location is not consistent across samples, but varies by more than the modelled error, it provides clear evidence that DIF exists. The presence of DIF between groups indicates that they cannot be compared meaningfully on the item. DIF is based on whether items have shifted in meaning for differing time points or groups (Bond, 2003; Bond & Fox, 2007).

Item response theory (IRT) is a parametric method for identifying DIF. Analysis of DIF is a powerful means of testing items for bias in IRT relative to CTT-based methods (Harvey & Hammer, 1999). Edelen, Thissen, Teresi, Kleinman, and Ocepek-Welikson (2006) agree that IRT and the likelihood-based model comparison approach are robust in their ability detect DIF in order to develop, refine and evaluate measures for use in ethnically diverse populations.

Rasch modelling, however, has advantages which make it more suitable for identifying DIF than IRT or CTT (Andrich, 2004a; Royal, 2010). Chiang et al (2009) assert that invariance analyses, although they can be conducted using CTT by examining differences in item means by group or time, are greatly simplified via use of Rasch modelling software. The separability of the item difficulty and person ability parameters is one advantage. This characteristic parameter separation is unique to the Rasch model (Andrich, 2004a). The parameters are derived independently and the item analysis is therefore not dependent on the sample from which it was taken. This provides fundamental person-free measurement and item-free calibration when the data adequately fits the Rasch model and persons and item can be mapped on a common invariant scale (Bond & Fox, 2001).

Two and three parameter IRT models control for factors like difficulty, discrimination and guessing. This means that the item response curves of different items can cross (Andrich 2004b). As a result the relative difficulty levels of items are not invariant across persons in the sample. This violates the assumptions of invariant measurement. Sample independent measurement is only feasible in a one-parameter model, like the Rasch model. The Rasch measurement model aims to deliver invariance in scientific measurement with estimates of item difficulty and estimates of person ability where the probability of a correct response is a function of the difference between item difficulty and person ability, and nothing else (Bond & Fox, 2007). Furthermore, Rasch analyses instantiates interval level measurement as opposed to ratio level measurement. The invariance of item and person estimate values therefore always remains relative (Bond & Fox, 2007).

In measurement, it is important that the values attributed to variables by a measurement system be independent of the particular measuring instrument that was used. The calibrations of the measuring instrument should also remain invariant when using an appropriate measuring instrument for the purpose intended (Bond & Fox, 2007). The Rasch model is based on a mathematical formulation of invariance, which is an operational criterion for fundamental measurement (Andrich, 2004a). The Rasch model therefore has significant advantages as a measurement model for the validation of tests and measuring instruments.

Proponents of Rasch modelling maintain that data is never pure or accurate and the data must therefore conform to the measurement model rather than the measurement model chosen to fit the data (as in two-parameter and three-parameter logistic IRT approaches). As a result, only data which adequately fit the Rasch model can satisfy the requirements of fundamental measurement.

Figure 7 plots the English and Afrikaans PDSS item location values (d) against each other. The diagonal dotted line is drawn through the points that represent the calibration mean of D.x and D.y (zero logit). It represents the precise modelled relation between the English and Afrikaans PDSS's sets of item estimates if they did not shift location, staying completely invariant in precise and error-free measurement conditions – a situation that is unachievable in practice (Bond & Fox, 2007, p. 73). Measurement error estimates are provided by Rasch modeling for all difficulty estimates which are used to construct 'quality control lines' on either side. These lines on the outside represent the 95% confidence band. This enables determining how close the distribution of the plotted ability points is to the modelled diagonal line for the measures to be considered sufficiently invariant. It also allows for distinguishing those items on the outside of the confidence 95% band which show significant shift. Measurement error estimates are always provided by Rasch modelling and therefore some shift in location is not unexpected.

Nearly one third of the items in the complete PDSS and Afrikaans PDSS exhibit differential item functioning indicating that those items functioned differently across the two language groups. Table 50 lists items that showed significant shift in the PDSS total item Rasch analysis.

Invariance (within measurement error) across the two language versions of the PDSS dimensions was supported for most items. This helps to affirm the integrity of the PDSS dimensions under Rasch analysis procedures. It further demonstrates that the PDSS dimensions maintain its measurement properties across both English and Afrikaans South African samples. The six items that showed significant shift in the PDSS dimensions are listed in Table 51.

Figure 8 to 14 plots the English and Afrikaans PDSS dimensions' item location values (d) against each other. Measurement error estimates, provided by Rasch modeling, are used to construct 'quality control lines' on either side and are represent by the 95 % confidence band on the outside. These figures provide a visual aid for distinguishing those items on the outside of the confidence 95% band which show significant shift.

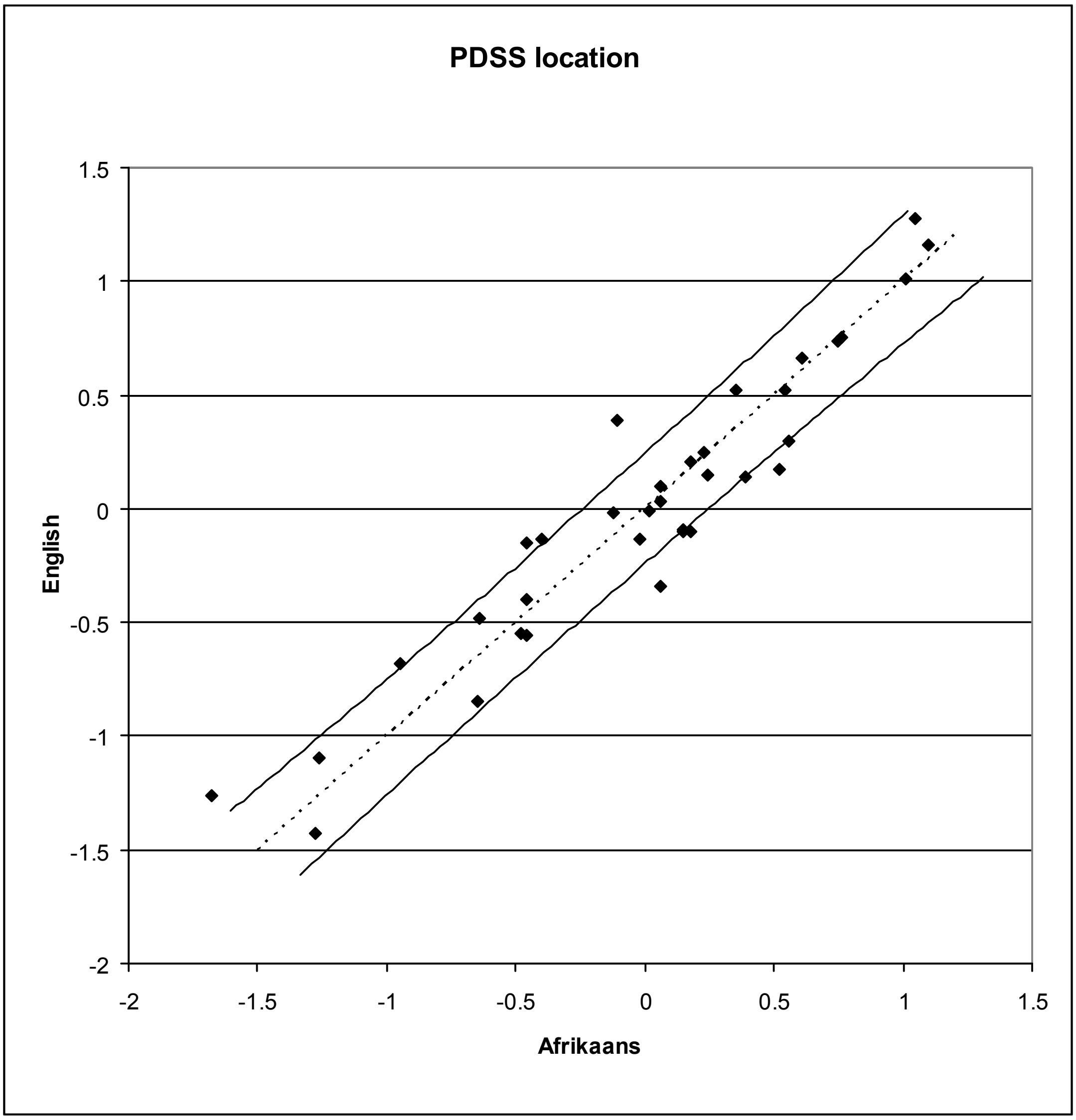

Figure 7 Differential Item Functioning of English and Afrikaans PDSS items.

Table 50 Items that Exhibit Differential Item Functioning in the PDSS Total Item Rasch Analysis

Item	Dim	Item content	Afrikaans PDSS	English PDSS	Afrikaans Model SE	English Model SE	z-value
1[a b]	SLP	I had trouble sleeping even when my baby was asleep. Al het my baba geslaap, het ek gesukkel om te slaap.	0.18	-0.10	0.09	0.09	2.20
9	ANX/INS	I felt really overwhelmed. Ek het heeltemal oorweldig gevoel.	-1.68	-1.26	0.10	0.09	-3.12
15[ab]	SLP	I woke up on my own in the middle of the night and had trouble getting back to sleep. Ek het in die middel van die nag vanself wakker geskrik en gesukkel om weer aan die slaap te raak.	0.52	0.17	0.09	0.08	2.91
18	MNT	I thought I was going crazy. Ek het gedink ek raak gek.	0.39	0.14	0.09	0.08	2.08
22[a]	SLP	I tossed and turned for a long time at night trying to fall asleep. Ek het snags lank rondgerol en gesukkel om aan die slaap te raak.	0.56	0.30	0.09	0.09	2.04
23	ANX/INS	I felt all alone. Ek het alleen gevoel.	-0.95	-0.68	0.09	0.08	-2.24
25	MNT	I had a difficult time making even a simple decision. Ek het dit moeilik gevind om die eenvoudigste besluite te neem.	0.06	-0.34	0.09	0.09	3.14
26	LOS	I felt like I was not normal. Ek het gevoel asof ek nie normaal was nie.	0.15	-0.10	0.09	0.08	2.08
30[b]	ANX	I felt like I had to keep moving or pacing. Ek het gevoel asof ek heeltyd aan die gang moes bly.	-0.11	0.39	0.09	0.09	-3.93
31	ELB	I felt full of anger ready to explode. Ek het baie kwaad gevoel en was gereed om te ontplof.	-0.46	-0.15	0.09	0.08	-2.57
34	GLT	I felt like a failure as a mother. Ek het gevoel asof ek as ma misluk.	-0.4	-0.13	0.09	0.08	-2.24

a Item also had problems in English PDSS total fit analysis

b Item also had problems in Afrikaans PDSS total fit analyis

Table 51 Items that Exhibit Differential Item Functioning in the PDSS Dimensions

Item	Dim	Item content	Afrikaans PDSS	English PDSS	Afrikaans Model SE	English Model SE	z-value
2	ANX/INS	I got anxious over even the littlest things that concerned my baby. Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak.	-0.05	-0.60	0.09	0.10	4.09
24	ELB	I have been very irritable. Ek was baie geïrriteerd.	-0.76	-1.10	0.11	0.11	2.19
25[a]	MNT	I had a difficult time making even a simple decision. Ek het dit moeilik gevind om die eenvoudigste besluit te neem.	-0.13	-0.47	0.12	0.12	2.00
30[ab]	ANX/INS	I felt like I had to keep moving or pacing. Ek het gevoel asof ek heeltyd aan die gang moes bly.	0.47	0.98	0.09	0.10	-3.79
32	MNT	I had difficulty focusing on a task. Ek het gesukkel om op 'n taak te konsentreer.	-0.20	0.15	0.12	0.12	-2.06
34[ab]	GLT	I felt like a failure as a mother. Ek het gevoel asof ek as ma misluk.	-0.36	0.03	0.12	0.11	-2.40

a Items showed significant shift in Rasch analysis of PDSS as a whole as well as in analysis of dimensions.

b Items also had misfit in PDSS dimensions

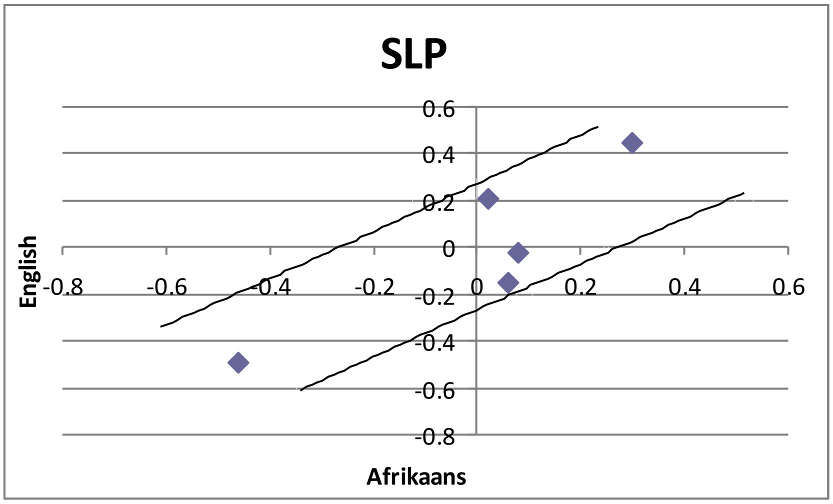

Figure 8 Differential item functioning of items in the Sleeping/Eating Disturbances (SLP) dimension.

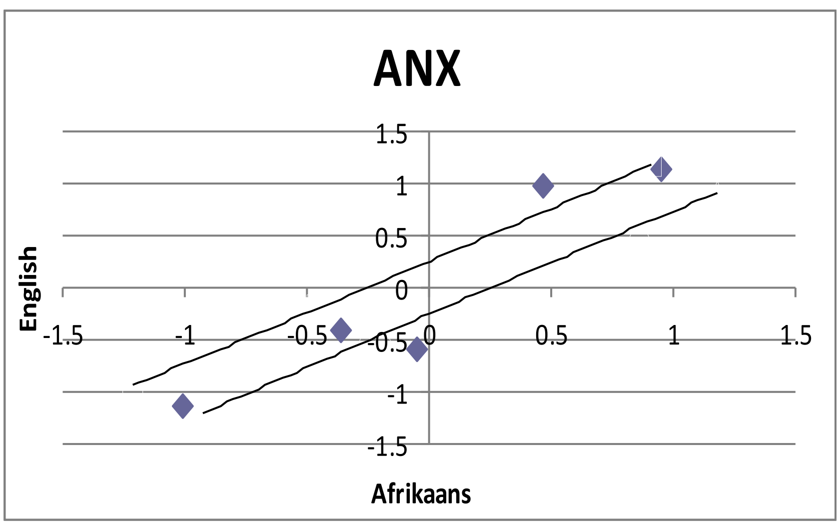

Figure 9. Differential item functioning of items in the Anxiety/Insecurity (ANX) dimension.

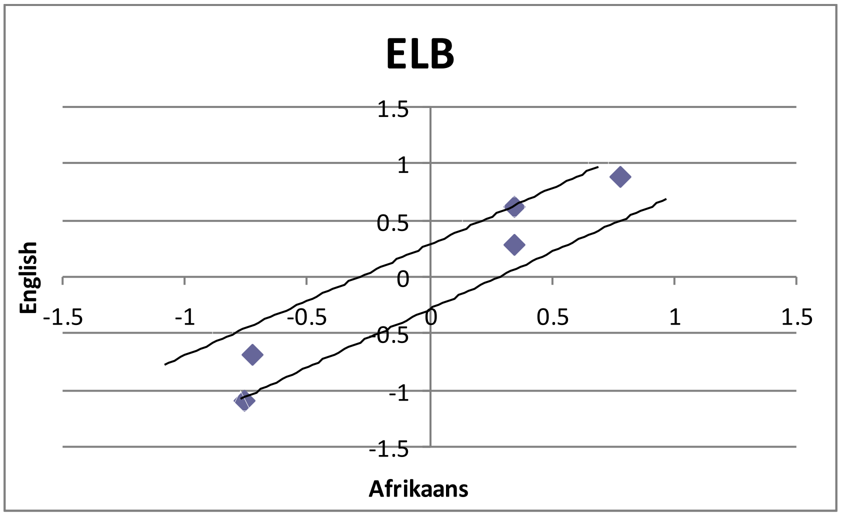

Figure 10. Differential item functioning of items in the Emotional Lability (ELB) dimension.

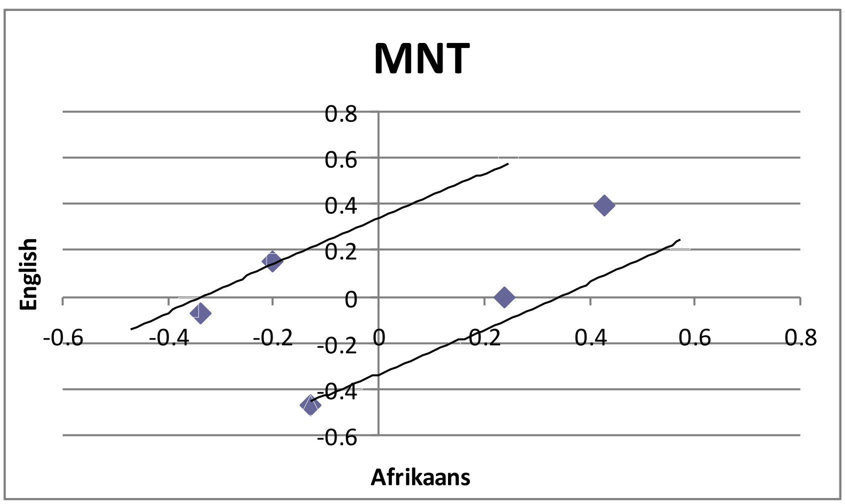

Figure 11. Differential item functioning of items in the Mental Confusion (MNT) dimension.

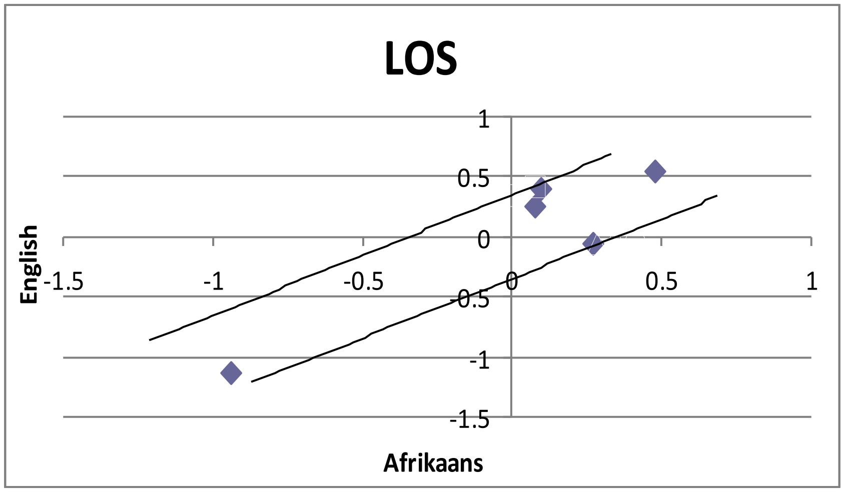

Figure 12. Differential item functioning of items in the Loss of Self (LOS) dimension.

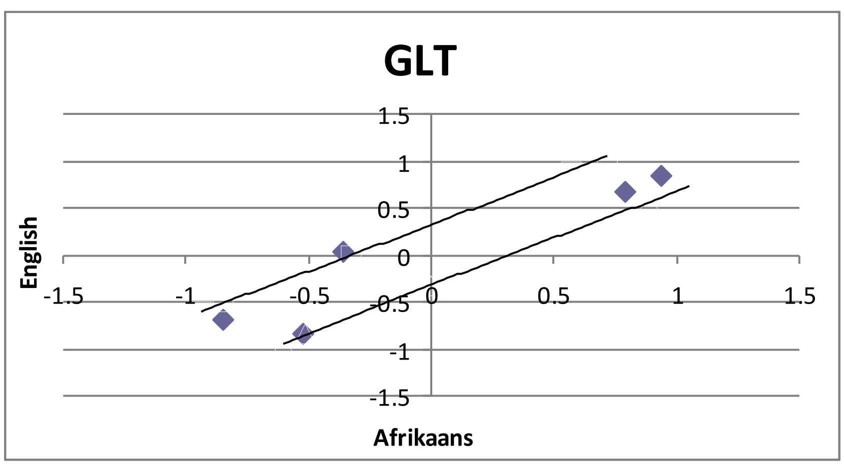

Figure 13. Differential item functioning of items in the Guilt/Shame (GLT) dimension.

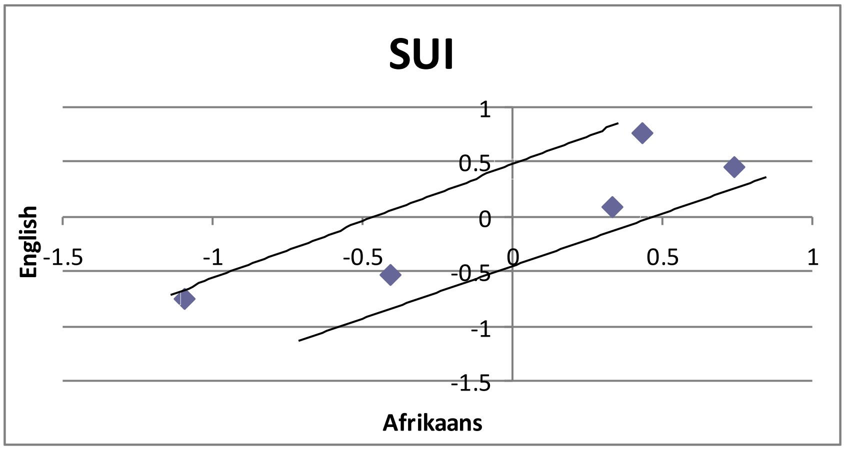

Figure 14. Differential item functioning of items in the Suicidal Thoughts (SUI) dimension.

8.7 Results of the Analysis of Risk Factors for PPD

The demographic and obstetric characteristics of the participants and their PDSS screening results across three screening outcome categories are presented in Table 52. The screening outcome categories, recommended by Beck and Gable (2002), are as follows: i) normal adjustment (total score of ≤59); ii) significant symptoms of PPD (total score of 60 to 79); and iii) positive screening for PPD (total score of ≥80). The prevalence of a positive screen for major PPD using the PDSS in this study was 47.9% (n = 175). Furthermore, screening identified an additional 17.3% (n = 63) of mothers with potential symptoms of PPD.

Table 52 Demographic and Obstetric Variables by PDSS Screening Result (N = 365)

Variable	Total	Normal Adjustment (≤59) n	%	Symptoms of PPD Present (60-79) n	%	Major PPD (≥80) n	%
	N = 365	127	34.8	63	17.3	175	47.9%
Marital status							
Married	324	120	37.0	57	17.6	147	45.4
Unmarried	24	2	8.3	4	16.7	18	75.0
Divorced	2	1	50.0	0	0.0	1	50.0
De facto relationship	15	4	26.7	2	13.3	9	60.0
Gestational age of baby at birth							
Before 28 weeks	7	4	57.1	0	0.0	3	42.9
29 - 33 weeks	11	2	18.2	3	27.3	6	54.5
34 - 37 weeks	62	14	22.6	12	19.4	36	58.1
38 - 40 weeks	217	83	38.2	37	17.1	97	44.7
Beyond 40 weeks	68	24	35.3	11	16.2	33	48.5
Type of delivery							
Normal vaginal	99	46	46.5	15	15.2	38	38.4
Traumatic vaginal	50	15	30.0	4	8.0	31	62.0
Elective caesarean	141	45	31.9	20	14.2	76	53.9
Emergency caesarean	74	21	28.4	24	32.4	29	39.2
Perception of care received during labour and delivery							
Excellent	215	94	43.7	38	17.7	83	38.6
Good	107	28	26.2	21	19.6	58	54.2
Unremarkable	21	4	19.1	2	9.5	15	71.4
Poor	22	1	4.5	2	9.1	19	86.4
Help and support received from baby's father							
Yes, most of the time	257	112	43.6	42	16.3	103	40.1
Not as often as needed	81	13	16.0	16	19.8	52	64.2
Hardly any	27	2	7.4	5	18.5	20	74.1
Help and support received from family							
Yes, most of the time	231	108	46.8	37	16.0	86	37.2
Not as often as needed	87	15	17.2	18	20.7	54	62.1
Hardly any	47	4	8.5	8	17.0	35	74.5
Help and support received from friends							
Yes, most of the time	129	73	56.6	20	15.5	36	27.9
Not as often as needed	75	19	25.3	18	24.0	38	50.7
Hardly any	161	35	21.7	25	15.5	101	62.7

Variable	**Total**	**Normal Adjustment (≤59)** n	%	**Symptoms of PPD Present (60-79)** n	%	**Major PPD (≥80)** n	%
Diagnosed with antenatal depression during recent pregnancy							
Yes	11	0	0.0	0	0.0	11	100.0
No	354	127	35.9	63	17.8	164	46.3
Postpartum blues after recent pregnancy							
Yes	256	51	19.9	49	19.1	156	60.9
No	109	76	69.7	14	12.8	19	17.4
Psychiatric history							
No history of depression	278	114	41.0	48	17.3	116	41.7
History of depression	87	13	14.9	15	17.2	59	67.8
History of PPD after previous pregnancy	24	3	12.5	3	12.5	18	75.0
History of antenatal depression during previous pregnancy	2	0	0.0	0	0.0	2	100.0
History of anxiety	30	6	20.0	5	16.7	19	63.3
History of obsessive compulsive disorder	1	0	0.0	0	0.0	1	100.0
History of eating disorders	12	1	8.3	1	8.3	10	83.3
Complicated pregnancy							
Yes	88	20	22.7	12	13.6	56	63.6
Fear of childbirth							
Yes	95	11	11.6	8	8.4	76	80.0
Difficulty conceiving							
Yes	52	16	30.8	8	15.4	28	53.8
No	312	111	35.6	54	17.3	147	47.1
Unplanned pregnancy	101	21	20.8	19	18.8	61	60.4
Planned pregnancy	264	106	40.2	44	16.7	114	43.2
Mother's feelings about expecting a baby							
Positive	269	117	43.5	48	17.8	104	38.7
Ambivalent, negative or anxious	96	10	10.4	15	15.6	71	74.0
Mother's perception of baby's temperament							
Good	242	119	49.2	42	17.4	81	33.5
Fussy, demanding, and/or difficult	123	8	6.5	21	17.1	94	76.4
Experience of specific concerns regarding baby:							
No concerns	159	87	54.7	25	15.7	47	29.6
Health problems	16	3	18.8	0	0.0	13	81.3
Colic	97	21	21.6	17	17.5	59	60.8
Sleeping concerns	93	6	6.5	14	15.1	73	78.5
Feeding concerns	81	5	6.2	11	13.6	65	80.2

Variable	Total	Normal Adjustment (≤59)		Symptoms of PPD Present (60-79)		Major PPD (≥80)	
		n	%	n	%	n	%
Allergies	15	3	20.0	6	40.0	6	40.0
Prematurity	39	8	20.5	8	20.5	23	59.0
Financial concerns							
Yes	216	56	25.9	40	18.5	120	55.6
No	149	71	47.7	23	15.4	55	36.9
Marital problems							
Yes	54	9	16.7	9	16.7	36	66.7
No	311	118	37.9	54	17.4	139	44.7

Multiple regression analysis with a stepwise selection method was employed to determine the variables that were statistically significant predictors of a positive screen for major PPD across the total sample. According to the multiple regression model assumptions, the minimum sample size should be at least 50 + 8k or 104 + k (k = number of predictors). Applied to this study with 11 predictor variables, the minimum sample size should be either 50 + 88 = 138, or 104 + 11 = 115. The larger of the two is selected, that is 138 (Field, 2005, p. 173). This number is smaller than the sample size in this study (N = 365). The sample size is therefore adequate.

The Durbin-Watson (1.947) is very close to two. This indicates that the assumption of independent residuals or errors is met (Field, 2005, p. 189). Values lower than one or larger than three are problematical (Field, 2005, pp. 170, 190).

The multiple correlation coefficient, *R* expresses the relationship between the total PDSS score and the set of predictor variables, which were selected based on the literature of risk factors for PPD. R^2 shows the proportion of variance in the positive screen for PPD which is accounted for, or explained by, the set of predictor variables (history of depression, etc). In other words, R^2 is an indication of how well the extent of PPD can be predicted when the predictor variables are known. According to Foster et al. (2006), R^2 is the most powerful indicator of how effective the prediction is. The Adjusted R^2 is calculated because R^2 is inclined to over-estimate the success of the prediction. Ideally the Adjusted R^2 should be the same or very close to the value of R^2 (Brace, Kemp, & Snelgar, 2009; Field, 2005). The Adjusted R^2 takes the number of predictor variables as well as the number of participants into account and is therefore a more accurate measure of the effectiveness of the prediction (Brace et al., 2009).

Table 53 presents the model summary. Stepwise regression analysis provided a model which indicates a very strong relationship between the predictor variables and a PDSS score (R = 72, R^2 = 0.52, Adjusted R^2 = 0.51). The model accounts for 50.8% of the overlap in variance between the variables. (Field, 2005, pp. 188-189) Table 54 presents the analysis of variance (ANOVA). The model is highly significant at $p \leq 0.001$ ($F_{(11,346)} = 35.53$; Field, 2005, p. 189).

Table 53 Model Summary of the Dependent Variable (PDSS score)

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
11	.07*	0.52	0.51	25.24

* Predictors: (Constant), Presence of postpartum blues, Felt negative or ambivalent about expecting this baby, Infant temperament, Psychiatric history, Fearful of birth, No father support, Infant's health problems Antenatal depression, No friend support, Difficulty falling pregnant, Life stress

Table 54 Analysis of Variance of the Dependent Variable (PDSS score)

Model		Sum of Squares	*df*	Mean Square	*F*	Sig.
11	Regression	242047.85	11	22004.35	34.53	.000*
	Residual	220497.15	346	637.28		
	Total	462545.01	357			

* Predictors: (Constant), Presence of postpartum blues, Felt negative or ambivalent about expecting this baby, Infant temperament, Psychiatric history, Fearful of birth, No father support, Infant's health problems Antenatal depression, No friend support, Difficulty falling pregnant, Life stress

The following variables were entered in the stepwise multiple regression: (a) baby's health problems; (b) infant temperament; (c) felt negative or ambivalent about expecting this baby; (d) rating of care received during labour and delivery; (e) traumatic birth experience; (f) fearful of birth; (g) premature baby; (h) complicated pregnancy; (i) difficulty conceiving; (j) unplanned pregnancy; (k) postpartum blues; (l) psychiatric history; (m) antenatal depression in recent pregnancy; (n) single marital status; (o) lack of support from baby's father; (p) lack of support from friends; (q) lack of support from family; and (r) life stress. Using the stepwise method, 11 significant predictor variables emerged:

$$\text{PDSS score} = c + m_1(x_1) + m_2(x_2) + m_3(x_3)...m_{11}(x_{11})$$

Table 55 presents the raw score (B) values of the predictor variables along with values for Beta (β), t, and the significance (p) for each of the predictors as provided by SPSS. β is the standardized regression coefficient. Its value is an indication of how strongly each predictor variable influences the criterion variable – in this case, the PDSS score. Larger β values have a greater influence on the PDSS score. The β value allows the predictor variables to be directly compared so that it can be seen which variables carry more weight in establishing the dependent variable, the PDSS score. Results indicate that postpartum blues (β = .24) and feeling ambivalent or negative towards the baby (β = .21) have the greatest influence on the PDSS score. Difficulty conceiving (β = .08), life stress (β = .09), and lack of support from friends (β = .09), although significant, have less impact.

Examination of the raw scores indicates that a diagnosis of antenatal depression during a recent pregnancy increases the predicted raw score by 24.67. Having postpartum blues increases the predicted raw score by 18.84. Both antenatal depression as well as postpartum blues increases the predicted score considerably by 43.51. Life stress is a significant predictor variable that has the smallest impact on the predicted score (it adds only 1.26 points). The significance of the contribution of each predictor variable to the model is also shown in the Table 55.

Table 55 Multiple Regression Analysis of the Association between Demographic and Obstetric Variables and Scores on the PDSS (N = 365)

Variable	Coefficients[a]						
	Unstandardized Coefficients		Standardized Coefficients	*t*	Sig.	Collinearity Statistics	
	B	*SE B*	*β*			Tolerance	VIF
(Constant)	34.85	3.47		10.05	0.000		
Presence of postpartum blues	18.84	3.20	.24	5.90	0.000	0.83	1.20
Felt negative or ambivalent about expecting this baby	16.84	3.45	.21	4.88	0.000	0.77	1.30
Infant temperament	10.61	3.48	.14	3.08	0.002	0.67	1.49
Psychiatric history	12.23	3.03	.16	4.04	0.000	0.87	1.15
Fearful of birth	12.49	3.28	.15	3.80	0.000	0.85	1.17
No father support	8.56	3.17	.11	2.71	0.007	0.84	1.19
Infant's health problems	8.36	3.14	.12	2.67	0.008	0.73	1.37
Antenatal depression	24.67	8.22	.11	3.00	0.003	0.97	1.03
No friend support	6.90	3.02	.09	2.29	0.023	0.86	1.17
Difficulty conceiving	8.33	3.92	.08	2.13	0.034	0.95	1.06
Life stress	1.26	0.61	.09	2.07	0.039	0.80	1.24

a. Dependent Variable: PDSS score

The following predictor variables were dropped from the model in the stepwise analysis as they did not significantly strengthen the model: single marital status, traumatic birth experience, rating of care received during labour and delivery, lack of support from family, unplanned pregnancy, complicated pregnancy, and having a baby born prematurely.

An assumption of regression analysis is that no multicollinearity is present in the data (Field, 2005, p. 196). SPSS 19 provides an indication of the presence of collinearity in the data by means of the variance inflation factor (VIF) and tolerance statistics.

The largest VIF should be less than 10 and the average VIF for all predictor variables should not be considerably greater than one as this may indicate that the regression is biased (Myers, and Bowerman & O'Connell as cited in Field, 2005, p. 196). The collinearity statistics in Table 55 shows the data meets this requirement. The largest VIF is well below ten (1.494). Furthermore, the average VIF is close to one (1.216) which means that the regression is not biased. Tolerance statistics below 0.1 are problematic, while those below 0.2 are potentially problematic (Menard as cited in Field, 2005, p. 196). The tolerance statistics (Table 55) for all the predictor variables in this study are well above 0.2. The VIF and tolerance statistics therefore indicate that no multicollinearity exists in the dataset.

Examination of the variance proportions may also be used to detect collinearity. Variance proportions should be spread equally across the dimensions. Furthermore, each dimension should have a unique high variance proportion (Field, 2005, pp. 196-197). Variance proportions are presented in the collinearity diagnostics table (Table 56) below. Dimension 3 shows a high variance proportion (72%) with antenatal depression and not with any other predictor variables. This suggests that antenatal depression does not correlate or overlap in variance with other predictor variables. "Infant's health problems" has most of its variance loading onto dimension 9 (62%) and does not overlap in variance with other predictor variables. A number of the other predictor variables have the majority of their variance distributed fairly equally onto two dimensions (e.g. "Life stress", "Felt negative or ambivalent about expecting this baby", "infant temperament", "psychiatric history", and "fearful of birth"). The majority of predictor variables, however, have unique and relatively high variance on unique dimensions. Given that the sample size is not very big and that the statistics above indicate no-multicollinearity, these overlapping variances are not overly problematic.

Table 56 Collinearity Diagnostics of the PDSS Scores

				Variance Proportions											
Model	Dimension	Eigenvalue	Condition Index	(Constant)	Postpartum blues	Negative or Ambivalent About expecting baby	Infant Temp*	Psychiatric History	Fearful of birth	Lack of support from father	Infant's health problems	Antenatal depression	Lack of support from friends	Difficulty Conceiving	Life stress
11	1	6.33	1.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00
	2	1.03	2.48	0.00	0.01	0.06	0.01	0.00	0.00	0.05	0.00	0.19	0.00	0.42	0.00
	3	0.95	2.58	0.00	0.00	0.05	0.00	0.02	0.00	0.02	0.00	0.72	0.00	0.08	0.00
	4	0.71	2.99	0.00	0.00	0.16	0.11	0.06	0.03	0.16	0.04	0.04	0.00	0.20	0.01
	5	0.66	3.09	0.01	0.00	0.07	0.04	0.19	0.40	0.06	0.00	0.01	0.01	0.08	0.01
	6	0.59	3.29	0.00	0.00	0.00	0.11	0.31	0.39	0.02	0.05	0.01	0.03	0.00	0.00
	7	0.47	3.67	0.01	0.01	0.26	0.09	0.01	0.01	0.64	0.01	0.00	0.00	0.06	0.02
	8	0.43	3.82	0.03	0.03	0.26	0.15	0.31	0.12	0.00	0.01	0.02	0.07	0.12	0.01
	9	0.28	4.77	0.00	0.01	0.00	0.19	0.03	0.00	0.01	0.62	0.00	0.31	0.02	0.03
	10	0.24	5.09	0.01	0.26	0.01	0.19	0.05	0.02	0.00	0.19	0.00	0.47	0.01	0.06
	11	0.21	5.55	0.02	0.52	0.10	0.09	0.00	0.01	0.03	0.06	0.00	0.00	0.00	0.41
	12	0.11	7.74	0.92	0.15	0.03	0.02	0.01	0.00	0.00	0.02	0.00	0.10	0.00	0.45

* Infant temperament

The residual statistics for extreme cases should be examined. For a fairly accurate model 95% of cases should have standardized residuals within ±2, and 99% of all cases should have standardized residuals within ±2.5. Only 1% of cases should like outside of these limits (Field, 2005, p. 199). Results from this dataset, reported in Table 57 below, indicate that only three observations were indicated as outliers with the casewise diagnostics. Three outliers in a sample of 358 (7 cases were excluded due to missing values) is merely 0.8 %, which is excellent. The sample therefore conforms to what is expected for a fairly accurate model. The three outliers (case 100, 142, and 179) have standardized residuals greater than three and should be investigated further.

