A Culturally Relevant and Responsive Approach to Screening for Perinatal Depression

Sarah Kye Price
Virginia Commonwealth University

Sandii Leland Handrick
St. Louis Healthy Start

Objectives: This study presents the design, implementation, and evaluation of a culturally relevant and responsive approach to screening for perinatal depression in low-income, predominantly African American women. **Method:** The study details the development of the community-informed instrument and subsequent evaluation of its psychometric properties relative to one specific community sample ($N = 139$). **Results:** The initial validity and reliability testing of the community-informed instrument reveal a higher internal consistency than the standardized instrument alone (alpha = .87 vs. .80). Subsequent factor analysis and inter-item correlation suggest consistency between most of the community-informed and CES-D items. **Conclusions:** While limited in scope to the community sample, the study reflects the perceived benefit of integrating community voice to enhance culturally relevant and empirically rigorous practice.

Keywords: *Depression Screening; Perinatal Depression; Community-based Participatory; Research; Maternal and Child Health; Healthy Start*

During the past decade, perinatal depression has emerged as a significant focus of attention for community-based organizations and programs providing case management to pregnant, post-partum, and newly parenting women. In the delivery of community-based social work in maternal and child health settings, perinatal depression screening has been widely implemented either by a mandate from funding sources or in response to a perceived need and benefit from standardized screening. However, all depression screening instruments are not necessarily created equal in their ability to detect perinatal depression, nor do they reflect the cultural and linguistic characteristics of communities at highest risk for fetal and infant mortality.

In order to develop a more culturally relevant and responsive approach to perinatal depression screening, we conducted a study which integrated the empirical information from existing literature with the cultural knowledge of one targeted community in order to establish relevant and rigorous practice and service delivery within an urban maternal and child health case management program. This article examines the culture and characteristics of the community sample, the process of developing a community-informed approach to screening, and the empirical evaluation of psychometric rigor using this approach as an aspect of enhancing social work and community maternal and child health practice.

Perinatal Depression in a Social and Cultural Context

The time during and around pregnancy has been identified as a period of elevated risk for emergent depressive symptoms (Beeghly et al., 2003; Evans, Heron, Francomb, Oke, & Golding, 2001; O'Hara & Swain, 1996). In spite of the well-documented prevalence, depression is less commonly detected in community obstetric/gynecology clinic settings than several other major mental health disorders (Smith et al., 2004), contributing to the concern that depression often remains undetected

in women served through the public health sector (Miranda, Azocar, Komaromy, & Golding, 1998). Community-based maternal and child health intervention programs, with a simultaneous focus on promoting positive health outcomes for low-income mothers and babies, foster the opportunity to impact public health practice by including maternal mental health as a part of the provision of comprehensive person-centered perinatal care. Indeed, research studies reinforce the importance of maximizing this time period to enhance the physical as well as mental health of mothers which, in turn, may improve the quality of parent–child relationships (DiPietro, Novak, Costigan, Atella, & Reusing, 2006; Milgrom, Ericksen, McCarthy, & Gemmill, 2006; Murray, Fiori-Cowley, Hooper, & Cooper, 1996; O'Hara & Swain, 1996).

Maternal and child health intervention programs such as Healthy Start, Bright Futures, and Healthy Beginnings as well as state-level maternal and child health Title V block grant programs target services to socioeconomically challenged communities and communities of color because of pervasive racial-ethnic and socioeconomic disparities in fetal and infant mortality. Traditionally, these maternal and child health programs focus on physical health and well-being, with less attention on maternal mental health and psychosocial well-being. Generally, the in-home case management and health education support services are provided by a combination of professional nurses (RN, LPN), social workers, and/or paraprofessional home visitors. While the elevated risk for negative birth outcomes in the target population has driven funding for these programs, women within these communities may also be at risk for elevated rates of perinatal depression.

There is a growing literature base focused on depression screening, prevalence, and intervention in low-income and ethnic minority populations, including the target group for our study: low-income, urban African American mothers. Indeed, Beeghly et al. (2003) report elevated prevalence rates of perinatal depression as high as 35% for low-income, African American mothers enrolled in their study. Similarly, Howell, Mora, Horowitz, and Leventhal (2005) report elevated depression symptoms in 43.9% of low-income, African American women, and Zayas, Cunningham, McKee, and Jankowski (2002) report elevated symptoms in 51% of low-income, African American and Latina mothers. The rates of depressive symptoms within these targeted communities far surpass the often quoted 10% to 15% prevalence rate estimated in the overall population (O'Hara & Swain, 1996). Furthermore, the combination of demographic and psychosocial risk factors further exacerbates the risk of depression for many low-income mothers (Beeghly et al., 2003; Zayas et al., 2002).

