A Psychometric Study of the Modified Bridges for Newborns Screening Tool

Brian D. Stucky, Victoria K. Ngo, Ashley M. Kranz, Chandra Garber, Gabriela Castro, Wenjing Huang, Joyce S. Marks

RAND
CORPORATION

Support RAND

Make a tax-deductible charitable contribution at

www.rand.org/giving/contribute

www.rand.org

Preface

Welcome Baby, a program that First 5 Los Angeles (First 5 LA) sponsors, provides new mothers with supportive services intended to create enriching environments for their children. To identify mothers in need of these services, First 5 LA employs hospital liaisons who administer a maternal risk assessment tool, the Modified Bridges for Newborns screening tool, during postpartum interviews of mothers. First 5 LA uses risk assessment scores from the Modified Bridges to classify mothers as low/moderate or high risk; high-risk mothers are eligible for additional supportive services that are not available to low- and moderate-risk mothers.

This report describes RAND Corporation work evaluating the psychometric characteristics of the Modified Bridges. We have arranged the report according to the key research questions that First 5 LA identified: (1) Do hospital liaisons assign similar scores when using the Modified Bridges? (2) Does the Modified Bridges identify a similar level of risk among the same women as other well-validated measures of psychosocial and/or medical risk? And (3) does the Modified Bridges accurately distinguish between low-/moderate-risk mothers and high-risk mothers? This document is the full report describing our psychometric evaluation of the Modified Bridges; an additional short report is also available upon request from First 5 LA that summarizes the findings included in this document. The work was conducted by RAND Health, a division of the RAND Corporation. A profile of RAND Health, abstracts of its publications, and ordering information can be found at www.rand.org/health.

The analyses on which we base this publication were performed under contract 08948, entitled, "A Psychometric Evaluation of the Modified Bridges for Newborns Screening Tool," which First 5 LA funded. The content of this publication does not necessarily reflect the views or policies of First 5 LA. We assume full responsibility for the accuracy and completeness of the ideas presented.

Contents

Figures and Tables

Figures

Tables

Summary

Welcome Baby, a Los Angeles County–based program that First 5 Los Angeles (First 5 LA) sponsors, provides free services during pregnancy and through a child's first nine months. It currently serves 14 hospitals and approximately 9 percent of families living in high-risk Los Angeles communities.

To identify mothers in need of more-focused support, First 5 LA employs hospital liaisons who administer a maternal risk assessment tool, the Modified Bridges for Newborns screening tool, during bedside, postpartum interviews of mothers. The Modified Bridges consists of 26 questions, which receive weights that correspond to the severity of risk that each question assesses, across three subscales (Medical, Psychosocial, and Demographics/Basic Needs).

First 5 LA requested that the RAND Corporation undertake a psychometric study of the Modified Bridges to assess whether the tool is accurately assessing risk among mothers in the participating hospitals. The research reported here aimed to provide information relating to both the reliability and validity of the Modified Bridges and whether the hospital liaisons can accurately and precisely score patients via the Modified Bridges. In particular, we examined the following research questions at the request of First 5 LA:

1. Do hospital liaisons assign similar Modified Bridges scores (overall and subscores) to mock patient interviews?
2. Does the Modified Bridges identify a similar level of risk among the same women as other well-validated measures of maternal/infant risk?
3. Does the Modified Bridges accurately distinguish between low-/moderate-risk women and high-risk women?

Findings

1. Do Hospital Liaisons Assign Similar Scores When Using the Modified Bridges?

- Liaisons were found to provide similar risk scores (inter-rater reliability), and their assessments were similar over time (test–retest reliability).

  - In particular, liaison inter-rater reliabilities are highest for the Demographics/Basic Needs subscale and somewhat lower for the Medical subscale.

- Overall, liaisons perform at acceptable levels. However, experience as a liaison and use of liaison training materials, such as the Modified Bridges protocol, improved performance.

2. Does the Modified Bridges Identify a Similar Level of Risk Among the Same Women as Other Measures of Psychosocial and/or Medical Risk?

- The Modified Bridges is most comparable to maternal risk assessments sponsored by the Florida Department of Health, which assess similar content to that of the Modified Bridges.
- Given the current scoring protocol, the Modified Bridges might overidentify mothers as being high risk.
- The Modified Bridges' ability to distinguish between high and low/moderate levels of risk would improve if a higher risk threshold were implemented.