The influence statistics for all the selected cases is shown in Table 58. None of the outliers have Cook's distance larger than one. This means that they do not influence or bias the regression model (Field, 2005, p. 200). The average leverage may be calculated as (k+1)/n or (12/358) = 0.03 and the recommended threshold should ideally be no bigger than three times this value (i.e. 0.09). All three outliers are well within this limit. The Mahalanobis distance is lower than the recommended threshold of 23 (in small sample sizes of 200 cases with five predictors, and a threshold of 25 in samples of 500 cases with five predictors; Field, 2005, p. 202; Stevens, 1984) for all three outliers. None of the cases exceeded this criterion. Case 179 has the largest Mahalanobis distance (10.99). These results indicate that it is unlikely that there were influential cases in the data.

Table 57 Casewise Diagnostics of the PDSS Score

Case Number	Standardized Residual	PDSS Score	Predicted Value	Residual
100	4.10	149	45.53	103.47
142	3.38	124	38.64	85.36
179	3.03	171	94.64	76.36

Table 58 Case Summaries

		Unstandardized Predicted Value	Mahalanobis Distance	Cook's Distance	Centered Leverage Value
100		45.53	5.27	0.03	0.02
142		38.64	4.19	0.01	0.01
179		94.64	10.99	0.03	0.03
Total	N	3	3	3	3

The histogram in Figure 15 (Appendix F) shows that the residuals are reasonably normally distributed as they should be (Field, 2005, p. 204). The normal distribution of residuals is confirmed by the straight line in the plot in Figure 16 (Appendix F). No deviation from normality is evident.

Some heteroscedacity is evident in the scatterplot of the residuals of the outcome variable and each PPD predictor variable when both variables are regressed separately on the remaining predictors (Figure 17 in Appendix F). The points should be random but a slight pattern that funnels out is apparent which indicates increasing variance across the residuals (Field, 2005, p. 203). Outliers on this plot represent cases that may have impacted excessively on a predictor's regression coefficient.

Two of the predictor variables in the multiple regression analysis were subjected to further analysis. These were life stressors and psychiatric history. Each of these predictor variables were composed of multiple characteristics items. Point biserial correlations (r_{pb}) were used to determine if certain life stressors and a history of specific psychiatric illnesses were more significantly associated with a high score on the PDSS. The point biserial correlation coefficient provides a measure of the association between a dichotomous variable and a continuous variable, such as the scores on a test (Ferguson, 1981). It is mathematically equivalent to the Pearson product-moment correlation (r), although the Pearson product-moment correlation can only be used when both variables are non-dichotomous. A p-value of ≤ 0.05 was used to indicate statistically significant results even though a less conservative alpha of $p < 0.15$ is commonly recommended in the literature for predictive models (as opposed to explanatory models; Bloch & Klein, 2005). The life stress variables and psychiatric history variables that were correlated with the total PDSS score are presented in Table 59. Point biserial correlations revealed that eight life stress variables were significantly associated with high PDSS scores, namely moving house, job loss of the mother's partner, change of jobs of the mother's partner, financial concerns, another pregnancy or birth, marriage, marital problems, and family problems. A history of depression was the only psychiatric history variable that was significantly associated with a high PDSS score indicative of major PPD in this study.

Table 59 Point Biserial Correlations of Psychiatric History and Life Stress Variables with Total PDSS Scores (N = 365)

Variables	r_{pb}		Sig	n
Psychiatric history				
Postpartum depression after a previous pregnancy	.100		0.057	365
Antenatal depression during a previous pregnancy	.015		0.769	365
Depression	.300	**	0.000	365
Anxiety	.087		0.096	365
Obsessive compulsive disorder	.036		0.487	365
Anorexia	.068		0.192	365
Bulimia	.078		0.136	365
No psychiatric history	-.338	**	0.000	365
Life stressors in past two years				
House alterations	.046		0.384	364
Moving house	.163	**	0.002	365
Moving city / migration	.071		0.179	364
Job changes: self	.079		0.134	365
Job changes: partner	.159	**	0.002	365
Job loss / retrenchment: self	.101		0.055	364
Job loss / retrenchment: partner	.178	**	0.001	365
Financial concerns	.170	**	0.001	365
Bereavement	.051		0.328	364
Loss of close friends / family relocating, emigrating, etc.	.051		0.334	364
Serious illness of a family member	-.031		0.554	365
Another pregnancy and birth	.124	*	0.018	365
Marriage	.112	*	0.033	364
Marital problems	.216	**	0.000	365
Family problems	.262	**	0.000	364
Been victimised by violence or crime	.102		0.052	365
Serious injury, illness, or personal health problems	.049		0.346	364

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

8.8 Results of the Comparison of the PDSS, the EPDS, and the QIDS-SR16

Descriptive statistics for the PDSS, the EPDS and the QIDS-SR16 were calculated and frequencies determined according to the participants' screening results at each of screening scales' recommended cut-off thresholds. Chi-square analysis was used to compare participants who scored positive for symptoms of PPD on the three measures. Pearson correlations were used to determine the relationship among the continuous scores on the screening scales.

The PDSS is intended to provide an overall score for PPD, but also considers the multidimensionality of postpartum depression and gives seven subscale scores. The summative scoring results in a total score range from 35 to 175. Participants in this study obtained scores that ranged from 35 to 173, with a mean of 82.04 (SD = 35.92). Descriptive statistics for the PDSS, the EPDS, and the QIDS-SR16 are presented in Table 60. The PDSS total score may be sorted into one of three categories: i) normal adjustment (total score of ≤59); ii) significant symptoms of PPD (total score of 60 to 79); and iii) positive screening for PPD (total score of ≥80; Beck & Gable, 2002). Beck and Gable (2001a) recommend a cut-off score of 80 for major PPD (sensitivity = 94%, specificity = 98%), and a cut-off score of 60 (sensitivity = 91%, specificity = 72%) for minor or major depression. Just over one third (n = 127; 34.8%) scored in the range classified as representing normal adjustment (score ≤59). There were 17.3% (n = 63) who obtained a score in the range classified as representing significant symptoms of PPD (score 60-79), while close to half of the participants in this study (n = 175; 47.9%) screened positive for major PPD with scores of 80 or more.

The EPDS is a 10-item self report measure with a 4-point Likert scale. Each of the 10 questions has 4 answer choices that are scored between 0 and 3. The EPDS total score, obtained by adding the scores for each item, may range from 0 to 30. Participants in this study obtained EPDS scores that ranged from 0 to 30, with a mean of 11.10 (SD = 7.20). The cut-off point of the EPDS is recommended at 12 or 13 for probable depression, and at 9 or 10 for possible depression (Cox et al 1987). Boyd et al (2005) have suggested, however, that different cut-off scores may be warranted for different cultural groups. In Beck and Gable's (2001a) comparative study of the PDSS, EPDS and BDI-II, the EPDS yielded a sensitivity of 78%, a specificity of 99% and a positive predictive value of 93% when using a cut-off score of ≥ 12. In this study 38.6% (n = 141) of the participants had scores ranging from 0 – 8 on the EPDS. A further 15.3% (n = 56) had scores ranging from 9 to 11, indicating possible depression, and 46% (n = 168) of the participants had scores ≥12, indicating probable depression.

The QIDS-C16 and the QIDS-SR16 total scores are obtained by adding scores for the nine criterion symptom domains: (1) sad mood; (2) concentration/decision-making; (3) self-outlook; (4) suicidal ideation; (5) involvement; (6) energy/fatigability ; (7) sleep (based on the highest score on any one of the four relevant items – sleep onset insomnia, mid-nocturnal insomnia, early morning insomnia, hypersomnia); (8) weight/appetite change (based on the highest score on any one of the four relevant items – weight increase, weight decrease, appetite increase, appetite decrease) ; and (9) psychomotor changes (based on the highest score on any one of the two relevant items – psychomotor slowing or psychomotor agitation). The total score ranges from 0 to 27. Participants in this study obtained QIDS-SR16 scores that ranged from 0 to 24, with a mean of 9.16 (SD = 5.34). The thresholds recommended when screening with the QIDS-SR16 are ≤5 for no depression, a score of 6 to 10 for mild depression, a score of 11 to 15 for moderate depression, a score of 16 to 20 for severe depression, and a score of ≥21 for very severe depression (Rush et al., 2003). There were 30.4% (n = 111) of participants in this sample had no depressive symptoms on the QIDS-SR16 with scores ≤5, 31.5% (n = 115) obtained scores ranging from 6 to 10, indicating mild depression, 25.5% (n = 93) obtained scores of 11 to 15 indicating moderate depression, 10.1% (n = 37) of participants were classified with severe depression with scores of 16 to 20, and a further 2.5% (n = 9) had scores indicative of very severe depression (≥21).

Table 60 Descriptive Statistics for the PDSS, EPDS, and QIDS-SR16

	N	Minimum	Maximum	Mean	Std. Deviation
QIDS-SR16	365	0	24	9.16	5.34
EPDS	365	0	30	11.10	7.20
PDSS	365	35	173	82.04	35.92

The published recommended cut-off scores for major depression for the three instruments are presented in Table 61. Based on these cut-off points, the PDSS identified 175 (47.9%) of the participants with major depression, while the EPDS identified 168 (46%), and the QIDS-SR16 identified 46 (12.6%).

Table 61 Cut-off Scores for Screening for the Diagnosis of Major Postpartum Depression for the PDSS, EPDS, and QIDS-SR16

Instrument	Cut-off score for major postpartum depression	n	Frequency
PDSS	≥ 80	175	47.9%
EPDS	≥ 12	168	46.0%
QIDS-SR16	≥ 16	46	12.6%

Cross-tabulation of the PDSS and the EPDS (Table 62) indicates that five mothers (1.4%) that were identified with major PPD by the PDSS were classified with no depression by the EPDS. Furthermore, the EPDS identified three mothers (0.8%) with probable depression that were classified as normal adjustment by the PDSS. Chi-square tests for categorical data indicate a significant correlation between these two measures at the $p < 0.05$ level (chi-square ($df = 4$) = 296.94, $p < 0.001$).

Table 62 Cross Tabulation of the Participants According to Cut-off Scores for the PDSS and EPDS

	PDSS			Total
	≤ 59	60 - 79	≥ 80	
EPDS				
No depression ≤ 8	112	24	5	141
Possible depression 9–11	12	26	18	56
Probable depression ≥ 12	3	13	152	168
Total	127	63	175	365

The cross-tabulation of the PDSS and the QIDS-SR16 (Table 63) shows that only 46 participants (12.6%) were classified by the QIDS-SR16 as presenting with severe or very severe depression, in comparison to the PDSS, which identified 47.9% (n = 175). One mother (0.3%) with a score ranging from 16 to 20 on the QIDS-SR16, indicative of severe depression, was classified by the PDSS as having minor depression. Furthermore, the PDSS identified two participants (0.6%) with major PPD and 16 participants (4.4%) with minor depression who all obtained low scores on the QIDS-SR16 suggesting that no depression was present. Chi-square analysis detected a significant correlation between the categorical data of these two measures (chi-square ($df = 8$) = 261.70, $p < 0.001$).

Table 63 Cross Tabulation of the Participants According to Cut-off Scores for the PDSS and QIDS-SR16

	PDSS			Total
	≤ 59	60 - 79	≥ 80	
QIDS-SR16				
No depression < 5	93	16	2	111
Mild 6 – 10	33	38	44	115
Moderate 11 – 15	1	8	84	93
Severe 16 – 20	0	1	36	37
Very Severe ≥ 21	0	0	9	9
Total	127	63	175	365

The EPDS identified 46% whereas the QIDS-SR16 only identified 12.6% of mothers with major depression. Cross tabulation of the EPDS and the QIDS-SR16 (Table 64) shows that the EPDS identified one participant (0.3%) at risk of probable depression and 13 participants (3.6%) at risk of possible depression who all were identified by the QIDS-

SR16 as having no depression. Comparisons of the categorical depression status of these two measures using chi-square tests indicate a significant correlation (chi-square ($df = 8$) = 251.92, $p < 0.001$).

Table 64 Cross Tabulation of the Participants According to Cut-off Scores for the EPDS and QIDS-SR16

	EPDS_Tot (Binned)			Total
	No depression ≤ 8	Possible depression 9 – 11	Probable depression ≥ 12	
QIDS-SR16				
No depression < 5	97	13	1	111
Mild 6 – 10	42	32	41	115
Moderate 11 – 15	2	11	80	93
Severe 16 – 20	0	0	37	37
Very Severe ≥ 21	0	0	9	9
Total	141	56	168	365

The Pearson product-moment correlation was used to determine how well the three instruments, the PDSS, the EPDS and the QIDS-SR16, correlate with each other. These results are reported in Table 65. The total scores obtained on the PDSS and the total scores obtained on the EPDS showed a statistically significant correlation ($r = 0.918$, N = 365, $p < 0.001$). The PDSS explains 84% of the variance in the EPDS ($r^2 = 0.84$). The correlation between the PDSS and the EPDS was slightly higher than the correlation between the PDSS and the QIDS-SR16, although both were strong. The QIDS-SR16 correlated equally well with both the EPDS and PDSS yielding a statistical significant results in both instances (r = 0.879, N = 365, $p < 0.001$). The QIDS-SR16 explains 77% of the variance in both the PDSS and the EPDS ($r^2 = 0.77$).

Table 65 Pearson Correlations between the Total Scores of the PDSS, EPDS, and QIDS-SR16 (N=365)

Scale	PDSS			EPDS		
	r	r^2	p	r	r^2	Sig
EPDS	0.918**	0.84	0.000	-	-	-
QIDS-SR16	0.879**	0.77	0.000	0.879**	0.77	0.000

** Correlation is significant at the 0.01 level (2-tailed).

In previous studies, Beck and Gable (2001a) examined the convergent validity of the PDSS. Correlations were calculated among the totals scores of the PDSS, the EPDS, the BDI-II, and diagnostic status as determined by the Structured Clinical Interview for DSM-IV Axis 1 Disorders (SCID). Their results indicated that the PDSS correlated strongly with these self-report depression measures as well as with the clinical interview. The PDSS' correlation with the EPDS was 0.79 (N = 150; $p < 0.001$). The correlation between these two screening measures in this study were very strong (r = 0.918; $p < 0.001$; N = 365).

8.9 Discussion

8.9.1 Discussion of Rasch analysis.

Results reveal excellent reliability for both the PDSS (person reliability estimate = .95, Cronbach α = .98) and the Afrikaans PDSS (person reliability estimate = .95, Cronbach α = .98). Person reliability estimates for the PDSS dimensions were very good and ranged from .71 to .85. The SLP dimension had the lowest person reliability estimate (.71) and person separation index (1.56) – the only dimension in the PDSS with a person separation index below 2.00. Person reliability estimates for the Afrikaans PDSS dimensions were generally good, ranging from .64 to .84. The Afrikaans SLP and Afrikaans ANX dimensions yielded the lowest person reliability estimates (.64 and .74 respectively). These two dimensions were also the only two dimensions in the Afrikaans PDSS with a person separation index below 2.00. (Afr SLP 1.34; Afr ANX 1.69).

Rasch analysis was performed with the PDSS and Afrikaans PDSS to evaluate how well the items contributed to underlying construct of PPD. The same analysis was also performed with the scales' dimensions. Average fit statistics for the PDSS and the Afrikaans PDSS as a whole were good with infit and outfit MNSQ statistics near 1.00.

Items 1, 8, 15, 22, and 29 in the English PDSS, and items 1, 15, 29, and 30 in the Afrikaans PDSS had infit MNSQ statistics greater than 1.40. This may indicate that these items did not fit the definition of the constructs they are measuring very well (thus forming another construct(s)). All the misfit items from the English PDSS total and three of the four misfit items from the Afrikaans PDSS total are from the Sleeping/Eating dimension. This may be a reflection that they form a separate dimension. No items were overfitted (i.e. < 0.60).

The majority of items in the PDSS dimensions as well as in the Afrikaans PDSS dimensions demonstrated fit statistics that supported the underlying constructs of each dimension. An analysis of the PDSS dimensions revealed that one of the 35 items (Item 28) had an infit and outfit MNSQ statistic beyond the acceptable range of 0.60 to 1.40 (Bond & Fox, 2007; Wright & Linacre, 1994), Item 34 had an outfit MNSQ statistic beyond the acceptable range, and Item 2 had a borderline outfit MNSQ statistic of 1.40. Two items from the Afrikaans PDSS demonstrated misfit for both infit and outfit MNSQ statistics. A summary of the misfit items are presented in Table 66 below.

Table 66 Infit and Outfit MNSQ Statistic for Misfit Items in the PDSS and Afrikaans PDSS Dimensions

Scale	Dimension	Item	Content	Infit MNSQ	Outfit MNSQ
English PDSS					
	ANX[c]	2	I got anxious over even the littlest things that concerned my baby		1.40
	SUI[a]	28	I felt that my baby would be better off without me.	1.85	1.66
	GLT[b]	34	I felt like a failure as a mother.		0.54
Afrikaans PDSS					
	SUI[a]	28	Ek het gevoel dat dit vir my baba beter sou wees sonder my.	1.48	1.45
	ANX[c]	30	Ek het gevoel asof ek heeltyd aan die gang moes bly.	1.59	1.54

[a] Suicidal Thoughts Dimension
[b] Guilt/Shame Dimension
[c] Anxiety/Insecurity Dimension

These items demonstrate poor fit to the Rasch model with their observed responses departing considerably from their expected responses. Item misfit occurs for any number of reasons, such as unclear or ambiguous items, items that are not closely related to the overall construct, items that load on another construct, or it may indicate item redundancy. Item 28 and Item 30, in particular, appear to be problematic items. They will be discussed in more detail in the next section.

Item person construct maps showing the positions of persons and items on the PDSS and Afrikaans PDSS were computed. The spread of the items on both questionnaires was fairly good, but there were still persons that scored higher than the items could measure and an overrepresentation of items at the mean level.

Item difficulty estimates indicated that suicidal thought symptoms were more difficult to endorse in both the English and Afrikaans samples. Item 3 (I felt like my emotions were on a roller coaster), Item 9 (I felt really overwhelmed), and Item 24 (I have been very irritable) were the more easily endorsed items from both samples.

The Rasch error estimates for the items in the PDSS as a whole were small with values less than 0.12 and a mean of 0.90. The Rasch error estimates on the PDSS dimensions were also small with the SUI dimension demonstrating the highest estimates and a mean of 0.15. The remaining PDSS dimensions had mean error estimates that ranged from 0.09 to 0.13 All Rasch error estimates for the Afrikaans PDSS items, as a whole, were also small with values less than 0.12 and a mean of 0.90. The Afrikaans PDSS dimensions revealed small error estimates, also with the higher estimates in the SUI dimension with a mean of 0.17. This suggests that the SUI dimensions in both samples had more haphazard responses than other dimensions.

The data was also examined to evaluate the effectiveness of the Likert response categories as this impacts on how well the response data defines the dimension. Except for item 29 on the PDSS, the average measure (in logits) for each item's response option in both the PDSS and the Afrikaans PDSS does increase with each higher response option, starting with a high negative, and increasing to a positive value. On the 5-point Likert scale, a higher response options therefore does correspond to a higher level of agreement with the item and "more" of the construct measured by the dimension.

The PDSS and the Afrikaans PDSS were compared to examine differential item functioning – i.e. if the items have significantly different meanings across the two samples.

8.9.2 Discussion of problematic items and items with differential item functioning.

Bond and Fox (2007) recommended that items which show DIF ought to be investigated thoroughly to determine what can be inferred about the underlying construct. Although statistical analyses are helpful to detect problematic items with DIF, they do not reveal the causes of item bias. The specific causes of cross-language DIF items that were identified statistically cannot be determined in this study. However, some potential sources of DIF are discussed below.

The more homogeneous the groups are the more accurate DIF detection is (Allalouf & Sireci, 1998). Pearson Chi-square statistics were used to determine if significant differences were present between the characteristics of the English and Afrikaans samples. The two samples were similar across most characteristics, but significant differences were noted for the following: number of weeks since birth ($p = 0.008$), support from father ($p = 0.006$), support from family ($p = 0.007$), gave birth prematurely ($p \leq 0.001$), and infant feeding method ($p = 0.045$), as well as for the following life stressors: moving house ($p = 0.004$), moving city or migrating ($p = 0.021$), mother changed jobs ($p = 0.009$), partner changed jobs ($p = 0.009$), bereavement ($p = 0.017$), and been victimised by violence or crime ($p = 0.024$). These significant differences make it difficult to determine whether DIF was due to differences in these sample characteristics, or whether bias could be attributed to translation or language issues. The presence of DIF in items that did not have misfit in the Rasch analysis may be a reflection of differences in the English and Afrikaans samples.

In Chapter 7 it was pointed out that DIF may have many explanations and be due to several factors, including differences in the item's meaning or item content due to an inaccurate translation or a word having more than one meaning in the target language, differences in the language, wording or format of items, differences in words or expressions which create problems in the interpretation of constructs due to cultural relevance, and so forth. According to Teresi (2006), there are a number of other factors that have received less attention in the literature that also influence the detection of DIF. These include model assumptions, model fit, the distribution of latent variables, sample size, and the length of the test or measuring instrument.

The Rasch measurement model, like most IRT models, assumes that the underlying trait being measured is unidimensional. A contentious issue is whether DIF is merely a reflection of multidimensionality or not. Roussos and Stout (as cited in Teresi, 2006, p. S154) suggest that the presence of multidimensionality is the general cause of DIF – that DIF items measure one or more dimensions apart from the primary dimension. It is important to examine the unidimensionality assumption of the model because multidimensionality can be mistaken for DIF (Teresi, 2006). A requirement of DIF analyses is that the two language versions demonstrate equivalence in their dimensionality structure. The results of this study indicate that the original PDSS and the Afrikaans translation of the PDSS demonstrated adequate equivalence in their dimensional structure through Rasch analysis. This indicates that the same psychological construct was measured for the seven PDSS content scales across both language groups.

The translation of an instrument is one of the critical factors that may contribute to measurement bias (Ramirez et al., 2006). Brislin's back-translation method together with the committee approach was selected for use in this study in an effort to improve the linguistic equivalence of translation of she PDSS. Despite efforts to arrive at a translation as close as possible to the original PDSS, a number of items were identified as showing DIF. The content of these items need to be examined to determine possible reasons for DIF across the two language groups.

Items with large DIF values, with a z-value beyond 1.96, indicate more problematic DIF. Items with borderline DIF values could be due to measurement error or sample idiosyncrasies. Items that did not present with significantly large DIF values in the analysis of the total PDSS and total Afrikaans PDSS could be as a result of multidimensionality.

Items that presented with fit problems and with large DIF in the total PDSS and the total Afrikaans PDSS were Item 30 (z = -3.93), Item 25 (z = 3.14), and Item 9 (z = -3.12). The performance of these items in the content scales was examined.

Item 30 (Ek het gevoel asof ek heeltyd aan die gang moes bly) presented with DIF (z = -3.79) as well as fit problems in the Afrikaans PDSS ANX content scale (infit MNSQ = 1.59; outfit MNSQ = 1.54). Furthermore, two participants who completed the Afrikaans PDSS marked item 30 as difficult to understand while three English participants marked this item as difficult to understand on the English PDSS. However, Item 30 only presented with fit problems in the Afrikaans PDSS ANX content scale, not in the English language version. This may indicate that the Afrikaans translation was not adequate, that Afrikaans respondents were not familiar with the item content, or that the item's content is not appropriate for this Afrikaans sample.

The researcher noticed when assessing some women in person that some English participants had read the word 'pacing' in item 30 as 'packing', and then interpreted 'moving' as relocating. This is likely due to poor reading skills in women who do not have English as a home language. It is uncertain how many women who participated online also misread this item. The terminology in this item may be more familiar to some participants than to others. Both the English and Afrikaans versions of this item should be revised so that an alternative may be found that demonstrates better fit to the Rasch model and with no DIF.

Item 25 (I had a difficult time making even a simple decision; Ek het dit moeilik gevind om die eenvoudigste besluit te neem) did not present with fit problems in either the

Afrikaans or the English PDSS content scales. It did, however, present with DIF in the PDSS MNT content scale (z = 2.00), although the DIF value was relatively small and could be due to measurement error or sample idiosyncrasies. Item 25 was not marked as an item that was difficult to fully understand. The performance of this item may need to be monitored in future studies.

Item 9 (I felt really overwhelmed; Ek het heeltemal oorweldig gevoel), which presented with DIF in the analysis of the total scale (z = -3.12) did not present with DIF or with fit problems in the ANX content scale (z = 0.85). This suggests that no translation problems are evident in this item and it fits the construct of the ANX content scale well. It was, however, marked as difficult to understand by three English participants and one Afrikaans participant. This item may be misunderstood by participants who are not proficient in either English or Afrikaans of these languages. Closer inspection of the item's Afrikaans translation (Ek het heeltemal oorweldig gevoel) reveals that the translated version indicates greater severity with the word "heeltemal". The use of this word implies "I felt completely overwhelmed" rather than "really overwhelmed". This changes the meaning of the item slightly and it may need to be revised.

Item 34 (I felt like a failure as a mother; Ek het gevoel asof ek as ma misluk) presented with borderline DIF in the total screening scale (z = -2.24) as well as borderline DIF in the GLT content scale (z = -2.40). Item 34 also had a low outfit MNSQ statistic in the English PDSS (0.54). Aberrant infit scores are generally a greater cause of concern than aberrant outfit scores (Bond & Fox, 2001). Outfit statistics are not weighted and are more sensitive to the influence of outlying scores. Nevertheless, some DIF together with some fit problems means that the Afrikaans version of this item may need to be monitored.

Relative bias may potentially be a cause for DIF in Item 34 (I felt like a failure as a mother). Relative bias has been identified as a possible source of DIF which occurs when a participant rates herself relative to others in the setting. An item may, for instance, require the respondent to rate herself in comparison to an imagined peer group. This type of item is therefore dependent on the respondent's frame of reference (Teresi, 2006). Item 34 may, to a certain extent, cause the mother to rate herself according to what she regards as failure.

Some items, which did not present with DIF in the analysis of the total scale, did present with DIF in the analysis of the content scales. These were Item 2 (z = 4.09; I got anxious over even the littlest things that concerned my baby; Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak), Item 24 (z = 2.19; I have been very irritable; Ek was baie geïrriteerd), and Item 32 (z = -2.06; I had difficulty focusing on a task; Ek het gesukkel om op 'n taak te konsentreer).

Of the items with DIF in the content scales, Item 24, Item 25, Item 32 and Item 34 presented with borderline DIF that did not seem highly significant, but should nevertheless be monitored in future studies. Only Item 2 and Item 30 had large DIF values in the content scales. Item 30 was discussed above. Item 2 (I got anxious over even the littlest things that concerned my baby; Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak) had a large DIF value (z = 4.09). Item 2 also presented with borderline fit problems in the English PDSS ANX content scale (outfit MNSQ = 1.40). Furthermore, seven participants (five English participants and two Afrikaans participants) indicated that they had difficulty fully understanding this item. DIF, fit results, and taking into account that this item was flagged as difficult to understand by some participants, particularly

English participants, indicates that the English version of this item was not well understood by the English participants of this sample.

No DIF was present for items from the SLP and LOS content scales. The SLP content scale is composed of three items which measure disruptions in normal sleeping habits (items 1, 15, and 22) and two items that measure disruptions in normal eating habits (items 8 and 29). All three items which measure sleep disruptions showed borderline DIF in the total PDSS. However, in the dimension analysis, not one of these three items showed DIF. Furthermore, all the items from the Sleeping/Eating content scale presented with good fit statistics within the content scale, supporting construct validity for the Sleeping/Eating content scale. Poor fit of items from this content scale in the analysis of the total PDSS and total Afrikaans PDSS may simply suggest that these items form a different construct.

Item 23 had borderline DIF in the analysis of the total scale ($z = -2.24$), which does not seem significant, especially considering that no DIF was evident for this item in the ANX content scale. Nevertheless, it may be argued that item 23 (I felt all alone) is slightly stronger in meaning than its Afrikaans translation (Ek het alleen gevoel) due to the word "all" in the original. This item did not present with fit problems and was not flagged as difficult to understand.

When Rasch analysis was performed with each respective content scale, item fit MNSQ statistics supported the measurement of a unidimensional construct in each content scale with the exception of two items, which had high MNSQ fit statistics, suggesting a lack of construct homogeneity. One item was Item 28 (I felt that my baby would be better off without me; Ek het gevoel dat dit vir my baba beter sou wees sonder my) in both the

English PDSS and the Afrikaans PDSS, and the other was Item 30, which was discussed earlier. Unlike Item 30, Item 28 did not present with DIF and was not indicated as an item that was difficult to understand. Poor fit of Item 28 may be an indication that it was consistently misunderstood by both English and Afrikaans respondents, but considering that the item demonstrated poor fit in both languages, it is more likely that it did not fit the construct of the SUI content scale very well. Pearson's correlation of the items with the PDSS content scales (Table 86 in Appendix F) shows that item 28 does not correlate better with another dimension in the PDSS. Item 28 correlates best with the dimension it purports to measure – the SUI content scale (r = .850; p < 001; N = 365). The language of this item may therefore need to be revised even though the language and sentence construction in both English and Afrikaans do not seem to indicate ambiguity. Alternatively, an additional equivalent item can be added to the screening scale and its performance, along with the original Item 28, can be determined in future studies with a wider sample. The additional item can be calibrated along with the other items and, if the additional item demonstrates better psychometric properties in a South African population, it may be considered a suitable alternative to replace the original Item 28.

The Afrikaans version of Item 31 (Ek het baie kwaad gevoel en was gereed om te ontplof) is only slightly different to the original (I felt full of anger ready to explode) due to the words "baie kwaad". This is likely to be translated back into English as "very angry" rather than "full of anger". In the translation process, two alternatives for this item were arrived at. The other alternative was "Ek was woedend en gereed om te ontplof". Future studies may consider substituting the items to see which performs better.

Angoff and Cook (as cited in Allalouf, 2003, p. 56) state that an item with less text (i.e. a shorter item) is more likely to have translation DIF. Furthermore, items with more text tended to retain their meaning and their psychometric characteristics. Allalouf (2003, p. 56) states that subsequent researchers have come to the same conclusion. All the PDSS items consist of relatively short statements, some slightly shorter than others. The length of the statements did not appear to impact on DIF.

8.9.3 Discussion of the risk factors for major PPD in this study.

A high score on the PDSS does not in itself confirm a depressive illness as it is screening instrument and not a diagnostic instrument. The PDSS has, however proved to be a reliable and valid screening instrument for the detection of PPD (Beck & Gable, 2002). It is therefore reasonable to assume that the risk factors (predictor variables) identified as significant in this study are important in the development of PPD.

The PDSS scores of almost two thirds (65%) of mothers in this study exceeded 59, indicating the presence of significant symptoms of PPD or a positive screen for major PPD. The prevalence of mothers who screened positively for major PPD between 4 and 16 weeks postpartum was 48%. A further 17% of mothers presented with symptoms that indicate a potential risk for PPD. This rate is not unexpected given that many mothers were recruited from antenatal and postnatal support groups, from magazine articles about postpartum depression, and from health practitioners who suspected that the mother may have PPD.

Statistically significant variables associated with major PPD in this study were a history of psychiatric illness – depression in particular, antenatal depression in recent pregnancy, postpartum blues, lack of support from the baby's father, lack of support from friends, life stress, infant temperament, difficulty conceiving, feeling negative or ambivalent about expecting this baby, fearful of childbirth, and concern about health related issues regarding the infant, like colic, sleeping and feeding problems, and allergies. Although multiple regression analysis did not reveal a statistically significant relationship between a previous diagnosis of PPD, mothers were slightly more likely to have a positive score of major PPD if they had previously been diagnosed with PPD. Furthermore, the incidence of major PPD was greater in mothers who reported greater dissatisfaction with the care they received during labour and delivery. This variable was, however, not statistically significant when multiple regression analysis was employed. Mothers presenting with these variables should be closely monitored by their health practitioners as they have an increased risk of developing PPD.

The following factors were not found to be associated with major PPD: marital status, gestational age of infant at birth, method of delivery, support from family, unplanned pregnancy, and complicated pregnancy.

Women with a previous history of depression were more likely to screen positive for major PPD. The incidence rate for major PPD in mothers who reported a past history of depression was 67.8% compared to 41.7% in mothers with no history of depression. This result replicates findings from numerous studies which indicated that a history of depression is a strong and significant risk factor for PPD. An antenatal history of anxiety disorders also slightly increased the likelihood that mothers may develop PPD, although no statistically significant relationship was noted. A history of psychiatric illness prior to becoming pregnant has also been associated with PPD, significantly increasing a woman's risk twofold (Forman et al., 2000).

Eleven mothers in this study (3%) indicated that they had been diagnosed with antenatal depression during their recent pregnancy. All these mothers screened positive for major PPD. The finding that antenatal depression is a risk factor for PPD is consistent with findings from other studies (e.g. Forman et al., 2000).

The significant relationship found between postpartum blues and PPD is consistent with findings from other studies. Postpartum blues is more prevalent than PPD. Results from this study are consistent with the literature that postpartum blues affects up to 70% percent of postpartum women. All mothers who experience postpartum blues will not necessarily develop PPD. The incidence of major PPD in this study was 60.9% in mothers who had postpartum blues PPD compared to 17.4% in mothers who reported not having had postpartum blues in their recent pregnancy.

Mothers who reported feeling ambivalent, negative or anxious about expecting a baby were significantly more likely to present with major PPD (74%) than those mothers who felt positive about expecting a baby (38.7%). Mothers whose recent pregnancy was unplanned were slightly more likely to present with major PPD (60.4%) than mothers whose pregnancy was planned.

Mothers who described their infants as demanding, fussy or difficult accounted for 32.1% of the sample. Infants with a difficult or irritable temperament have been implicated as a factor that contributes to maternal depression. Results from this study also indicate a significant relationship between these infant temperament characteristics and major PPD.

The incidence of major PPD in mothers who described their infants as demanding, fussy or difficulty was 76.4% compared to a 33.5% incidence of major PPD in mothers who did not report these infant characteristics.

Results indicate a significant relationship between major PPD and mothers' reports of infant health concerns, such as concerns with feeding and sleeping, colic, reflux and infant illness. Maternal reports of depression have been associated with infant sleep problems. A quarter (25.5%) of the mothers in this study indicated that their infants were sleeping poorly. More than three quarters (78.5%) of the mothers who screened positive for major PPD reported that their infants were sleeping poorly. Maternal sleep quality may act as an important mediator in the relationship between depression and infant sleep problems. It is therefore important to ensure that mothers who present with PPD and who report to be sleeping poorly themselves, receive assistance in teaching their infants to settle independently.

Infantile colic is a common problem of early infancy and has been reported to be associated with early postpartum depressive symptoms (Akman et al., 2006; Howell et al., 2006). More than a quarter of the mothers in this study (26.6%) reported that their infants suffered from colic. The incidence of major PPD in these mothers was 60.8%.

Surprisingly, the incidence of major PPD in mothers who reported concern about their infants' health and feeding problems was even higher at 81.3% and 80.2% respectively. A participation requirement was that mothers gave birth to a healthy baby without a disability. It may therefore be reasonable to assume that the health concerns the mothers had about their infants were not major health issues. This was, however, not determined. Anxiety (Beck, 1992, 1993) and a negative cognitive attributional style, when assessed through self-report, is strongly related to high levels of PPD symptoms (O'Hara & Swain, 1996). These variables were not explored in this study but have led the researcher to wonder whether they have an impact on mothers who present with major PPD and express concern regarding their infants' feeding, appropriate weight gain, and health. This may be explored in future studies.

Fear of childbirth is not uncommon in pregnant women. In this study 26% of mothers reported feeling intensely anxious or fearful prior to delivering their baby. It has been found that fear of childbirth is a risk factor for both PPD and postpartum post-traumatic stress (Soderquist et al, 2009). Eighty percent of mothers who screened positive for major PPD in this study reported fear of childbirth in their pregnancy.

Low levels of social support and lack of support from the mother's partner are among the strongest predictors of PPD (e.g. Forman et al., 2000). Findings from this study indicate that lack of support from the mother's partner and from friends are significant variables associated with a high PDSS score. The incidence of major PPD in this study was 74.1% in mothers who reported that they did not receive any support from their partners and 62.7% in mothers who reported not receiving support from friends.

Life stress has been shown to be a significant predictor of PPD. Mothers who had a high PDSS score were significantly more likely to have moved house, had a partner who changed jobs or lost his job, had financial concerns, and experienced marital and family problems. Having another baby and getting married in the last two years were also associated with high scores on the PDSS, although somewhat less highly significant.

More than half of mothers in this study indicated that they were concerned about their finances in the previous two years. The prevalence of PPD has been reported to be significantly higher in women who experience financial stress or who are financially poor (e.g. Segre et al., 2007). The percentage of women who screened positive for major PPD who indicated that they were experiencing financial stress was 55.6%. In comparison, 36.9% of mothers screened positive for major PPD who did not report experiencing financial stress. This result replicates findings in other studies which indicated that financial stress is a strong and significant risk factor for PPD.

The results of this study confirm findings from other studies that marital conflict is a strong and significant predictor of PPD. The prevalence of major PPD in women who reported to be experiencing marital problems was 66.7% Results from this study also indicate that family problems is associated with major PPD. A limitation of this study is that it was not determined what family the mother was referring to, and whether family problems were experienced within the nuclear family, with extended family, or problems in the daughter-in-law-mother-in-law relationship.

Difficulty conceiving was found to be significantly associated with major PPD in this study. This variable is not generally regarded as a risk factor for PPD. The amount of mothers who indicated that they had difficulty conceiving was 14.2%. While 7.4% of the mothers in this study sought assistance with conception, seeking treatment for infertility was not significantly associated with major PPD. Yet, research has shown that assisted conception is a risk factor for postpartum mood disturbance (Fisher, Hammarberg, & Baker, 2005). A potential reason that Fisher et al (2005) cites is that women who struggled to conceive may feel they have a lowered sense of entitlement to seek help or to complain because the infant was so highly desired. This reason potentially also applies to women who struggled to conceive who did not opt for – or who could not afford – assisted reproductive technologies. Furthermore, other factors that were not explored in this study, but that may have been related to both difficulty in conception as well as predictive of postpartum mood disturbance, may be an area for future research.

8.9.4 Discussion of the correlation of the PDSS, the EPDS, and the QIDS-SR16.

Using multiple screening scales to determine convergent validity is, according to Campbell and Fiske (as cited in Beck & Gable, 2002, p. 39) a preferred approach to demonstrate that a measure has construct validity. Convergent validity indicates whether a test correlates positively with other tests that claim to measure the same construct. It is therefore an important part of construct validity.