Low-income women also face additional challenges in attempting to access services for perinatal depression when health care, counseling, and specialty mental health services are not available, accessible, or affordable (Cook, Selig, Wedge, & Gohn-Baube, 1999; Miranda et al., 2003; Rosen, Warner, & Tolman, 2006; Song, Sands, & Wong, 2004). Within this context, the provision of depression screening in community-based maternal and child health programs serving low-income women is not only about detection of symptoms for those at potentially elevated risk but also about providing a gateway to mental health services for women who may not otherwise interface with this sector of care.

Successful introduction of perinatal depression screening into a community is also dependent on how receptive community members are to the idea of discussing and identifying depression. Race, ethnicity, and culture may impact the ways in which depressive symptoms are experienced and presented (U.S. Department of Health and Human Services, 2001) as well as influence willingness to seek mental health treatment based on institutional barriers of mistrust and stigma that have emerged over time (Fortier & Bishop, 2003). Knowledge about how members of the community define and describe depression as well as what factors influence whether or not to seek help are vital to program planning that engages the community in a culturally meaningful way to address a mutually defined social problem. As discussed by Zayas, McKee, and Jankowski (2004), an inherent willingness to engage the community as a full partner in collaborative decision making around process, goals, and desired outcomes may be vital to the lasting success of adapting evidence-based interventions within community settings. A social work perspective on the identification of perinatal depression suggests that we should be attuned to the cultural experiences of the community in which maternal and child health programs reside, as well as focusing attention on empirically supported approaches for screening and identification that resonate with the cultural experiences within that community.

Practice Standards in Perinatal Depression Screening

There are a myriad of self-report screening instruments that have been validated for use in assessment for depressive symptoms during pregnancy and/or the postpartum period. In their review of screening instruments for postpartum depression, Boyd, Le, and Somberg (2005) illustrate that there is no ideal screening instrument for

postpartum depression that can be universally applied. Summarizing the empirical studies conducted using these standardized measures, screening instruments such as the Edinburgh Postnatal Depression Scale (EPDS) and Postpartum Depression Screening Scale (PDSS) have a greater sensitivity to postpartum depression and/or emotional distress that does not necessarily meet a diagnostic threshold for major mental health disorders, while others such as the Beck Depression Inventory (BDI-II) and the Center for Epidemiological Studies–Depression Instrument (CES-D) demonstrate a higher specificity for detecting clinically significant levels of depression but less sensitivity to symptoms that may differentiate postpartum depression from major depression (Boyd et al., 2005).

Similarly, some instruments are intended to be diagnostic (i.e., BDI-II, PDSS), while others have been developed to capture epidemiological prevalence (i.e., CES-D). A few instruments, such as the 9-item Patient Health Questionnaire (PHQ-9) derived from the Primary Care Evaluation of Mental Disorders (Spitzer et al., 1994), have even been implemented and tested in multiple settings and diverse populations with scoring guidelines that accurately reflect diagnostic levels of depression severity (Kroenke, Spitzer, & Williams, 2001). Finally, there is the consideration of whether instruments focusing more predominantly on somatic symptoms (i.e., fatigue, weight loss/gain, sleep changes) can adequately differentiate between medical symptoms of pregnancy or somatic symptoms that may reflect mental health challenges (Kroenke, Spitzer, & Williams, 2002).

Applying this conceptualization of selection of screening instruments to direct practice settings, a case management program seeking to support and intervene with any and all women potentially at risk for postpartum depression may, for example, wish to select a highly sensitive instrument that errs toward over-detection of potentially distressing symptoms in their clientele. In this first case, an instrument such as the EPDS that is sensitive to the levels of psychosocial distress that could suggest risk for postpartum depression might be a good choice, in spite of its relatively low positive predictive value for clinical levels of depression which may be as low as 50% (Eberhard-Gran, Eskild, Tambs, Opjordsmoen, & Samuelsen, 2001). Conversely, a community-wide intervention program seeking to isolate women at highest risk and provide targeted linkage with specialized mental health services may prefer a highly specific instrument that can be briefly administered to a large number of women so that limited services are most effectively triaged. In this later case where greater specificity is needed to identify probable clinical depression in a large cross-section of women, a brief screening tool such as the 2-item Primary Care Evaluation of Mental

Disorders (PRIME-MD) may be an effective and valid choice (Howell et al., 2005; Kronke, Spitzer & Williams, 2003; Whooley, Avins, Miranda & Browner, 1997).