3. Does the Modified Bridges Accurately Distinguish Between Low-/Moderate-Risk Mothers and High-Risk Mothers?

- Because nearly half of mothers qualify as being at risk, the current risk threshold score of 50 cannot distinguish between the highest levels of risk.
- The current reliability of the Modified Bridges indicates, in practical terms, that First 5 LA can have only limited confidence that the given risk score for a mother reflects that mother's *true* level of risk.
  - The current scoring protocol and item weights used for the Modified Bridges have a moderate level of reliability (0.65), below the widely used reliability threshold of 0.70, while scoring the tool without the current item weights would improve its reliability to a more acceptable level (0.74).
- Given the lower reliability of the Modified Bridges, we considered potential revisions to the item weights in order to improve the Modified Bridges' ability to distinguish between high- and low-risk mothers.
  - Given the relationship between the item weight and items' ability to distinguish between risk levels, we propose increasing the weight of seven items and decreasing the weight of three items.
  - In addition to these changes, we recommend removing two items: infant medical problems, which is redundant with another item on the Modified Bridges that performs better, and mother's use of English, which does not distinguish between maternal risk levels.
- Implementing these changes would improve the Modified Bridges' ability to distinguish between risk levels and increase the tool's reliability from 0.65 to 0.73 (standardized reliability = 0.75).

Acknowledgments

Abbreviations

AUROCC area under the receiver operating characteristic curve

CFA confirmatory factor analysis

CFI comparative fit index

EFA exploratory factor analysis

First 5 LA First 5 Los Angeles

LABBN LA Best Babies Network

LAC Los Angeles County

LSP Life Skills Progression

MCH maternal and child health

PNC prenatal care

PPD postpartum depression

RMSEA root mean square error of approximation

ROC receiver operating characteristic

SD standard deviation

SHV Select Home Visitation

T1 time 1

T2 time 2

TLI Tucker–Lewis Index

Chapter One. Introduction

The Welcome Baby Program

Welcome Baby is the direct service component of the place-based approach that First 5 Los Angeles (First 5 LA) developed to support and strengthen 14 Los Angeles County (LAC) communities, referred to as Best Start communities. First 5 LA selected these communities because they face several critical issues, including poverty, high unemployment rates, and high teen birth rates. Welcome Baby currently provides access and support to approximately 9 percent of families living in LAC. At full implementation, it is expected that Welcome Baby will serve up to 25 percent of LAC families. As a voluntary program, Welcome Baby provides free services during pregnancy and through a child's first nine months. The supportive services are primarily provided through home-visiting programs. These programs offer family support and coaching and include a broad range of services, such as consultations for pregnant women, newborn visits, and health and education services.

The Modified Bridges for Newborns Screening Tool

A critical component of the Welcome Baby program is the universal screening instrument, the Modified Bridges for Newborns screening tool. A precursor to the Modified Bridges, the Bridges for Newborns Screening Tool has been used to determine risk thresholds for pregnant women and new mothers in Orange County hospitals for more than ten years. The risk assessment tool is intended to provide a universal assessment of risk and includes risk factors related to financial hardship, known medical conditions, psychosocial concerns (including family support, history of drug and alcohol abuse, and history of domestic violence), and demographic issues (e.g., mother's age, use of English, and other children in the home).

During the expansion of the Welcome Baby program in 2012, First 5 LA conducted a small, qualitative pilot study of the Bridges to evaluate its possible use for the Welcome Baby program. Based on expert opinion of the Bridges' item wording and scoring procedures, several revisions to the instrument were identified, including revisions of double-barreled questions and underrepresentation of psychosocial risk factors. As a response to these issues, the revised or modified instrument adjusted the points and weights assigned to each question, separated the double-barreled questions, and created three subscale domains (Medical, Psychosocial, and Demographics/Basic Needs). Table 1.1 presents the 26 questions and subscales arranged according to the relative weight that each question receives. Higher scores on the items indicate higher levels of risk. The items receive weights that range from three points each to 18 points each and correspond to the severity of risk that each item measures.