Comparisons of the categorical depression status of the PDSS, EPDS, and the QIDS-SR16 with each other using chi-square tests indicate significant correlation between all three measures (all $p \leq 0.001$). Parametric correlation of the continuous scores on the PDSS, the EPDS, and the QIDS-SR16 also indicate that the three measures were highly correlated (all $p \leq 0.001$). In this case the correlation was slightly stronger between the PDSS and the EPDS than between the PDSS and the QIDS-SR16. The QIDS-SR16 correlated equally well with both the PDSS and the EPDS. All three instruments therefore identified the same women as likely to have post-partum depression. The finding that the

PDSS was correlated strongly with both the EPDS and the QIDS-SR16 provides evidence of its convergent validity, and hence its construct validity.

CHAPTER 9

CONCLUSION, LIMITATIONS, AND RECOMMENDATIONS FOR FUTURE RESEARCH

Postpartum depression is a highly prevalent complication of childbirth that often goes unrecognised. It has an impact on the health of the mother, her baby, and on other members of her family. The main objective of this study was to develop, and psychometrically evaluate, the properties of the Afrikaans version of the PDSS. The data generated from this study suggests that the Afrikaans PDSS is an effective screening measure that health practitioners in South Africa can use to identify mothers with PPD.

The Afrikaans PDSS demonstrates good psychometric properties in this study when compared to the English PDSS. Reliability indices for both the PDSS and the Afrikaans PDSS were excellent. Rasch analysis confirmed the presence of subdimensions (known as content scales) in the PDSS, which represents a multidimensional construct of PPD. Examination of fit indices for the English PDSS total screening scale reveal that all the items except the items from the SLP content scale had acceptable fit indices. A similar trend was noted in the analysis of the Afrikaans PDSS where the majority of items that had fit indices beyond the acceptable range of 1.40 were from the SLP content scale. This seems to indicate that these items clearly form a separate construct. These items did not present with fit problems within the SLP content scale.

Results from confirmatory factor analysis and IRT techniques during the development and psychometric testing of the English PDSS by Beck and Gable (2000) demonstrated the presence of seven dimensions and provided support for the construct validity of the PDSS. Beck and Gable (2002) also investigated how well each PDSS item loads on all extracted factors using exploratory factor analysis with their diagnostic sample. The analysis yielded 7 factors with eigenvalues greater than 1.00. Sleep-specific items from the SLP content scale and appetite-specific items from the SLP content scale loaded on separate factors. Item analysis and internal consistency estimates in the diagnostic sample reveal, however, that the reliability of the SLP content scale (.85) was slightly higher when all five items were included in the content scale than if either of the sleep- or appetite-specific items were removed.

Construct validity of the PDSS and Afrikaans PDSS content scales in this study is confirmed by Rasch analysis and Pearson's correlation. Correlation coefficients of the PDSS and Afrikaans PDSS content scales indicate that items correlate higher with the factor it purports to measure. This provides support for the presence of different subscales. Findings from Rasch analysis of fit indices in the PDSS and Afrikaans PDSS content scales, with item fit indices that demonstrate good item performance on the separate dimensions, confirm the unidimensionality of the content scales and provide support for their construct validity. This suggests that both language versions accurately capture and quantify the multidimensional construct of PPD in English and Afrikaans-speaking South African postpartum women.

This discussion therefore focuses on the unidimensionality of the seven content scales. When Rasch analysis was performed with each respective content scale, item fit

MNSQ statistics supported the measurement of a unidimensional construct in each content scale with the exception of two items: Item 28 (I felt that my baby would be better off without me; Ek het gevoel dat dit vir my baba beter sou wees sonder my) in both the English PDSS and the Afrikaans PDSS SUI content scale, and Item 30 (Ek het gevoel asof ek heeltyd aan die gang moes bly) in the Afrikaans PDSS ANX content scale.

Poor fit of Item 28 may be an indication that it was consistently misunderstood by both English and Afrikaans respondents, but considering that the item demonstrated poor fit in both languages, it is more likely that it did not fit the construct of the SUI content scale very well.

Item 30 had fit problems in the total Afrikaans PDSS, borderline fit problems in the total English PDSS, and also demonstrated poor fit in the Afrikaans ANX content scale. In addition to fit problems, Item 30 was also flagged as an item that was found difficult to understand by both English and Afrikaans participants, and presented with DIF in the ANX content scale. This may indicate that the Afrikaans translation was not adequate, that Afrikaans respondents were not familiar with the item content, or that the item's content is not appropriate for this Afrikaans sample.

In the development of the English PDSS (Beck & Gable, 2000), using data from the diagnostic sample, the deletion of a particular item did not increase the reliability of the respective content scales, with only one exception – Item 28. Deleting this item increased the reliability of the content scale (.90). Item analysis and internal consistency estimates revealed that Item 28 had the lowest correlation (.39) of the five items that comprise the SUI content scale, but the reliability of the scale still remained sufficiently high when the item was not deleted (alpha = .86). Clinical experts judged the content of Item 28 to be a good fit with the operational definition of the SUI content scale (Beck & Gable, 2001b). Item 30 had moderate correlations with the diagnostic sample (.55) and the reliability of the ANX content scale (.80) did not improve if this item was deleted (.78 if Item 30 was deleted). Results from confirmatory factor analysis of the PDSS using the development sample (Beck & Gable, 2000) revealed that both Item 28 and Item 30 (along with all other PDSS items) had a good fit with the hypothesized model (.75 and .66 respectively).

Item 2 (I got anxious over even the littlest things that concerned my baby; Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak) had large DIF in the ANX content scale, borderline fit problems in the English PDSS ANX content scale, and was flagged by seven participants (five English participants and two Afrikaans participants) as an item that was difficult to understand. This suggests that the English version of this item was not well understood by the English participants of this sample. In future studies, it may be explored whether the term ‘anxious’ is appropriate for the broader English-speaking population in South Africa, particularly because some South Africans refer to the term ‘nerves’ to describe anxiety. Therefore, Item 2 may not be considered appropriate for all English-speaking South Africans.

Although no fit problems were evident for Item 25 (I had a difficult time making even a simple decision; Ek het dit moeilik gevind om die eenvoudigste besluit te neem), this item did present with large DIF in the total PDSS and less DIF in the PDSS MNT content scale. The DIF in the content scale was relatively small (z = 2.00) and could be due to measurement error or sample idiosyncrasies. Item 25 was not marked as an item that was difficult to fully understand. The performance of this item may, however, need to be monitored in future studies.

In summary, both the English PDSS and the Afrikaans PDSS performed fairly well in these English and Afrikaans South African samples. The items that were identified as most problematic were Item 2, Item 25, Item 28, and Item 30. Other items that may require minor revision of their Afrikaans translation, although they did not present with fit or DIF problems, are Item 9, Item 23, and Item 31. Furthermore, together with the problematic items listed above, the following items need to be monitored in future studies: Item 24, Item 25, Item 32, and Item 34.

Analysis of the 5-point Likert response categories provided evidence for meaningful score interpretations. The rating scale analysis showed that responses to the different categories are separated sufficiently so as not to warrant combining some of the response categories. The response data therefore defines the dimension well, with higher responses corresponding to higher levels of agreement with the construct being measured.

A limitation of this study is that the reading level of the Afrikaans participants was not established prior to completing the Afrikaans PDSS. All participants were asked, however, to indicate whether there were items that they did not understand. This was done in order to account for comprehensibility of the PDSS in English participants and of the Afrikaans PDSS in Afrikaans participants. During the screening process some items were identified that a number of mothers found difficult to fully understand. Both the Afrikaans and the original English PDSS should therefore be used with caution in South African women who are not proficient in either of these languages. Future studies may consider using larger samples where the literacy or education level of respondents can function as strata which can be compared for item location shift.

The translation methodology employed in this study was chosen to help ensure adequate semantic translation of the PDSS into Afrikaans. Analysis revealed, however, that DIF was present in a number of items. Some significant differences were evident between the English and Afrikaans samples, which may account for some of the DIF found. It is also possible that DIF could be attributed to the translation, possible cultural differences in the verbal expression of emotional symptoms, or differences in the manifestation of depression symptoms.

Allalouf, Hambleton, and Sireci (as cited in Allalouf, 2003, p. 56) found different causes for DIF in different item types like analogies, sentence completion, and reading comprehension in verbal reasoning items. Cultural relevance was found to contribute to DIF in reading comprehension items. In order to eliminate the possibility that DIF is due to cultural relevance, a complementary study using only English and Afrikaans speaking participants from the same cultural group should be conducted. This may be an area of future research.

Van de Vijver and Hambleton (as cited in Allalouf, 2003, p. 56) recommend that a panel of psychologists with knowledge about potential causes of DIF and linguistic experts be involved in the translation and revision process. Future studies may attempt to improve the Afrikaans translation of the PDSS by changing the wording of the items with DIF in the target language and then retesting for DIF to determine if a different translation resulted in non-DIF or lower DIF items. Synonym questionnaires may also help to determine whether some words that appeared in items with DIF were well understood, such as anxious, failure, irritable, angstig, misluk, konsentreer, and geïrriteerd.

The Inconsistent Responding Index (INC) of the PDSS is a useful basic measure of a respondent's consistency in completing the PDSS, providing an indication of response validity. The INC score is derived from ten pairs of PDSS items for which ratings are typically very similar. The INC score is a count of the amount of item pairs for which there is a discrepancy of more than one point in the respondent's rating. When there is an INC score of four (in other words, four pairs of items showed inconsistent responding), then there is an 85% likelihood that the PDSS items were not completed in a way that consistently reflected the screening scale's content. An INC score of five results in a 94% likelihood, while an INC score of six results in a 97% likelihood that the PDSS items were not completed consistently. Beck and Gable (2000) recommend that the examiner regard an INC score of four or more as an indication that the respondent did not complete the PDSS consistently and that the PDSS cannot be interpreted accurately. Inconsistent responding may be due to a respondent misreading an item, marking a response other than the one intended, an inability to maintain sufficient focus during the time taken to complete the PDSS, or it may reflect a misunderstanding of the test instructions or of the item content. Consulting the respondent about the discrepancy in scores may provide some additional information about the comprehensibility of the items. A total of 16 participants (seven Afrikaans and nine English) had INC scores of 4 or more. The researcher did not follow up with these participants to determine the reason for the discrepancy in their responses. These participants were also not eliminated from the sample. In this study, items that participants marked as difficult to understand did not contribute greatly to high INC scores.

A useful screening questionnaire should be able to correctly identify, with consistent reliability, the illness it purports to measure. The screening questionnaire should also be able to correctly identify those persons who do not have the illness. In order to assess a screening questionnaire's accuracy (such as measures of sensitivity, specificity, and positive predictive value), it is necessary to know whether the respondent has been diagnosed with the illness or not. This may achieved by means of comparison to a reference standard, such as a diagnostic interview. At the time of data collection, this study was not funded. Due to limited resources a decision was made not to use the DSM-IV diagnostic interview as the gold standard to determine how many women would have been diagnosed with major PPD, minor PPD, and with no depression. Obtaining a sample of sufficient size from across South Africa who would be willing to participate through detailed psychiatric interviews in the early postpartum period was thought to be an insurmountable effort for one person to undertake. The researcher rather sought to determine how well the PDSS correlated with two other brief self-report screening measures which were easily administered, and could be completed online.

Not using a diagnostic interview, like the Structured Clinical Interview for DSM (SCID), was a major limitation of this study which resulted in verification bias. Firstly because no comparison could be made between the scores of the three screening measures used in this study with the diagnostic classification of membership to the non-depressed, or to the minor or major depressed groups to determine the accuracy of the measures in an English and Afrikaans South African population. And secondly, it was not possible to determine whether some participants' scores on the screening scales were influenced by co-morbidity of other disorders, like eating disorders for example, which, according to

Nishizono-Maher et al (as cited in Nishizono-Maher et al., 2004, p.189) is not uncommon in persons with very high scores on the EPDS. Verification bias often occurs in clinical research, frequently as a consequence of either resource limitations or ethical dilemmas, or both (Hanusa, Scholle, Haskett, Spadaro, & Wisner, 2008). It is recommended that future studies establish the screening accuracy of the PDSS and the Afrikaans PDSS in South African postpartum women by comparing the screening outcomes to the SCID.

The three instruments that were selected for use in this study were considered suitable instruments for identifying major PPD. Beck and Gable (2001a; 2001c) report that the ability of the PDSS to detect PPD is comparable to the ability of the Structured Clinical Interview for DSM-IV Axis 1 Disorders (SCID) to detect major depressive disorder. The PDSS, while still a relatively new measure, has also been translated into a number of languages in recent years. At present English and Spanish versions are readily available. It has also been used in different socioeconomic and ethnic groups. The PDSS is easily administered and, although not as brief as the EPDS, can be administered in five to ten minutes. The PDSS Short Form comprises the first seven items of the complete 35 item PDSS. This version may be completed in as little as one to two minutes and yields a total score which is comparable to the full PDSS total score as it provides an index of the general severity of the mother's PPD symptoms. The Short Form does not allow scoring of the PDSS dimensions (symptom content scales) but gives an indication whether the mother is in need of formal psychiatric evaluation. An advantage that the 35 item PDSS offers over the seven item PDSS Short Form, the EPDS, and the QIDS-SR16 is that it allows investigators and clinicians to identify the dimensions of PPD in which mothers have elevated symptoms, like emotional lability, anxiety and insecurity, and loss of self. This is particularly helpful to clinicians and therapists for determining specific areas that need to be addressed and for selecting the more suitable treatment options. The PDSS is copyrighted and must be purchased for use in clinical as well as research settings.

The QIDS-SR16 is a relatively new, but increasingly used, brief 16-item self-report measure (Rush et al., 2003; Trivedi, Rush, Ibrahim, et al., 2004). The QIDS-SR16 has demonstrated favourable psychometric profiles across paediatric, adult, and elderly populations with a major depressive disorder. The QIDS is regarded as an effective screening measure for depression in a variety of settings as it is available in both clinician and self-rated versions, it is not time-consuming, and it is administered easily (Rush et al., 2005; Rush et al., 2006; Trivedi, Rush, Crismon, et al., 2004). Although it has not been used extensively in PPD studies, Bernstein et al. (2008) regard the QIDS-SR16 a useful measure in the assessment of PPD as it screens for symptoms of agitation and restlessness as well as decision-making and concentration ability which they found to be symptomatically different between postpartum depression and other depression samples. The IDS, which is a longer format and comprises all 16 items of the QIDS, has demonstrated excellent sensitivity, good specificity and moderate PPV in English and Spanish speaking postpartum women (Yonkers et al., 2001).

The EPDS has been used extensively in postpartum studies across the world, been translated and validated in other cultures, and has been found to be a reliable and valid measure for the detection of PPD (e.g. Navarro et al., 2007). The EPDS has been found to be a valid screening instrument for PPD in low income, socially disadvantaged urban South African women (Lawrie et al., 1998) and an isiXhosa version of the EPDS has shown reliable scores in isiXhosa-speaking postpartum women in South Africa (De Bruin et al., 2004). Hanusa et al (2008) report that the advantages of using the EPDS is its brevity, ease of administration, it can be used free of charge by investigators, it has many translated versions and has been used with several different socioeconomic and ethnic groups.

Results from this study of continuous screening scores for the PDSS, the EPDS, and the QIDS-SR16 indicated that all three measures were highly correlated and effective in their ability to place women in order of increasing risk for a major depressive disorder. The QIDS-SR16 was translated into Afrikaans for the purpose of this study, but the effectiveness of the Afrikaans translation was not established. Correlation statistics with the QIDS-SR16 must therefore be interpreted cautiously.

This study had several limitations that affect its generalizability. The sample was relatively homogenous (85% white, 89% married, 67% with tertiary education) and the screening measures were not administered randomly. The effect of all the participants completing the questionnaires in the same order may have impacted on internal validity.

All postpartum women who met the research criteria were encouraged to participate in the study. Contact was also established with antenatal women at antenatal classes and if they were willing to participate in the study then they were contacted after their babies were born. There were, however, many participants that were referred from a variety of health practitioners, some of whom suspected the presence of PPD. Many mothers were also eager to participate and be screened for symptoms of PPD after reading articles the researcher wrote about PPD in popular baby magazines. The means used to recruit mothers for participation was another limitation of this study. The participants were asked to volunteer and were not selected randomly which may have resulted in selection bias and accounted for the higher prevalence rate (47.9%) of major PPD as assessed by the PDSS. This prevalence rate can therefore not be generalised across a similar South African sample.

It was not determined whether there were significant differences in the psychometrics of the PDSS when participants completed the questionnaires online compared to those participants who completed the questionnaires in the presence of the researcher using pen-paper administration. This may be investigated in a future study.

Baker and Oswalt (2008) screened women from a rural community for PPD and found a significant relationship between race, ethnicity and depression rates. The homogeneity of the sample in this study, however, did not make it possible to determine the influence that these particular maternal demographic characteristics may have had on the rate of PPD. Future studies with concentrated South African ethnic groups may provide a clearer indication of the factors associated with PPD within particular groups in the culturally diverse South African population.

The risk factors for PPD in this South African sample are generally consistent with those reported in other studies. Multiple regression analysis revealed eleven risk factors that were significantly associated with screening positive for major PPD using the PDSS. These were a history of psychiatric illness – depression in particular, antenatal depression in recent pregnancy, postpartum blues, lack of support from the baby's father, lack of support from friends, life stress, infant temperament, difficulty conceiving, feeling negative or ambivalent about expecting this baby, fearful of childbirth, and concern about health related issues regarding the infant, like colic, sleeping and feeding problems, and allergies. When the life stress variables were correlated with the participants' scores on the PDSS, the following variables were found to have a strong relationship to a positive screen for major PPD: moving house, partner changing jobs, partner's job loss, financial concerns, marital problems, family problems, and to a lesser degree, getting married within the past two years.

Mothers with a previous diagnosis of PPD and those who expressed dissatisfaction with the care received during labour and delivery were slightly more likely to have a positive score of major PPD. These variable were, however, not statistically significant when multiple regression analysis was employed.

Health practitioners who have contact with antenatal as well as postpartum women should be made aware of the risk factors for PPD as well as the symptoms of PPD in an effort to care for the mother's well-being in a comprehensive manner as opposed to focussing only on obstetric factors. Health practitioners should be particularly vigilant about those women who have an existing depression or a history of depression. Antenatal women who become depressed during their pregnancies, or are negative or ambivalent about becoming pregnant, or who seem especially anxious about the delivery of their babies should be closely monitored by their health practitioners for symptoms of PPD after delivery. Women need to be educated about postpartum blues and encouraged to seek advice from their health practitioner if their symptoms of postpartum blues persist for longer than two weeks. New mothers should be asked whether they are feeling isolated since the birth of their baby, and whether their partners provide them with adequate support. It is not uncommon for a new mother to complain that her infant sleeps poorly, or that she struggles with feeding her infant, or that her infant is difficult to settle, or suffers from colic or other health-related concerns. These concerns should not be taken lightly.

These mothers may need emotional support as well as support with their infants presenting problems.

The significant predictors of PPD that were identified in this study may serve as red flags to health practitioners that a women is at risk of developing PPD. Health practitioners should monitor these women carefully so that specific interventions can be implemented before PPD takes its toll on the mother, her infant, and her family.

Very few obstetricians, family practitioners and midwives in South Africa screen postpartum women routinely for symptoms of PPD. If routine screening is not practiced an effort should be made for the routine distribution of information to pregnant and postpartum women about PPD, what the symptoms are, where they may be screened for symptoms of PPD, and where they may obtain treatment. Creating awareness about PPD may potentially encourage women to discuss their symptoms with their health practitioners.

In conclusion, it is evident that PPD is an illness which touches the lives of many South African women. It is likely that many of these women would suffer in silence if their symptoms were not recognised. Screening for symptoms of PPD by using a screening scale like the PDSS or the EPDS will assist in improving the PPD diagnostic rate. It is imperative though, that when screening is done for PPD, that efficient systems are in place to provide treatment and follow-up for mothers with positive results.

APPENDIX A

ITC Guidelines for Test Adaptation

The International Test Commission (ITC), in collaboration with the European Test Publishers Group, the European Association of Psychological Assessment, the International Association of Applied Psychology, the International Association for Cross-Cultural Psychology, the International Association for the Evaluation of Educational Achievement, the International Language Testing Association, and the International Union of Psychological Science, prepared guidelines for translating and adapting tests and psychological instruments, and establishing score equivalence across language and/or cultural groups. After several years of preparation and field-testing, the following guidelines were approved by the ITC for distribution to national psychological societies, researchers, and test publishers (Hambleton, Merenda, & Spielberger, 2005; International Test Commission, 2010, pp. 2-3). The guidelines are classified into four categories as follows:

Context

C.1 Effects of cultural differences which are not relevant or important to the main purposes of the study should be minimized to the extent possible.

C.2 The amount of overlap in the construct measured by the test or instrument in the populations of interest should be assessed.

Test Development and Adaptation

D.1 Test developers/publishers should insure that the adaptation process takes full account of linguistic and cultural differences among the populations for whom adapted versions of the test or instrument are intended.

D.2 Test developers/publishers should provide evidence that the language use in the directions, rubrics, and items themselves as well as in the handbook are appropriate for all cultural and language populations for whom the test or instrument is intended.

D.3 Test developers/publishers should provide evidence that the choice of testing techniques, item formats, test conventions, and procedures are familiar to all intended populations.

D.4 Test developers/publishers should provide evidence that item content and stimulus materials are familiar to all intended populations.

D.5 Test developers/publishers should implement systematic judgmental evidence, both linguistic and psychological, to improve the accuracy of the adaptation process and compile evidence on the equivalence of all language versions.

D.6 Test developers/publishers should ensure that the data collection design permits the use of appropriate statistical techniques to establish item equivalence between the different language versions of the test or instrument.

D.7 Test developers/publishers should apply appropriate statistical techniques to (1) establish the equivalence of the different versions of the test or instrument, and (2) identify problematic components or aspects of the test or instrument which may be inadequate to one or more of the intended populations.

D.8 Test developers/publishers should provide information on the evaluation of validity in all target populations for whom the adapted versions are intended.

D.9 Test developers/publishers should provide statistical evidence of the equivalence of questions for all intended populations.

D.10 Non-equivalent questions between versions intended for different populations should not be used in preparing a common scale or in comparing these populations. However, they may be useful in enhancing content validity of scores reported for each population separately.

Administration

A.1 Test developers and administrators should try to anticipate the types of problems that can be expected, and take appropriate actions to remedy these problems through the preparation of appropriate materials and instructions.

A.2 Test administrators should be sensitive to a number of factors related to the stimulus materials, administration procedures, and response modes that can moderate the validity of the inferences drawn from the scores.

A.3 Those aspects of the environment that influence the administration of a test or instrument should be made as similar as possible across populations of interest.

A.4 Test administration instructions should be in the source and target languages to minimize the influence of unwanted sources of variation across populations.

A.5 The test manual should specify all aspects of the administration that require scrutiny in a new cultural context.

A.6 The administrator should be unobtrusive and the administrator-examinee interaction should be minimized. Explicit rules that are described in the manual for administration should be followed.

Documentation/Score Interpretations

I.1 When a test or instrument is adapted for use in another population, documentation of the changes should be provided, along with evidence of the equivalence.

I.2 Score differences among samples of populations administered the test or instrument should not be taken at face value. The researcher has the responsibility to substantiate the differences with other empirical evidence.

I.3 Comparisons across populations can only be made at the level of invariance that has been established for the scale on which scores are reported.

I.4 The test developer should provide specific information on the ways in which the socio-cultural and ecological contexts of the populations might affect performance, and should suggest procedures to account for these effects in the interpretation of results.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

APPENDIX B

Informed Consent and Research Information Provided to Mothers

Dear Mother

I am busy with a PhD in Clinical Psychology at the University of Pretoria. The topic of my research falls within the realm of the assessment of postpartum depression. The purpose of this study is:

- Firstly to address the problem of the unavailability of suitable postpartum depression screening measures for the majority of Afrikaans-speakers,
- And secondly, to determine the validity and the reliability of the Postpartum Depression Screening Scale for English and Afrikaans speaking South African mothers.

For the purposes of the study I need to screen new mothers for postpartum depression, whether they have symptoms of postpartum depression or not.

Mothers who wish to participate must:

- Be a South African citizen, residing in South Africa
- Be able to speak and read English or Afrikaans fluently
- Be between 4 and 16 weeks postpartum
- Have a baby without a disability.

Individuals who suffer from disorders that affect their ability to complete self-report measures reliably and validly should not volunteer for this study.

Participation is voluntary and screening is done free of charge. The participants' information will be treated with utmost confidentiality. A participant's data will be destroyed if she should decide to withdraw.

I sincerely hope that you will consider participating in this study and kindly request that you complete the questionnaires. If there are any queries please do not hesitate to contact me.

Thank you,
Melony Struik
Researcher
(Contact details).

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Consent form

I ____ hereby acknowledge that I am aware of this study and give my consent to participate. I am aware that the results will be treated with the utmost confidentiality and will be used for research purposes only. I may withdraw from participation at any time without adverse consequences and all my data will be destroyed should I decide to withdraw.

Mother's details:

Name: ____

Contact no: cel no.: ____

Tel (h): ____

Home language(s): ____

____

Signature

____

Date

____

Place

____

Researcher

Thank you,
Melony Struik
Researcher

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Appendix: Purpose and Procedure of the Research

UNIVERSITY OF PRETORIA PSYCHOLOGY DEPARTMENT

Pretoria 0002 Republic of South Africa

Research title: Validation of the Postpartum Depression Screening Scale in English and Afrikaans South African postpartum women.

Postpartum depression (PPD) is not uncommon – with up to 20 percent of all mothers, in all circumstances suffering from this type of depression. PPD is not always easy to identify without screening measures and may develop slowly any time during the first year of the baby's life. Every mother is different and may have a different combination of symptoms. Some may be more anxious or irritable than sad. It may be mild or severe. Some mothers have been depressed ever since the pregnancy, and sometimes "The Blues" just don't seem to go away. Some mothers manage well initially and then their mood becomes darker and darker. If untreated, it can adversely affect a mother's functioning as well as her infant's development. Screening all mothers after birth is therefore very important to ensure that they get the necessary help and support they need.

Purpose of the study:

- Firstly to address the problem of the unavailability of suitable postpartum depression screening measures for the majority of Afrikaans-speakers,
- And secondly, to determine the validity and the reliability of the Postpartum Depression Screening Scale for English and Afrikaans speaking South African mothers.

The Postpartum Depression Screening Scale (PDSS)

The PDSS is a brief 35 item, self-report instrument that can be administered in just 5 minutes. The PDSS screens for PPD and assesses the presence, severity and type of PPD symptoms. It enables health practitioners to identify mothers at risk, mothers who feel unhappy or overwhelmed, so that they may be referred for definitive diagnosis and treatment, thereby getting the necessary help and support they need.

Sample:

Eligibility for sample inclusion: All postpartum mothers, whether they present with symptoms of depression or not must:

- Be a South African citizen, residing in South Africa
- Be able to speak and read English or Afrikaans fluently
- Be between 4 and 16 weeks postpartum
- Have a baby without a disability.

Individuals who suffer from disorders that affect their ability to complete self-report measures reliably and validly will not be asked to volunteer for this study.

Procedure

Mothers who meet the above criteria and who are interested in participating in the research will be screened, either in person (if resident in Port Elizabeth) or online on a secure password protected website. Mothers who wish to participate online must contact the researcher to obtain the required password.

The mothers will be required to complete a form for statistics purposes and three brief mood questionnaires: the PDSS (described above), the Quick Inventory for Depressive Symptomatology - 16-Item - Self Report (QIDS-SR16), and the Edinburgh Postnatal Depression Scale (EPDS). The QIDS-SR16 is a short 16 item multiple choice questionnaire which usually takes no more than 5 minutes to complete. The EPDS is a brief 10 item rating scale and is also quick to complete.

Participants will be required to provide their name and contact number as mothers who present with symptoms of PPD will be contacted by the researcher for referral to their doctor. These mothers will also be advised to join a support group or seek psychological counselling. Only the researcher will have access to participants' personal details.

Participation is voluntary and screening is done free of charge. The participants' information will be treated with utmost confidentiality. A participant's data will be destroyed if she should decide to withdraw.

For queries or further information, please contact:

Melony Struik	or	Research Supervisor: Name
Researcher		Contact details
Contact details		

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Afrikaans Version:

Ingeligte Toestemming en Navorsing Inligting wat aan Moeders Voorsien is

Liewe Moeder

Ek is tans besig met ‘n doktorale skripsie in Sielkunde aan die Universiteit van Pretoria. Die onderwerp van my navorsing val binne die raamwerk van die evaluasie van nageboortelike depressie by moeders. Die doel van hierdie navorsing is:

- Eerstens om die gebrek aan geskikte nageboortelike depressie siftingsvraelyste vir die meerderheid Afrikaanssprekendes aan te spreek.
- En tweedens, om die geldigheid en betroubaarheid van die “Postpartum Depression Screening Scale (PDSS)” onder beide Engels- en Afrikaanssprekende moeders te bepaal.

Vir die doeleindes van die studie is dit nodig om nuwe moeders te toets vir simptome van nageboortelike depressie deur vraelyste te voltooi om te bepaal vir nageboortelike depressie.

Moeders wat graag wil deelneem moet:
- Suid-Afrikaanse burgers wees, tans woonagtig in Suid-Afrika
- óf Engels óf Afrikaans goed kan lees en praat.
- tussen 4 en 16 weke ná geboorte wees
- ‘n ongestremde baba hê

Individuëe wie nie in staat is om self die vraelyste te voltooi nie word gevra om nie deel te neem nie.

Deelname aan hierdie navorsing sal met die grootste vertroulikheid hanteer word, is ook totaal vrywillig en gratis. Deelnemers kan ter enige tyd gedurende die navorsing onttrek, sonder nagevolge, waarna alle data vernietig sal word.

Ek hoop dat ek op u samewerking kan staatmaak en vra dat u die vraelyste voltooi. Indien u enige navrae het kan u my gerus kontak.

Baie dankie,
Melony Struik
(Kontak besonderhede van navorser)

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Toestemming Vorm

Ek ____ erken hiermee dat ek bewus is van die navorsingstudie en gee hiermee my toestemming om daaraan deel te neem. Ek is bewus daarvan dat deelname aan die navorsing met die grootste vertroulikheid hanteer sal word en dat data slegs vir die doeleindes van die studie gebruik sal word. Ek is ook bewus daarvan dat ek ter enige tyd gedurende die navorsing mag onttrek, sonder nagevolge, waarna alle data vernietig sal word.

Moeder se inligting:

Naam: ____

Kontak no.: Sel: ____

Tel (h): ____

Huistaal: ____

______________________ ______________________

Handtekening Datum

____

Plek

____

Navorser

Baie dankie,
Melony Struik
Navorser

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Bylaag: Doel en Prosedure van die Navorsing

UNIVERSITEIT VAN PRETORIA SIELKUNDE DEPARTMENT

Pretoria 0002 Republiek van Suid Afrika

Geldigheid van die 'Postpartum Depression Screening Scale' by Engels- en Afrikaanssprekende moeders.

Nageboortelike depressie is nie ongewoon nie – tot 20% van alle moeders, uit alle omstandighede lei aan nageboortelike depressie. Nageboortelike depressie is nie altyd maklik identifiseerbaar sonder noukeurige siftingsvraelyste nie en kan enige tyd gedurende die eerste jaar van die baba se lewe ontwikkel. Elke moeder is anders en toon 'n verskillende kombinasie van simptome. Sommige mag meer angstig of geïrriteerd as neerslagtig wees. Dit mag matig of ernstig wees. Party moeders mag neerslagtig wees van die begin van die swangerskap af en die "blues" wil net nie wyk nie. Sommige moeders hanteer die situasie aanvanklik goed, maar mettertyd vererger hul gemoedstoestand. Onbehandeld kan dit die moeder se daaglikse optrede en die baba se ontwikkeling nadelig beïnvloed. Toets vir nageboortelike depressie aan alle moeders na bevalling is dus van die uiterste belang om die nodige hulp en bystand te kan verleen.

Doel van die navorsing:

- Eerstens om die gebrek aan geskikte nageboortelike depressie siftingsvraelyste vir die meerderheid Afrikaanssprekende moeders aan te spreek,
- En tweedens, om die geldigheid en betroubaarheid van die "Postpartum Depression Screening Scale (PDSS)" onder beide Engels- en Afrikaanssprekende moeders te bepaal.

Die 'Postpartum Depression Screening Scale' (PDSS)

Die PDSS is 'n 35-punt selfverslag instrument wat slegs ongeveer 5 minute neem om te voltooi. Die PDSS is 'n siftingsvraelys wat die teenwoordigheid, erns en tipe nageboortelike depressie simptome vasstel. Dit stel gesondheidsdeskundiges in staat om moeders met 'n hoë risiko van nageboortelike depressie, moeders wat ongelukkig of oorweldig voel, vroegtydig en maklik te identifiseer vir vroë diagnose en behandeling sodat hulle die nodige hulp en bystand mag kry wat hulle nodig het.

Steekproef

Moeders, of hulle simptome het van nageboortelike depressie het of nie, wat graag wil deel neem moet:

- Suid-Afrikaanse burgers wees, tans woonagtig in Suid-Afrika
- óf Engels óf Afrikaans goed kan lees en praat.
- tussen 4 en 16 weke ná geboorte wees
- 'n ongestremde baba hê

Individue wie nie in staat is om self die vraelyste te voltooi nie word gevra om nie deel te neem nie.

Prosedure

Moeders wat aan die bogenoemde vereistes voldoen en graag wil deelneem aan die studie, kan persoonlik deur die navorser getoets word indien woonagtig in Port Elizabeth. Anders kan moeders deur middel van die “secure password protected website”op die internet deelneem. Indien die moeder op hierdie manier wil deelneem sal sy die navorser moet kontak om die “password” vir deelname aan die studie te kry.

Die moeders sal ‘n vraelys vir statistiek doeleindes en drie kort gemoedsvraelyste moet voltooi: die PDSS (hierbo beskryf), die ‘Quick Inventory for Depressive Symptomatology - 16-Item - Self Report’ (QIDS-SR16), en die ‘Edinburgh Postnatal Depression Scale’ (EPDS). Die QIDS-SR16 is ‘n kort 16 item meervoudige keuse vraelys wat gewoonlik nie meer as 5 minute neem om te voltooi nie. Die EPDS is ‘n kort 10 item vraelys en is ook vinnig om te voltooi.

Deelnemers sal hul naam en ‘n kontak nommer moet voorsien aangesien moeders wat simptome van nageboortelike depressie toon na hul geneesheer verwys word. Die moeders sal ook aangemoedig word om deel te word van ‘n ondersteuningsgroep of om met ‘n sielkundige kontak te maak vir berading. Slegs die navorser sal toegang hê tot die moeder se persoonlike inligting.

Deelname aan hierdie navorsing sal met die grootste vertroulikheid hanteer word, is ook totaal vrywillig en gratis. Deelnemers kan ter enige tyd gedurende die navorsing onttrek, sonder nagevolge, waarna alle data vernietig sal word.