In spite of the acknowledgment that cultural experiences and expressions of depression may vary among racial and ethnic groups, very few studies have examined the cross-cultural validation of these various depression screening instruments either within the United States or in cross-cultural samples specific to pregnant and/or postpartum women (Boyd et al., 2005). For social workers in maternal and child health practice settings, this presents a challenge for grounding empirically supported screening within a cultural framework that is relevant to a specific population perceived to be at risk. Working with the community to select a screening instrument that does not falsely inflate or diminish the reported presence of depression within a community also advances practice standards by mutually addressing the social problem of perinatal depression without creating or perpetuating stigma about mental health within that community. Communication between researcher and agency regarding the benefits and limitations of each instrument is essential in order to strike a balance between community preferences and advantages and/or challenges to research rigor.

Method

Agency and Community Setting

This study was conducted in an urban, Midwestern city, housed in an agency providing services to low-income, predominantly African American women and families through a Healthy Start project. Healthy Start is a federally funded program providing outreach, health education, and case management to women who are pregnant, post-partum, or in the 2 years after the birth of a baby. Healthy Start projects are competitively funded and strategically situated in geographic areas with elevated racial-ethnic disparities in fetal and infant mortality (see http://mchb.hrsa.gov/about/dhsps.htm). In 2004, depression screening was included as a core service component in Healthy Start projects around the country, although the selection of depression screening instrumentation was left to the discretion of each funded project.

The Healthy Start project site which served as the basis for this study encompasses three specific zip code areas in the metropolitan area with the highest levels of disparity between African American and Caucasian infant mortality rates. Healthy Start consumers enrolled in the project were all low-income women of reproductive age, with over 95% self-identifying as African American and many of whom were adolescent parents.

At the time of project start-up, the first author was contracted by the agency (represented by the second author) to provide consultation around initiating perinatal depression screening, provider education, and the development of risk-management approaches for women identified with depression.

Informing Screening From Scholarly Literature

The initial task within this community-research partnership was to conduct a literature review on the major screening instruments available for consideration as a universal screening tool for Healthy Start consumers. The literature review focused on screening instruments that had been demonstrated to have respectable scientific validity, reliability, and internal consistency in the research literature. As discussed previously, the literature review included instruments that were either sensitive to detection of any mood disorder or psychosocial distress in the perinatal period, as well as those specific to identification of women at highest clinical risk for depression and, therefore, at greatest need for intervention. The goal of the broad literature search was to allow the agency to explore a range of options in developing a screening and risk assessment protocol.

A specific question raised by the community agency was whether any of the instruments in the literature had been tested or validated for use specifically with low-income and/or African American women. Therefore, the literature review was expanded to include recent articles where screening and/or intervention with these groups made mention of the validity of a specific screening instrument for that sample. The author provided an examination copy of each potential screening instrument to the community agency, along with a summary of findings from the empirical literature that outlined the strengths, limitations, and cultural applicability (if any) for each of these screening instruments. Potential screening instruments included the CES-D full version, the 10-item CES-D short form, the PRIME-MD, the BDI-II, and the EPDS.

A major concern identified in this process was the lack of documented cultural relevance or population-specific validity of these standardized screening instruments in the published literature. The concern was reinforced for the agency staff when they reviewed the wording of the instruments and perceived it to be less than culturally relevant to the women and families served by the agency. Issues emerged around topics such as the reading/language level of the questions, the use of expressions that were inconsistent with the language used by consumers in the program, and the use of emotional terms that felt non-representative of what was described as the community's overt, behavioral manifestations of potential depression. The team made the recommendation to pursue independent research to determine if there was a way to culturally inform and enhance an existing screening instrument and whether this maintained or enhanced the scientific rigor of the instrumentation. The team recommended a strategy of community participation in the development of questionnaire items that reflected the community's own perceptions about symptoms suggesting potential perinatal depression which could be "added on" to a standardized screening instrument in order to maximize both relevance and rigor.

Informing Screening From Community and Cultural Voice

As a first step, face-to-face meetings were conducted with various stakeholder groups and individuals representing "insider" perspectives on the project, including participation from management and administration, supervisory program staff, direct service staff, sub-contract agency staff, paraprofessional outreach staff (indigenous to the community), and consumers. During individual and group meetings, the central question discussed was how depression was expressed and/or experienced by the community, cultural values pertaining to depression, how to determine whether a woman's level of depression was normal versus a problem, and what would create a need or desire to seek help. Emergent themes and specific wording about depression were brought back to the overall team for discussion. Simultaneously, the agency's leadership solicited feedback from key stakeholders regarding their preferred choice among the standardized screening instruments that had been reviewed.