Table 1.1. The Modified Bridges for Newborns Screening Tool Questions and Scoring, in Descending Order by Weight

Medical		Psychosocial		Demographics/Basic Needs	
Question	Item Weight	Question	Item Weight	Question	Item Weight
18. Infant medical problems	18	23. Mental health problem	18	1. Mother's age	18
22. Mother's medical problem	18	24. History of domestic violence	18	12. Current housing conditions	18
14. Adequate and timely PNC	12	25. History of child abuse or neglect	18	3. Mother's highest education level	9
20. Infant feeding issues	9	26. History of excessive alcohol or other drug use by people who will impact infant's well-being	18	5. Annual household income level	9
6. Mother's health coverage	3	21. Strength of maternal bond with infant	15	7. Adequate food in the house	9
13. Tobacco smoke in home	3	10. Level of family support	6	4. Mother's marital status	6
15. Infant health coverage	3			2. Mother's use of English	3
16. Source of medical care for infant	3			8. Mother's demonstrated awareness of available resources	3
17. Mother's intent to remain current with well-baby care and immunizations for her infant	3			9. Transportation a barrier	3
19. Mother's worry about infant's health	3			11. Other children at home	3
Subscale total	75		93		81

NOTE: PNC = prenatal care.

The Welcome Baby program uses the Modified Bridges to identify the level of support a family might need and refers the family to an appropriate home-visiting program. At a minimum, all mothers within a Best Start community can enroll in Welcome Baby before their children are born. The services that Welcome Baby staff—who are specialists, parent coaches, hospital liaisons, and nurses—provide include a total of up to nine engagement points that begin during the first or second trimester and continue up to nine months following the baby's birth. These services, provided primarily in Spanish or English, include phone calls and home visits designed to cover developmentally appropriate topics through a family-centered approach. There is an understanding that the needs of each family are unique, so the services are personalized to those needs.

Following the 2012 expansion of the Welcome Baby program, First 5 LA launched a new initiative, the Select Home Visitation (SHV) program, which provides more-intensive support for

high-risk families living within Best Start communities. First 5 LA uses the Modified Bridges to identify high-risk mothers who are eligible for the additional supportive services that the SHV program provides (specifically, either Healthy Families America or Parents as Teachers).

Guided by a script, trained hospital liaisons complete the Modified Bridges instrument. Hospital liaisons are people who approach prospective postpartum mothers in the hospital to introduce the Welcome Baby program, verify willingness to participate, and obtain consent. If the mother provides consent for services, the hospital liaison conducts an interview with the mother in which the liaison administers the Modified Bridges for Newborns screening tool. Because of the sensitive and personal nature of many of the Modified Bridges questions (e.g., the conversation includes a discussion of possible history of child abuse or neglect, history of domestic violence, incarceration of close family members), the hospital liaison will initially begin by developing rapport with the mother through conversation before proceeding with the interview. The interviews typically take liaisons 30 minutes to complete but can occasionally require as long as 90 minutes to complete, depending on the risk factors present.

The hospital liaison then immediately scores the Modified Bridges following the postnatal interview. The placement of families within either the Welcome Baby program (low-/moderate-risk mothers) or SHV program (high-risk mothers) depends on the mother's residence within a Best Start community, as well as her score on the screening instrument. Currently, First 5 LA employs a cutoff score of 50 to classify mothers as either high risk (50 or above) or low/moderate risk (below 50). A mother who scores 50 or above (high risk) and resides in a Best Start community is offered the most-intensive level of services, the SHV program. A mother who scores below 50 (low/moderate risk) and lives in a Best Start community is offered enrollment in the less intensive Welcome Baby program.

Conducting a Psychometric Study of the Modified Bridges

First 5 LA asked the RAND Corporation to undertake a psychometric study of the Modified Bridges to assess whether the tool is accurately assessing risk among mothers giving birth in Welcome Baby hospitals near or in First 5 LA's 14 Best Start communities. The research reported here aimed to provide information as to whether the hospital liaisons are scoring patients reliably. We also examined both the reliability and the validity of the Modified Bridges. In particular, we examined the following research questions at the request of First 5 LA:

- Do hospital liaisons assign similar scores (overall and subscores) to mock screening interviews when using the Modified Bridges for Newborns screening tool?
- Does the Modified Bridges for Newborns screening tool identify a similar level of risk among the same women as other well-validated measures of psychosocial and/or medical risk?
- Does the Modified Bridges for Newborns screening tool accurately distinguish between low-/moderate-risk women and high-risk women?