Indien u enige navrae het kontak gerus:

Melony Struik	of	Navorsing Opsiener: Naam
Navorser		(Kontak besonderhede)
(Kontak besonderhede)		

APPENDIX C

Demographic Questionnaire

Please select your answer by making a tick in the appropriate block

1. Today's date: ……………………………………………

2. Name (optional) …………………………………………

3. Telephone or celphone number where you may be contacted by the researcher if you present with symptoms of postpartum depression ……………………………………

4. Home language

☐ Afrikaans	☐ Tsonga	☐ Dutch
☐ English	☐ Tswana	☐ French
☐ Ndebele	☐ Venda	☐ German
☐ Northern-Sotho	☐ Xhosa	☐ Greek
☐ Southern-Sotho	☐ Zulu	☐ Portuguese
☐ Swazi	☐ Chinese	Other …………………

5. Did you have English as a subject at high school?
   - ☐ Yes, as 1st language
   - ☐ Yes, as 2nd language
   - ☐ Yes, as 3rd language
   - ☐ No

6. Are you fluent in English (i.e. can speak and read English well)
   - ☐ Yes
   - ☐ No

7. Are you a South African citizen and currently live in South Africa?
   - ☐ Yes
   - ☐ No

8. Indicate your race/ethnic group

☐ White	☐ Coloured
☐ Asian	☐ Other
☐ Black	

9. Current marital status
   - ☐ Married
   - ☐ Unmarried
   - ☐ Widowed
   - ☐ Divorced
   - ☐ Separated
   - ☐ In a de facto relationship (live together as if married)

10. Indicate the highest level of education you have attained:

☐	Degree or Diploma		☐	Grade 7	(Standard 5)
☐	Trade certificate		☐	Grade 6	(Standard 4)
☐	Grade 12	(Standard 10)	☐	Grade 5	(Standard 3)
☐	Grade 11	(Standard 9)	☐	Grade 4	(Standard 2)
☐	Grade 10	(Standard 8)	☐	Grade 3	(Standard 1)
☐	Grade 9	(Standard 7)	☐	Grade 2	
☐	Grade 8	(Standard 6)	☐	Grade 1	

11. Employment status
- ☐ Full-time
- ☐ Part-time
- ☐ Unemployed
- ☐ Self-employed

12. Age (in years) ……………….

13. Baby's date of birth …………………………..

14. Baby's age in weeks

☐	4 weeks	☐	9 weeks	☐	14 weeks
☐	5 weeks	☐	10 weeks	☐	15 weeks
☐	6 weeks	☐	11 weeks	☐	16 weeks
☐	7 weeks	☐	12 weeks	Other	………
☐	8 weeks	☐	13 weeks		

15. Baby's sex
- ☐ male
- ☐ female

16. Gestational age of baby at birth
- ☐ before 28 weeks
- ☐ 29-33 weeks
- ☐ 34-37 weeks
- ☐ 38-40 weeks
- ☐ Beyond 40 weeks

17. For your most recent birth - what type of delivery did you have?
- ☐ Normal vaginal birth
- ☐ Traumatic vaginal birth (e.g. complicated breech delivery, forceps delivery or ventouse (suction) assisted delivery)
- ☐ Elective caesarean (scheduled caesarean)
- ☐ Emergency caesarean (mother was already in labour and experienced complications which necessitated a caesarean delivery)

18. Rate your care during labour and delivery
- ☐ Excellent
- ☐ Good
- ☐ Unremarkable
- ☐ Poor
- ☐ Very poor

19. How have you been feeding your baby? (note: bottle feeding implies formula milk)
- ☐ Bottle feeding – from birth
- ☐ Breast feeding – from birth
- ☐ Initially breastfed but now bottle feed only
- ☐ Combination of breast and bottle

20. Indicate if you received help and support from the following people after you came home with your baby:

Baby's father	☐	Yes, most of the time when I needed it
	☐	Not as often as I needed
	☐	Hardly any
Family	☐	Yes, most of the time when I needed it
	☐	Not as often as I needed
	☐	Hardly any
Friends	☐	Yes, most of the time when I needed it
	☐	Not as often as I needed
	☐	Hardly any
Other	☐	Yes, most of the time when I needed it
	☐	Not as often as I needed
	☐	Hardly any

21. How many times have you been pregnant?

☐ 1	☐ 5
☐ 2	☐ 6
☐ 3	☐ More than 6
☐ 4	

22 How many biological children do you have?

☐ 1	☐ 5
☐ 2	☐ 6
☐ 3	☐ More than 6
☐ 4	

23. Does your baby have any serious illnesses or disabilities?
- ☐ No
- ☐ Yes

24. Did your caregiver enquire whether you were depressed at your postnatal checkup?
    - ☐ No
    - ☐ Yes
    - ☐ Have not yet had a postnatal checkup

25. Has a doctor or other health practitioner diagnosed you with postpartum depression after this recent pregnancy?
    - ☐ No
    - ☐ Yes

26. Has a health practitioner diagnosed you with antenatal depression during this recent pregnancy?
    - ☐ No
    - ☐ Yes

27. If you answered yes to the above two questions, are you receiving counseling or psychotherapy?
    - ☐ No
    - ☐ Yes
    - ☐ Not applicable

28. Are you currently using any medication for depression or anxiety?
    - ☐ Yes
    - ☐ No

29. Please indicate if you have ever been diagnosed with any of the following by a doctor or health practitioner:
    - ☐ Postpartum depression after a previous pregnancy
    - ☐ Antenatal depression during a previous pregnancy
    - ☐ Depression
    - ☐ Anxiety
    - ☐ Obsessive compulsive disorder
    - ☐ Anorexia
    - ☐ Bulimia
    - ☐ None

30. Please read the following statements and choose one which describes you best:
    - ☐ I think I may have some symptoms of postpartum depression
    - ☐ I think I may have postpartum depression
    - ☐ I don’t really know what postpartum depression is
    - ☐ I know what postpartum depression is and I don’t think I am suffering from it
    - ☐ I feel uncertain about whether or not I may have postpartum depression

31.	Did you have postpartum blues? (Also referred to as 'baby blues' - tearfulness, sadness, lack of concentration, feelings of dependency, and anxiety or irritability – these symptoms typically peak on the fourth or fifth day after delivery and may last for a few hours or a few days)	No	Yes
32.	Was this a planned pregnancy?	No	Yes
33.	Did you have difficulty falling pregnant?	No	Yes
34.	Did you have fertility treatment?	No	Yes
35.	Was this a complicated pregnancy? (e.g. pre-eclampsia, threatening miscarriage)	No	Yes
36.	Were you fearful of childbirth – a great deal more anxious and fearful than would be considered normal?	No	Yes
37.	Do you normally suffer from PMS (pre-menstrual syndrome – a condition with symptoms of mild depression, tension, irritability, headache, a feeling of bloatedness, with some evidence of edema, that usually begins in the week prior to menstruation and resolves completely the day after the onset of menstruation)?	No	Yes
38.	Do you consider yourself a perfectionist?	No	Yes

39.	In the past two years, have you experienced any of the following major life stresses?			
a	House alterations		No	Yes
b	Moving house		No	Yes
c	Moving city / immigrate		No	Yes
d	Job changes: self		No	Yes
e	Job changes: partner	Not applicable	No	Yes
f	Job loss / retrenchment: self		No	Yes
g	Job loss / retrenchment: partner	Not applicable	No	Yes
h	Financial concerns		No	Yes
i	Bereavement		No	Yes
j	Loss of close friends / family relocating, emigrating, etc.		No	Yes
k	Serious illness of a family member		No	Yes
l	Another pregnancy and birth		No	Yes
m	Marriage		No	Yes
n	Marital problems		No	Yes
o	Family problems		No	Yes
p	Been victimised by violence or crime		No	Yes
q	Serious injury, illness, or personal health problems		No	Yes

40. How did you feel about expecting a baby?

- ☐ Positive
- ☐ Ambivalent
- ☐ Negative
- ☐

Other:____

41. Do you experience your baby as:

- ☐ Good
- ☐ Fussy
- ☐ Demanding
- ☐ Difficult
- ☐

Other:____

42. Have you experienced any specific problems with your baby?

- ☐ No problems
- ☐ Health problems
- ☐ Colicky
- ☐ Sleep
- ☐ Feeding
- ☐ Allergies
- ☐ Premature
- ☐ Other:

____

Thank you for completing this form.

Three mood questionnaires follow. They are brief and each takes only a few minutes to complete. Please complete both on the same day – preferably one after the other as one's mood can vary considerably from day to day. Thank you. Your participation in this study is greatly appreciated.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Demografiese Vraelys

Merk asseblief u respons in die toepaslike blok

1. Vandag se datum ……………………………………….

2. Naam (opsioneel) ……………………………………...

3. Telefoon- of selfoonnommer waar navorser u mag kontak indien u simptome van nageboortelike depressie toon ……………………………………………………….

4. Huistaal

☐ Afrikaans	☐ Tsonga	☐ Hollands
☐ Engels	☐ Tswana	☐ Frans
☐ Ndebele	☐ Venda	☐ Duits
☐ Northern-Sotho	☐ Xhosa	☐ Grieks
☐ Southern-Sotho	☐ Zoeloe	☐ Portugees
☐ Swazi	☐ Chinees	Ander: ………………….

5. Het u Afrikaans as vak op hoërskool geneem?
   - ☐ Ja, as 1ste taal
   - ☐ Ja, as 2de taal
   - ☐ Ja, as 3de taal
   - ☐ Nee

6. Is u vlot in Afrikaans? (kan Afrikaans goed praat en lees)
   - ☐ Ja
   - ☐ Nee

7. Is u 'n Suid-Afrikaanse burger en tans woonagtig in Suid Afrika?
   - ☐ Ja
   - ☐ Nee

8. Dui u ras / etniese groep aan

☐ Blank	☐ Kleurling
☐ Asiaties	☐ Ander
☐ Swart	

9. Huwelikstatus
   - ☐ Getroud
   - ☐ Ongetroud
   - ☐ Weduwee
   - ☐ Geskei
   - ☐ Vervreemd
   - ☐ Woon saam asof getroud

10. Dui hoogste vlak opvoeding aan wat u verwerf het:

☐ Graad of Diploma		☐ Graad 7	(Standerd 5)
☐ Ambag sertifikaat		☐ Graad 6	(Standerd 4)
☐ Graad 12	(Standerd 10)	☐ Graad 5	(Standerd 3)
☐ Graad 11	(Standerd 9)	☐ Graad 4	(Standerd 2)
☐ Graad 10	(Standerd 8)	☐ Graad 3	(Standerd 1)
☐ Graad 9	(Standerd 7)	☐ Graad 2	
☐ Graad 8	(Standerd 6)	☐ Graad 1	

11. Werkstatus
- ☐ Voltyds
- ☐ Deeltyds
- ☐ Werkloos
- ☐ In eie diens

12. Ouderdom (in jaar) …………..

13. Baba se geboortedatum ………..

14. Baba se ouderdom in weke

☐ 4 weke	☐ 9 weke	☐ 14 weke
☐ 5 weke	☐ 10 weke	☐ 15 weke
☐ 6 weke	☐ 11 weke	☐ 16 weke
☐ 7 weke	☐ 12 weke	☐ Ander
☐ 8 weke	☐ 13 weke	

15. Baba se geslag
- ☐ manlik
- ☐ vroulik

16. Op hoeveel weke is u baba gebore?
- ☐ Voor 28 weke
- ☐ 29-33 weke
- ☐ 34-37 weke
- ☐ 38-40 weke
- ☐ Na 40 weke

17. Met die mees onlangse geboorte – watter tipe bevalling het u gehad?
- ☐ Normale vaginale verlossing
- ☐ Traumatiese vaginale verlossing (bv. Gekompliseerde stuitverlossing,tangverlossing of ventouse (suierverlossing).
- ☐ Elektiewe keisersnee (beplande keisersnee)
- ☐ Nood keisersnee (keisersnee as gevolg van komplikasies tydens kraam)

18. Beoordeel u sorg tydens u kraam en bevalling
- ☐ Uitstekend
- ☐ Goed
- ☐ Nie noemenswaardig
- ☐ Swak
- ☐ Baie swak

19. Hoe word u baba gevoed? (let op: bottelvoed impliseer formule melk)
- ☐ Bottelvoed – vanaf geboorte
- ☐ Borsvoed – vanaf geboorte
- ☐ Aanvanklik geborsvoed, maar bottelvoed nou uitsluitlik
- ☐ Beide bors- en bottelvoeding

20. Dui aan of u hulp en ondersteuning ontvang het van die volgende mense nadat u met u baba tuis gekom het:

Baba se vader	☐	Ja, meeste van die tyd soos wat ek nodig gehad het
	☐	Nie so dikwels soos wat ek nodig gehad het nie
	☐	Amper niks nie
Familie	☐	Ja, meeste van die tyd soos wat ek nodig gehad het
	☐	Nie so dikwels soos wat ek nodig gehad het nie
	☐	Amper niks nie
Vriendinne	☐	Ja, meeste van die tyd soos wat ek nodig gehad het
	☐	Nie so dikwels soos wat ek nodig gehad het nie
	☐	Amper niks nie
Ander mense	☐	Ja, meeste van die tyd soos wat ek nodig gehad het
	☐	Nie so dikwels soos wat ek nodig gehad het nie
	☐	Amper niks nie

21. Hoeveel keer was u al swanger?
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ Meer as 6

22. Hoeveel biologiese kinders het u?
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ Meer as 6

23. Het u baba enige ernstige siektes of gestremdheid?
- ☐ Nee
- ☐ Ja

24. Tydens u nageboortelike ondersoek, was u deur u verloskundige of geneesheer gevra of u depressief voel?
- ☐ Nee
- ☐ Ja
- ☐ Het nog nie 'n nageboortelike ondersoek gehad nie.

25. Na die mees onlangse swangerskap, was u deur 'n geneesheer of ander professionele gesondheidsdeskundige met nageboortelike depressie gediagnoseer?
- ☐ Nee
- ☐ Ja

26. Tydens die mees onlangse swangerskap, was u deur 'n geneesheer of ander professionele gesondheidsdeskundige met voorgeboortelike depressie gediagnoseer?
- ☐ Nee
- ☐ Ja

27. Indien u 'ja'op bogenoemde twee vrae beantwoord het, ontvang u berading of psigoterapie?
- ☐ Nee
- ☐ Ja
- ☐ Nie van toepassing

28. Gebruik u tans enige medikasie vir depressie of angs?
- ☐ Ja
- ☐ Nee

29. Dui asseblief aan of u al ooit met enige van die voldgende deur 'n geneesheer of gesondheidsdeskundige gediagnoseer is:
- ☐ Nageboortelike depressie na 'n vorige swangerskap
- ☐ Voorgeboortelike depressie tydens 'n vorige swangerskap
- ☐ Depressie
- ☐ Angs
- ☐ Obsessiewe kompulsiewe versteuring
- ☐ Anoreksie
- ☐ Bulimie
- ☐ Geen

30. Lees asseblief die volgende stellings en kies een wat u die beste beskryf:

- ☐ Ek dink ek het sommige simptome van nageboortelike depressie
- ☐ Ek dink ek het nageboortelike depressie
- ☐ Ek weet nie regtig wat nageboortelike depressie is nie
- ☐ Ek weet wat nageboortelike depressie is en ek dink nie ek ly daaraan nie
- ☐ Ek voel onseker of ek nageboortelike depressie het of nie

31.	Het u nageboorte "blues"gehad? (Word ook "baby blues"genoem - huilerig, hartseer, moeg, sukkel om te konsentreer, gevoel van afhanklikheid, angstig of geïrriteerd – die simptome bereik tipies 'n hoogtepunt op die 4de of 5de dag na geboorte en mag 'n paar uur of 'n paar dae duur.)	Nee	Ja
32.	Was die swangerskap beplan?	Nee	Ja
33.	Het u gesukkel om swanger te raak?	Nee	Ja
34.	Was u behandel vir onvrugbaarheid?	Nee	Ja
35.	Het u komplikasies tydens u swangerskap ondervind? (bv. preëklampsie, dreigende miskraam)	Nee	Ja
36.	Was u vreesbevange oor die geboorte – heelwat meer angstig en vreesbevange as wat normaal beskou sou word?	Nee	Ja
37.	Ly u gewoonlik aan "PMS" ("pre-menstrual syndrome"- 'n toestand met simptome van matige depressie, spanning, geïrriteerdheid, hoofpyne, en 'n gevoel van opgeblaasdheid met enige tekens van edeem wat gewoonlik so 'n week voor menstruasie begin, en na menstruasie weer verdwyn.	Nee	Ja
38.	Beskou u uself as 'n perfeksionis?	Nee	Ja

39. In die afgelope 2 jaar, het u enige van die volgende belangrike spanning situasies ervaar?

39.	In the past two years, have you experienced any of the following major life stresses?			
a	Huisverbeterings		Nee	Ja
b	Verhuis		Nee	Ja
c	Na 'n ander stad verhuis / immigreer		Nee	Ja
d	Van werk verander: self		Nee	Ja
e	Van werk verander: eggenoot	Nie van toepassing	Nee	Ja
f	Werk verloor / afgedank: self		Nee	Ja
g	Werk verloor / afgedank: eggenoot	Nie van toepassing	Nee	Ja
h	Finasiële kommer		Nee	Ja
i	'n Familielid of vriend verloor		Nee	Ja
j	Intieme vriende of familie wat weggetrek het.		Nee	Ja
k	Familielid wat ernstig siek is		Nee	Ja
l	Nog 'n swangerskap of geboorte		Nee	Ja

39.	In the past two years, have you experienced any of the following major life stresses?		
m	In die huwelik getree	Nee	Ja
n	Huweliksprobleme	Nee	Ja
o	Familie probleme	Nee	Ja
p	Geviktimiseer of 'n slagoffer van misdaad	Nee	Ja
q	Ernstige ongeluk, siekte of persoonlike gesondheidsprobleem.	Nee	Ja

40. Hoe het u gevoel oor u swangerskap?
- ☐ Positief
- ☐ Ambivalent (partykeer meer positief; ander kere effens negatief)
- ☐ Negatief
- ☐ Ander:

____

41. Ervaar u u baba as:
- ☐ Soet
- ☐ Puntenerig
- ☐ Veeleisend
- ☐ Moeilik
- ☐ Ander:

____

42. Het u enige spesifieke probleme met u baba ervaar?
- ☐ Geen probleme
- ☐ Gesondheidsprobleme
- ☐ Koliek
- ☐ Slaap
- ☐ Voeding
- ☐ Allergieë
- ☐ Prematuur
- ☐ Ander:

____

Baie dankie dat u die vraelys voltooi het.

Drie, kort gemoedsvraelyste volg wat slegs 'n paar minute elk neem om te voltooi. Dit moet op dieselfde dag voltooi word, verkieslik direk na mekaar aangesien 'n mens se gemoed van dag tot dag aansienlik kan verskil. Dankie. U deelname aan die studie word opreg waardeer.

APPENDIX D

Postpartum Depression Screening Scale (PDSS)
Cheryl Tatano Beck, D.N.Sc., and Robert K. Gable, Ed.D.

- Afrikaans Version / Afrikaanse Weergawe -

Hieronder is ‘n lys van stellings wat beskryf hoe ‘n ‘moeder kan voel na die geboorte van haar baba. Dui asseblief aan hoe veel jy met elke stelling saamstem of verskil. Beantwoord die vrae soos dit ooreenstem met hoe jy oor die afgelope twee weke gevoel het. Lees elke item versigtig. Omkring dan die nommer wat jou antwoord die beste beskryf. Gee asseblief slegs een antwoord vir elke stelling. Gebruik die volgende skaal om jou antwoorde aan te dui:

1 = Verskil sterk
2 = Verskil
3 = Verskil nie, maar stem ook nie saam nie
4 = Stem saam
5 = Stem beslis saam

Indien u u antwoord wil verander, trek ‘n “X” deur u eertse antwoord. Omkring dan die nommer wat u nuwe keuse die beste beskryf. Indien daar in die vraelys ‘n stelling is wat u moeilik vind om te verstaan, dui asseblief die nommer(s) van die stelling(s) aan in die toepaslike spasie aan die einde van die PDSS vraelys.

1 = Verskil sterk 2 = Verskil 3 = Verskil nie, maar stem ook nie saam nie 4 = Stem saam
5 = Stem beslis saam

Oor die afgelope twee weke,

		1	2	3	4	5
1	Al het my baba geslaap, het ek gesukkel om te slaap.	1	2	3	4	5
2	Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak.	1	2	3	4	5
3	Ek het gevoel asof my emosies wipplank ry.	1	2	3	4	5

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

1 = Verskil sterk 2 = Verskil 3 = Verskil nie, maar stem ook nie saam nie 4 = Stem saam
5 = Stem beslis saam

Oor die afgelope twee weke,

		1	2	3	4	5
4	Ek het gevoel of ek van my verstand af raak.	1	2	3	4	5
5	Ek was bang dat ek nooit weer my normale self sou wees nie.	1	2	3	4	5
6	Ek het gevoel asof ek nie die ma is wat ek wou wees nie.	1	2	3	4	5
7	Ek het gedink die dood sou die enigste uitweg uit hierdie nagmerrie wees.	1	2	3	4	5
8	Ek het my eetlus verloor.	1	2	3	4	5
9	Ek het heeltemal oorweldig gevoel.	1	2	3	4	5
10	Ek was bang dat ek nooit weer gelukkig sou wees nie.	1	2	3	4	5
11	Ek kon op niks konsentreer nie.	1	2	3	4	5
12	Ek het soos 'n vreemde vir myself gevoel.	1	2	3	4	5
13	Ek het gevoel asof baie ander ma's beter as ek was.	1	2	3	4	5
14	Ek het begin dink dat dit beter sou wees as ek dood was.	1	2	3	4	5
15	Ek het in die middel van die nag vanself wakker geskrik en gesukkel om weer aan die slaap te raak.	1	2	3	4	5
16	Ek was so angstig ek het gevoel asof ek uit my vel wou spring.	1	2	3	4	5
17	Ek het sonder enige rede baie gehuil.	1	2	3	4	5
18	Ek het gedink ek raak gek.	1	2	3	4	5
19	Ek het myself nie meer geken nie.	1	2	3	4	5
20	Ek het skuldig gevoel omdat dit vir my gevoel het asof ek nie my baba lief genoeg het nie.	1	2	3	4	5
21	Ek wou myself seermaak.	1	2	3	4	5
22	Ek het snags lank rondgerol en gesukkel om aan die slaap te raak.	1	2	3	4	5

1 = Verskil sterk 2 = Verskil 3 = Verskil nie, maar stem ook nie saam nie 4 = Stem saam
5 = Stem beslis saam

Oor die afgelope twee weke,

		1	2	3	4	5
23	Ek het alleen gevoel.	1	2	3	4	5
24	Ek was baie geïrriteerd.	1	2	3	4	5
25	Ek het dit moeilik gevind om die eenvoudigste besluite te neem.	1	2	3	4	5
26	Ek het gevoel asof ek nie normaal was nie.	1	2	3	4	5
27	Dit het gevoel asof ek my ware gevoelens en gedagtes oor my baba moes wegsteek.	1	2	3	4	5
28	Ek het gevoel dat dit vir my baba beter sou wees sonder my.	1	2	3	4	5
29	Ek het geweet ek moes eet, maar kon nie.	1	2	3	4	5
30	Ek het gevoel asof ek heeltyd aan die gang moes bly.	1	2	3	4	5
31	Ek het baie kwaad gevoel en was gereed om te ontplof.	1	2	3	4	5
32	Ek het gesukkel om op 'n taak te konsentreer.	1	2	3	4	5
33	Ek het nie eg gevoel nie.	1	2	3	4	5
34	Ek het gevoel asof ek as ma misluk.	1	2	3	4	5
35	Ek wou eenvoudig hierdie wêreld agterlaat.	1	2	3	4	5

Dui asseblief in die spasie aan ____________ die nommer(s) van die stelling(s) in die PDSS vraelys wat moeilik was om te verstaan.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

APPENDIX E

The Quick Inventory of Depressive Symptomatology (Self-Report) (QIDS-SR16)

- Afrikaans Version / Afrikaanse Weergawe –

Merk een stelling vir elke item wat die beste beskryf hoe jy die afgelope 7 dae gevoel het.

1 Aan die slaap raak:

☐ 0 Dit neem my nooit langer as 30 minute om aan die slaap te raak nie.
☐ 1 Dit neem my minder as die helfte van die tyd minstens 30 minute om aan die slaap te raak.
☐ 2 Dit neem my meer as die helfte van die tyd minstens 30 minute om aan die slaap te raak.
☐ 3 Dit neem my meer as die helfte van die tyd meer 60 minute om aan die slaap te raak.

2 Slaap gedurende die nag:

☐ 0 Ek word nie snags wakker nie.
☐ 1 Ek slaap rusteloos en lig, en word elke aand ‘n paar keer kort-kort wakker.
☐ 2 Ek word snags minstens een keer wakker, maar raak weer maklik aan die slaap.
☐ 3 Meer as helfte van die tyd word ek meer as een keer snags wakker en bly ten minste 20 minute of langer wakker.

3 Word te vroeg wakker:

☐ 0 Die meeste van die tyd word ek nie meer as 30 minute voor opstaantyd wakker nie.
☐ 1 Ek word meer as die helfte van die tyd meer as 30 minute voor opstaantyd wakker.
☐ 2 Ek word feitlik altyd minstens sowat een uur voor opstaantyd wakker, maar raak uiteindelik weer aan die slaap.
☐ 3 Ek word minstens een uur voor opstaantyd wakker en kan dan nie weer aan die slaap raak nie.

4 Slaap te veel:

☐ 0 Snags slaap ek nie langer as 7 tot 8 ure nie, en ek slaap nie bedags nie.
☐ 1 Ek slaap nie meer as 10 ure in `n 24 uur tydperk nie, met insluiting van middagslapies.
☐ 2 Ek slaap nie meer as 12 ure in `n 24 uur tydperk nie, met insluiting van middagslapies.
☐ 3 Ek slaap meer as 12 ure in `n 24 uur tydperk, met insluiting van middagslapies.

5 Hartseer voel:

☐ 0 Ek voel nie hartseer nie.
☐ 1 Ek voel minder as die helfte van die tyd hartseer.
☐ 2 Ek voel meer as die helfte van die tyd hartseer.
☐ 3 Ek voel feitlik heeltyd hartseer.

Voltooi asseblief of 6 of 7 (nie beide nie)

6 Afname in eetlus:

☐ 0 Daar is geen verandering in my gewone eetlus nie.
☐ 1 Ek eet effens minder gereeld of minder hoeveelhede kos as gewoonlik.
☐ 2 Ek eet baie minder as gewoonlik en slegs as ek `n poging aanwend.
☐ 3 Ek eet selde binne `n tydperk van 24 uur, en slegs met uiterste moeite of wanneer ander mense my aanmoedig om te eet.

- OF -

7 Toename in eetlus:

☐ 0 Daar is geen verandering in my gewone eetlus nie.
☐ 1 Ek het `n behoefte om meer gereeld as gewoonlik te eet.
☐ 2 Ek eet gereeld meer dikwels en/of groter hoeveelhede kos as gewoonlik.
☐ 3 Ek voel gedwing om tydens maaltye en tussen maaltye te ooreet.

Voltooi asseblief of 8 of 9 (nie beide nie)

8 Afname in gewig (in die afgelope twee weke):
☐ 0 My gewig het nie verander nie.
☐ 1 Dit voel asof ek `n bietjie gewig verloor het.
☐ 2 Ek het 1 of meer kilogram gewig verloor.
☐ 3 Ek het 2 of meer kilogram gewig verloor.

- OF –

9 Toename in gewig (in die afgelope twee weke):
☐ 0 My gewig het nie verander nie.
☐ 1 Dit voel asof ek effens gewig opgetel het.
☐ 2 Ek het 1 kilogram of meer opgetel.
☐ 3 Ek het 2 kilogram of meer opgetel.

10 Konsentrasie/Besluitnemingsvermoë:
☐ 0 Daar is geen verandering in my normale vermoë om te konsentreer of besluite te neem nie.
☐ 1 Ek voel af en toe besluiteloos of dat my aandag afgelei word.
☐ 2 Ek sukkel die meeste van die tyd om my aandag te fokus of om besluite te maak.
☐ 3 Ek kan nie goed genoeg konsentreer om te lees nie en kan selfs nie klein besluite neem nie.

11 Hoe ek myself beskou:
☐ 0 Ek beskou myself as ewe waardevol en verdienstelik as ander mense.
☐ 1 Ek blameer myself meer as gewoonlik.
☐ 2 Ek glo hoofsaaklik dat ek probleme vir ander veroorsaak.
☐ 3 Ek dink feitlik heeltyd oor my groot en klein tekortkominge.

12 Gedagtes oor die dood of selfmoord:
☐ 0 Ek dink nie aan selfmoord of oor die dood nie.
☐ 1 Ek ervaar die lewe as leeg en twyfel of die lewe die moeite werd is.
☐ 2 Ek dink verskeie kere per week vir etlike minute aan selfmoord of die dood.
☐ 3 Ek dink verskeie kere per dag in besonderhede aan selfmoord of die dood, of ek het spesifieke planne vir selfmoord pleeg of het voorheen probeer om my lewe te neem.

13 Algemene belangstelling:

- ☐ 0 My belangstelling in ander mense en aktiwiteite het nie verander nie.
- ☐ 1 Ek kom agter dat ek minder in mense of aktiwiteite belangstel.
- ☐ 2 Ek vind dat ek slegs in een of twee aktiwiteite waarmee ek my voorheen besig gehou het belangstel.
- ☐ 3 Ek het feitlik geen belangstelling in die aktiwiteite waarmee ek my voorheen besig gehou het.

14 Energievlak:

- ☐ 0 Daar is geen verandering in my normale energievlak nie.
- ☐ 1 Ek raak makliker as gewoonlik moeg.
- ☐ 2 Ek moet `n hewige poging aanwend om my daaglikse aktiwiteite te begin of te voltooi (bv. Inkopies of huiswerk doen, kook of werk toe gaan).
- ☐ 3 Ek kan die meeste van my daaglikse aktiwiteite glad nie uitvoer nie omdat ek envoudig nie die energie het nie.

15 Gevoel van traagheid:

- ☐ 0 Ek dink, praat en beweeg teen my gewone tempo.
- ☐ 1 Ek kom agter dat ek stadiger dink of dat my stem flou of afgestomp klink.
- ☐ 2 Dit neem etlike sekondes voordat ek op die meeste vrae reageer en ek is oortuig dat my denke traag is.
- ☐ 3 Ek kan dikwels nie op vrae reageer sonder om `n uiterste poging aan te wend nie.

16 Gevoel van rusteloosheid:

- ☐ 0 Ek voel nie rusteloos nie.
- ☐ 1 Ek is dikwels kriewelrig, wring my hande, of moet my sitposisie verander.
- ☐ 2 Ek het `n drang om rond te beweeg en voel taamlik rusteloos.
- ☐ 3 By tye is dit vir my onmoontlik om te bly sit en ek voel ek moet rond beweeg.

APPENDIX F

Additional Tables and Figures for Chapter 8

Table 67 Association of sample characteristics with English and Afrikaans samples

Sample Characteristics	Pearson Chi-Square		
	X^2	*df*	*P*
Current marital status	3.06	3	0.383
Indicate the highest level of education you have attained	5.75	7	0.569
Employment status	3.62	3	0.305
Age (in years)	18.07	24	0.800
Baby's age (in weeks)	27.07	12	0.008**
Baby's sex	0.36	1	0.549
Gestational age of baby at birth	6.68	4	0.154
For your most recent birth - what type of delivery did you have?	4.66	3	0.198
Rate your care during labour and delivery	5.25	3	0.154
How have you been feeding your baby? (bottle feeding implies formula milk)	2.49	3	0.476
Indicate if you received help and support from the following people after you came home with your baby:			
Father	10.09	2	0.006**
Family	10.05	2	0.007**
Friends	2.34	2	0.311
Other	4.24	2	0.120
How many times have you been pregnant?	3.00	5	0.700
How many biological children do you have?	3.38	4	0.497
Does your baby have any serious illnesses or disabilities?	1.05	1	0.305
Did your caregiver enquire whether you were depressed at your postnatal check-up?	0.56	2	0.755
Has a doctor or other health practitioner diagnosed you with postpartum depression after this recent pregnancy?	0.01	1	0.910
Has a health practitioner diagnosed you with antenatal depression during this recent pregnancy?	0.05	1	0.823
If you answered yes to the above two questions, are you receiving counseling or psychotherapy?	1.30	2	0.523
Are you currently using any medication for depression or anxiety?	2.53	1	0.112
Please indicate if you have ever been diagnosed with any of the following by a doctor or health practitioner:			
Postpartum depression after a previous pregnancy	0.01	1	0.910
Antenatal depression during a previous pregnancy	0.05	1	0.823

Sample Characteristics	Pearson Chi-Square		
	X^2	*df*	*P*
Please read the following statements and choose one which describes you best:			
I think I may have some symptoms of postpartum depression I think I may have postpartum depression I don't really know what postpartum depression is I know what postpartum depression is, I don't think I am suffering from it I feel uncertain about whether or not I may have postpartum depression	10.90	4	0.028*
Did you have postpartum blues?	1.79	1	0.181
Was this a planned pregnancy?	0.99	1	0.319
Did you have difficulty falling pregnant?	0.18	1	0.669
Did you have fertility treatment?	2.35	1	0.125
Was this a complicated pregnancy?	1.55	1	0.213
Were you fearful of childbirth – a great deal more anxious and fearful than would be considered normal?	0.28	1	0.600
Do you normally suffer from PMS	2.51	1	0.113
Do you consider yourself a perfectionist?	2.75	1	0.097
In the past two years, have you experienced any of the following major life stresses?			
House alterations	1.44	1	0.229
Moving house	8.52	1	0.004**
Moving city / immigrate	5.34	1	0.021*
Job changes: self	6.77	1	0.009**
Job changes: partner	9.38	2	0.009**
Job loss / retrenchment: self	1.91	1	0.167
Job loss / retrenchment: partner	0.47	2	0.791
Financial concerns	0.08	1	0.776
Bereavement	5.73	1	0.017*
Loss of close friends / family relocating, emigrating, etc.	0.11	1	0.737
Serious illness of a family member	2.28	1	0.131
Another pregnancy and birth	1.93	1	0.164
Marriage	0.51	1	0.475
Marital problems	0.16	1	0.694
Family problems	0.02	1	0.897
Been victimised by violence or crime	5.07	1	0.024*
Serious injury, illness, or personal health problems	0.01	1	0.935
How did you feel about expecting a baby?			
Positive Ambivalent Negative Other	4.38	3	0.223

Sample Characteristics	Pearson Chi-Square		
	X^2	*df*	*P*
Specify other: Anxious; Anxious overwhelmed; Anxious losing baby; Anxious pregnancy; Anxious responsibility; Anxious motherhood & weight gain.	3.75	5	0.586
Did you experience your baby as:			
Good			
Fussy	5.78	4	0.216
Demanding			
Difficult			
Other			
Have you experienced any specific problems with your baby?			
No problems	0.26	1	0.613
Health problems	0.002	1	0.964
Colicky	2.01	1	0.157
Sleep	0.23	1	0.633
Feeding	4.03	1	0.045*
Allergies	3.25	1	0.071
Premature	13.21	1	0.000***

* $p \leq 0.05$
** $p \leq 0.01$
*** $p \leq 0.001$

Table 68a Crosstabulation of Support Recived from the Baby's Father and Questionnaire Language

			PDSS Language		Total
			English	Afrikaans	
Support from Baby's Father	No	Count	14	13	27
		Expected Count	13.8	13.2	27.0
		Std. Residual	0.0	0.0	
	Yes	Count	144	113	257
		Expected Count	131.7	125.3	257.0
		Std. Residual	1.1	-1.1	
	Some	Count	29	52	81
		Expected Count	41.5	39.5	81.0
		Std. Residual	-1.9	2.0	
Total		Count	187	178	365
		Expected Count	187.0	178.0	365.0

Table 68a Chi-Square Statistics from Crosstabulation of Support Recived from the Baby's Father and Questionnaire Language

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.09[a]	2	.006
Likelihood Ratio	10.19	2	.006
Linear-by-Linear Association	6.40	1	.011
Number of Valid Cases	365		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.17.

Table 69a Crosstabulation of Support Recived from Family and Questionnaire Language

			PDSS Language		Total
			English	Afrikaans	
Support from Family	No	Count	24	23	47
		Expected Count	24.1	22.9	47.0
		Std. Residual	.0	.0	
	Yes	Count	131	100	231
		Expected Count	118.3	112.7	231.0
		Std. Residual	1.2	-1.2	
	Some	Count	32	55	87
		Expected Count	44.6	42.4	87.0
		Std. Residual	-1.9	1.9	
Total		Count	187	178	365
		Expected Count	187.0	178.0	365.0

Table 69b Chi-Square Statistics from Crosstabulation of Support Recived from Family and Questionnaire Language

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.05[a]	2	.007
Likelihood Ratio	10.13	2	.006
Linear-by-Linear Association	4.81	1	.028
Number of Valid Cases	365		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 22.92.

Table 70 Summary Statistics of 187 Extreme and Non-Extreme Participants for the English PDSS.