Women in the community reflected strong cultural values surrounding depression and depression screening. Most notably, women expressed that "[African Americans] do not have the privilege of being 'depressed' because we have to be strong, go forward, take care of our own." Participants also noted that depression was often expressed through anger and irritability, rather than a state of feeling "down" or depressed. Another important point echoed by women in the community regarding their cultural experiences of being urban, African American mothers was expressed in the words of one participant: "We look out for our own . . . and if my sister says she's depressed then I think, yeah, I know, we all are. But if she thinks that she might hurt herself or that something bad is going to happen to her, I know it's a real problem, and I've got her back."

When all stakeholder feedback was considered and after receiving detailed summaries of empirical findings

regarding the benefits and limitations of each potential instrument, the organization chose the CES-D short form as the foundation for screening. Statements made by the team were that the CES-D was easy to understand, readable, and specific for identifying women at highest risk and had respectable validity in widespread and diverse (although not necessarily similar) community epidemiology samples. The team then used the community's own expressions, which were presented via content analysis of community meetings, to develop five additional "tag-on" screening items based on community-derived language describing experiences of depression, operationalized to fit within the measurement format of the CES-D items. The CES-D plus the five community-informed items comprised the depression screening questionnaire that was implemented with all project consumers. A full listing of the questionnaire items (CES-D plus community-informed tag-on items) is shown in Table 1.

In order to sensitize staff members to the screening instrument that was to be administered as well as the culturally relevant and community-informed approach that guided its development, the first author conducted a full-day seminar to provide an overview of knowledge regarding perinatal depression, review the development of the screening protocol, discuss depression within a cultural context, and engage in open-ended exercises to break down cultural barriers in communication about depression. The training was held twice, with mandatory participation from all staff members of the agency and its subcontractors (approximately 45 attendees) and with input and active participation (and leadership) of paraprofessional outreach workers who were culturally representative of the local community.

Psychometric Evaluation of Community-Informed Screening

In order to assess the validity and internal consistency of the additional community-informed items, an evaluation protocol was established to collect two waves of data. The first wave of data collected baseline screening on all new and existing clients. The second wave of data followed the newly enrolled clients who received screening with a 3-month repetition of the screening. Data from both the baseline and follow-up administration were analyzed to describe the patterns of symptoms in clients served and to assess for validity, internal consistency, and test–retest reliability of the instrument in this specific community setting.

Screening data were recorded electronically along with other assessment data and maintained as part of the sub-contract agency's records. Data for Wave 1 and

Table 1
Center for Epidemiological Studies-Depression Instrument (CES-D) Items (1-10) Plus Community-Informed Tag-On Questions (11-15)

1. I was bothered by things that don't usually bother me.
2. I had trouble keeping my mind on what I was doing.
3. I felt depressed.
4. I felt that everything I did was an effort.
5. I felt hopeful about the future (reverse coded).
6. I felt fearful.
7. My sleep was restless.
8. I was happy (reverse coded).
9. I felt lonely.
10. I could not get going.
11. I felt like everyone was against me.
12. I felt angry or irritable.
13. I felt like giving up.
14. I felt like hurting myself, my pregnancy/baby, or others.
15. I worried that something might happen to me or my baby.
    Scale (all items), based on the past week:
    0 = rarely or none (less than 1 day)
    1 = some or a little (1-2 days)
    2 = occasionally or moderate (3-4 days)
    3 = all of the time (5–7 days)

Wave 2 were released to the first author for baseline assessment of need and psychometric testing at designated time intervals (6 months and 12 months) after implementation. The data were deidentified when transferred, with only a time/date stamp and non-identifiable evaluation ID number. Non-identifiable demographic data (age, race-ethnicity, status as pregnant or post-partum) were transmitted along with depression screening data. This study was reviewed and designated as exempt by the Human Subjects Review Committee at the first author's host institution based on participant anonymity in the analysis of secondary data from agency records. Aggregated results from the data analysis were provided to the organization for internal review and team discussion of findings.

For each wave of data, psychometric tests were conducted on the administration of the CES-D items alone, and the CES-D items combined with community-informed tag-on questions. Descriptive statistics considered depressive symptoms by client age, pregnancy versus post-partum status, and race-ethnicity. Statistical analysis used SPSS 14.0 and included computations of variables, split file operations to compare demographic groups and report descriptive differences in response to CES-D and tag-on questions, and calculation of CES-D and total screening score comparisons using independent sample t tests, one-way analyses of variance, and chi-square procedures. Statistical significance was reported using 95% confidence intervals ($p < .05$).

