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.