	Raw Score	Count	Measure	Model Error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
Mean	48.30	35.00	-0.95	0.30				
S.D.	36.90	0.10	1.84	0.30				
Max	138.00	35.00	4.19	1.85				
Min	0.00	34.00	-6.29	0.16	0.28	-4.30	0.16	-3.90

Real RMSE	0.46	True S.D.	1.78	Separation	3.86	Particip Reliability	0.94
Model RMSE	0.43	True S.D.	1.79	Separation	4.15	Particip Reliability	0.95

S.E. of participant mean = 0.13

Participant raw score-to-measure correlation = 0.91

Cronbach Alpha (KR-20) Participant raw score reliability = 0.98

Table 71 Summary Statistics of 178 Extreme and Non-Extreme Participants for the Afrikaans PDSS

	Raw Score	Count	Measure	Model Error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
Mean	45.70	35.00	-1.21	0.32				
S.D.	34.70	0.10	1.72	0.36				
Max	129.00	35.00	2.60	1.83				
Min	0.00	34.00	-5.84	0.16				

Real RMSE	0.50	True S.D.	1.64	Separation	3.27	Particip Reliability	0.91
Model RMSE	0.49	True S.D.	1.65	Separation	3.39	Particip Reliability	0.92

S.E. of participant mean = 0.13

Participant raw score-to-measure correlation = 0.91

Cronbach Alpha (KR-20) Participant raw score reliability = 0.98

Table 72 **Item Option and Distractor Frequencies for English PDSS Sleeping/Eating Disturbances Content Scale: Measure Order (N = 187)**

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
29	0	0	109	58	-2.19	.14	1.0	-.61	PDSS_29	0 Strongly Disagree
	1	1	36	19	-.85	.12	1.0	.12		1 Disagree
	2	2	17	9	0.39	.18	0.5	.30		2 Neither Disagree nor Agree
	3	3	13	7	0.04*	.23	1.9	.20		3 Agree
	4	4	12	6	1.94	.40	1.2	.48		4 Strongly Agree
8	0	0	93	50	-2.38	.15	1.1	-.62	PDSS_8	0 Strongly Disagree
	1	1	38	20	-1.02	.14	1.4	.07		1 Disagree
	2	2	17	9	-0.41	.20	1.2	.16		2 Neither Disagree nor Agree
	3	3	28	15	0.26	.18	1.1	.37		3 Agree
	4	4	11	6	1.84	.42	0.9	.44		4 Strongly Agree
22	0	0	87	47	-2.63	.14	0.8	-.72	PDSS_22	0 Strongly Disagree
	1	1	36	19	-1.02	.10	0.6	.07		1 Disagree
	2	2	20	11	-0.04	.17	0.8	.24		2 Neither Disagree nor Agree
	3	3	29	16	0.20	.16	0.8	.36		3 Agree
	4	4	15	8	1.43	.34	0.9	.45		4 Strongly Agree
15	0	0	84	45	-2.66	.14	0.9	-.71	PDSS_15	0 Strongly Disagree
	1	1	36	19	-1.03	.11	0.8	.07		1 Disagree
	2	2	18	10	-0.25	.16	0.7	.19		2 Neither Disagree nor Agree
	3	3	30	16	-0.08	.11	0.7	.30		3 Agree
	4	4	18	10	1.63	.29	0.7	.54		4 Strongly Agree
	MISSING	***	1	1#	-1.24			.00		
1	0	0	61	33	-3.03	.17	1.3	-.69	PDSS_1	0 Strongly Disagree
	1	1	49	26	-1.35	.10	0.6	-.02		1 Disagree
	2	2	17	9	-0.51	.16	0.8	.14		2 Neither Disagree nor Agree
	3	3	42	22	0.07	.17	1.2	.41		3 Agree
	4	4	18	10	0.98	.32	1.1	.42		4 Strongly Agree

* Average ability does not ascend with category score
\# Missing % includes all categories. Scored % only of scored categories

Table 73 Item Option and Distractor Frequencies for English PDSS Anxiety/Insecurity Content Scale: Measure Order (N = 187)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
16	0	0	85	45	-2.31	.19	1.1	-.68	PDSS_16	0 Strongly Disagree
	1	1	44	24	-0.31	.11	0.4	.08		1 Disagree
	2	2	22	12	0.57	.20	1.0	.20		2 Neither Disagree nor Agree
	3	3	26	14	1.52	.20	0.8	.39		3 Agree
	4	4	10	5	3.62	.46	0.9	.46		4 Strongly Agree
30	0	0	84	45	-2.29	.19	1.0	-.67	PDSS_30	0 Strongly Disagree
	1	1	44	24	-0.28	.14	0.7	.09		1 Disagree
	2	2	22	12	0.48	.18	0.8	.19		2 Neither Disagree nor Agree
	3	3	24	13	1.18	.25	2.0	.32		3 Agree
	4	4	13	7	3.40	.38	0.7	.50		4 Strongly Agree
23	0	0	52	28	-3.07	.23	1.0	-.68	PDSS_23	0 Strongly Disagree
	1	1	34	18	-1.13	.18	1.7	-.10		1 Disagree
	2	2	20	11	-0.66	.12	0.4	.00		2 Neither Disagree nor Agree
	3	3	42	22	0.23	.12	0.6	.21		3 Agree
	4	4	39	21	2.05	.23	0.7	.62		4 Strongly Agree
2	0	0	33	18	-3.62	.29	1.1	-.62	PDSS_2	0 Strongly Disagree
	1	1	53	28	-1.50	.16	1.3	-.24		1 Disagree
	2	2	22	12	-0.13	.17	0.9	.08		2 Neither Disagree nor Agree
	3	3	46	25	0.54	.19	1.7	.31		3 Agree
	4	4	33	18	1.68	.30	1.7	.49		4 Strongly Agree
9	0	0	25	13	-4.20	.31	1.4	-.63	PDSS_9	0 Strongly Disagree
	1	1	34	18	-1.98	.17	0.7	-.28		1 Disagree
	2	2	30	16	-0.85	.16	0.8	-.04		2 Neither Disagree nor Agree
	3	3	62	33	0.04	.14	1.0	.22		3 Agree
	4	4	36	19	2.05	.24	0.8	.59		4 Strongly Agree

Table 74 Item Option and Distractor Frequencies for English PDSS Emotional Lability Content Scale: Measure Order (N = 187)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
10	0	0	73	39	-2.49	.19	0.8	-.72	PDSS_10	0 Strongly Disagree
	1	1	40	21	-0.41	.13	0.5	-.02		1 Disagree
	2	2	23	12	0.74	.18	0.7	.17		2 Neither Disagree nor Agree
	3	3	30	16	1.47	.18	0.9	.33		3 Agree
	4	4	21	11	3.57	.35	1.1	.58		4 Strongly Agree
31	0	0	80	43	-2.28	.19	1.4	-.70	PDSS_31	0 Strongly Disagree
	1	1	21	11	-0.70	.19	0.6	-.05		1 Disagree
	2	2	21	11	0.60	.12	0.3	.14		2 Neither Disagree nor Agree
	3	3	39	21	0.97	.19	2.5	.28		3 Agree
	4	4	26	14	3.28	.31	1.0	.60		4 Strongly Agree
17	0	0	62	33	-2.71	.20	1.0	-.69	PDSS_17	0 Strongly Disagree
	1	1	36	19	-0.77	.13	0.5	-.09		1 Disagree
	2	2	20	11	0.36	.20	0.8	.10		2 Neither Disagree nor Agree
	3	3	38	20	1.04	.20	1.9	.29		3 Agree
	4	4	31	17	2.81	.32	1.2	.58		4 Strongly Agree
3	0	0	33	18	-3.45	.27	1.2	-.60	PDSS_3	0 Strongly Disagree
	1	1	38	20	-1.76	.17	0.8	-.30		1 Disagree
	2	2	20	11	-0.68	.25	1.2	-.05		2 Neither Disagree nor Agree
	3	3	53	28	0.54	.14	0.8	.23		3 Agree
	4	4	43	23	2.41	.26	1.0	.62		4 Strongly Agree
24	0	0	23	12	-3.69	.37	1.8	-.52	PDSS_24	0 Strongly Disagree
	1	1	40	21	-2.11	.18	0.9	-.39		1 Disagree
	2	2	25	13	-0.79	.18	0.7	-.07		2 Neither Disagree nor Agree
	3	3	48	26	0.08	.14	0.9	.10		3 Agree
	4	4	51	27	2.41	.21	0.7	.70		4 Strongly Agree

Table 75 **Item Option and Distractor Frequencies for English PDSS Mental Confusion Content Scale: Measure Order (N = 187)**

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
18	0	0	87	47	-3.50	.20	0.9	-.73	PDSS_18	0 Strongly Disagree
	1	1	35	19	-0.87	.14	0.5	.06		1 Disagree
	2	2	19	10	-0.07	.15	0.5	.13		2 Neither Disagree nor Agree
	3	3	25	13	0.76	.23	1.3	.27		3 Agree
	4	4	21	11	4.16	.35	0.6	.66		4 Strongly Agree
32	0	0	59	32	-4.21	.22	1.1	-.69	PDSS_32	0 Strongly Disagree
	1	1	53	28	-1.60	.16	0.9	-.08		1 Disagree
	2	2	22	12	-0.30	.15	0.5	.12		2 Neither Disagree nor Agree
	3	3	38	20	1.03	.20	0.7	.39		3 Agree
	4	4	15	8	4.71	.37	0.8	.60		4 Strongly Agree
11	0	0	49	26	-4.58	.21	1.1	-.68	PDSS_11	0 Strongly Disagree
	1	1	55	29	-1.89	.17	1.2	-.14		1 Disagree
	2	2	27	14	-0.09	.16	0.6	.16		2 Neither Disagree nor Agree
	3	3	43	23	0.96	.22	0.9	.41		3 Agree
	4	4	13	7	4.54	.59	1.0	.54		4 Strongly Agree
4	0	0	64	34	-4.00	.22	1.3	-.68	PDSS_4	0 Strongly Disagree
	1	1	46	25	-1.53	.17	1.0	-.06		1 Disagree
	2	2	24	13	-0.32	.20	1.1	.12		2 Neither Disagree nor Agree
	3	3	31	17	0.60	.24	1.3	.28		3 Agree
	4	4	22	12	3.88	.38	0.7	.64		4 Strongly Agree
25	0	0	49	26	-4.68	.20	1.1	-.70	PDSS_25	0 Strongly Disagree
	1	1	44	24	-1.94	.15	0.6	-.13		1 Disagree
	2	2	32	17	-0.11	.25	1.9	.17		2 Neither Disagree nor Agree
	3	3	42	22	0.34	.21	1.4	.29		3 Agree
	4	4	20	11	3.67	.49	1.2	.58		4 Strongly Agree

Table 76 Item Option and Distractor Frequencies for English PDSS Loss of Self Content Scale: Measure Order (N = 187)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
33	0	0	78	42	-4.60	.18	0.8	-.79	PDSS_33	0 Strongly Disagree
	1	1	45	24	-1.07	.14	0.4	.07		1 Disagree
	2	2	24	13	0.38	.25	1.0	.21		2 Neither Disagree nor Agree
	3	3	21	11	1.46	.22	0.7	.31		3 Agree
	4	4	19	10	4.67	.40	1.5	.62		4 Strongly Agree
19	0	0	82	44	-4.37	.19	0.9	-.76	PDSS_19	0 Strongly Disagree
	1	1	37	20	-1.36	.21	1.2	.02		1 Disagree
	2	2	19	10	0.08	.20	0.6	.16		2 Neither Disagree nor Agree
	3	3	29	16	1.22	.22	0.8	.35		3 Agree
	4	4	20	11	4.67	.32	0.8	.64		4 Strongly Agree
12	0	0	71	38	-4.74	.19	1.1	-.76	PDSS_12	0 Strongly Disagree
	1	1	42	22	-1.37	.19	0.9	.02		1 Disagree
	2	2	28	15	-0.30	.18	0.8	.15		2 Neither Disagree nor Agree
	3	3	27	14	1.37	.29	1.3	.35		3 Agree
	4	4	19	10	4.61	.37	1.1	.61		4 Strongly Agree
26	0	0	69	37	-4.86	.17	0.9	-.77	PDSS_26	0 Strongly Disagree
	1	1	45	24	-1.47	.18	0.8	.00		1 Disagree
	2	2	23	12	-0.12	.16	0.5	.15		2 Neither Disagree nor Agree
	3	3	26	14	1.14	.29	1.3	.32		3 Agree
	4	4	24	13	4.03	.40	1.2	.63		4 Strongly Agree
5	0	0	49	26	-5.25	.21	1.5	-.67	PDSS_5	0 Strongly Disagree
	1	1	46	25	-2.60	.22	1.0	-.19		1 Disagree
	2	2	23	12	-0.94	.23	1.0	.06		2 Neither Disagree nor Agree
	3	3	36	19	0.30	.23	1.6	.26		3 Agree
	4	4	33	18	3.33	.39	1.1	.67		4 Strongly Agree

Table 77 Item Option and Distractor Frequencies for English PDSS Guilt/Shame Content Scale: Measure Order (N = 187)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
27	0	0	92	49	-3.42	.22	1.1	-.71	PDSS_27	0 Strongly Disagree
	1	1	38	20	-0.90	.17	1.0	.06		1 Disagree
	2	2	18	10	0.45	.31	1.5	.18		2 Neither Disagree nor Agree
	3	3	19	10	1.90	.35	1.2	.35		3 Agree
	4	4	20	11	3.59	.40	2.3	.55		4 Strongly Agree
20	0	0	98	52	-3.35	.21	1.6	-.73	PDSS_20	0 Strongly Disagree
	1	1	30	16	-0.71	.15	0.5	.08		1 Disagree
	2	2	12	6	0.47	.24	0.5	.15		2 Neither Disagree nor Agree
	3	3	25	13	1.23	.28	1.2	.32		3 Agree
	4	4	22	12	3.67	.37	1.8	.59		4 Strongly Agree
34	0	0	76	41	-4.11	.19	0.7	-.78	PDSS_34	0 Strongly Disagree
	1	1	42	22	-0.93	.12	0.5	.06		1 Disagree
	2	2	12	6	-0.45	.14	0.5	.07		2 Neither Disagree nor Agree
	3	3	28	15	0.77	.17	0.4	.28		3 Agree
	4	4	29	16	3.52	.27	0.7	.67		4 Strongly Agree
13	0	0	52	28	-4.78	.19	1.3	-.72	PDSS_13	0 Strongly Disagree
	1	1	39	21	-2.09	.20	0.9	-.14		1 Disagree
	2	2	20	11	-0.46	.19	0.9	.09		2 Neither Disagree nor Agree
	3	3	50	27	0.32	.20	1.1	.31		3 Agree
	4	4	26	14	3.45	.31	0.9	.62		4 Strongly Agree
6	0	0	49	26	-4.82	.21	1.5	-.70	PDSS_6	0 Strongly Disagree
	1	1	43	23	-2.29	.19	0.8	-.19		1 Disagree
	2	2	22	12	-0.46	.13	0.5	.10		2 Neither Disagree nor Agree
	3	3	43	23	0.18	.17	0.8	.26		3 Agree
	4	4	30	16	3.45	.25	0.6	.68		4 Strongly Agree

Table 78 Item Option and Distractor Frequencies for English PDSS Suicidal Thoughts Content Scale: Measure Order (N = 187)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDDS	
21	0	0	136	73	-4.01	.08	0.7	-.79	PDSS_21	0 Strongly Disagree
	1	1	20	11	-1.17	.25	1.1	.23		1 Disagree
	2	2	16	9	-0.23	.32	0.8	.31		2 Neither Disagree nor Agree
	3	3	7	4	0.96	.74	1.5	.30		3 Agree
	4	4	8	4	5.34	.35	0.8	.69		4 Strongly Agree
7	0	0	140	75	-3.93	.08	0.8	-.79	PDSS_7	0 Strongly Disagree
	1	1	17	9	-1.57	.17	0.7	.15		1 Disagree
	2	2	10	5	-0.04	.35	0.6	.26		2 Neither Disagree nor Agree
	3	3	11	6	1.12	.46	0.6	.39		3 Agree
	4	4	9	5	4.76	.66	1.4	.68		4 Strongly Agree
14	0	0	138	74	-3.97	.08	0.8	-.79	PDSS_14	0 Strongly Disagree
	1	1	13	7	-1.72	.18	0.6	.12		1 Disagree
	2	2	12	6	-0.66	.21	0.4	.22		2 Neither Disagree nor Agree
	3	3	14	7	0.44	.32	0.7	.37		3 Agree
	4	4	10	5	4.95	.40	0.6	.74		4 Strongly Agree
35	0	0	130	70	-4.07	.07	0.9	-.78	PDSS_35	0 Strongly Disagree
	1	1	21	11	-1.65	.13	0.4	.16		1 Disagree
	2	2	8	4	-1.10	.35	1.2	.14		2 Neither Disagree nor Agree
	3	3	12	6	-0.16	.26	0.7	.28		3 Agree
	4	4	15	8	3.78	.54	0.6	.78		4 Strongly Agree
	MISSING	***	1	1#	-4.15			-.04		
28	0	0	127	68	-4.03	.09	1.9	-.72	PDSS_28	0 Strongly Disagree
	1	1	21	11	-1.92	.26	2.0	.12		1 Disagree
	2	2	13	7	-1.57	.26	1.5	.13		2 Neither Disagree nor Agree
	3	3	9	5	-0.63	.33	1.2	.19		3 Agree
	4	4	17	9	3.26	.61	1.6	.77		4 Strongly Agree

Missing % includes all categories. Scored % only of scored categories

Table 79 Item Option and Distractor Frequencies for Afrikaans PDSS Sleeping/Eating Disturbances Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
29	0	0	112	63	-2.83	.13	1.1	-.72	PDSS_29	0 Strongly Disagree
	1	1	26	15	-0.79	.16	1.0	.25		1 Disagree
	2	2	9	5	-0.14	.22	0.5	.22		2 Neither Disagree nor Agree
	3	3	26	15	0.14	.13	1.0	.47		3 Agree
	4	4	5	3	0.93	.37	0.8	.27		4 Strongly Agree
22	0	0	107	60	-2.96	.12	0.7	-.77	PDSS_22	0 Strongly Disagree
	1	1	29	16	-0.91	.15	0.9	.24		1 Disagree
	2	2	7	4	-0.21	.22	0.4	.19		2 Neither Disagree nor Agree
	3	3	27	15	0.29	.10	0.4	.51		3 Agree
	4	4	8	4	0.84	.26	0.7	.33		4 Strongly Agree
15	0	0	110	62	-2.91	.13	0.8	-.75	PDSS_15	0 Strongly Disagree
	1	1	28	16	-0.81	.12	0.5	.26		1 Disagree
	2	2	6	3	0.13	.28	0.6	.21		2 Neither Disagree nor Agree
	3	3	24	13	0.19	.13	0.7	.46		3 Agree
	4	4	10	6	0.63	.24	0.8	.34		4 Strongly Agree
8	0	0	94	53	-3.13	.13	1.2	-.76	PDSS_8	0 Strongly Disagree
	1	1	40	22	-1.00	.14	1.0	.26		1 Disagree
	2	2	8	4	-0.37	.30	1.1	.18		2 Neither Disagree nor Agree
	3	3	30	17	0.00	.12	1.1	.47		3 Agree
	4	4	6	3	1.20	.27	0.7	.32		4 Strongly Agree
1	0	0	82	46	-3.32	.13	1.3	-.76	PDSS_1	0 Strongly Disagree
	1	1	36	20	-1.46	.16	1.0	.11		1 Disagree
	2	2	15	8	-0.29	.27	1.8	.27		2 Neither Disagree nor Agree
	3	3	34	19	-0.18	.14	1.4	.46		3 Agree
	4	4	11	6	0.42	.17	0.9	.33		4 Strongly Agree

Table 80 Item Option and Distractor Frequencies for Afrikaans PDSS Anxiety/Insecurity Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
16	0	0	81	46	-1.82	.16	0.8	-.70	PDSS_16	0 Strongly Disagree
	1	1	37	21	-0.09	.12	0.6	.14		1 Disagree
	2	2	17	10	0.26	.16	0.9	.16		2 Neither Disagree nor Agree
	3	3	37	21	1.10	.13	0.7	.50		3 Agree
	4	4	6	3	1.65	.21	0.9	.24		4 Strongly Agree
30	0	0	67	38	-1.91	.19	1.5	-.64	PDSS_30	0 Strongly Disagree
	1	1	30	17	-0.40	.20	1.7	.04		1 Disagree
	2	2	25	14	0.38	.13	0.5	.22		2 Neither Disagree nor Agree
	3	3	44	25	0.51	.16	2.2	.36		3 Agree
	4	4	12	7	1.04	.36	1.6	.25		4 Strongly Agree
2	0	0	47	26	-2.35	.22	1.1	-.65	PDSS_2	0 Strongly Disagree
	1	1	40	22	-0.94	.15	1.0	-.13		1 Disagree
	2	2	24	13	-0.34	.13	0.6	.05		2 Neither Disagree nor Agree
	3	3	43	24	0.55	.10	0.6	.37		3 Agree
	4	4	24	13	1.53	.15	0.7	.49		4 Strongly Agree
23	0	0	52	29	-2.42	.19	0.7	-.72	PDSS_23	0 Strongly Disagree
	1	1	19	11	-1.16	.19	0.9	-.13		1 Disagree
	2	2	12	7	0.05	.19	1.0	.09		2 Neither Disagree nor Agree
	3	3	57	32	0.22	.10	0.9	.31		3 Agree
	4	4	38	21	1.02	.15	1.0	.49		4 Strongly Agree
9	0	0	21	12	-3.48	.29	1.2	-.64	PDSS_9	0 Strongly Disagree
	1	1	31	17	-1.67	.16	0.7	-.31		1 Disagree
	2	2	26	15	-0.69	.14	0.6	-.04		2 Neither Disagree nor Agree
	3	3	66	37	0.26	.12	1.2	.37		3 Agree
	4	4	34	19	0.88	.16	1.1	.41		4 Strongly Agree

Table 81 Item Option and Distractor Frequencies for Afrikaans PDSS Emotional Lability Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
10	0	0	76	43	-2.51	.22	1.0	-.72	PDSS_10	0 Strongly Disagree
	1	1	27	15	-0.24	.14	0.5	.05		1 Disagree
	2	2	15	8	0.55	.20	0.8	.14		2 Neither Disagree nor Agree
	3	3	43	24	1.18	.16	1.7	.41		3 Agree
	4	4	17	10	2.67	.40	1.1	.44		4 Strongly Agree
17	0	0	58	33	-2.95	.24	1.1	-.71	PDSS_17	0 Strongly Disagree
	1	1	30	17	-1.02	.21	1.3	-.09		1 Disagree
	2	2	25	14	0.59	.20	1.4	.19		2 Neither Disagree nor Agree
	3	3	44	25	1.02	.12	0.7	.37		3 Agree
	4	4	21	12	2.33	.38	1.2	.44		4 Strongly Agree
31	0	0	57	32	-3.18	.22	0.9	-.77	PDSS_31	0 Strongly Disagree
	1	1	33	19	-0.54	.12	0.4	.00		1 Disagree
	2	2	26	15	0.41	.18	1.1	.16		2 Neither Disagree nor Agree
	3	3	40	22	0.96	.14	0.9	.34		3 Agree
	4	4	22	12	2.58	.32	1.0	.49		4 Strongly Agree
3	0	0	35	20	-3.95	.25	2.2	-.72	PDSS_3	0 Strongly Disagree
	1	1	24	13	-1.85	.25	1.1	-.22		1 Disagree
	2	2	17	10	-0.22	.14	0.7	.04		2 Neither Disagree nor Agree
	3	3	72	40	0.38	.13	1.2	.31		3 Agree
	4	4	30	17	2.20	.27	0.9	.52		4 Strongly Agree
24	0	0	33	19	-4.15	.22	1.4	-.73	PDSS_24	0 Strongly Disagree
	1	1	32	18	-1.61	.19	0.7	-.22		1 Disagree
	2	2	15	8	-0.56	.26	0.9	.00		2 Neither Disagree nor Agree
	3	3	63	35	0.48	.10	0.7	.31		3 Agree
	4	4	35	20	2.11	.26	0.9	.55		4 Strongly Agree

Table 82 Item Option and Distractor Frequencies for Afrikaans PDSS Mental Confusion Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
18	0	0	86	48	-3.82	.19	0.9	-.70	PDSS_18	0 Strongly Disagree
	1	1	47	26	-1.32	.14	0.8	.12		1 Disagree
	2	2	16	9	-0.22	.22	0.7	.19		2 Neither Disagree nor Agree
	3	3	20	11	1.21	.23	0.5	.40		3 Agree
	4	4	9	5	4.27	.82	0.9	.52		4 Strongly Agree
11	0	0	50	28	-4.86	.21	1.0	-.69	PDSS_11	0 Strongly Disagree
	1	1	65	37	-2.09	.13	0.7	-.06		1 Disagree
	2	2	21	12	-0.34	.16	0.4	.21		2 Neither Disagree nor Agree
	3	3	38	21	0.76	.24	1.0	.50		3 Agree
	4	4	4	2	6.36	.71	0.7	.46		4 Strongly Agree
25	0	0	57	32	-4.71	.20	0.9	-.72	PDSS_25	0 Strongly Disagree
	1	1	58	33	-1.85	.11	0.5	.00		1 Disagree
	2	2	16	9	-0.45	.24	0.8	.16		2 Neither Disagree nor Agree
	3	3	40	22	0.48	.22	1.1	.47		3 Agree
	4	4	7	4	4.62	1.06	0.7	.48		4 Strongly Agree
32	0	0	58	33	-4.42	.24	1.7	-.66	PDSS_32	0 Strongly Disagree
	1	1	52	29	-2.08	.14	0.8	-.05		1 Disagree
	2	2	24	13	-0.87	.22	1.1	.14		2 Neither Disagree nor Agree
	3	3	36	20	0.59	.24	1.0	.45		3 Agree
	4	4	8	4	4.12	.96	1.1	.48		4 Strongly Agree
4	0	0	67	38	-4.28	.21	1.1	-.69	PDSS_4	0 Strongly Disagree
	1	1	45	25	-1.74	.16	1.2	.03		1 Disagree
	2	2	18	10	-0.65	.27	1.3	.15		2 Neither Disagree nor Agree
	3	3	36	20	0.26	.25	1.5	.39		3 Agree
	4	4	12	7	2.99	.87	2.4	.48		4 Strongly Agree

Table 83 Item Option and Distractor Frequencies for Afrikaans PDSS Loss of Self Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
33	0	0	86	48	-3.79	.19	1.1	-.72	PDSS_33	0 Strongly Disagree
	1	1	40	22	-1.30	.18	1.1	.08		1 Disagree
	2	2	14	8	-0.18	.27	1.3	.16		2 Neither Disagree nor Agree
	3	3	26	15	1.33	.26	1.3	.44		3 Agree
	4	4	12	7	3.71	.58	1.3	.51		4 Strongly Agree
26	0	0	80	45	-4.13	.17	0.8	-.78	PDSS_26	0 Strongly Disagree
	1	1	35	20	-1.33	.16	0.8	.06		1 Disagree
	2	2	18	10	-0.01	.15	0.3	.20		2 Neither Disagree nor Agree
	3	3	34	19	1.19	.27	1.0	.49		3 Agree
	4	4	11	6	3.21	.72	1.6	.44		4 Strongly Agree
19	0	0	79	44	-4.02	.19	1.1	-.74	PDSS_19	0 Strongly Disagree
	1	1	37	21	-1.57	.16	0.9	.02		1 Disagree
	2	2	15	8	-0.13	.27	0.8	.17		2 Neither Disagree nor Agree
	3	3	33	19	0.73	.20	0.9	.41		3 Agree
	4	4	14	8	3.78	.43	0.7	.56		4 Strongly Agree
12	0	0	71	40	-4.39	.16	0.8	-.78	PDSS_12	0 Strongly Disagree
	1	1	41	23	-1.53	.14	0.6	.03		1 Disagree
	2	2	14	8	-0.42	.18	0.4	.13		2 Neither Disagree nor Agree
	3	3	42	24	0.98	.22	0.9	.52		3 Agree
	4	4	10	6	3.81	.60	1.0	.47		4 Strongly Agree
5	0	0	56	31	-4.86	.13	0.9	-.76	PDSS_5	0 Strongly Disagree
	1	1	41	23	-1.95	.18	1.2	-.05		1 Disagree
	2	2	15	8	-1.11	.27	1.1	.06		2 Neither Disagree nor Agree
	3	3	44	25	0.26	.19	0.9	.39		3 Agree
	4	4	22	12	2.55	.48	1.2	.56		4 Strongly Agree

Table 84 Item Option and Distractor Frequencies for Afrikaans PDSS Guilt/Shame Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
27	0	0	97	54	-3.28	.21	1.1	-.72	PDSS_27	0 Strongly Disagree
	1	1	27	15	-0.82	.17	1.0	.07		1 Disagree
	2	2	15	8	0.96	.29	1.0	.24		2 Neither Disagree nor Agree
	3	3	24	13	1.40	.21	0.9	.37		3 Agree
	4	4	15	8	3.71	.53	3.0	.52		4 Strongly Agree
20	0	0	97	54	-3.35	.20	0.9	-.75	PDSS_20	0 Strongly Disagree
	1	1	25	14	-0.70	.17	0.8	.09		1 Disagree
	2	2	5	3	0.90	.47	1.1	.13		2 Neither Disagree nor Agree
	3	3	36	20	1.21	.18	0.8	.43		3 Agree
	4	4	15	8	3.85	.44	1.1	.53		4 Strongly Agree
34	0	0	61	34	-4.57	.16	0.9	-.79	PDSS_34	0 Strongly Disagree
	1	1	36	20	-1.49	.19	1.1	-.03		1 Disagree
	2	2	15	8	-0.30	.25	0.9	.11		2 Neither Disagree nor Agree
	3	3	44	25	0.69	.14	0.5	.39		3 Agree
	4	4	22	12	3.18	.42	1.0	.57		4 Strongly Agree
6	0	0	61	34	-4.59	.16	1.4	-.80	PDSS_6	0 Strongly Disagree
	1	1	28	16	-1.58	.15	0.5	-.04		1 Disagree
	2	2	15	8	-0.77	.26	1.0	.06		2 Neither Disagree nor Agree
	3	3	51	29	0.48	.14	0.7	.39		3 Agree
	4	4	23	13	3.23	.38	0.8	.60		4 Strongly Agree
13	0	0	56	31	-4.68	.19	4.0	-.77	PDSS_13	0 Strongly Disagree
	1	1	24	13	-1.85	.23	1.0	-.07		1 Disagree
	2	2	22	12	-1.17	.15	0.5	.02		2 Neither Disagree nor Agree
	3	3	50	28	0.43	.15	0.7	.37		3 Agree
	4	4	26	15	2.84	.40	1.0	.58		4 Strongly Agree

Table 85 Item Option and Distractor Frequencies for Afrikaans PDSS Suicidal Thoughts Content Scale: Measure Order (N = 178)

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	r_{it}	PDSS	
7	0	0	135	76	-4.86	.12	1.0	-.79	PDSS_7	0 Strongly Disagree
	1	1	22	12	-1.56	.20	0.9	.30		1 Disagree
	2	2	10	6	0.45	.38	0.6	.38		2 Neither Disagree nor Agree
	3	3	6	3	2.28	.44	0.5	.42		3 Agree
	4	4	5	3	3.32	1.30	5.1	.45		4 Strongly Agree
21	0	0	138	78	-4.82	.11	0.8	-.80	PDSS_21	0 Strongly Disagree
	1	1	18	10	-1.46	.23	1.0	.28		1 Disagree
	2	2	4	2	0.18	.78	1.0	.22		2 Neither Disagree nor Agree
	3	3	13	7	1.15	.38	0.8	.52		3 Agree
	4	4	5	3	3.86	.94	0.8	.49		4 Strongly Agree
14	0	0	131	74	-4.98	.10	0.9	-.82	PDSS_14	0 Strongly Disagree
	1	1	22	12	-1.76	.15	0.6	.28		1 Disagree
	2	2	9	5	-0.11	.26	0.3	.31		2 Neither Disagree nor Agree
	3	3	11	6	1.53	.46	0.8	.51		3 Agree
	4	4	5	3	3.76	.84	1.2	.48		4 Strongly Agree
35	0	0	123	69	-5.15	.09	0.9	-.83	PDSS_35	0 Strongly Disagree
	1	1	24	14	-1.99	.19	0.8	.26		1 Disagree
	2	2	8	5	-0.80	.34	0.7	.24		2 Neither Disagree nor Agree
	3	3	14	8	0.38	.37	1.1	.45		3 Agree
	4	4	8	5	3.45	.59	0.7	.59		4 Strongly Agree
	MISSING	***	1	1#	-2.01			.05		
28	0	0	112	63	-5.34	.07	2.1	-.82	PDSS_28	0 Strongly Disagree
	1	1	25	14	-2.45	.26	1.9	.19		1 Disagree
	2	2	10	6	-1.51	.26	0.7	.20		2 Neither Disagree nor Agree
	3	3	23	13	-0.29	.32	1.2	.50		3 Agree
	4	4	8	4	3.08	.82	1.0	.56		4 Strongly Agree

Missing % includes all categories. Scored % only of scored categories

Table 86 Item Correlations with PDSS Dimensions (N = 365)

PDSS Item	Sleeping / eating disturbances	Anxiety / insecurity	Emotional lability	Mental confusion	Loss of self	Guilt / shame	Suicidal thoughts
Sleeping/Eating Disturbances (SLP)							
PDSS 1	.786**	.488**	.449**	.531**	.457**	.412**	.419**
PDSS 8	.768**	.614**	.582**	.553**	.594**	.584**	.454**
PDSS 15	.832**	.551**	.496**	.563**	.516**	.446**	.453**
PDSS 22	.860**	.584**	.537**	.622**	.584**	.499**	.514**
PDSS 29	.750**	.599**	.560**	.540**	.570**	.563**	.430**
Anxiety/Insecurity (ANX)							
PDSS 2	.616**	.813**	.656**	.623**	.638**	.603**	.433**
PDSS 9	.511**	.815**	.794**	.672**	.665**	.704**	.468**
PDSS 16	.644**	.814**	.708**	.699**	.697**	.615**	.554**
PDSS 23	.564**	.847**	.788**	.674**	.735**	.723**	.524**
PDSS 30	.507**	.726**	.576**	.534**	.536**	.508**	.366**
Emotional Lability (ELB)							
PDSS 3	.552**	.738**	.851**	.727**	.709**	.662**	.447**
PDSS 10	.612**	.768**	.844**	.752**	.815**	.784**	.660**
PDSS 17	.616**	.784**	.850**	.715**	.704**	.700**	.538**
PDSS 24	.496**	.740**	.867**	.688**	.668**	.676**	.479**
PDSS 31	.510**	.714**	.856**	.664**	.659**	.693**	.575**
Mental Confusion (MNT)							
PDSS 4	.559**	.682**	.722**	.847**	.796**	.666**	.609**
PDSS 11	.645**	.693**	.712**	.881**	.759**	.665**	.575**
PDSS 18	.617**	.738**	.741**	.883**	.825**	.686**	.643**
PDSS 25	.627**	.678**	.740**	.864**	.738**	.650**	.563**
PDSS 32	.610**	.671**	.694**	.875**	.725**	.622**	.551**
Loss of Self (LOS)							
PDSS 5	.601**	.708**	.751**	.756**	.876**	.748**	.591**
PDSS 12	.626**	.745**	.774**	.821**	.906**	.755**	.624**
PDSS 19	.583**	.761**	.750**	.785**	.909**	.744**	.598**
PDSS 26	.646**	.729**	.744**	.813**	.905**	.763**	.641**
PDSS 33	.587**	.717**	.719**	.800**	.894**	.745**	.653**
Guilt/Shame (GLT)							
PDSS 6	.586**	.733**	.769**	.700**	.772**	.907**	.606**
PDSS 13	.509**	.720**	.743**	.638**	.717**	.875**	.534**
PDSS 20	.515**	.630**	.664**	.622**	.716**	.866**	.651**

PDSS Item	Sleeping / eating disturbances	Anxiety / insecurity	Emotional lability	Mental confusion	Loss of self	Guilt / shame	Suicidal thoughts
PDSS 27	.549**	.628**	.663**	.673**	.717**	.847**	.708**
PDSS 34	.591**	.759**	.797**	.711**	.772**	.922**	.689**
Suicidal Thoughts (SUI)							
PDSS 7	.499**	.494**	.530**	.600**	.606**	.585**	.919**
PDSS 14	.509**	.506**	.549**	.625**	.606**	.629**	.941**
PDSS 21	.471**	.490**	.559**	.598**	.579**	.594**	.910**
PDSS 28	.553**	.593**	.617**	.625**	.700**	.768**	.850**
PDSS 35	.536**	.555**	.611**	.627**	.637**	.673**	.934**

** Correlation is significant at the 0.01 level (2-tailed).

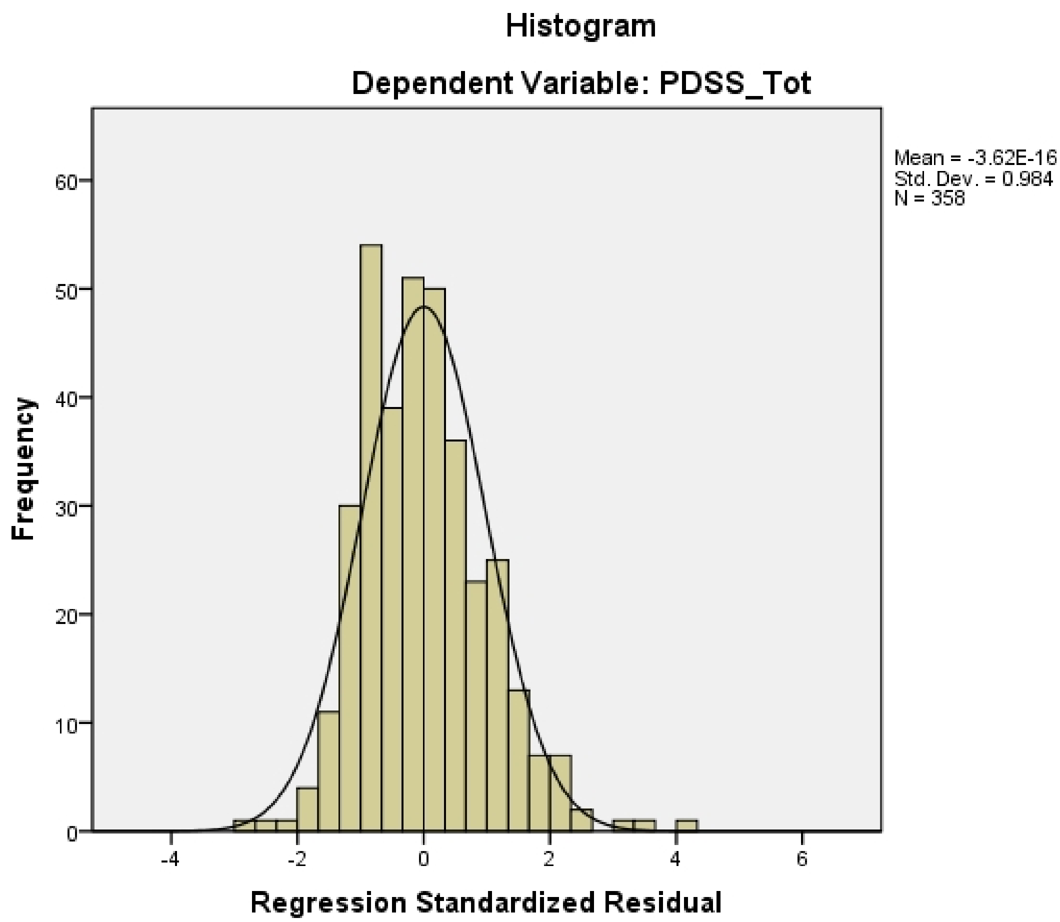

Figure 15 Histogram showing the distribution of the regression standardized residuals.

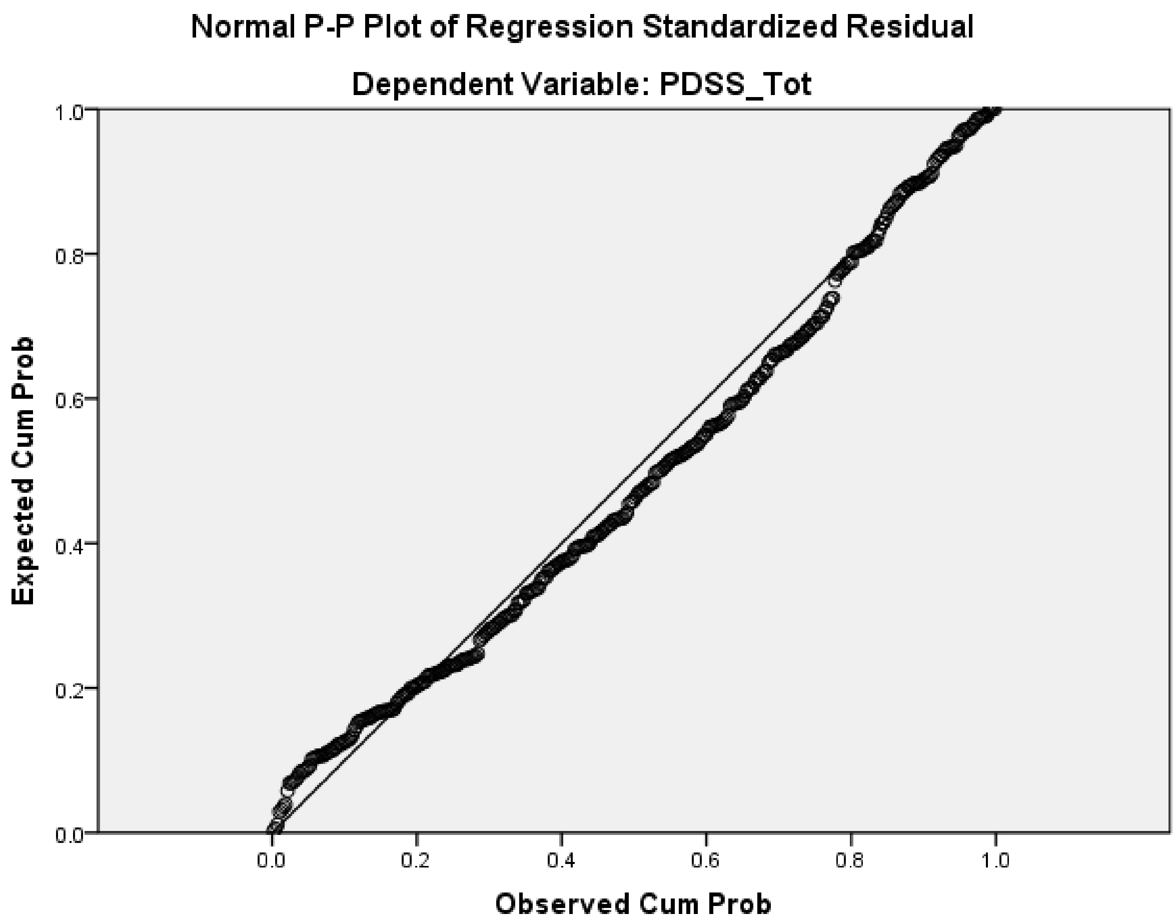

Figure 16 Normal probability plot showing the distribution of the regression standardized residuals.