Table 2
Demographic Categories of Baseline and Follow-Up Participants

	Baseline (*N* = 139)		Follow-Up (*N* = 40)	
Maternal race by category				
African American or Black, non-Hispanic	133	95.7%	37	94.9%
Caucasian or White, non-Hispanic	6	4.3%	2	5.1%
Maternal age by category				
Under 18	62	44.6%	21	53.8%
19-23	36	25.9%	8	20.5%
24-28	26	18.7%	8	20.5%
29 and over	15	10.8%	2	5.1%
Pregnancy and prenatal status				
Prenatal	37	26.6%	11	27.5%
Post-partum	24	17.3%	16	40.0%
Missing/unknown	78	56.1%	13	32.5%

Validity testing of the screening instrument consisted of initial inter-item correlation testing and factor analysis of the CES-D items alone as well as the total screening score including tag-on questions. Next, two criterion validity tests were conducted. The first used the suggested cut-off score of 10 for the CES-D short-form items (Carpenter et al., 1998; Radloff, 1977) indicative of elevated risk for depression as a criterion, while the second used the agency's comprehensive assessment of risk level that integrated specific groupings of symptoms, as well as presence of psychosocial risk factors and life events to assess each person as meeting high-, medium-, or low-risk criteria. Assessment of the client in a high-risk category served as the criteria for the second criterion validity test. The internal consistency of the CES-D screener alone and the CES-D plus tag-on item combined screener was conducted using Cronbach's alpha coefficients, again comparing the CES-D alone and the CES-D with community tag-on items. Finally, test–retest reliability was ascertained through a series of paired sample *t* tests from baseline to follow-up administration in the newly enrolled Healthy Start consumers.

Results

Study Sample Demographics and Depressive Symptom Prevalence

For this study, baseline screening results were available for 139 women; 40 participants had repeat screens completed at approximately 3 months after baseline. Incomplete screens or those with missing items were removed from the final analysis for calculation of total CES-D scores and combined total scores, as well as in reliability and validity testing to avoid skewing the results. The study sample was reflective of the overall composition of this Healthy Start project. As noted in Table 2, over 95% of baseline participants were African American (*N* = 133), and 44% (*N* = 62) were under the age of 18. These sample characteristics were relatively consistent at follow-up (*N* = 40). The prevalence of depressive symptoms in women served by the agency was examined first using the CES-D alone (range = 0-30), using the suggested cut-off score of 10 to indicate possible major depression. The mean baseline CES-D score in the sample was 7.51 (*SD* = 5.54), with a median score of 7. The mean follow-up CES-D score was 7.63 (*SD* = 6.72), with a median score of 6. For participants in the baseline evaluation who completed all 10 CES-D items (*N* = 135), 44 (31.7%) were above the published cut-off score on the CES-D; for participants in the follow-up evaluation who completed all 10 CES-D items (*N* = 40), 14 (35%) were above this cut-off score. Adding the community-informed tag-on items increased the total scoring range to 0 to 40. The mean baseline CES-D plus tag-on question (Total Score) was 9.69 (*SD* = 7.89), with a median score of 8. The mean follow-up baseline CES-D plus tag-on questions (Total Score) was also 9.69 (*SD* = 9.28), with a median score of 8 for the 39 women completing these items. In both standard instrumentation and community-informed screening, there was a wide degree of variability in item response across the range of possible scores (evidenced in the standard deviation), with the mean score suggesting the presence of at least a mild-moderate depressive symptom level for many women.

Psychometric Evaluation of Community-Informed Screening

The next phase of evaluation compared the psychometric properties of standardized screening alone and community-informed screening in the study sample.

Table 3
Correlation Coefficients and Probability Between Center for Epidemiological
Studies–Depression Instrument (CES-D) and Tag-On Items