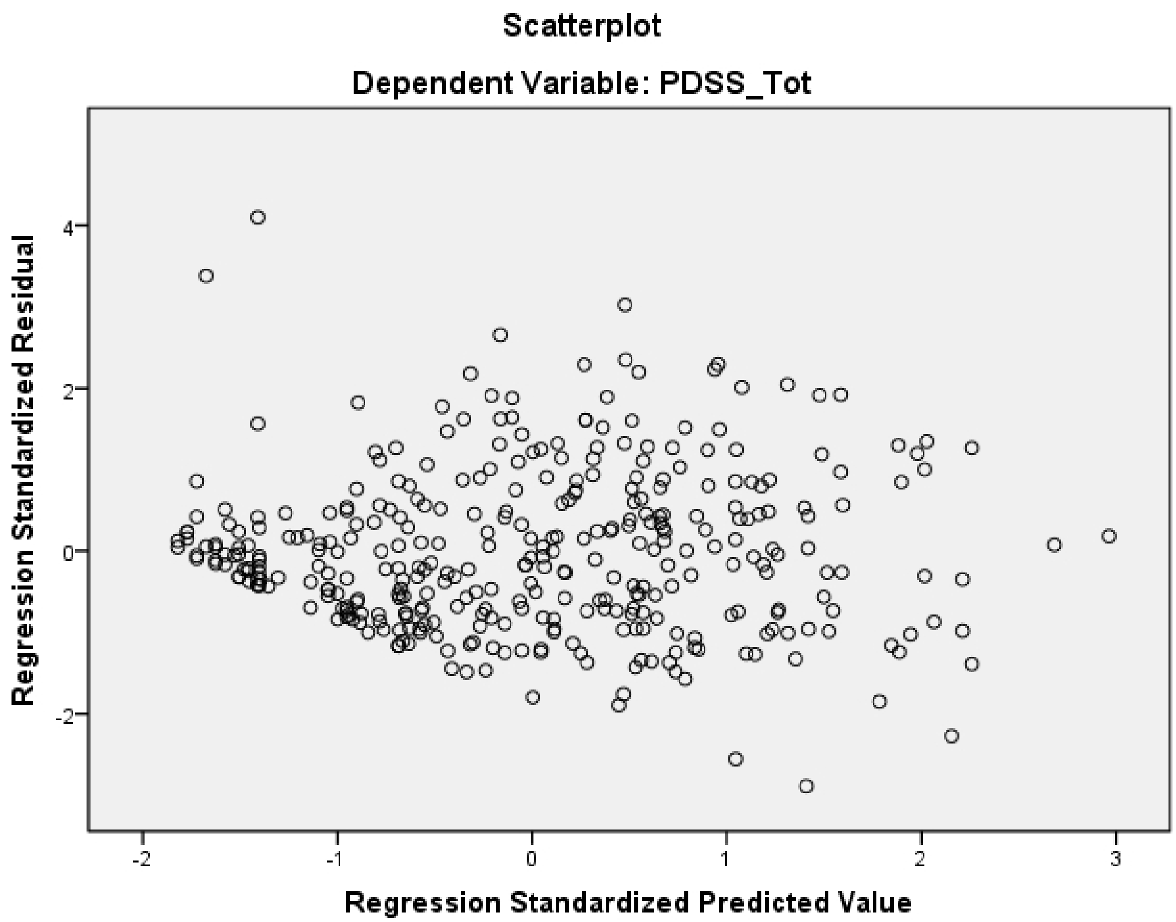

Figure 17 Scatterplot.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

REFERENCES

Abramowitz, J. S., Schwartz, S. A., Moore, K. M., & Luenzmann, K. R. (2003). Obsessive-compulsive symptoms in pregnancy and the puerperium: A review of the literature. *Journal of Anxiety Disorders, 17*, 461-478.

Ægisdóttir, S., Gerstein, L. H., & Çinarbas, D. C. (2008). Methodological Issues in Cross-Cultural Counseling Research : Equivalence, Bias, and Translations. *The Counseling Psychologist, 36*, 188-219.

Affonso, D. D., De, A. K., Horowitz, J. A., & Mayberry, L. J. (2000). An international study exploring levels of postpartum depressive symptomatology. *Journal of Psychosomatic Research, 49*, 207-216.

Ahokas, A., Kuakoranta, J., & Aito, M. (1999). Effect of oestradiol on postpartum depression. *Psychopharmacology, 146*(1), 108-110.

Akman, I., Kuscu, K., Ozdemir, N., Yurdakul, Z., Solakoglu, M., Orhan, L., Karabekiroglu, A., & Ozek, E. (2006). Mothers' postpartum psychological adjustment and infantile colic. *Archives of Disease in Childhood, 91*, 417-419.

Albertsson-Karlgren, U., Graff, M., & Nettelbladt, P. (2001). Mental disease postpartum and parent-infant interaction: Evaluation of videotaped sessions. *Child Abuse Review, 10*, 5-17.

Alkar, O. Y., & Gencoz, T. (2005). Critical factors associated with early postpartum depression among Turkish women. *Contemporary Family Therapy: An International Journal, 27*, 263-275.

Allalouf, A. (2003). Revising translated differential item functioning items as a tool for improving cross-lingual assessment. *Applied Measurement in Education, 16*(1), 55-73. DOI: 10.1207/S15324818AME1601_3

Allalouf, A., & Sireci, S.G. (1998, April). *Detecting Sources of DIF in translated verbal item*. Paper presented at the annual meeting of the American Education Research Association, San Diego, CA.

Altemus, M., Fong, J., Yang, R., Damast, S., Luine, V., & Ferguson, D. (2004). Changes in cerebrospinal fluid neurochemistry during pregnancy. *Biological Psychiatry, 56*(6), 386-92.

American College of Obstetricians and Gynecologists (ACOG) (2002). Clinical updates in women's health care. *Depression in women, 1* (2), 1-82.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, D. C.: Author.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. Text Revision). Washington, DC: Author.

American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations*. Washington, D. C.: Author.

American Psychological Association. (1993). *Responsible test use: Case studies for assessing human behaviour*. Washington, D. C.: Author.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Mac Millan Publishing Co., Inc.

Anastasi, A., & Urbina, S. (1997). *Psychological testing (7th ed.)*. New Jersey: Prentice Hall, Inc.

Andrich, D. (2004a). Controversy and the Rasch model: A characteristic of incompatible paradigms? In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: Theory, models and applications* (pp. 143-166). Maple Grove, Minnesota: JAM Press.

Andrich, D. (2004b). Understanding resistance to the data-model relationship in Rasch's paradigm: a reflection for the next generation. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: Theory, models and applications* (pp. 143-166). Maple Grove, Minnesota: JAM Press.

Archibald, D. (1969). The Afrikaners as an emergent minority. *British Journal of Sociology, 20*, 416-426.

Ardila, A. (1995). Cross-cultural neuropsychology. *Journal of Clinical and Experimental Psychology, 17*, 143-150.

Areias, M. E. G., Kumar, R., Barros, H., & Figueiredo, E. (1996). Correlates of postnatal depression in mothers and fathers. *The British Journal of Psychiatry, 169*, 36-41

Arizmendi, T. G., & Affonso, D. D. (1987). Sressful events related to pregnancy and postpartum. *Journal of Psychosomatic Research, 31*, 743-756.

Armstrong, K., O'Donnell, H., McCallum, R., & Dadds, M. (1998). Childhood sleep problems: Association with prenatal factors and maternal distress/depression. *Journal of Pediatric and Child Health, 34*, 263-266.

Armstrong, K., van Haeringen, A., Dadds, M., & Cash, R. (1998). Sleep deprivation or postnatal depression in later infancy: Separating the chicken from the egg. *Journal of Pediatric and Child Health, 34*, 260-262.

Asten, P., Marks, M. N., & Oates, M. R. (2004). Aims, measures, study sites and participant samples of the Transcultural Study of Postnatal Depression. *British Journal of Psychiatry, 184*(46), s3-9.

Austin, M. P. (2003). Psychosocial assessment and management of depression and anxiety in pregnancy: Key aspects of antenatal care for general practice. *Aust Family Physician, 32*, 119-126.

Austin, M. P., & Lumley J. (2003). Antenatal screening for postnatal depression: A systematic review. *Acta Psychiatrica Scandinavica, 107*, 10-17.

Austin, M. P., & Lumley, J. (2002). Antenatal screening for women at risk of postnatal depression. *Acta Psychiatrica Scandinavica, 106*, 1-8.

Austin, M. P., & Priest, S. R. (2005). Clinical issues in perinatal mental health: New developments in the detection and treatment of perinatal mood and anxiety disorders. *Acta Psychiatrica Scandinavica, 112*, 97-104.

Austin, M. P., Hadzi-Pavlovic, D., Leader, L., Saint, K., & Parker, G. (2005). Maternal trait anxiety, depression and life event stress in pregnancy: Relationships with infant temperament. *Early Human Development, 81*, 183-190.

Aydin, N., Inandi, T., & Karabulut, N. (2005). Depression and associated factors among women within their first postnatal year in Erzurum province in Eastern Turkey. *Women and Health, 41*, 1-12.

Aydin, N., Inandi, T., Yigit, A., & Hodoglugil, N. N. S. (2004). Validation of the Turkish version of the Edinburgh Postnatal Depression Scale among women within their first postpartum year. *Social Psychiatry and Psychiatric Epidemiology, 39*, 483-486.

Baker, L., & Oswalt, K. (2008). Screening for postpartum depression in a rural community. *Community Mental Health Journal, 44*, 171-180.

Baker, S. R., Henshaw, C., & Choi, P. Y. L. (2003). "I felt as though I'd been in jail": Women's experiences of care during labour and birth. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109.

Ballard, C. G., Stanley, A. K., & Brockington, I. F. (1995). Post-traumatic stress disorder (PTSD) after childbirth. *British Journal of Psychiatry, 166*, 525-528.

Banti, S., Mauri, M., Oppo, A., Borri, C., Rambelli, C., Ramacciotti, D., … Cassano, G. B. (2011). From the third month of pregnancy to 1 year postpartum. Prevalence, incidence, recurrence, and new onset of depression. Results from the Perinatal Depression–Research & Screening Unit study. *Comprehensive Psychiatry, 52*(4), 343-351.

Barnett, B., Joffe, A., Duggan, A. K., Wilson, M. D., Repke, J. T. (1996). Depression symptoms, stress, and social support in pregnant and postpartum adolescents. *Archives of Pediatric & Adolescent Medicine, 150*, 64-69.

Barnett, B., Matthey, S., & Gyaneshwar, R. (1999). Screening for postnatal depression in women of non-English speaking background. *Archives of Womens Mental Health, 2*, 67-74.

Baron, R. A., & Byrne, D. (1994). *Social psychology: Understanding human interaction.* New York: Allyn and Bacon.

Barr, J. (2003). Women with PND share their darkest secrets: women talk about thoughts of infanticide. *Archives of Womens Mental Health, 6* (Suppl.2), s83-s109

Barr, J. A. (2008). Postpartum depression, delayed maternal adaptation, and mechanical infant caring: A phenomenological hermeneutic study. *International Journal of Nursing Studies, 45*, 362-369.

Barrett, K. H., & George, W. H. (Eds.). (2004). *Race, culture, psychology, & law.* Thousand Oaks, CA: Sage Publications.

Bashiri, N., & Spielvogel, A. M. (1999) Postpartum depression: a cross-cultural perspective. *Primary Care Update for OB/GYNS, 6*(3), 82-87.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

Beck, C. T. (1992). The lived experience of postpartum depression: A phenomenological study. *Nursing Research, 41*, 166-170.

Beck, C. T. (1993). Teetering on the edge: A substantive theory of postpartum depression. *Nursing Research, 42*, 42-48.

Beck, C. T. (1995). The effects of postpartum depression on maternal-infant interaction: A meta-analysis. *Nursing Research, 42*, 42-48.

Beck, C. T. (1996a). A meta-analysis of predictors of postpartum depression. *Nursing Research, 45*, 297-303.

Beck, C. T. (1996b). A meta-analysis of the relationships between postpartum depression and infant temperament. *Nursing Research, 45*, 225-230.

Beck, C. T. (1996c). Postpartum Depressed Mothers' Experiences Interacting With Their Children. *Nursing Research, 45*, 98-104.

Beck, C. T. (1998a). Postpartum onset of panic disorder. *Image: Journal of Nursing Scholarship, 30* (2), 131-135.

Beck, C. T. (1998b). The effects of postpartum depression on child development: A meta-analysis. *Archives of Psychiatric Nursing, 12*, 12-20.

Beck, C. T. (2001). Predictors of postpartum depression: an update. *Nursing Research, 50*, 275-285.

Beck, C. T. (2003). Recognizing and screening for postpartum depression in mothers of NICU infants. *Advances in Neonatal Care, 3*(1), 37-46.

Beck, C. T. (2006). Postpartum Depression: It isn't just the blues. *American Journal of Nursing, 106* (5), 40-50.

Beck, C. T., & Gable, R. K. (2000). Postpartum Depression Screening Scale: development and psychometric testing. *Nursing Research, 49*, 272-282.

Beck, C. T., & Gable, R. K. (2001a). Comparative analysis of the performance of the Postpartum Depression Screening Scale with two other depression instruments. *Nursing Research, 50*, 242-250.

Beck, C. T., & Gable, R. K. (2001b) Ensuring content validity: an illustration of the process. *Journal of Nursing Measurement, 9*, 201-215.

Beck, C. T., & Gable, R. K. (2001c). Further validation of the Postpartum Depression Screening Scale. *Nursing Research, 50*, 155-164.

Beck, C. T., & Gable, R. K. (2001d). Item response theory in affective instrument development: An illustration. *Journal of Nursing Measurement, 9*(1), 5-22.

Beck, C. T., & Gable, R. K. (2002). *Postpartum depression screening scale (PDSS) manual*. Los Angeles: Western Psychological Services.

Beck, C. T., & Gable, R. K. (2003). Postpartum Depression Screening Scale: Spanish version. *Nursing Research, 52*(5), 296-306.

Beck, C. T., & Indman, P. (2005). The many faces of postpartum depression. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 34* (5), 569-576.

Beck, C. T., Bernal, H., & Froman, R. D. (2003). Methods to document semantic equivalence of a translated scale. *Research in Nursing and Health, 26*, 64-73.

Behling, O. & Law, K. S. (2000). *Translating Questionnaires and Other Research Instruments: Problems and Solutions*. United Kingdom: SAGE Publications Inc.

Belle, D. (1990). Poverty and women's mental health. *The American Psychologist, 45,* 385-389.

Bender, E. (2003). Depression in pregnancy often goes untreated. *Psychiatric News, 38*(40), p. 20.

Bennett, S. S., & Indman, P. (2003). *Beyond the Blues: A guide to understanding and treating prenatal and postpartum depression*. Retrieved May 25, 2006, from MFT Moodswings Press, Postpartum Support International Web site: www.postpartum.net/brief.html

Benvenuti, P., Ferrara, M., Niccolai, C., Valoriani, V., & Cox, J. L. (1999). The Edinburgh Postnatal Depression Scale: Validation for an Italian sample. *Journal of Affective Disorders, 53*, 137-141.

Berle, J. O., Aarre, T. F., Mykletun, A., Dahl, A. A., & Holsten, F. (2003). Screening for postnatal depression: Validation of the Norwegian version of the Edinburgh Postnatal Depression Scale, and assessment of risk factors for postnatal depression. *Journal of Affective Disorders, 76*, 151-156.

Bernstein, I. H., Rush, A. J., Yonkers, K., Carmody, T. J., Woo, A., McConnell, K., & Trivedi, M. H. (2008). Symptom features of postpartum depression: Are they distinct? *Depression and Anxiety, 25*, 20-26.

Beydoun, H. A., Al-Sahab, B., Beydoun, M. A., & Tamim, H. (2010). Intimate partner violence as a risk factor for postpartum depression among Canadian women in the Maternity Experience Survey. *Annals of Epidemiology, 20*, 575-583.

Bhamjee, R. A. (1991). *A comparison of the performance of normal British and South African Indian children on the Griffiths Scales of Mental Development.* Unpublished doctoral thesis, University of Port Elizabeth, Port Elizabeth, South Africa.

Bielawska-Batorowicz, E., & Kossakowska-Petrycka, K. (2006). Depressive mood in men after the birth of their offspring in relation to a partner's depression, social support, fathers' personality and prenatal expectations. *Journal of Reproductive and Infant Psychology, 24*, 21-29.

Bina, R. (2008). The Impact of Cultural Factors Upon Postpartum Depression: A Literature Review. *Health Care for Women International, 29*, 568-592

Biondi, M., Picardi, A., Pasquini, M., Gaetano, P., & Pancheri, P. (2005). Dimensional psychopathology of depression: Detection of an 'activation' dimension in unipolar depressed outpatients. Journal of Affective Disorders, 84(2-3), 133-139.

Birbili, M. (2000). Translating from one language to another. *Social Research Update, 31*. Retrieved from http://www.soc.surrey.ac.uk/sru/SRU31.html

Birkeland, R., Thompson, J. K., & Phares, V. (2005). Adolescent motherhood and postpartum depression. *Journal of Clinical Child and Adolescent Psychology, 34*, 292-300.

Bloch, M., & Klein, E. (2005). Risk factors associated with the development of postpartum mood disorders. *Journal of Affective Disorders, 88*, 9-18.

Bloch, M., Daly, R. C., & Rubinow, D. R. (2003). Endocrine factors in the etiology of postpartum depression. *Comprehensive Psychiatry, 44*(3), 234-246.

Bloch, M., Rotenberg, N., Koren, D., & Klein, E. (2005). Risk factors associated with the development of postpartum mood disorders. *Journal of Affective Disorders, 88*, 9-18.

Bloch, M., Schmidt, P. J., Danaceau, M., Murphy, J., Nieman, L., & Rubinow, D. R. (2000). The effects of gonadal steroids in women with a history of postpartum depression. *American Journal of Psychiatry, 157*(6), 924-930.

Bokhari, R., Bhatara, V. S., Bandettini, F., & McMillin, J. M. (1998). Postpartum psychosis and postpartum thyroiditis. *Psychoneuroendocrinology, 23*(6), 643-650.

Bond , T. G. , & Fox , C. (2007). *Applying the Rasch model* (2nd ed.). Mahwah, NJ: Erlbaum.

Bond, T. G. (2003). Validity and assessment: a Rasch measurement perspective. *Metodología de las Ciencias del Comportamiento, 5*(2), 179-194.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* London: Lawrence Erlbaum Associates.

Boyce, P., & Hickey, A. (2005). Psychosocial risk factors to major depression after childbirth. *Social Psychiatry and Psychiatric Epidemiology, 40*, 605-612.

Boyce, P., Stubbs, J., & Todd, A. (1993). The Edinburgh Postnatal Depression Scale: Validation for an Australian sample. *The Australian and New Zealand Journal of Psychiatry, 27*, 472-476.

Boyd, R. C., Le, H. N., & Somberg, R. (2005). Review of screening instruments for postpartum depression. *Archives of Women's Mental Health, 8*, 141-153

Boyd, R. C., Pearson, J. L., & Blehar, M. C. (2002). Prevention and treatment of depression in pregnancy and the postpartum periodsummary of maternal depression roundtable: a US perspective. *Archives of Womens Mental Health, 4*, 79-82.

Boyd-Bragadeste, K. L. (1998). A biopsychosocial approach to understanding postpartum adjustment. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 59* (6-B), p. 3048.

Brace, N., Kemp, R., & Snelgar, R. (2009). *SPSS for Psychologists* (4th ed.). London, UK: Palgrave Macmillan.

Bracken, B. A., & Barona, A. (1991). State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International, 12*, 119-132.

Brennan, P. A., Hammen, C., Andersen, M. J., Bor, W., Najman, J. M., & Williams, G. M. (2000). Chronicity, severity, and timing of maternal depressive symptoms: Relationship with child outcomes at age 5. *Developmental Psychology, 36*(6), 759-766.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1* (3), 185-216.

Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology, 11*, 215-229.

Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In H. C. Triandis & H. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 389-444). Boston: Allyn & Bacon.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137-164). Newbury Park, CA: Sage Publications.

Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Crosscultural research methods*. New York: Wiley.

British Psychological Society (2006). *Code of ethics and conduct*. Leicester: British Psychological Society. Retrieved from http://www.bps.org.uk

British Psychological Society (2007, August 3). *Conducting research on the internet: Guidelines for ethical practice in psychological research online*. Leicester: British Psychological Society. Retrieved from http://www.bps.org.uk/sites/default/files/documents.conducting_research_on_the_internet_guidelines_for_ethical_practice_in_psychological_research_online.pdf

British Psychological Society (2009, September 25). *Ethical Principles for Conducting Research with Human Participants*. Leicester: British Psychological Society. Retrieved from http://www.bps.org.uk

Brockington, I. F., Cernik, K. F., Scho¢eld, E. M., Downing, A. R., Francis, A. F., & Keelan, C. (1981). Puerperal psychosis: Phenomena and diagnosis. *Archives of General Psychiatry, 38*, 829-833.

Brockington, I. F., Macdonald, E., & Wainscott, G. (2006). Anxiety, obsessions and morbid preoccupations in pregnancy and the puerperium. *Archives of Women's Mental Health, 9*, 253-263.

Brockington, I. F., Oates, J., George, S., Turner, D., Vostanis, P., Sullivan, M., Loh, C., & Murdoch, C. (2001). A screening questionnaire for mother-infant bonding disorders. *Archives of Women's Mental Health, 3*, 133-140.

Brockington, I. F., Schofield, E. M., Donnelly, P., & Hyde, C. (1978). A clinical study of post-partum psychosis. In M. Sandler (Ed.), *Mental illness in pregnancy and the puerperium* (pp.40-51). Oxford, London: Oxford University Press.

Brockington, I., 2004. Postpartum psychiatric disorders. *Lancet, 363*(9405), 303-310.

Brouwers, E. P., van Baar, A. L., & Pop, V. J. (2001). Does the Edinburgh Postnatal Depression Scale measure anxiety? *Journal of Psychosomatic Research, 51*(5), 659-63.

Buist, A., Barnett, B. E., Milgrom, J., Pope. S., Condon, J. T., Ellwood, D. A., Boyce, P. M., Austin, M. P., & Hayes, B. A. (2002). To screen or not to screen – that is the question in perinatal depression. *Medical Journal of Australia, 177*(suppl.), 101-105.

Burke, L. (2003). The impact of maternal depression on familial relationships. *International Review of Psychiatry, 15*, 243-255.

Byrne, B. M., & Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology, 30*(5), 555-574.

Campbell, S. B., & Cohn, J. F. (1991). Prevalence and correlates of postpartum depression in first-time mothers. *Journal of Abnormal Psychology, 100*, 594-599.

Cantilino, A., Carvalho, J. A., Maia, A., Albuquerque, C., Cantilino, G., & Sougey, E. B. (2007). Translation, validation and cultural aspects of Postpartum Depression Screening Scale in Brazilian Portuguese. *Transcultural Psychiatry, 44*(4), 672-684. doi:10.1177/1363461507083904

Canuso, R. (2008). Maternal depression: The "dual" diagnosis of mother and child. *Issues in Mental Health Nursing, 29*, 785-787.

Carter, A. S., Baker, C. W., & Brownell, K. D. (2000). Body mass index, eating attitudes, and symptoms of depression and anxiety in pregnancy and the postpartum period. *Psychosomatic Medicine, 62*, 264-270.

Carter, A. S., Garrity-Rokous, F. E. J. D., Chazan-Cohen, R., Little, C., & Briggs-Gowan, M. J. (2001). Maternal depression and comorbidity: Predicting early parenting, attachment security, and toddler social-emotional problems and competencies. *American Academy of Child and Adolescent Psychiatry, 40*, 18-26.

Carter, F. A., Frampton, C. M. A., Mulder, R. T. (2006). Cesarean Section and Postpartum Depression: A Review of the Evidence Examining the Link. *American Psychosomatic Society, 68*, 321-330.

Chabrol, H., Teissedre, F., Saint-Jean, M., Teisseyre, N., Roge, B., & Mullet, E. (2002). Prevention and treatment of post-partum depression: A controlled randomized study on women at risk. *Psychological Medicine, 32*, 1039-1047.

Chan, S., & Levy, V. (2004). Postnatal depression: A qualitative study of the experiences of a group of Hong Kong Chinese women. *Journal of Clinical Nursing, 13*, 120-123.

Chandra, P. S., Vankatasubramanian, G., & Thomas, T. (2002). Infanticidal ideas and infanticidal behavior in Indian women with severe postpartum psychiatric disorders. *Journal of Nervous and Mental Disease, 190*, 457-461.

Chandran, M., Tharyan, P., Muliyil, J., & Abraham, S. (2002). Post-partum depression in a cohort of women from a rural area of Tamil Nadu, India. Incidence and risk factors. *British Journal of Psychiatry, 181*, 499-504.

Chee, C. Y. I., Lee, D. T. S., Chong, Y. S., Tan, L. K., Ng, T. R, & Fones, C. S. L. (2005). Confinement and other psychosocial factors in perinatal depression: A transcultural study in Singapore. *Journal of Affective Disorders, 89*, 157-166.

Chiang, K. S., Green, K. E., & Cox, E. O. (2009). Rasch analysis of the Geriatric Depression Scale – Short Form. *The Gerontologist, 49*(2), 262-275.

Chisholm, D., Conroy, S., Glangeaud-Freudenthal, N., Oates, M. R., Asten, P., Barry, S., Figueiredo, B., Kammerer, M. H., Klier, C. M., Seneviratne, G., & Sutter-Dallay, A. (2004). Health services research into postnatal depression: results from a preliminary cross-cultural study. *The British Journal of Psychiatry, 184*, s45-52.

Chou, Y., & Wang, W. (2010). Checking dimensionality in item response models with principal component analysis on standardized residual. *Educational and Psychological Measurement, 70*(5), 717-731.

Chung, T., Lau, T. K., Alexander, S. K., Chiu, H., & Lee, D. (2001). Antepartum depressive symptomatology is associated with adverse obstetric and newnatal outcomes. *Psychosomatic Medicine, 63*, 830-834.

Clay, E. C., & Seehusen, D. A. (2004). A review of postpartum depression for the primary care physician. *Southern Medical Journal, 97*(2), 157-161.

Clemmens, D., Driscoll, J. W., & Beck, C.T. (2004). Postpartum depression as profiled through the depression screening scale. *The American Journal of Maternal/Child Nursing, 29*, 180-185.

Condon, J. T. & Corkindale, C. J. (1997). The assessment of depression in the postnatal period: a comparison of four self-report questionnaires. *Australian and New-Zealand Journal of Psychiatry, 31*, 353-359.

Cooper, M., Turpin, G., Bucks, R., & Kent, G. (2005). *Good practice guidelines for the conduct of psychological research within the NHS*. Leicester: The British Psychological Society.

Cooper, P. J., & Murray, L. (1998). Postnatal depression. *British Medical Journal (Clinical research ed.). 316*,(7148), 1884-1886.

Cooper, P. J., Murray, L., & Hooper, R., & West, A. (1996). The development and validation of a predictive index for postpartum depression. *Psychological Medicine, 26*(3), 627-634.

Cooper, P. J., Tomlinson, M., Swartz, L., Woolgar, M., Murray, L., & Molteno, C. (1999). Post-partum depression and the mother-infant relationship in a South African peri-urban settlement. *The British Journal of Psychiatry, 175*, 554-558.

Coplan, R. J., O'Neil, K., & Arbeau, K. A. (2005). Maternal anxiety during and after pregnancy and infant temperament at three months of age. *Journal of Prenatal and Perinatal Psychology and Health, 19*, 199-215.

Cornish, A. M., McMahon, C. A., Ungerer, J. A., Barnett, B., Kowalenko, N., & Tennant, C. (2005). Postnatal depression and infant cognitive and motor development in the second postnatal year: The impact of depression chronicity and infant gender. *Infant Behavior and Development, 28*, 407-417.

Cortina, J.M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 96-104.

Cox, J. (2005). Postnatal depression in fathers. *Lancet, 366*, 982.

Cox, J. L. (1996). Perinatal mental disorder: A cultural approach. *International Review of Psychiatry, 8*, 9-16.

Cox, J. L. (1999). Perinatal mood disorders in a changing culture: A transcultural European and African perspective. *International Review of Psychiatry, 11*, 103-110.

Cox, J. L. (2002). Commentary towards a more integrated international system of psychiatric classification. *Psychopathology, 35*, 195-196.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry, 150*, 782-786.

Craig, M. (2004). Perinatal risk factors for neonaticide and infant homicide: Can we identify those at risk? *Journal of the Royal Society of Medicine, 97*(2), 57-61.

Cramer, D. (2003). *Advanced Quantitative Data Analysis*. Glasgow: Bell and Bain Ltd.

Crockenberg, S. C., & Leerkes, E. M. (2003). Parental acceptance, postpartum depression, and maternal sensitivity: Mediating and moderating processes. *Journal of Family Psychology, 17*, 80-93.

Cryan, E., Keogh, F., Connolly, E., Cody, S., Quinlan, A., & Daly, I. (2001). Depression among postnatal women in an urban Irish community. *Irish Journal of Psychological Medicine, 18*(1), 5-10.

Cushner, K., &, Brislin, R. W. (1996). *Intercultural interactions: A practical guide* (2nd ed.). USA: Sage Publications.

Czarkowski, K. (1999). Postpartum Depression and the "Baby Blues". *American Family Physician, 15*, 1259-1263.

Czarnocka, J., & Slade, P. (2000). Prevalence and predictors of post-traumatic stress symptoms following childbirth. *The British Journal of Clinical Psychology, 39*, 35-51.

Da Costa, D., Larouche, J., Dritsa, M., & Brender, W. (2000). Psychosocial correlates of prepartum and postpartum depressed mood. *Journal of Affective Disorders, 59*, 31-40.

Da Silva Magalhães, P. V., Pinheiro, R. T., Horta, B. L., Pinheiro, K. A. T., & Da Silva, R. A. (2008). Validity of the Beck Depression Inventory in the postpartum period. *International Journal of Psychiatry in Clinical Practice, 12*, 81-84.

Dalton, K. (1996). *Depression after childbirth*. New York: Oxford Unversity Press.

Dalton, K., & Holton, W. (2001). *Depression after childbirth* (4th ed). New York: Oxford University Press.

Dankner, R., Goldberg, R. P., Fisch, R. Z., & Crum, R. M. (2000). Cultural elements of postpartum depression. A study of 327 Jewish Jerusalem women. *Journal of Reproductive Medicine, 45*(2), 97-104.

Davey, S. J., Dziurawiec, S., & O'Brien-Malone, A. (2006). Men's voices: Postnatal depression from the perspective of male partners. *Qualitative Health Research, 16*, 206-220.

De Bruin, G. P., Swartz, L., Tomlinson, M., Cooper, P. J., & Molteno, C. (2004). The factor structure of the Edinburgh Postnatal Depression Scale in a South African peri-urban settlement. *South African Journal of Psychology, 34*(1), 113-121.

De Klerk, W. (1984). *The second (r)evolution. Afrikanerdom and the crisis of identity.* Johannesburg: Jonathan Ball.

Dearing, E., Taylor, B. A., & McCartney, K. (2004). Implications of family income dynamics for women's depressive symptoms during the first 3 years after childbirth. *American Journal of Public Health, 94*, 1372-1377.

Dennis, C. L. E., Janssen, P. A., & Singer, J. (2004). Identifying women at-risk for postpartum depression in the immediate postpartum period. *Acta Psychiatrica Scandinavica, 110*, 338-346.

Dennis, C. L., & Boyce, P. (2004). Further psychometric testing of a brief personality scale to measure vulnerability to postpartum depression. *Journal of Psychosomatic Obstetrics and Gynecology, 25*, 305-311.

Dennis, C. L., & Ross, L. (2006). Women's perceptions of partner support and conflict in the development of postpartum depressive symptoms. *Journal of Advanced Nursing, 56*(6), 588-599.

De-Vriese, S. R., Christophe, A. B., & Maes, M. (2003). Lowered serum n-3 polyunsaturated fatty acid (PUFA) levels predict the occurrence of postpartum depression: Further evidence that lowered n-PUFAs are related to major depression. *Life Sciences, 73(25)*, 3181-3187.

Diego, M. A., Field, T., Hernandez-Reif, M., Cullen, C., Schanberg, S., & Kuhn, C. (2004). Prepartum, postpartum, and chronic depression effects on newborns. *Psychiatry: Interpersonal and Biological Processes, 67*, 63-80.

Doucet, S., Dennis, C., Letourneau, N., & Blackmore, E. R. (2009). Differentiation and clinical implications of postpartum depression and postpartum psychosis. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 38*(3), 269-279.

Downing, S. M., & Haladyna, T. M. (2004). Validity threats: Overcoming interference with proposed interpretation of assessment data. *Medical Education, 38*, 327-333.

Drewett, R., Blair, P., Emmett, P., & Emond, A. (2004). Failure to thrive in the term and preterm infants of mothers depressed in the postnatal period: A population-based birth cohort study. *Journal of Child Psychology and Psychiatry, 45*, 359-366.

Eberhard-Gran, M., Eskild, A., Tambs, K., Opjordsmoen, S., & Samuelson, S. O. (2001). Review of the validation studies of the Edinburgh Postnatal Depression Scale. *Acta Psychiatrica Scandinavica, 104*, 243-249.

Eberhard-Gran, M., Eskild, A., Tambs, K., Samuelsen, S. O., & Opjordsmoen, S. (2002). Depression in postpartum and nonpostpartum women: Prevalence and risk factors. *Acta Psychiatrica Scandinavica, 106*(6), 426-433.

Edelen, M. O., Thissen, D., Teresi, J. A., Kleinman, M., & Ocepek-Welikson, K. (2006). Identification of differential item functioning using Item Response Theory and the Likelihood-Based Model Comparison Approach Application to the Mini-Mental State Examination. *Medical Care, 44*(11), Suppl 3, S134-S142.

Edge, D., Baker, D., & Rogers, A. (2004). Perinatal depression among black Caribbean women. *Health and Social Care in the Community, 12*(5), 430-438.

Edhborg, M., Seimyr, L., Lundh, W., & Widstroem, A. M. (2000). Fussy child – difficult parenthood? Comparisons between families with a "depressed" mother and non-depressed mother 2 months postpartum. *Journal of Reproductive and Infant Psychology, 18*, 225-238.

England, S. J., Ballard, C., & George, S. (1994). Chronicity in postnatal depression. *European Journal of Psychiatry, 8*, 93-96.

Epperson, C. N. (1999). Postpartum Major Depression: Detection and Treatment. *American Family Physician, 15*, 2247-2259.

Epperson, C. N., Gueorguieva, R., Czarkowski, K. A., Stiklus, S., Sellers, E., Krystal, J. H., Rothman, D. L., & Mason, G. F. (2006). Preliminary evidence of reduced occipital GABA concentrations in puerperal women: A 1H-MRS study. *Psychopharmacology, 186*(3), 425-433.

Epperson, N., Mason, G., Sanacora, G., Haga, K., Appel, M., Czarkowski, K., Rothman, D. L., & Krystal, J. H. (2003). Cortical GABA levels in healthy and depressed puerperal and non-puerperal women: A 1H-MRS study. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109.

Eremenco, S. L., Cella, D., & Arnold, B. J. (2005). A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation & the Health Professions, 28*, 212-232.

Essex, M. J., Klein, M. H., Miech, R., & Smider, N. A. (2001). Timing of initial exposure to maternal major depression and children's mental health symptoms in kindergarten. *British Journal of Psychiatry,* 179, 151-156.

Evans, J., Heron, J., Francomb, H., Oke, S., & Golding, J. (2001). Cohort study of depressed mood during pregnancy and after childbirth. *British Medical Journal, 323*, 257-260.

Evins, G. G., Theofrastous, J. P., & Galvin, S. L. (2000). Postpartum depression: a comparison of screening and routine clinical evaluation. *American Journal of Obstetrics and Gynecology, 182*, 1080-1082.

Faisal-Cury, A., Tedesco, J. J. A., Kahhale, S., Menezes, P. R., & Zugaib, M. (2004). Postpartum depression: In relation to life events and patterns of coping. *Archives of Women's Mental Health, 7*, 123-131.

Fenwick, J. & Rosenhain, C. (1991). *South Africa: From settlement to self-determination.* Australia: Oxford University Press.

Ferber, R. (1995). Sleeplessness in children. In: R. Ferber & M. Kryger (Eds.), *Principles and practice of sleep medicine in the child* (pp. 79-89). Philadelphia, PA: WB, Saunders Company.

Ferguson, G. A. (1981). *Statistical Analysis in Psychology and Education* (5th ed.). Tokyo: McGraw-Hill Kogakusha Ltd.

Ferraro, R. F. (Ed.). (2002). *Minority and cross-cultural aspects of neuropsychological assessment.* Lisse, the Netherlands: Swets & Zeitlinger.

Field, A. P. (2005). *Discovering statistics using SPSS: (and sex, drugs and rock 'n' roll)* (2 ed.). London: Sage Publications.

Field, T. (1995). Infants of depressed mothers. *Infant Behaviour and Development, 18*, 1-13.

Fisher, J. R. W., Feekery, C. J., & Rowe-Murray, H. J. (2002). Nature, severity and correlates of psychological distress in women admitted to a private mother-baby unit. *Journal of Paediatrics and Child Health, 38*, 140-145.

Fisher, J. R. W., Hammarberg, K., & Baker, H. W. G. (2005). Assisted conception is a risk factor for postnatal mood disturbance and early parenting difficulties. *Fertility and Sterility, 84*(2), 426-430.

Fitch, K., D. (2002). The influence of culture and social support on postpartum depression. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 63*(5-B).

Fletcher-Janzen, E., Strickland, T. L., & Reynolds, C. R. (Eds.). (2000). *Handbook of Cross-Cultural Neuropsychology.* Netherlands: Kluwer Academic Publishers Group.

Flores, L. (2006). Translation of English materials in Spanish [PDF document]. *NCTSN Culture & Trauma Briefs, 1*(3). Retrieved from http://www.nctsnet.org/nctsn_assets/pdfs/culture_and_trauma_brief_translations.pdf

Flynn, H. A., Sexton, M., Ratliff, S., Porter, K., & Zivin, K. (2011). Comparative performance of the Edinburgh Postnatal Depression Scale and the Patient Health Questionnaire-9 in pregnant and postpartum women seeking psychiatric services. *Psychiatry Research, 187*, 130-134.

Fones, C. (1996). Posttraumatic stress disorder occurring after painful childbirth. *Journal of Nervous and Mental Disease, 184*, 195-196.

Fontaine, K. R., & Jones, L. C. (1997). Self-esteem, optimism, and postpartum depression. *Journal of Clinical Psychology, 53*, 59-63.

Forman, D. N., Videbech, P., Hedegaard, M., Salvid, J. D., & Secher, N. J. (2000). Postparum depression: identification of women at risk. *British Journal of Obstetrics and Gynaecology, 107*, 1210-1217.

Foster, J., Barkus, E., & Yavorsky, C. (2006). *Understanding and Using Advanced Satistics*. London: Sage Publications.

Fouche, J. P., Heyns, P., Malan, F., Miemie, C., Schoon, M. G., & Bam, R. H. (1998). The psychological needs of African women in labour in western-orientated obstetrical care. *South African Journal of Psychology, 28*(2), 71-73.

Fowles, E. R. (1998). The relationship between maternal role attainment and postpartum depression. *Health Care for Women International, 19*, 83-94.

Foxcroft, C. & Roodt, G. (Eds.) (2006). *An introduction to psychological assessment in the South African context (2nd ed.)*. South Africa, Cape Town: Oxford University Press.

Foxcroft, C., Roodt, G & Abrahams, F. (2006). Psychological assessment: a brief retrospective overview. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (2nd ed.; p. 8-23). Oxford: Oxford University Press.

Freeman, M. P., Wright, R., Watchman, M., Wahl, R. A., Sisk, D. J., Fraleigh, L., & Weibrecht, J. M. (2005). Postpartum depression assessments at well-baby visits: Screening feasibility, prevalence, and risk factors. *Journal of Women's Health, 14*, 929-935.

Fuggle, P., Glover, L., Khan, F., & Haydon, K. (2002). Screening for postnatal depression in Bengali women: Preliminary observations from using a translated version of the Edinburgh Postnatal Depression Scale (EPDS). *Journal of Reproductive and Infant Psychology, 20*, 71-82.

Furnham, A. & Malik, R. (1994). Cross-cultural beliefs about "depression". *International Journal of Social Psychiatry, 40*, 106-123.

Furr, R. M. & Bacharach, V. R. (2007). *Psychometrics: An Introduction*. Thousand Oaks, CA: Sage Publications.

Galler, J. R., Ramsey, F. C., Harrison, R. H., Taylor, J., Cumberbatch, G., & Forde, V. (2004). Postpartum maternal moods and infant size predict performance on a national high school entrance examination. *Journal of Child Psychology and Psychiatry, 45*, 1064-1075.

Garcia-Esteve, L., Ascaso, C., Ojuel, J., & Navarro, P. (2003). Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Spanish mothers. *Journal of Affective Disorders, 75*, 71-76.

Gaynes, B. N., Gavin, N., Meltzer-Brody, S., Lohr, K. N., Swinson, T., Gartlehner, G., Brody, S., & Miller, W. C. (2005). Perinatal depression: Prevalence, screening accuracy, and screening outcomes. *Evidence Report/Technology Assessment, 119.* (Prepared by the RTI-University of North Carolina Evidence-based Practice Center, under Contract No. 290-02-0016.) AHRQ Publication No. 05- E006-2. Rockville, MD: Agency for Healthcare Research and Quality. February 2005.

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*, 304-312.

Geller, D., Vinokurov, A., & Martin, T. L. (2004). *Cross-cultural issues in survey translation: Translation of meaning and meaning of translation* [PDF document]. Retrieved July 31, 2009, from http://www.fcsm.gov/05papers/Geller_Vinokurov_Martin_VIIIA.pdf

Georgiopoulos, A. M., Bryan, T. L., Wollan, P., & Yawn, B. P. (2001). Routine screening for postpartum depression. *Journal of Family Practice, 50*, 117-122.

Ghubash, R., & Abou-Saleh, M. T. (1997). Postpartum psychiatric illness in Arab culture: Prevalence and psychosocial correlates. *The British Journal of Psychiatry, 171*(7), 65-68.

Gibson, J., McKenzie-McHarg, K., Shakespeare, J., Price, J., & Gray, R. (2009). A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. *Acta Psychiatrica Scandinavica, 119*, 350-364.

Giliomee, H. (1975). The development of the Afrikaner's self-concept. In: H. W. Van Der Merwe (Ed.), *Looking at the Afrikaner today.* Cape Town: Tafelberg.:

Giliomee, H. (2003). *The Afrikaners: Biography of a people*. Cape Town: Tafelberg.

Goldbort, J. (2006). Transcultural Analysis of Postpartum Depression. *The American Journal of Maternal/Child Nursing, 31*(2), 121-126.

Golding, J. M., & Lipton, R. I. (1990). Depressed mood and major depressive disorder in two ethnic groups. *Journal of Psychiatric Research, 24*(1) 65-82.

Goldman, J. M. (1986). Postpartum thyroid function. *Archives of Internal Medicine, 146*(7), 1296-1299.

Goldsmith, M. E. (2007). Postpartum depression screening by family nurse practitioners. *Journal of the American Academy of Nurse Practitioners, 19*, 321-327.

Goodman, J. H. (2005). Influences of maternal postpartum depression on fathers and the father-infant relationship. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66*(1-B), 199.

Gorman, L. L., O'Hara, M. W., Figueiredo, B., Hayes, S., Jacquemain, F., Kammerer, M., … Sutter-Dallay, A. L. (2004). Adaptation of the Structured Clinical Interview for DSM-IV Disorders for assessing depression in women during pregnancy and post-partum across countries and cultures. *British Journal of Psychiatry, 184*(46), s17-23.

Gotlib, I. H., Whiffen, V. E., Mount, J. H., Milne, K., & Cordy, N. I. (1989). Prevalence rates and demographic characteristics associated with depression in pregnancy and the postpartum. Journal of Consulting and Clinical Psychology, *57*(2), 269-274.

Gottlieb, S., & Barrett, D. (1986). Effects of unanticipated cesarean section on mothers, infants, and their interaction in the first month of life. *Journal of Developmental and Behavioral Pediatrics, 7*(3), 180-185.

Grace, S. L., Evindar, A., & Stewart, D. E. (2003). The effect of postpartum depression on child cognitive development and behavior: A review and critical analysis of the literature. *Archives of Women's Mental Health, 6*, 263-274.

Graham, J. E., Lobel, M., & DeLuca, R. S. (2002). Anger after childbirth: An overlooked reaction to postpartum stressors. *Psychology of Women Quarterly, 26*, 222-233.

Grazioli, R., & Terry, D. J. (2000). The role of cognitive vulnerability and stress in the prediction of postpartum depressive symptomatology. *The British Journal of Clinical Psychology, 39*, 329-347.

Green, J. M. (1990). 'Who is unhappy after childbirth?' Antenatal and intrapartum correlates from a prospective study. *Journal of Reproductive and Infant Psychology, 8*, 175-183.

Grieve, K. W. (2006). Factors affecting assessment results. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (2nd ed., p. 224-241). South Africa, Cape Town: Oxford University Press.

Guedeney, N., Fermanian, J., Guelfi, J. D. & Kumar, R. C. (2000). The Edinburgh Postnatal Depression Scale (EPDS) and the detection of major depressive disorders in early postpartum: some concerns about false negatives. *Journal of Affective Disorders, 61*, 107-12.

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46*, 1417-1432.

Guthery, D., & Lowe, B. A. (1992), Translation problems in international marketing research. *Journal of Language for International Business, 4*, 1-14.

Halbreich, U. (2003). The etiology, biology and evolving pathology of premenstrual syndromes. *Psychoneuroendocrinology, 28*, 55-99.

Halbreich, U. (2004). Prevalence of mood symptoms and depressions during pregnancy: Implications for clinical practice and research. *CNS Spectrums, 9*(3), 177-184.

Halbreich, U. (2005). Postpartum disorders: multiple interacting underlying mechanisms and risk factors. *Journal of Affective Disorders, 88*, 1-7.

Halbreich, U., & Karkun, S. (2006). Cross-cultural and social diversity of prevalence of postpartum depression and depressive symptoms. *Journal of Affective Disorders, 91*, 97-111.

Hall, L. A., Kotch, J. B., Browne, D., & Rayens, M. K. (1996). Self-esteem as a mediator of the effects of stressors and social resources on depressive symptoms in postpartum mothers. *Nursing Research, 45*, 231-238.

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A Progress Report. *European Journal of Psychological Assessment, 10*, 229-244.

Hambleton, R. K. (2001). The next generation of ITC test translation and adaptation guidelines. *European Journal for Psychological Assessment, 17*, 164-172.

Hambleton, R. K. (2004). Adapting achievement tests in multiple languages for international assessments. In A. C. Porter and A. Gamoran (Eds.), *Methodological* *advances in cross-national surveys of educational achievement* (pp. 3-23). Washington, D. C.: Board of International Comparative Studies in Education, National Academy Press.

Hambleton, R. K., & de Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing, 20*, 127-134.

Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology, 1*, 1-30.

Hambleton, R. K., Merenda, P., & Spielberger, C. (Eds.). (2005). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. Hillsdale, NJ: Lawrence S. Erlbaum Publishers.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.

Hamilton, J. A., & Harberger, P. N. (1992). *Postpartum psychiatric illness: A picture puzzle*. Philadelphia: University of Pennsylvania Press.

Hanlon, C., Whitley, R., Wondimagegn, D., Alem, A., & Prince, M. (2009). Postnatal mental distress in relation to the sociocultural practices of childbirth: An exploratory qualitative study from Ethiopia. *Social Science and Medicine, 69*(8), 1211-1219.

Hanna, B., Jarman, J., Savage, S., & Layton, K. (2004). The early detection of postpartum depression: Midwives and nurses trial a checklist. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 33*, 191-7.

Hanusa, B. H., Scholle, S. H., Haskett, R. F., Spadaro, K., & Wisner, K. L. (2008). Screening for depression in the postpartum period: A comparison of three instruments. *Journal of Women's Health, 17*(4), 585-596. doi:10.1089/jwh.2006.0248

Harkness, S. (1987). The cultural mediation of postpartum depression. *Medical Anthropology Quarterly, 1*(2), 194-209.

Harpham, T. (1994) Urbanisation and mental health in developing countries. *Social Science and Medicine, 39*, 133-145.

Harris, B., Huckle, P., Thomas, R., Johns, S., & Fung, H. (1989). The use of rating scales to identify post-natal depression. *British Journal of Psychiatry, 154*, 813-817.

Harris. B., Lovett, L., Smith, J., Read, G., Walker, R., & Newcombe, R. (1996). Cardiff puerperal mood and hormone study. III. Postnatal depression at 5 to 6 weeks postpartum, and its hormonal correlates across the peripartum period. *British Journal of Psychiatry, 168*, 739-744.

Hart, S., Field, T., & Nearing, G. (1998). Depressed mothers' neonates improve following the MABI and a Brazelton demonstration. *Journal of Pediatric Psychology, 23*(6), 351-356.

Harvey, R. J., & Hammer, A. L. (1999). Item resonse theory. *The Counseling Psychologist, 27*(3), 353-383.

Hau, F. W. & Levy, V. A. (2003). The maternity blues and Hong Kong Chinese women: An exploratory study. *Journal of Affective Disorders, 75*(2), 197-203.

Hay, H. (2002). *The effect of home language on performance in an English assessment instrument*. Unpublished masters thesis, University of Pretoria, Pretoria, South Africa.

Hearn, G., Iliff, Jones, Kirby, Ormiston, Parr, Rout, & Wardman, (1998). Postnatal depression in the community. *British Journal of General Practice, 48*(428), 1064-1066.

Helms, J. E., Henze, K. T., Sass, T. L. & Mifsud, V. A. (2006). Treating Cronbach's alpha reliability coefficients as data in counseling research. *The Counseling Psychologist*, 34(5), 630-660.

Hendrick, V., Altshuler, L. L., & Suri, R. (1998). Hormonal changes in the postpartum and implications for postpartum depression. *Psychosomatics, 39*, 93-101.

Henshaw, C. (2003). Mood disturbance in the early puerperium: a review. *Archives of Womens Mental Health, 6*, Suppl. 2, s33-s42.

Heron, J., McGuinness, M., Blackmore, E. R., Craddock, N., & Jones, I. (2008). Early postpartum symptoms in puerperal psychosis. *BJOG: An International Journal of Obstetrics and Gynaecology, 115*(3), 348-353.

Hickey, A. R., Boyce, P. M., Ellwood, D., & Morris-Yates, A. D. (1997). Early discharge and risk of postnatal depression. *The Medical Journal of Australia, 167*(5), 244-247.

Hiscock, H., & Wake, M. (2001). Infant sleep problems and postnatal depression: A community-based study. *Pediatrics, 107*, 1317.

Hiscock, H., & Wake, M. (2002). Randomised controlled trial of behavioural infant sleep intervention to improve infant sleep and maternal mood. *British Medical Journal (Clinical research ed.). 324*, 1062-1065.

Hobfoll, S. E., Ritter, C., & Lavin, J. (1995) Depression prevalence and incidence among inner-city pregnant and postpartum women. *Journal of Consulting and Clinical Psychology, 3*, 445-453.

Hofberg, K. (2003). Tokophobia: a profound dread and avoidance of childbirth: A review. *Archives of Womens Mental Health, 6*, [Suppl.2], s83-s109.

Hoffman, S., & Hatch, M. C. (2000). Depressive symptomatology during pregnancy: evidence for an association with decreased fetal growth in pregnancies of lower social class women. *Health Psychology, 19*, 535-543.

Hogan, T. P. (2007). *Psychological Testing: A practical introduction* (2nd ed.). New Jersey: John Wiley & Sons, Inc.

Holding, P., Taylor, H., Kazungu, S., Mkala, T., Gona, J., Mwamuye, B., … Stevenson, J. (2004). Assessing cognitive outcomes in a rural African population. *Journal of the International Neuropsychological Society, 10*, 261-270.

Honey, K. L., Bennett, P., & Morgan, M. (2003). Predicting postnatal depression. *Journal of Affective Disorders, 76*, 201-210.

Hong, S. & Wong, E. C. (2005). Rasch rating scale modeling of the Korean version of the Beck Depression Inventory. *Educational and Psychological Measurement, 65*, 124-139.

Honikman, J. (2008). Putting prevention into practice: Closing the gap between what we know and what we don't in women's healthcare. In Report of the LD 792

Workgroup: Efforts concerning postpartum mental health education, January 2008. Retrieved from http://www.maine.gov/dhhs/boh/documents.LD_792_Workgroup_Report_Final_January_2008.doc

Horowitz, J. A., & Goodman, J. H. (2005). Identifying and treating postpartum depression. *Journal of Gynecologic, Obstetric, & Neonatal Nursing, 34*(2), 264-273.

Howell, E. A., Mora, P., & Leventhal, H. (2006). Correlates of early postpartum depressive symptoms. *Maternal and Child Health Journal, 10*(2), 149-157.

Huang, Y. C., & Mathers, N. (2001). Postnatal depression-biological or cultural? A comparative study of postnatal women in the UK and Taiwan. *Journal of Advanced Nursing, 33*(3), 279-287.

Huffman, L. C., Lamour, M., Byran, Y. E., & Pederson, F. A. (1990) Depressive symptomatology during pregnancy and the postpartum period: Is the Beck Depression Inventory applicable? *Journal of Reproductive and Infant Psychology, 8*, 87-97.

Hui, C. H. & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16*(2), 131-152.

Huysman, A. (1998). *A mother's tears: Understandig the mood swings that follow childbirth.* New York: Seven Stories Press.

Ichida, M. (1996). A case of posttraumatic stress disorder (PTSD), the onset of which was the complication of childbirth. *Japanese Journal of Psychosomatic Medicine, 36*, 431-434.

Inandi, T., Elci, O. C., Ozturk, A., Egri, M., Polat, A., & Sahin, T. K. (2002). Risk factors for depression in postnatal first year, in eastern Turkey. *International Journal of Epidemiology, 31*, 1201-1207.

International Test Commission (2000). *International guidelines for test use*. Retrieved from http://www.intestcom.org

International Test Commission (2001). International guidelines for test use. *International Journal of Testing, 1*(2), 93-114.

International Test Commission (2010). *International Test Commission Guidelines for Translating and Adapting Tests*. Retrieved from http://www.intestcom.org./Guidelines/Adapting+Tests.php

Jacobsen, T. (1999). Effects of postpartum disorders on parenting and on offspring. In L. J. Miller (Ed.). *Postpartum mood disorders* (pp. 119-139). Washington, DC, USA: American Psychiatric Association.

Jardri, R., Pelta, J., Maron, M., Thomas, P., Delion, P., Codaccioni, X., & Goudemand, M. (2006). Predictive validation study of the Edinburgh Postnatal Depression Scale in the first week after delivery and risk analysis for postnatal depression. *Journal of Affective Disorders, 93*, 169-176.

Jenkin, W., & Tiggemann, M. (1997). Psychological effects of weight retained after pregnancy. *Women and Health, 25*, 89-98.

Jennings, K. D., Ross, S., Popper, S., & Elmore, M. (1999). Thoughts of harming infants in depressed and nondepressed mothers. *Journal of Affective Disorders, 54*, 21-28.

Jolley, S. N., & Betrus, P. (2007). Comparing postpartum depression and major depressive disorder: Issues in assessment. *Issues in Mental Health Nursing, 28*, 765-780.

Jones, I., & Craddock, N. (2001). Familiarity of the puerperal trigger in bipolar disorder: Results of a family study. *American Journal of Psychiatry, 158*, 913-917.

Josefsson, A. (2003). Postpartum depression: Epidemiological and biological aspects. *Nordic Journal of Psychiatry, 57*, 395-396.

Josefsson, A., Berg, G., Nordin, C., & Sydsjö, G. (2001). Prevalence of depressive symptoms in late pregnancy and postpartum. *Acta Obstetricia et Gynecologica Scandinavica, 80*(3), 251-255.

Kanjee, A. (2006). Cross-cultural test adaptation and translation. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (2nd ed., pp. 57-66). South Africa, Cape Town: Oxford University Press.

Karaçam, Z., & Kitiş, Y. (2008). The Postpartum Depression Screening Scale: Its reliability and validity for the Turkish population. *Turkish Journal of Psychiatry, 19*(2), 1-10.

Kelly, R., Zatzick, D., & Anders, T. (2001). The detection and treatment of psychiatric disorders and substance use among pregnant women cared for in obstetrics. *American Journal of Psychiatry, 158*, 213-219.

Kendell, R., Chalmers, J., & Platz, C. (1987). Epidemiology of puerperal psychoses. *British Journal of Psychiatry, 150*, 662-673.

Kendler, K. S., Walters, E. E., & Kessler, R. C. (1997). The prediction of length of major depressive episodes: Results from an epidemiological sample of female twins. *Psychological Medicine, 27*, 107-117.

Keshen, A., & MacDonald, J. L. (2004). Postpartum depression: Making the case for routine screening. *McGill Journal of Medicine, 7*(1). Retrieved from http://www.medicine.mcgill.ca/mjm/issues/v07n01/case_rep/case_rep.htm

Kestler, L. P. (2006). The impact of peripartum maternal depression and mother-infant interaction on infant cortisol. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67* (1-B), 577.

Kim, J., & Buist, A. (2005). Postnatal depression: A Korean perspective. *Australasian-Psychiatry, 13*, 68-71.

King, J. F., Slaytor, E.K., & Sullivan, E. A. (2004). Maternal deaths in Australia, 1997-1999. *Medical Journal of Australia, 181*, 413-414.

Kinloch, G.C. (1974). *The dynamics of race relations: A sociological analysis.* New York: McGraw-Hill.

Kirschenbaum, S. (1995, Spring). More than blue. *Mothering*, pp. 73-78.

Kitayama, S., & Cohen, D. (Eds.). (2007). *Handbook of cultural psychology*. Guilford Publications.

Kleinman, A. (2004). Culture and depression. *New England Journal of Medicine, 351*(10), 951-953.

Kline, P. (1993). *The handbook of psychological testing.* London: Routledge.

Klopper, S. (1998). The Zulu. In P. Magubane (Ed.), *Vanishing cultures of South Africa: Changing customs in a changing world*. Cape Town, South Africa: Struik Publishers.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the conduct of research on the internet. *American Psychologist, 59*(2), 105-117.

Kruckman, L., & Smith, S. (2006). *An Introduction to Postpartum Depression.* Retrieved Jaunary 29, 2006, from Postpartum Net Web site: http://www.postpartum.net/in-depth.html

Küçükdeveci, A. A., Sahin, H., Ataman, S., Griffiths, B., & Tennant, A. (2004). Issues in cross-cultural validity: Example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. *Arthritis Care and Research, 51*, 14-19.

Kumar, R. (1994). Postnatal mental illness: a transcultural perspective. *Social Psychiatry and Psychiatric Epidemiology, 29*, 250-264.

Kumar, R., & Robson, K.M. (1984). A prospective study of emotional disorders in childbearing women. *British Journal of Psychiatry, 144*, 35-47.

Kurstjens, S., & Wolke, D. (2001). Effects of maternal depression on cognitive development of children over the first 7 years of life. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 623-636.

Kyriakides, L., Kaloyirou, C., & Lindsay, G. (2006). An analysis of the Revised Olweus Bully/Victim Questionnaire using the Rasch measurement model. *British Journal of Educational Psychology, 76*, 781-801

Laizner, A. M., & Jeans, M. E. (1990). Identification of predictor variables of a postpartum emotional reaction. *Health Care Women International, 11*, 191-207.

Lane, A., Keville, R., Morris, M., Kinsella, A., Turner, M., & Barry, S. (1997). Postnatal depression and elation among mothers and their partners: Prevalence and predictors. *The British Journal of Psychiatry, 171*, 550-555.

Larsen, K. E., & O'Hara, M. W. (2002). The effects of postpartum depression on close relationships. In J.H. Harvey & A. Wenzel (Eds.), *A clinician's guide to maintaining and enhancing close relationships* (pp. 157-176). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Lavigne, J. V., Arend, R., Rosenbaum, D., Smith, A., Weissbluth, M., Binns, H. J., & Christoffel, K. K. (1999). Sleep and behavior problems among preschoolers. *Journal of Developmental and Behavioral Pediatrics, 20*(3), 164-169.

Lawrie, T. A., Hofmeyr, G. J., De Jager, M., & Berk, M. (1998). Validation of the Edinburgh Postnatal Depression Scale on a cohort of South African women. *South African Medical Journal, 88*, 1340-1344.

Le May, G. H. L. (1995). *The Afrikaners: An historical interpretation*. Oxford: Blackwell.

Lee, D. T. S., Yip, A. S. K., Chiu, H. F. K., & Chung, T. K. H. (2000). Screening for postnatal depression using the double-test strategy. *Psychosomatic Medicine, 62*, 258-263.

Lee, D. T. S., Yip, A. S. K., Chiu, H. F. K., Leung, T. Y. S., Chung, T. K. H. (2001). A psychiatric epidemiological study of postpartum Chinese women. *American Journal of Psychiatry, 158*(2), 220-226.

Lee, D. T., Yip, S. K., Chiu, H. F., Leung, T. Y., Chan, K. P., Chau, I. O., … Chung, T. K. (1998). Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry, 172*, 433-437.

Lee, D., Alexander T. S., Yip, S. K., Leung, T. Y. S., & Chung, T. K. H. (2004). Ethnoepidemiology of postnatal depression: Prospective multivariate study of sociocultural risk factors in a Chinese population in Hong Kong. *British Journal of Psychiatry, 184*, 34-40.

Lehohla, P. (2009). *South African Statistics, 2009*. Statistics South Africa: Pretoria. Retrieved from www.statssa.gov.za

Leiferman, J., (2002). The effect of maternal depressive symptomatology on maternal behaviors associated with child health. *Health Education and Behavior, 29*, 596-607.

Lesser, J., Koniak-Griffin, D., & Anderson, N. L. R. (1999). Depressed adolescent mothers' perceptons of their own maternal role. *Issues in Mental Health Nursing, 20*, 131-149.

Leung, S. S. K. (2002). Postpartum depression: Perceived social support and stress among Hong Kong Chinese women (Doctoral dissertation, Hong Kong Polytechnic University, 2002). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 62*, 4469.

Leung, S. S. K., Martinson, I. M., & Arthur, D. (2005). Postpartum depression and related psychosocial variables in Hong Kong Chinese women: findings from a prospective study. *Research in Nursing and Health, 28*(1), 27-38.

Li, L., Liu, F., Zhang, H., Wang, L., & Chen, X. (2011). Chinese version of the Postpartum Depression Screening Scale: Translation and Validation. *Nursing Research, 60*(4), 231-239.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions, 16(2), p. 878. Retrieved from http://www.rasch.org/rmt/rmt162f.htm

Linacre, J. M. (2004). Optimizing rating scale category effectiveness. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: Theory, models and applications* (pp. 258-278). Maple Grove, Minnesota: JAM Press.

Linacre, J. M. (2008). The expected value of a point-biserial (or similar) correlation. *Rasch Measurement Transactions, 22*(1), 1154. Retrieved from: http://www.rasch.org/rmt/rmt221e.htm

Linacre, J. M. (2009). Winsteps® (Version 3.69.0) [Computer Software]. Beaverton, Oregon: Winsteps.com.. Retrieved August 1, 2009. Available from http://www.winsteps.com/

Lison, S. M. (1977). *An MMPI-based study of the personality characteristics of three groups of South African students.* Unpublished master’s thesis, University of Cape Town, Cape Town, South Africa.

Llewellyn, A. M., Stowe, Z. N., & Nemeroff, C. B. (1997). Depression during pregnancy and the puerperium. *Journal of Clinical Psychiatry, 58*(suppl 15), 26-32.

Locke, R., Baumgart, S., Locke, K., Doodstein, M., Thies, C., & Greenspan, J. (1997). Effect of maternal depression on premature infant health during initial hospitalization. *The Journal of the American Osteopathic Association, 97*(3), 145-149.

Logsdon, M. C., & Usui, W. (2003). Psychosocial predictors of postpartum depression in diverse groups of women. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109.

Logsdon, M. C., Birkimer, J. C., & Usui, W. M. (2000). The link of social support and postpartum depressive symptoms in African-American women with low incomes. The *American Journal of Maternal Child Nursing, 25*(5), 262-266.

Logsdon, M. C., Birkimer, J. C., Simpson, T., & Looney, S. (2005). Postpartum depression and social support in adolescents. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 34*(1), 46-54.

Lucio, G. M., Reyes-Lagunes, I., & Scott, R. L. (1994). MMPI-2 for Mexico: Translation and adaptation. *Journal of Personality Assessment, 63*, 105-116.

Luoma, I., Tamminen, T., Kaukonen, P., Laippala, P., Puura, K., Salmelin, R., & Almqvist, F. (2001). Longitudinal study of maternal depressive symptoms and child well-being. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1367-1374.

Lussier, V., David, H., Saucier, J., & Borgeat, F. (1996). Self-rating assessment of postnatal depression: A comparison of the Beck Depression Inventory and the Edinburgh Postnatal Depression Scale. *Pre- and Perinatal Psychology Journal, 11*(2), 81-91.

Malpass, R. S., & Poortinga, Y. H. (1986). Strategies for design and analysis. In J. W. Berry & J. W. Lonner (Eds.), *Cross-cultural research and methodology series: Field methods in cross-cultural research* (Vol 8. pp. 47-83). Beverly Hills, CA: Sage Publications.

Manaster, G. J. & Havighurst, R. J. (1972). *Cross national research: Social-psychological methods and problems*. Boston: Houghton Mifflin Company.

Mann, J. (2007). Status: The marginal reaction: Mixed bloods and Jews. In P. Watson (Ed.), *Psychology and race* (pp. 213-223). U.S.A.: Aldine Transaction.

Maree, D. J. F. (2004, October). Basic IRT tutorial for analyzing data. Tutorial 2 - Interpreting output for dichotomous data. University of Pretoria, Pretoria: D. J. F. Maree.

Marks, M., Wieck, A., Checkley, S., & Kumar, C. (1996). How does marriage protect women with histories of affective disorder from post-partum relapse? *British Journal of Medical Psychology, 69*, 329-342.

Marrs, C. R., Durette, R. T., Ferraro, D. P., & Cross, C. L. (2009). Dimensions of ostpartum psychiatric distress: Preliminary evidence for broadening clinical scope. *Journal of Affective Disorders, 115*(1-2), 100-111.

Martinez-Schallmoser, L., Telleen, S., & MacMullen, N. J. (2003). The effect of social support and acculturation on postpartum depression in Mexican American women. *Journal of Transcultural Nursing, 14*, 329-338.

Mason, W. A., Rice, M. J., & Records, K. (2005). Lived experience of postpartum depression in a psychiatric population. *Perspectives in Psychiatric Care, 41*(2), 52-61.

Matthey, S., Barnett, B., Howie, P., & Kavanagh, D. J. (2003). Diagnosing postpartum depression in mothers and fathers: Whatever happened to anxiety? *Journal of Affective Disorders, 74*, 139-147.

Matthey, S., Barnett, B., Ungerer, J., & Waters, B. (2000). Paternal and maternal depressed mood during the transition to parenthood. *Journal of Affective Disorders, 60*(2), 75-85.

Maxted, A. E., Dickstein, S., Miller-Loncar, C., High, P., Spritz, B., Liu, J., & Lester, B. M. (2005). Infant colic and maternal depression. *Infant Mental Health Journal, 26*, 56-68.

Mazzeo, S. E., Landt, M. C. T. S., Jones, I., Mitchell, K., Kendler, K. S., Neale, M. C., Aggen, S. H., & Bulik, C. M. (2006). Associations among postpartum depression, eating disorders, and perfectionism in a population-based sample of adult women. *International Journal of Eating Disorders, 39*, 202-211.

McGorry, S. Y. (2000). Measurement in a cross-cultural environment: survey translation issues. *Qualitative Market Research: An International Journal, 3*, 74-81.

McLearn, K. T., Minkovitz, C. S., Strobino, D. M., Marks, E., & Hou, W. (2006). Maternal depressive symptoms at 2 to 4 months post partum and early parenting practice. *Archives of Pediatrics & Adolescent Medicine, 160*, 279-284.

Mehta, A., & Sheth, S. (2006). Postpartum depression: How to recognize and treat this common condition. *Medscape Psychiatry & Mental Health, 11*(1). Retrieved May, 10, 2006, from http://www.medscape.com/viewarticle/529390

Merchant, D. C., Alfonso, D. D., & Mayberry, L. J. (1995). Influence of marital relationship and child-care stress on maternal depression symptoms in the postpartum. *Journal of Psychosomatic Obstetrics and Gynaecology, 16*(4), 193-200.

Mesthrie, R. (1995). *Language and social history: Studies in South African sociolinguistics.* Cape Town: David Philip.

Metz, A., Sichel, D. A., & Goff, D. C. (1988). Postpartum panic disorder. *The Journal of Clinical Psychiatry, 49*(7), 278-279.

Milgrom, J., & Westley, D. T. (2003). Postnatal depression and child development: Is the early mother-infant interaction a mediating factor? *Archives of Women's Mental Health, 6* (Suppl. 2), s83-s109.

Milgrom, J., Mendelsohn, J., & Gemmill, A. W. (2011). Does postnatal depression screening work? Throwing out the bathwater, keeping the baby. *Journal of Affective Disorders, 132*, 301-310.

Miller, L. J. (2002). Postpartum depression. *JAMA: Journal of the American Medical Association, 287*(6), 762-765.

Misri, S. (2000). The impact of partner support in the treatment of postpartum depression. *Canadian Journal of Psychiatry, 45*, 554-558.

Misri, S., Kostaras, X., Fox, D., & Kostaras, D. (2000). The impact of partner support in the treatment of postpartum depression. *Canadian Journal of Psychiatry, 45*, 554-558.

Moehler, E., Brunner, R., Wiebel, A., Reck, C., & Resch, F. (2006). Maternal depressive symptoms in the postnatal period are associated with long-term impairment of mother-child bonding. *Archives of Women's Mental Health, 9*, 273-278.

Moleman, N., Van der Hart, O., & Van der Kolk, B. A. (1992). The partus reaction: A neglected etiological factor in postpartum psychiatric disorder. *Journal of Nervous and Mental Disease, 180*, 271-272.

Montgomery, D. E. (2001). The relationship between postpartum depression and maternal confidence in Hispanic mothers. *Dissertation Abstracts International: Section B: the Sciences and Engineering, 61* (11-B), 5799.

Morgan, J. F., Lacey, J. H., & Chung, E. (2006). Risk of postnatal depression, miscarriage, and preterm birth in bulimia nervosa: retrospective controlled study. *Psychosomatic Medicine, 68*(3), 487-492.

Morton, J. (2000). The relationship of various demographic factors and health locus of control to postpartum depression in young adults. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 60*, 8-B, 3853.

Munk-Olsen, T., Laursen, T. M., Pedersen, C. B., Mors, O., & Mortensen, P. B. (2006). New parents and mental disorders: A population-based register study. *Journal of the American Medical Association, 296*, 2582-2589.

Munoz, R. F., Le, H. N., & Ippen, C. G. (2000). We should screen for major depression. *Applied and Preventive Psychology, 9*(2), 123-133.

Murray, D., & Cox, J. L. (1990). Screening for depression during pregnancy with the Edinburgh Depression Scale (EDS). *Journal of Reproductive and Infant Psychology, 8*, 99-107.

Murray, L. (1992). The impact of postnatal depression on infant development. *Journal of Child Psychology and Psychiatry, 33*, 543-561.

Murray, L. (2001, October). How postnatal depression can affect children and their families. CPHVA conference proceedings, pp. 20-23.

Murray, L., & Carothers, A. D. (1990). The validation of the Edinburgh Post-natal Depression Scale on a community sample. *British Journal of Psychiatry, 157*, 288-290.

Murray, L., & Cartwright, W. (1993). The role of obstetric factors in postpartum depression. *Journal of Reproductive and Infant Psychology, 11*, 215-219.

Murray, L., Fiori-Cowley, A., Hooper, R., & Cooper, P. (1996). The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Development, 67*(5), 2512-2526.

Murray, L., Hipwell, A., & Hooper, R. (1996). The cognitive development of 5-year-old children of postnatally depressed mothers. *Journal of Child Psychology and Psychiatry, 37*, 927-935.

Murray, L., Sinclair, D., Cooper, P., Ducournau, P., & Turner, P. (1999). The socioemotional development of 5-year-old children of postnatally depressed mothers. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 1259-1271.

Murray, L., Stanley, C., Hooper, R., King, F., Fiori-Cowley, A. (1996). The role of infant factors in postnatal depression and mother-infant interactions. *Developmental Medicine and Child Neurology, 38*, 109-119.

Murray, L., Woolgar, M., Murray, J., & Cooper, P. (2003). Self-exclusion from health care in women at high risk for postpartum depression. *Journal of Public Health Medicine, 25*(2), 131-137.

Mushquash, C. J. & Bova, D. L. (2007). Cross-Cultural Assessment and Measurement Issues. *Journal on Developmental Disabilities, 13*, 53-66.

Muzik, M., Klier, C. M., Rosenblum, K. L., Holzinger, A., Umek, W., & Katschnig, H. (2000). Are commonly used self-report inventories suitable for screening postpartum depression and anxiety disorders? *Acta Psychiatrica Scandinavica, 102*, 71-73.

Mwamwenda, T. S. (1995). *Educational Psychology: An African perspective* (2nd ed.). Durban: Butterworths.

Nasser, R. (2005). A method for social scientists to adapt instruments from one culture to another: The case of the Job Descriptive Index. *Journal of Social Sciences, 1*, 232-237.

Nath, S. R. (2005). Narrative coherence as a predictor of postpartum depression: At-risk women's stories of the transition to motherhood. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 65*, 10-B, p 5415.

Navarro, P., Ascaso, C., Garcia-Esteve, L., Aguado, J., Torres, A., & Martín-Santos, R. (2007). Postnatal psychiatric morbidity: A validation study of the GHQ-12 and the EPDS as screening tools. *General Hospital Psychiatry, 29*, 1-7.

Nell, V. (2000). *Cross-Cultural Neuropsychological Assessment: Theory and practice*. Lawrence Erlbaum Associates, Publishers, NJ: Mahwah.

Nishizono-Maher, A., Kishimoto, J., Yoshida, H., Urayama, K., Miyato, M., Otsuka, Y., & Matsui, H. (2004). The role of self-report questionnaire in the screening of postnatal depression: A community sample survey in central Tokyo. *Social Psychiatry and Psychiatric Epidemiology, 39*, 185-190.

O'Hara, M. W. & Zekoski, E. M. (1988). Postpartum depression: A comprehensive review. In R. Kumar & I. F. Brockington (Eds.), *Motherhood and Mental Illness* (pp. 17-63). London: Wright.

O'Hara, M. W. (1995). *Postpartum Depression: Causes and Consequences.* New York: Springer-Verlag.

O'Hara, M. W., & Swain, A. M. (1996). Rates and risk of postpartum depression - a meta-analysis. *International Review of Psychiatry, 8*, 37-54.