	Bothered by Things	Trouble Keeping Mind	Felt Depressed	Everything Was an Effort	Felt Hopeful (reversed)	Felt Fearful	Sleep Restless	Felt Happy (reversed)	Felt Lonely	Could Not Get Going	Everyone Is Against Me	Angry or Irritable	Felt Like Giving Up	Felt Like Hurting Self/Baby	Worried Something Might Happen
Bothered by things	1	.42*	.61*	.39*	.15	.29*	.41*	.35*	.46*	.39*	.55*	.68*	.41*	.52*	−.047
		.000	.000	.000	.097	.001	.000	.000	.000	.000	.000	.000	.000	.000	.604
Trouble keeping mind		1	.39*	.24*	.16	.27*	.23*	.32*	.26*	.54*	.39*	.39*	.56*	.46*	−.022
			.000	.006	.062	.002	.012	.000	.004	.000	.000	.000	.000	.000	.809
Felt depressed			1	.42*	.14*	.38*	.41*	.46*	.60*	.35*	.56*	.62*	.43*	.53*	.041
				.000	.114	.000	.000	.000	.000	.000	.000	.000	.000	.000	.651
Everything was an effort				1	−.11	.27*	.32*	.128	.34*	.26*	.23*	.37*	.144	.34*	.020
					.208	.002	.000	.163	.000	.004	.011	.000	.122	.000	.829
Felt hopeful (reversed)					1	.17	.053	.34*	.09	.10	.03	−.03	.07	.05	.04
						.053	.561	.000	.303	.268	.718	.000	.446	.592	.640
Felt fearful						1	.45*	.16	.38*	.26*	.23*	.37*	.24*	.37*	.03
							.000	.087	.000	.004	.011	.000	.010	.000	.640
Sleep restless							1	.19*	.30*	.13	.30*	.43*	.12	.48*	.02
								.035	.001	.146	.001	.000	.186	.000	.86
Felt happy (reversed)								1	.37*	.28*	.38*	.34*	.20*	.21*	−.04
									.000	.000	.000	.000	.028	.02	.677
Felt lonely									1	.25*	.60*	.43*	.53*	.36*	.03
										.005	.000	.000	.000	.000	.784
Could not get going										1	.34*	.39*	.42*	.36*	.05
											.000	.000	.000	.000	.567
Everyone is against me											1	.57*	.56*	.38*	.026
												.000	.000	.000	.778
Angry or irritable												1	.41*	.46*	−.01
													.000	.000	.925
Felt like giving up													1	.44*	.03
														.000	.727
Felt like hurting self/baby														1	.02
															.811
Worried something might happen to self/baby															1

*Statistically significant at $p < .05$.

Psychometric testing of baseline participants ($N = 139$) revealed that the community-informed screener with additional tag-on items had a higher internal consistency than the CES-D alone (Cronbach's alpha = .87 vs. .80). As shown in the inter-item correlation matrix in Table 3, there is also considerable correlation between the individual tag-on items and the CES-D items, with the exception of the last item, "I'm worried that something may happen to myself or to my baby." Additionally, an examination of inter-item correlations suggests that, when reverse coded, responses to the question, "I felt hopeful about the future," are not highly correlated with other questions, except for the other reverse-coded question, "I felt happy."

Initial validity testing was conducted using factor analysis to determine whether both versions of the screening were addressing the main construct of depression. The principal component analysis for the CES-D alone and the CES-D combined with tag-on questions revealed only one principal factor (Eigenvalue > 3) for each version of the scale. This single factor is consistent with other published reports of the CES-D and its shortened forms

Table 4
Criterion Validity Factor Loadings for the
Combined Screener Using the Center for
Epidemiological Studies–Depression Instrument
(CES-D) Cut-Off Score (A) and the Project's
High-Risk Definition (B) (*N* = 139)

	A	B
I was bothered by things that don't usually bother me.	.612	.838
I had trouble keeping my mind on what I was doing.	.722	.888
I felt depressed.	.635	.844
I felt that everything I did was an effort.	.596	.806
I felt fearful.	.697	.729
My sleep was restless.	.814	.788
I felt lonely.	.670	.868
I could not get going.	.615	.796
I felt hopeful about the future (reverse coded).	.683	.943
I was happy (reverse coded).	.698	.860
I feel like everyone is against me.	.619	.878
I feel angry or irritable.	.651	.787
I feel like giving up.	.739	.780
I feel like hurting myself, my pregnancy/baby, or others.	.736	.868
I'm worried that something might happen to myself/my baby.	.690	.613

(Carpenter et al., 1998). This single factor was used in subsequent criterion validity tests. The first test of criterion validity used the published CES-D cut-off score; factor loadings for the community-informed items were in the range of .61 to .74 (see Table 4, column A). The second validity test used the project's designation of a psychosocial high-risk category based on comprehensive assessment; in this analysis, the community-informed items presented with factor loadings ranging from .61 to .87 (see Table 4, column B). In both analyses, factor loadings were relatively consistent between the community-informed and CES-D items, although the project's high-risk criterion produced noticeably higher factor loadings for most scale items.

An additional test–retest reliability analysis was conducted using 3-month follow-up data (*N* = 39); a series of paired-sample *t* tests for each baseline and follow-up administration indicated that there were no statistically significant differences in item response on either the CES-D or community-informed questions when each baseline item response was considered individually in comparison with its repeat administration item response.

Discussion and Applications to Practice

Study Sample Demographics and Depressive Symptom Prevalence

The mean scores on both the CES-D items and the tag-on items illustrate that symptoms of possible depression occur with great frequency in consumers served by this program. The fact that over 30% of participants had screening scores above the suggested cut-off score for depression on the CES-D short-form items alone in both baseline and follow-up administration suggests that symptoms of possible depression are a frequent and common occurrence in the daily lives of women served through this Healthy Start project. The presence of possible perinatal depression at nearly double the prevalence estimates of the overall population (O'Hara & Swain, 1996) is consistent with other studies involving urban, low-income, African American mothers (Beeghly et al., 2003; Howell et al., 2005; Zayas et al., 2002).