O'Hara, M. W., Neunaber, D. J., & Zekoski, E. M. (1984). Prospective study of postpartum depression: Prevalence, course, and predictive factors. *Journal of Abnormal Psychology, 93*, 158-171.

O'Hara, M. W., Zekoski, E. M., Philipps, L. H., & Wright, E. J. (1990). Controlled prospective study of postpartum mood disorders: Comparison of childbearing and nonchildbearing women. *Journal of Abnormal Psychology, 99*(1), 3-15.

O'Hara, M.W. (1994).. Postpartum depression: identification and measurement in a cross-cultural context. In: J. L. Cox and J. Holden (Eds.), *Perinatal psychiatry: Use and misuse of the Edinburgh Postnatal Depression Scale* (pp. 145-164). London: American Psychiatric Press.

Oates, M. R., Cox, J. L., Neema, S., Asten, P., Glangeaud-Freudenthat, N., Figueiredo, B., Gorman, L. L., Hacking, S., Hirst, E., Kammerer, M. H., Klier, C. M., Seneviratne, G., Smith, M., Sutter-Dallay, A. L., Valoriani, V., Wickberg, B., Youshida, K., & the TCS-PND Group. (2004). Postnatal depression across countries and cultures: A qualitative study. *British Journal of Psychiatry, 184*(46), s10-14.

Okano, T., Nomura, J., Kumar, R., Kaneko, E., Tamaki, R., Hanafusa,I., Hayashi, M., & Matsuyama, A. (1998). An epidemiological and clinical investigation of postpartum psychiatric illness in Japanese mothers. *Journal of Affective Disorders, 48*(2-3), 233-240.

Omond, R. (1985). *A handbook of apartheid: A guide to South Africa's everyday racial policies.* Harmondsworth: Penquin Books.

Onkvisit, S. & Shaw, J. J. (2004). *International Marketing: Analysis and strategy* (4th ed.). United Kingdom: Taylor & Francis LTD.

Onozawa, K., Kumar, R. C., Adams, D., Dore, C., & Glover, V. (2003). High EPDS scores in women from ethnic minorities living in London. *Archives of Women's Mental Health, 6*, Suppl.2, s51-s55.

Oren, D. A., Wisner, K. L., Spinelli, M., Epperson, C. N., Pwindl, K. S., Terman, J. S., & Terman, M. (2002). An open tiral of morning light therapy for treatment of antepartum depression. *American Journal of Psychiatry, 159*, 666-669.

Orpen, C. (1970). Authoritarianism in an authoritarian culture: The case of Afrikaans-speaking South Africa. *Journal of Social Psychology, 81*, 119-120.

Orr, S. T., James, S. S., & Blackmore Prince, C. (2002). Maternal prenatal depressive symptoms and spontaneous preterm births among African-American women in Baltimore, Maryland. *American Journal of Epidemiology, 156*, 797-802.

Owoeye, A. O., Aina, O. F., & Morakinyo, O. (2006). Risk factors of postpartum depression and EPDS scores in a group of Nigerian women. *Tropical Doctor, 36*, 100-103.

Pallant, J. F., Miller, R. L., & Tennant, A. (2006). Evaluation of the Edinburgh Post Natal Depression Scale using Rasch analysis. BMC Psychiatry, 6(28). doi:10.1186/1471-244X-6-28

Pan, Y., & De La Puente, M. (2005, August 24). *Census bureau guideline for the translation of data collection instruments and supporting materials: Documentation on how the guideline was developed.* Research report series. Washington DC: U.S. Census Bureau. Retrieved from http://www.census.gov/srd/papers/pdf/rsm2005-06.pd

Park, E., & Dimigen, G. (1995). A cross-cultural comparison: Postnatal depression in Korean and Scottish mothers. *Psychologia, 38*(3), 199-207.

Parry, B. L., Sorenson, D. L., Zirpoli, G. G., Basavaraj, N., Patterson, A. E., Hauger, R., & Gillin, J. C. (2003). Sleep and endocrine studies of pregnancy and postpartum depression. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109.

Parry, C. D. H. (1996). A review of psychiatric epidemiology in Africa: Strategies for increasing validity when using instruments transculturally. *Transcultural Psychiatry, 33*, 173-188.

Patel, R. R., Murphy, D. J., & Peters, T. J. (2005). Operative delivery and postnatal depression: A cohort study. *British Medical Journal, 330*, 879-889.

Patel, V., Rahman, A., Jacob, K. S., & Hughes, M. (2004). Effect of maternal health on infant growth in low income countries: New evidence from South Asia. *British Medical Journal, 328*, 820-823.

Patel, V., Rodriquez, M., and DeSouza, N. (2002). Gender, poverty, and postnatal depression: A study of mothers in Goa, India. *American Journal of Psychiatry, 159*, 43-47.

Peindl, K. S., Zolnik, E. J., Wisner, K. L., & Hanusa, B. H. (1995). Effects of postpartum psychiatric illness on family planning. *International Journal of Psychiatry in Medicine, 25*(3), 291-300.

Perinatal Mood Disorders. Retrieved August, 27, 2004, from Massachusetts General Hospital, The MADI Resource Center – Mood and Anxiety Disorders Institute Web site: http://www.mgh.harvard.edu/madiresourcecenter/adult.asp

Pesonen, A. K., Raikkonen, K., Strandberg, T., Kelitikangas, J. L., & Jarvenpaa, A. (2004). Insecure adult attachment style and depressive symptoms: Implications for parental perceptions of infant temperament. *Infant Mental Health Journal, 25*, 99-116.

Philipps, L. H., & O'Hara, M.W. (1991). Prospective study of postpartum depression: Four and a half year follow-up of women and children. *Journal of Abnormal Psychology, 100*, 151-155.

Pierce, K. M., Strauman, T. J., & Lowe-Vandell, D. (1999). Self-discrepancy, negative life events, and social support in relation to dejection in mothers of infants. *Journal of Social and Clinical Psychology, 18*, 490-501.

Pinheiro, R. T., Magalhaes, P. V. S., Horta, B. L., Pinheiro, K. A. T., Da-Silva, R. A., & Pinto, R. H. (2006). Is paternal postpartum depression associated with maternal postpartum depression? Population-based study in Brazil. *Acta Psychiatrica Scandinavica, 113*, 230-232.

Plank, G. A. (2001). Plank, G. A. (2001). Application of the cross battery approach in the assessment of American Indian children: A viable alternative. In Mushquash, C. J. & Bova, D. L. (2007). Cross-Cultural Assessment and Measurement Issues. *Journal on Developmental Disabilities, 13*, 53-66.

Pogge, T. W. (1997). Group rights and Ethnicity. In I. Shapiro & W. Kymlicka (Eds.), *NOMOS: Ethnicity and group rights* (Vol. 39, pp. 187-221). New York: New York University Press.

Ponterotto, J. G., Casas, J. M., Suzuki, L. A., & Alexander, C. M. (Eds.). (1995). *Handbook of multicultural counseling* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Poortinga, Y. H., & Malpass, R. S. (1986). Making inferences from cross-cultural data. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross cultural psychology* (pp. 17-46). Beverly Hills, CA: Sage Publications.

Pop, V. J., De Rooy, H. A., Vader, H. L., Van der Heide, D., Van Son, M. M., Komproe, I. H. (1993). Microsomal antibodies during gestation in relation to postpartum thyroid dysfunction and depression. *Acta Endocrinology,129*, 26-30.

Pop, V. J., De Rooy, H. A., Vader, H. L., Van der Heide, D., Van Son, M. M, Komproe, I. H., Essed, G. G., & de Geus, C. A. (1991). Postpartum thyroid dysfunction and depression in an unselected population. *New England Journal of Medicine, 324*, 1815-1816.

Posmontier, B., & Horowitz, J. A. (2004). Postpartum practices and depression prevalences: Technocentric and ethnokinship cultural perspectives. *Journal of Transcultural Nursing, 15*(1), 34-43.

Posner, N. A., Unterman, R. R., Williams, K. N., & Williams, G. H. (1997). Screening for postpartum depression: An antepartum questionnaire. *Journal of Reproductive Medicine, 42*, 207-215.

Postpartum depression consensus statement (2002). Retrieved March 15, 2006, from Indiana Perinatal Web site: http://www.indianaperinatal.org

Pound, G. A. (2006). The relationship between postpartum depression and infant development. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66* (11-B), 6315.

Preston-Whyte, E., & Zondi, M. (1989). To control their own reproduction: The agenda of black teenage mothers in Durban. Agenda, 4, 47-68. URL: http://www.jstor.org/stable/4547930

Priel, B., & Besser, A. (1999). Vulnerability to postpartum depressive symptomatology: Dependency, self-criticism and the moderating role of antenatal attachment. *Journal of Social and Clinical Psychology, 18*, 240-253.

Priest, S. R., Henderson, J., Evans, S. F., & Hagan, R. (2003). Stress debriefing after childbirth: A randomised controlled trial. *Medical Journal of Australia, 178*, 542-545.

Prieto, A. J. (1992). A method for translation of instruments to other languages. *Adult Education Quarterly, 43*, 1-14.

Puckering, C. (2005). Mind the gap! Helping the children of mothers with postnatal depression. *Child: Care, Health and Development, 31*, 7-9.

Rahman, A., Iqbal, Z., Harrington, R. (2003). Life events, social support and depression in childbirth: Perspectives from a rural community in the developing world. *Psychological Medicine, 33*(7), 1161-1167.

Raidt, E. H. (1989). Ontwikkeling van vroeë Afrikaans. In T. J. R. Botha, F. A. Ponelis, J. G. H. Combrink, & F. F. Odendal (Eds.), *Inleiding tot die Afrikaanse Taalkunde* (pp. 96-126). Pretoria: Academica.

Ramchandani, P. G., Richter, L. M., Stein, A., & Norris, S. A. (2009). Predictors of postnatal depression in an urban South African cohort. *Journal of Affective Disorders, 113*, 279-284.

Ramirez, M., Teresi, J. A., Holmes, D., Gurland, B., & Lantigua, R. (2006). Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE): Overview, sample, and issues of translation. *Medical Care, 44*(11), Suppl 3, S95-S106.

Ramsay, R. (1993). Postnatal depression. *Lancet, 341*, 1358.

Rebelsky, F., & Daniel, P. A. (1976). Cross-cultural studies of infant intelligence. In M. Lewis (Ed.), *Origins of intelligence: Infancy and early childhood* (pp. 279-297). New York: Plenum Press.

Records, K. A., & Rice, M. (2002). Childbearing experiences of abused Hispanic women. *Journal of Midwifery and Women's Health, 47*(2), 101-110.

Reichenheim, M. E., & Harpham, T. (1991). Maternal mental health in a squatter settlement in Rio de Janeiro. *British Journal of Psychiatry, 159*, 683-690.

Reid, A. J., Biringer, A., Carroll, J. D., Midmer, D., Wilson, L. M., Chalmers, B., & Stewart, D. E. (1998). Using the ALPHA form in practice to assess antenatal psychosocial health. Antenatal Psychosocial Health Assessment. *Canadian Medical Association Journal , 159*(6), 677-684.

Retief, A. (1988). *Method and theory in cross-cultural psychological assessment. Research report No. 6*. Pretoria: Human Sciences Research Council.

Rich-Edwards, J. W., Kleinman, K., Abrams, A., Harlow, B. L., Thomas, J., Joffe, H., & Gillman, M. W. (2006). Sociodemographic predictors of antenatal and postpartum depressive symptoms among women in a medical group practice. *Journal of Epidemiology and Community Health, 60*, 221-227.

Righetti-Veltema, M., Conne-Perreard, E., Bousquet, A., & Manzano, J. (1998). Risk factors and predictive signs of postpartum depression. *Journal of Affective Disorders, 49*(3), 167-180.

Rissik, D. (1994). *Culture Shock! South Africa.* London: Kuperard.

Roan, S. L. (1997). *Postpartum Depression: Every woman's guide to diagnosis, treatment & prevention.* Massachusetts: Adams Media Corporation.

Roberts, S. L., Bushnell, J. A., Collings, S. C., & Purdie, G. L. (2006). Psychological health of men with partners who have post-partum depression. *Australian and New Zealand Journal of Psychiatry, 40*, 704-711.

Robertson, E., Grace, S., Wallington, T., & Stewart, D. E. (2004). Antenatal risk factors for postpartum depression: A synthesis of recent literature. *General Hospital Psychiatry, 26*, 289-295

Ross, L. E., Gilbert Evans, S. E., Sellers, E. M., & Romach, M. K. (2003). Measurement issues in postpartum depression part 1: Anxiety as a feature of postpartum depression. *Archives of Women's Mental Health, 6*, 51-57.

Ross, L. E., Murray, B. J., & Steiner, M. (2006). Sleep and perinatal mood disorders: A critical review. *Journal of Psychiatry and Neuroscience, 30*, 247-256.

Ross, L. E., Sellers, E. M., & Romach, M. K. (2003). Interactions between biological and psychosocial risk factors in postpartum depression and anxiety. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109.

Rowe, H. J., Fisher, J. R. W., & Loh, W. M. (2008). The Edinburgh Postnatal Depression Scale detects but does not distinguish anxiety disorders from depression in mothers of infants. *Archives of Women's Mental Health, 11*(2), 103-108.

Rowe, H., Fisher, J., & Feekery, C. (2003). Nature and correlates of unsettled infant behaviour in a cohort of mother-infant pairs admitted to a specialised early parenting service. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109

Royal, K. D. (2010). Making meaningful measurement in survey research: A demonstration of the utility of the Rasch model. *IR Applications, 28*, 1-15.

Rubertsson, C., Waldenstrom, U., Wickberg, B., Radestad, I., & Hildingsson, I. (2005). Depressive mood in early pregnancy and postpartum: Prevalence and women at risk in a national Swedish sample. *Journal of Reproductive and Infant Psychology, 23*, 155-166.

Rubertsson, C., Wickberg, B., Gustavsson, P., & Radestad, I. (2005). Depressive symptoms in early pregnancy, two months and one year postpartum-prevalence and psychosocial risk factors in a national Swedish sample. *Archives of Women's Mental Health, 8*, 97-104.

Rush, A. J., Bernstein, I. H., Trivedi, M. H., Carmody, T. J., Wisniewski, S., Mundt, J. C., … Fava, M. (2006). An evaluation of the Quick Inventory of Depressive

Symptomatology and the Hamilton Rating Scale for Depression: A sequenced treatment alternatives to relieve depression trial report. *Biological Psychiatry, 59*(6), 493-501.

Rush, A. J., Giles, D. E., Schlesser, M. A., Fulton, C. L., Weissenburger, J. E., & Burns, C. T. (1986). The Inventory for Depressive Symptomatology (IDS): Preliminary findings. *Psychiatry Research, 18*, 65-87.

Rush, A. J., Gullion, C. M., Basco, M. R., Jarrett, R. B. & Trivedi, M. H. (1996). The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychological Medicine, 26*, 477-486.

Rush, A. J., Trivedi, M. H., Carmody, T. J., Ibrahim, H. M., Markowitz, J. C., Keitner, G. I., … Keller, M. B. (2005). Self-reported depressive symptom measures: Sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. *Neuropsychopharmacology, 30*(2), 405-416.

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., Keller, M. B. (2003). The 16-item Quick Inventory of Depressive Symptomatology (QIDS) Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry, 54*, 573-583.

Ryan, D., Milis, L. & Misri, N. (2005). Depression during pregnancy. *Canadian Family Physician, 51*(8), 1087-1093.

Sagami, A., Kayama, M., & Senoo, E. (2004). The relationship between postpartum depression and abusive parenting behaviour of Japanese mothers: A survey of mothers with a child less than one year old. *Bulletin of the Menninger Clinic, 68*, 174-187.

Saisto, T., & Halmesmäki, E. (2003). Fear of childbirth: A neglected dilemma. *Acta Obsteticia et Gynecologica Scandinavica, 82*, 201-208.

Salzberger, T. (2010). Does the Rasch model convert an ordinal scale into an interval scale? *Rasch Measurement Transactions, 24*(2), 1273-1275. Retrieved from http://www.rasch.org/rmt/rmt242.pdf

Schlemmer, L. (1974). The Afrikaners: Youth and change. *Optima, 34*, 56-65.

Schreiber, S., & Galai-Gat, T. (1993). Uncontrolled pain following physical injury as the core-trauma in post-traumatic stress disorder. *Pain, 54*, 107-110.

Schumacher, M., Zubaran, C., & White. G. (2008). Bringing birth-related paternal depression to the fore. *Women and Birth, 21*(2), 65-70.

Schumacker, R. E. (2004). Rasch measurement: The dichotomous model. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: theory, models and applications* (pp. 226-253). Maple Grove, Minnesota: JAM Press.

Schumacker, R. E. (2010). Classical test analysis. Applied Measurement Associates LLC. Retrieved July 29, 2011, from http://www.appliedmeasurementassociates.com/articles/CLASSICAL%20TEST%20ANALYSIS.pdf

Schumacker, R. E., & Smith, E. V. (2007). Reliability: A Rasch perspective. *Educational and Psychological Measurement, 67*(3), 394-409.

Science Daily (n.d.). *Science reference: Ethnic group.* Retrieved from http://www.sciencedaily.com/articles/e/ethnic_group.htm

Seehusen, D. A., Baldwin, L. M., Runkle, G. P., & Clark, G. (2005). Are family physicians appropriately screening for postpartum depression? *Journal of the American Board of Family Practice, 18*, 104-12.

Segre, L. S., O'Hara, M. W., Arndt, S., & Stuart, S. (2007). The prevalence of postpartum depression. The relative significance of three social status indices. *Social Psychiatry and Psychiatric Epidemiology, 42*, 316-321.

Seguin, L., Potvin, L., St-Denis, M., & Loiselle, J. (1999). Socio-environmental factors and postnatal depressive symptomatology: A longitudinal study. *Women and Health, 29*, 57-72.

Seng, J., & Hassinger, J. (1998). Relationship strategies and interdisciplinary collaboration: Improving maternity care with survivors of childhood sexual abuse. *Journal of Nurse-Midwifery, 43*, 287-295.

Sharma, V. (2003). Pharmacotherapy of postpartum psychosis: Expert Opinion. *Pharmacotherapy, 4*, 1651-1658.

Sheeder, J., Kabir, K., & Stafford, B. (2009). Screening for postpartum depression at well-child visits: Is once enough during the first 6 months of life? *Pediatrics, 123*, 982-988.

Sheinkopf, S. J., Lester, B. M., LaGasse, L. L., Seifer, R., Bauer, C. R., Shankaran, S., Bada, H. S., Poole, W. K., & Wright, L. L. (2006). Interactions between maternal characteristics and neonatal behavior in the prediction of parenting stress and perception of infant temperament. *Journal of Pediatric Psychology, 31*, 27-40.

Sherman-Slate, E. (2005). Attachment as predictor of postpartum depression in first-time mothers. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 65* (10-B), 5422.

Shillington, S. J. (1988). *An Afrikaans MMPI: Some variables which affect the responses of a normal Afrikaans population*. Unpublished doctoral dissertation, University of Port Elizabeth.

Shuttleworth-Edwards, A., Kemp, R., Rust, A., Muirhead, J., Hartman, N. & Radloff, S. (2004). Vanishing ethnic differences in the wake of acculturation: A review and preliminary indications on WAIS-III test performance in South Africa. *Journal of Clinical and Experimental Neuropsychology, 26*, 903-920.

Shuttleworth-Jordan, A. B. (1996). On not reinventing the wheel: a clinical perspective on culturally relevant test usage. *South African Journal of Psychology, 26*(2), 96-102.

Sichel, D. & Driscoll, J. W. (1999). *Women's moods: What every woman must know about hormones, the brain, and emotional health.* New York: Harper Collins Publishers.

Sit, D., Rothschild, A.J., & Wisner, K.L. (2006). A review of postpartum psychosis. *Journal of Womens Health, 15*, 352-368

Smith, E. V. (2004). Evidence for the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: theory, models and applications* (pp. 93-122). Maple Grove, Minnesota: JAM Press.

Smith, E. V., Conrad, K., Chang, K., & Piazza, J. (2002). An introduction to Rasch measurement for scale development and person assessment. *Journal of Nursing Measurement, 10*(3), 189-206.

Smith, E. V., Wakely, M. B., De Kruif, R. E. L., & Swartz, C. W. (2003). Optimizing rating scales for self-efficacy (and other) research. *Educational and Psychological Measurement, 63*, 369-391.

Smith, M., Brunetto, W., & Yonkers, K. A. (2004). Identifying perinatal depression: Sooner is better. *Contemporary Obstetrics and Gynecology, 49*(4), 58-64.

Smith, R. M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement, 1*(2), 199-218.

Smith, R. M., Schumacker, R. E., & Bush, M. J. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement, 2*, 66-78.

Sobey, W. (2002). Barriers to postpartum depression prevention and treatment: A policy analysis. *Journal of Midwifery and Womens Health, 47,* 331-6.

Soderquist, J., Wijma, B., Thorbert, G., & Wijma, K. (2009). Risk factors in pregnancy for post-traumatic stress and depression after childbirth. *British Journal of Obstetrics and Gynaecology, 116*, 672-680.

Soderquist, J., Wijma, B., Wijma, K. (2006). The longitudinal course of post-traumatic stress after childbirth. *Journal of Psychosomatic Obstetrics and Gynaecology, 27*, 113-9.

Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Australian Occupational Therapy Journal, 56*(3), 150-155.

Spinelli, M. G. (2004). Maternal infanticide associated with mental illness: Prevention and the promise of saved lives. *American Journal of Psychiatry, 161*, 1548-1557.

Spinelli, M. G. (2005). Infanticide: Contrasting views. *Archives of Womens Mental Health, 8*, 15-24.

Spinelli, M. G. (2009). Postpartum psychosis: Detection of risk and management. *The American Journal of Psychiatry, 166*(4), 405-408.

Spinelli, M. G., & Endicott, J. (2003). Controlled clinical tiral of interpersonal psychotherapy versus parenting education program for depressed pregnant women. *American Journal of Psychiatry, 160*, 555-562.

Srisaeng, P. (2004). Self-esteem, stressful life events, social support, and postpartum depression in adolescent mothers in Thailand. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 6410-B*, 4867.

Stagnaro-Green, A. (2004). Postpartum thyroiditis. *Best Practice & Research Clinical Endocrinology & Metabolism, 18*, 303-316.

Stanley, C., Murray, L., & Stein, A. (2004). The effect of postnatal depression on mother-infant interaction, infant response to the Still-face perturbation, and performance on an Instrumental Learning task. *Development and Psychopathology, 16*, 1-18.

Stansfield, C., & Bowles, M. (2007, February). *Test Translation Matters: Back translation*. Issue 2. Retrieved from Second Language Testing, Inc website http://www.2LTI.com

Stats SA, Mid-year population estimates, South Africa, 2009 (Statistical release P0302) in Lehohla, P. (2009). South African Statistics, 2009. Statistics South Africa: Pretoria. Retrieved from www.statssa.gov.za

Stein, G. (2007). Perinatal Psychiatric Disorders. In G. Stein & G. Wilkinson (Eds), *Seminars in General Adult Psychiatry (2^{nd} ed.)* (pp.635-667). London: Royal College of Psychiatrists Publications.

Stein, G., & Van den Akker, O. (1992) The retrospective diagnosis of postnatal depression by questionnaire. *Journal of Psychosomatic Research, 36*, 67-75.

Steiner, M. (1990). Postpartum psychiatric disorders. *Canadian Journal of Psychiatry, 35*, 89-95.

Stern, G., & Kruckman, L. (1983). Multidisciplinary perspectives on postpartum depression: An anthropological critique. *Social Science and Medicine, 17*, 1027-1041.

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95*(2), 334-344.

Stowe, Z. N., Levy, S. T., & Nemeroff, C. B. (1997). Comorbid Postpartum Depression and Bereavement: A Complicated Case. *American Journal of Psychiatry, 154*, 418-422.

Stuart, S., Couser, G., Schilder, K., O'Hara, M. W., & Gorman, L. (1998). Postpartum anxiety and depression: Onset and comorbidity in a community sample. *Journal of Nervous Mental Disorders, 186*, 420-424.

Stuchbery, M., Matthey, S., & Barnett, B. (1998). Postnatal depression and social supports in Vietnamese, Arabic and Anglo-Celtic mothers. *Social Psychiatry and Psychiatric Epidemiology, 33*(10), 483-490.

Sugawara, M., Sakamoto, S., Kitamura, T., Toda, M. A., & Shima, S. (1999). Structure of depressive symptoms in pregnancy and the postpartum period. *Journal of Affective Disorders, 54*, 161-169.

Suri, R. (2004, October 1). Postpartum Depression: Risk Factors and Treatment Options. *Psychiatric Times*, pp. 64-68.

Sutter-Dallay, A. L., Giaconne-Marcesche, V., Glatigny-Dallay, E., & Verdoux, H. (2004). Women with anxiety disorders during pregnancy are at increased risk of intense postnatal depressive symptoms: A prospective survey of the MATQUID cohort. *European Psychiatry, 19*, 459-463.

Sved-Williams, A. (2003). Perfectionism, older first time mothers and their infants. *Archives of Womens Mental Health, 6*, Suppl.2, s83-s109.

Swendsen, J. D., & Mazure, C. M. (2000). Life stress as a risk factor for postpartum depression: Current research and methodological issues. *Clinical Psychology: Science and Practice, 7*, 17-31.

Tamaki, R., Murata, M. & Okano, T. (1997). Risk factors for postpartum depression in Japan. *Psychiatry and Clinical Neuroscience, 51*, 93-98.

Tentoni, S. C. & High, K. A. (1980). Culturally induced postpartum depression: Theoretical position. *JOGN Nursing: Journal of Obstetric, Gynecologic and Neonatal Nursing, 9*(4), 246-249.

Teresi, J. A. (2006). Different approaches to differential item functioning in health applications: Advantages, disadvantages and some neglected topics. *Medical Care, 44*(11), Suppl 3, S152-S170.

Theron Report (The) (1976). *Report of the Commission of Inquiry into matters relating to the coloured population group* (R.P. 38-1976). Pretoria: Government Printer.

Thio, I. M. (2004). Resilience against stress and depression in the postpartum period. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 65*,1 B, 454.

Thurtle, V. (1995). Post-natal depression: The relevance of sociological approaches. *Journal of Advanced Nursing, 22*, 416-424.

Trivedi, M. H., Rush, A. J., Crismon, M. L., Kashner, T. M., Toprac, M. G., Carmody, T. J., …Shon, S. P. (2004). Clinical results for patients with major depressive disorder in the Texas Medication Algorithm Project. *Archives of General Psychiatry, 61*(7), 669-680.

Trivedi, M. H., Rush, A. J., Ibrahim, H. M., Carmody, T. J., Biggs, M. M., Suppes, T., Crismon, M. L., Shores-Wilson, K., Toprac, M. G., Dennehy, E. B., Whitte, B., & Kashner, T. M. (2004). The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), the Quick Inventory of Depressive Symtomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: A psychometric evaluation. *Psychological Medicine, 34*, 73-82.

Trochim, W. M. K. (2006). Measurement Validity Types. Retrieved March 29, 2011, from Research Methods Knowledge Base Web site: http://www.socialresearchmethods.net/kb/measval.php

Troutman, B. R., & Cutrona, C. E. (1990). Nonpsychotic postpartum depression among adolescent mothers. *Journal of Abnormal Psychology, 99*, 69-78.

Turton, P., Hughes, P., Evans, C. D. H., & Fainman, D. (2001). Incidence, correlates and predictors of post-traumatic stress disorder in the pregnancy after stillbirth. *British Journal of Psychiatry, 178*, 556-560.

Ugarriza, D. N. (2000). Screening for postpartum depression. *Journal of Psychosocial Nursing and Mental Health Services, 38*(12), 44-53.

Van de Vijver, F. J. R. (2001). The evolution of cross-cultural research methods. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 77-97). NY: Oxford University Press.

Van de Vijver, F. J. R. (2002). Cross-Cultural Assessment: Value for Money? *Applied Psychology: An International Review, 51*, 545-566.

Van de Vijver, F. J. R., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89-99.

Van de Vijver, F. J. R., & Leung, K. (1997a). Methods and data analysis of comparative research. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology* (2nd ed., pp. 257-300). Boston: Allyn & Bacon.

Van de Vijver, F. J. R., & Leung, K. (1997b). *Methods and data analysis for cross-cultural research*. New York: Sage Publications.

Van de Vijver, F. J. R., & Leung, K. (2001). Methodological issues in psychological research on culture. *Journal of Cross-Culture Psychology, 31*, 33-51.

Van de Vijver, F. J. R., & Lonner, W. J. (1995). A bibliometric analysis of the Journal of Cross-Cultural Psychology. *Journal of Cross-Cultural Psychology, 26*, 591-602.

Van de Vijver, F. J. R., & Poortinga, Y. H. (1997). Towards an Integrated Analysis of Bias in Cross-Cultural Assessment. *European Journal of Psychological Assessment, 13*, 29-37.

Van de Vijver, F. J. R., & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds). (pp.39-64). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. London: Lawrence Erlbaum Associates, Publishers.

Van de Vijver, F. J. R., & Rothmann, S. (2004). Assesment in multicultural groups: The South African case. *South African Journal of Industrial Psychology, 30*(4), 1-7.

Van de Vijver, F. J. R., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology, 47*, 263-280.

Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée, 54*(2), 119-135 .

Van der Ross, R. E. (1979). *Myths and attitudes: An inside look at the coloured people.* Cape Town: Tafelberg.

Van der Ven, A. H. G. S., & Ellis, J. L. (2000). A Rasch analysis of Raven's Standard Progressive Matrices. *Personality and Individual Differences, 29*(1), 45-64. doi:10.1016/S0191-8869(99)00177-4

Van Ede, D. M. (1996). How to adapt a measuring instrument for use with various culture groups: A practical step-by-step introduction. *SAJHE/SATHO, 10*, 153-160.

Van Eeden, R. & Prinsloo, C. H. (1997). Using the South African version of the 16PF in a multicultural context. *South African Journal of Psychology, 27*, 151-159.

Van Widenfelt B. M., Treffers P. D., De Beurs E., Siebelink B. M., & Koudijs E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review, 8*(2), 135-147.

Venter, A. J. (1974). *Coloured: A profile of two million South Africans.* Cape Town: Human & Rousseau.

Verkerk, G. J. M., Denollet, J., Van-Heck, G. L., Van-Son, M. J. M., & Pop, V. J. M. (2005). Personality factors as determinants of depression in postpartum women: A prospective 1-year follow-up study. *Psychosomatic Medicine, 67*, 632-637.

Verkerk, G. J. M., Pop, V. J. M., Van Son, M. J. M., & Van Heck, G. L. (2003). Prediction of depression in the postpartum period: A longitudinal follow-up study in high-risk and low-risk women. *Journal of Affective Disorders 77*, 159-166.

Viljoen, A. G., & Grobler, E. (1972). A comparison between the moral codes of American, Korean and a group of Afrikaans-speaking South African students. *Journal of Social Psychology, 86*, 147-149.

Vittayanont, A., Liabsuetrakul, T., & Pitanupong, J. (2006). Development of Postpartum Depression Screening Scale (PDSS): A Thai version for screening postpartum depression. *Journal of the Medical Association of Thailand, 89*(1), 1-7.

Waldenström, U., Hildingsson, I., & Ryding, E. L. (2006). Antenatal fear of childbirth and its association with subsequent caesarean section and experience of childbirth. *BJOG: An International Journal of Obstetrics & Gynaecology, 113*, 638-646

Walker, L. O. (1997). Weight and weight-related distress after childbirth: Relationships to stress, social support, and depressive symptoms. *Journal of Holistic Nursing, 15*, 389-405.

Walther, V. N. (1997). Postpartum depression: A review for perinatal social workers. *Social Work in Health Care, 24*(3-4), 99-111.

Wang, S., Jiang, X., Jan, W., & Chen, C. (2003). A comparative study of postnatal depression and its predictors in Taiwan and mainland China. *American Journal of Obstetrics and Gynecology, 189*(5), 1407-1412.

Warner, R., Appleby, L., Whitton, A., & Faragher, B. (1996). Demographic and obstetric risk factors for postnatal psychiatric morbidity. *British Journal of Psychiatry, 168*, 607-611.

Watson, C. G., Juba, M. P., Manifold, V., Kucala, T., & Anderson, E. D. (1991). The PTSD interview: Rationale, description, reliability, and concurrent validity of a DSM-III based technique. *Journal of Clinical Psychology, 47*, 179-188.

Weinberg, M. K., Tronick, E. Z. (1998). The impact of maternal psychiatric illness on infant development. *The Journal of Clinical Psychiatry, 59*(Suppl 2), 53-61.

Weinberg, M. K., Tronick, E. Z., Beeghly, M., Olson, K. L., Kernan, H.,& Riley, J. M. (2001). Subsyndromal depressive symptoms and major depression in postpartum women. *American Journal of Orthopsychiatry, 71*, 87-97.

Whiffen, V. E. (1988). Vulnerability of postpartum depression: A prospective multivariate study. *Journal of Abnormal Psychology, 97*, 467-474.

Whiffen, V. E. (1992). Is postpartum depression a distinct diagnosis? Clinical Psychology Review, 12, 485-508.

Whiffen, V. E., & Gotlib, I. H. (1989). Infants of postpartum depressed mothers: temperament and cognitive status. *Journal of Abnormal Psychology, 98*(3), 274-279.

White, L.A. (1959). *The evolution of culture: The development of civilization to the fall of Rome.* New York: McGraw-Hill.

Whitton, A., Warner, R., & Appleby, L. (1996). The pathway to care in post-natal depression: Women's attitudes to post-natal depression and its treatment. *British Journal of General Practice, 46* (408), 427-428.

Wickberg, B., & Hwang, C. P. (1997). Screening for postnatal depression in a population-based Swedish sample. *Acta Psychiatrica Scandinavica, 95*, 62-66.

Wijma, K. (2003). Why focus on ‘fear of childbirth’? *Journal of Psychosomatic Obstetrics and Gynecology, 24*, 3, 141-143.

Wile, J. & Arechigo, M. (1999). Sociocultural aspects of postpartum depression. In L. J. Miller (Ed.), *Postpartum Mood Disorders* (pp. 83-98). Washington, D.C.: American Psychiatric Press.

Wilkinson, M. G. (2001). Postnatal depression. *Psychiatry Update, 17*, 28-31

Wisner, K. L., & Stowe, Z. N. (1997). Psychobiology of postpartum mood disorders. *Brain, Behaviour and Reproductive Function, 15*, 77-90.

Wisner, K. L., Gelenberg, A. J., Leonard, H., Zarin, D., & Frank, E. (1999). Pharmacologic treatment of depression during pregnancy. *The Journal of the American Medical Association, 282*(13), 1264-1269.

Wisner, K. L., Zarin, D., Holmboe, E., Appelbaum, P., Gelenberg, A. J., Leonard, H. L., & Frank, E. (2000). Risk-benefit decision making for treatment of depression during pregnancy. *American Journal of Psychiatry, 157*, 1933-1940.

Wolf, A.W., De Andraca, I., & Lozoff, B. (2002). Maternal depression in three Latin American samples. *Social Psychiatry and Psychiatric Epidemiology, 37*(4), 169-176.

Wright, B. D. (1995). Which Standard Error? *Rasch Measurement Transactions, 9*(2), p. 436. Retrieved from www.rasch.org/rmt/rmt92n.htm

Wright, B. & Stone. M. (1999). *Measurement essentials* (2nd ed.). Wilmington, Delaware: Wide Range, Inc.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), p. 370. Retrieved from www.rasch.org/rmt/rmt83b.htm

Wright, B. D., & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: theory, models and applications* (pp. 1-24). Maple Grove, Minnesota: JAM Press.

Wu, P. & Chang, L. (2008). Psychometric properties of the Chinese version of the Beck Depression Inventory-II using the Rasch model. *Measurement and Evaluation in Counseling and Development, 41*(1), 13-31.

Yonkers, K. A. (2003). Special issues related to the treatment of depression in women. *Journal of Clinical Psychiatry, 64* (Suppl. 18), 8-13.

Yonkers, K. A., & Sampson, J. (2000). Mood disorder measures: American Psychiatric Association. In: J. A. Rush (Ed.), *Handbook of Psychiatric Measures* (pp. 515-548). Washington, DC: American Psychiatric Association.

Yonkers, K. A., Ramin, S. M., Rush, A. J., Navarrete, C. A., Carmody, T., March, D., Heartwell, S. F. & Leveno, K. J. (2001). Onset and persistence of postpartum depression in an inter-city maternal health clinical system. *American Journal of Psychiatry, 158*, 1856-1863.

Yoshida, K., Marks, M. N., Kibe, N., Kumar, R., Nakano, H., & Tashiro, N. (1997). Postnatal depression in Japanese women who have given birth in England. *Journal of Affective Disorders, 43*(1), 69-77.

Yoshida, K., Yamashita, H., Ueda, M., & Tashiro, N. (2001). Postnatal depression in Japanese mothers and the reconsideration of 'Satogaeri bunben.' *Pediatrics International, 43*(2), 189-193.

Yu, C. H. (2010). *A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling.* Retrieved from http://www.creative-wisdom.com/computer/sas/IRT.pdf

Zapata, L. B. (2005). Associations between maternal depression and child social competence and display of problem behaviors: A longitudinal investigation of direct, indirect and moderating effects. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66* (3-B), 1431.

Zeidner, M., Matthews, G., & Roberts, R. D. (2004). Intelligence theory, assessment and research: The Israeli experience. In R. J. Sternberg (Ed.), *International handbook of intelligence* (pp. 212-247). UK: Cambridge University Press.

Zelkowitz, P., & Milet, T. H. (1995). Screening for post-partum depression in a community sample. *Canadian Journal of Psychiatry, 40*, 80-86.

Zelkowtiz, P., & Milet, T. H. (1996). Postpartum psychiatric disorders: Their relationship to psychological adjustment and marital dissatisfaction in the spouses. *Journal of Abnormal Psychology, 105*, 281-285.

Zuckerman, B., Bauchner, H., Parker, S., & Cabral, H. (1990). Maternal depressive symptoms during pregnancy and newborn irritability. *Journal of Developmental and Behavioural Pediatrics, 114*, 190-194.