Symptoms on the community-informed screening items were endorsed with similar frequency as the original CES-D items, according to the inter-item correlation analysis conducted in this study. This suggests that the symptoms described by the community (feeling angry or irritable; feeling like everyone is against me; feeling like giving up; feeling like hurting myself, my baby, or others; worried that something might happen to me or my baby) are also prevalent in this community sample. For practitioners, these symptoms of psychosocial distress provide an opportunity for further discussion and follow-up about the impact of these specific symptoms on the woman's daily life, in addition to measuring her degree of depressive symptomatology. Culturally specific expressions that augment standardized instruments emphasize the role of screening as not a route to a "score" but a gateway through which symptoms can be expressed and explored and through which consumers can be made aware of options for intervention, support, and/or empowered decision making for mental health promotion.

Psychometric Evaluation of Community-Informed Screening

The findings from the psychometric evaluation of the community-informed screener in comparison with traditional screening methods suggest that the community's own words that define and operationalize depression may actually strengthen and contribute to an understanding of how perinatal depression is experienced in the real world where practitioners interface with the community. In this study, depression is acknowledged in both the language

of an empirically supported screener as well as the language of community experience. Through this approach, validity of the instrument was enhanced and stakeholders had an active voice in the process of integrating mental health assessment within an existing maternal and child health program. Similarly, practitioners may choose to integrate community and culturally specific questions concurrent with standardized screening instruments in order to respond to the experiences of the local community. The intention of this study was not to create a new, validated screening instrument for widespread use but rather to respond in a culturally relevant way to the community's own perceptions of how depression may present itself or be described within a cultural context and offer that framework for other community-research collaborations.

Examining the results from the factor analysis using two separate criteria also suggests that the agency's high-risk definition may be measuring more than depression alone. The consistently higher factor loadings with the agency's high-risk definition when compared with the suggested cut-off score from the CES-D warrant further discussion and subsequent research. It is possible that the agency's conceptualization of high risk as a combination of perinatal depression symptoms along with the presence of family and psychosocial stressors may be more appropriate (and accurate) in guiding referrals than meeting the criteria of a single depression screening score alone. However, this phase of the study did not incorporate a specific intervention and outcome component, which precludes assessment of its clinical results and impact. An important consideration for future evaluation and research would be the examination of referral outcomes and service utilization based on screening scores alone as compared with the more broad psychosocial stress present in the agency's high-risk definition.

An unanticipated finding from the inter-item correlation is the lack of association between "I feel hopeful about the future" (when reverse coded) and other individual screening questions. While response bias cannot be ruled out, it is less likely that a set response pattern is an explanation since the other reverse-coded item ("I feel happy") does correlate with most other CES-D and tag-on items when reverse coded. One possible consideration is that in spite of major obstacles, stressors, and challenging life situations, women in this community sample do remain hopeful. The time during and around pregnancy is one that does convey a great deal of hope for many people, and the organizations that exist to provide support and increase the likelihood of a positive birth outcome clearly attempt to instill a great degree of hope in their clients. More work with the community members

themselves may be warranted in order to understand the meaning of *hopeful* within the context of their experiences and life events and assess whether it has a perceived relationship with the experience of perinatal depression. This finding has clinical relevance to practitioners who may focus simultaneously on hopefulness about life options and parenting success while simultaneously being aware of the mental health and psychosocial challenges faced by many mothers.

Considered together, the findings from this study, while limited in their scope and generalizability, also offer hope to practitioners and researchers seeking to bridge the gap between the voice of science and the voice of the community. Through the deliberate process of integrating community voice and preferences into the development of an empirically supported screening and risk assessment protocol, a culturally relevant and responsive tool was developed that enhanced the scientific rigor of the screening instrument in this community sample. The tool's applicability to the community was confirmed by the inclusion of their own voices and words into instrumentation development and further supplemented by empirical evaluation of the psychometrics of the community-informed screening instrument.

For practitioners and researchers seeking to address issues of cultural relevance and responsiveness, the final message from this study is not the product of the instrument produced. It is the documentation of a process by which other community-informed measures and protocols can be developed, implemented, tested, and jointly evaluated through both the lens of science and the community's own regard and receptivity to being included in program design and implementation. The contribution of this study is to begin the initial groundwork needed to establish larger and more comprehensive community-based mental health services for women of reproductive age. Relevance, responsiveness, and rigor may be enhanced when all stakeholders are brought into the partnership and multiple voices and sources of knowledge come together to address the concern of perinatal depression.

References

Beeghly, M. B., Olson, K. L., Weinberg, M. K., Pierre, S. C., Downey, N., & Tronick, E. Z. (2003). Prevalence, stability and socio-demographic correlates of depressive symptoms in Black mothers during the first 18 months postpartum. *Maternal and Child Health Journal*, *7*, 157-168.

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.

Carpenter, J. S., Andrykowski, M. A., Wilson, J., Hall, L. A., Rayens, M. K., Sachs, B., et al. (1998). Psychometrics for two short forms

of the Center for Epidemiologic Studies–Depression Scale. *Issues in Mental Health Nursing*, *19*, 481-494.

Cook, C. A., Selig, K. L., Wedge, B. J., & Gohn-Baube, E. (1999). Access barriers and the use of prenatal care by low-income, inner-city women. *Social Work*, *44*, 129-134.

DiPietro, J. A., Novak, M., Costigan, K., Atella, L., & Reusing, S. (2006). Maternal psychological distress during pregnancy in relation to child development at age two. *Child Development*, *77*, 573-587.

Eberhard-Gran, M., Eskild, A., Tambs, K., Opjordsmoen, S., & Samuelsen, S. O. (2001). Review of validation studies of the Edinburgh Postnatal Depression Scale. *Acta Psychiatrica Scandinavica*, *104*, 243-249.

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.

Fortier, J. P., & Bishop, D. (2003). *Setting the agenda for research on cultural competence in health care: Final report*. Rockville, MD: U.S. Department of Health and Human Services, Office of Minority Health.

Howell, E. A., Mora, P. A., Horowitz, C. R., & Leventhal, H. (2005). Racial and ethnic differences in factors associated with early postpartum depressive symptoms. *Obstetrics and Gynecology*, *105*, 1442-1450.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606-613.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*, *64*, 258-266.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire–2: Validity of a two-item depression screener. *Medical Care*, *41*, 1284-1292.

Milgrom, J., Ericksen, J., McCarthy, R., & Gemmill, A. W. (2006). Stressful impact of depression on early mother-infant relations. *Stress and Health*, *22*, 229-238.

Miranda, J., Azocar, F., Komaromy, M., & Golding, J. M. (1998). Unmet mental health needs of women in public-sector gynecologic clinics. *American Journal of Obstetrics and Gynecology*, *178*, 212-217.

Miranda, J., Chung, J., Green, B., Krupnick, J., Siddique, J., Revicki, D., & Belin, T. (2003). Treating depression in predominantly low-income young minority women: A randomized controlled trial. *Journal of the American Medical Association*, *290*, 57-65.

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*, 2512-2526.

O'Hara, M., & Swain, A. (1996). Rates and risk of postpartum depression—A meta-analysis. *International Review of Psychiatry*, *8*, 37-54.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385-401.

Rosen, D., Warner, L., & Tolman, R. M. (2006). Comparing psychiatric service use among low-income women and women in a general household population. *Social Work Research*, *30*, 223-232.

Smith, M. V., Rosenheck, R. A., Cavaleri, M. A., Howell, H. B., Poschman, K., & Yonkers, K. A. (2004). Screening for and detection of depression, panic disorder, and PTSD in public-sector obstetric clinics. *Psychiatric Services*, *55*, 407-414.

Song, D., Sands, R. G., & Wong, Y. I. (2004). Utilization of mental health services by low-income pregnant and postpartum women on medical assistance. *Women and Health*, *39*, 1-24.

Spitzer, R. L., Williams, J. B., Kroenke, K., Linzer, M., deGruy, F. V., III, Hahn, S. R., et al. (1994). Utility of a new procedure for diagnosing mental disorders in primary care: The PRIME-MD 1000 study. *Journal of the American Medical Association*, *272*, 1749-1756.

U.S. Department of Health and Human Services. (2001). Mental Health: Culture, race, and ethnicity—a supplement to Mental Health: A report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services, Public Health Office, Office of the Surgeon General.

Whooley, M. A., Avins, A. L., Miranda, J., & Browner, W. S. (1997). Case-finding instruments for depression: Two questions are as good as many. *Journal of General Internal Medicine*, *12*, 439-445.

Zayas, L. H., Cunningham, M., McKee, M. D., & Jankowski, K. R. (2002). Depression and negative life events among pregnant African-American and Hispanic women. *Women's Health Issues*, *12*, 16-22.

Zayas, L. H., McKee, M. D., & Jankowski, K. R. (2004). Adapting psychosocial intervention research to urban primary care environments: A case example. *Annals of Family Medicine*, *2*, 504-508.