The primary purpose of this study is to examine the psychometric properties of the Modified Bridges. We have arranged this report according to the key research question addressed in each chapter. Chapter Two answers the question, "Do hospital liaisons assign similar scores when using the Modified Bridges?" To address this issue, we developed video vignettes that depict Modified Bridges interviews between hospital liaisons and mothers. The hospital liaisons then viewed the video vignettes and scored the Modified Bridges accordingly. From these scores, we assessed hospital liaison inter-rater reliability and test–retest reliability and identified factors related to hospital liaison performance. Chapter Three answers the research question, "Does the Modified Bridges identify a similar level of risk among the same women as other well-validated measures of psychosocial and/or medical risk?" To answer this question, we first conducted a review of 23 potential measures that assess aspects of maternal risk that could be comparable to the Modified Bridges. Our review identified two instruments sponsored and administered by the Florida Department of Health that had characteristics similar to the Modified Bridges (see Florida Department of Health, undated [a]). Using the Florida instruments, comparisons to the Modified Bridges were made using assessments of correlations at the item, scale, and risk threshold levels. Compared to the Florida measure, we also evaluated the extent to which the Modified Bridges is able to correctly identify high-risk mothers (i.e., sensitivity analyses) and low-risk mothers (i.e., specificity analyses). Chapter Four addresses the research question, "Does the Modified Bridges accurately distinguish between low-/moderate-risk mothers and high-risk mothers?" This question is evaluated through factor analyses and an analysis of the Modified Bridges' score reliability. This chapter also evaluates potential revisions to the Modified Bridges that might improve the instrument's score reliability. The final chapter provides an overview of key study findings.

Chapter Two. Do Hospital Liaisons Assign Similar Scores When Using the Modified Bridges for Newborns Screening Tool?

In this chapter, we first provide an overview of our approach for collecting data to assess hospital liaisons' ability to consistently score the Modified Bridges. This approach involved the creation of three video vignettes showing a trained hospital liaison interviewing a mother (played by an actor). Participating hospital liaisons then screened the video vignettes. After viewing each video, the hospital liaisons then completed the Modified Bridges. The hospital liaisons' scores for the Modified Bridges, based on each video vignette, provided the data used to evaluate the similarity of the scores across the liaisons. We also review the quality of the resulting data (e.g., response rates and missing data) and characteristics of the participating hospital liaisons. Finally, we provide evidence of the performance of hospital liaisons, and we report indices of inter-rater reliability and test–retest reliability, as well as liaison outlier detection and prediction.

Development and Administration of the Hospital Liaison Data Collection Tool

Vignette Scripts

To assess the hospital liaisons' inter-rater reliability, we developed three vignette videos, based on three patient cases, each of which the population of hospital liaisons would score. The three patient cases were selected by the oversight entity, LA Best Babies Network (LABBN), which is tasked with overseeing and supporting the standardization of Welcome Baby, conducting hospital liaison training, and providing technical assistance to all Welcome Baby sites. LABBN identified cases that are representative of mothers with moderate and high risk. LABBN selected the three cases, described as Aimee, Brooke, and Clare, to correspond to mothers who would likely score *near* the threshold for high risk[1] (i.e., threshold risk), above the threshold (i.e., high risk), and those representative of severer aspects of risk (i.e., very high risk), respectively. Although some risk factors were consistent across the vignettes, we chose the vignettes so that each one depicted a unique element of risk. For example, Aimee (i.e., threshold risk) is a 20-year-old first-time mother with a high school diploma who is currently unemployed and lives with her boyfriend. Her child was born preterm at 36 weeks but is otherwise healthy. Aimee had little or no evidence of psychosocial risk. Brooke (i.e., high risk) is a 30-year-old

---

[1] During the early phase of this study, the threshold for high risk was a score of 60 or higher. This threshold score was subsequently lowered to 50.

first-time mother who was admitted to the hospital at 22 weeks to delay delivery. She gave birth at 27 weeks via a cesarean section when dilation could no longer be delayed. Among her other risk factors, Brooke cannot provide breast milk to the child because she tested positive for marijuana while staying in the hospital—there is concern that she was exposed to marijuana from the baby's father, who visited her in the hospital. Finally, Clare depicts the most at-risk case (i.e., very high risk). In this instance, the child and mother appear healthy, but there is a history of domestic violence, drug and alcohol abuse, and allegations of child abuse.

Working from the patient cases that LABBN provided, we created scripts for the videos to guide the actors' responses to the liaisons' questions. Each script consisted of a narrative section with an overview of the key elements of each case and separate sections that each correspond to a particular Modified Bridges question that the hospital liaison might ask—this was intended to assist the actors during their interviews with the liaisons. Note that we developed the narratives to provide a relevant description of mothers' lives in order to shape the conversations with the liaisons. Our approach was intended to result in a more natural conversation between the actors and the liaisons.

Vignette Recording

The vignettes, which were recorded at our Santa Monica facility in November 2015, featured sets decorated with reclining chairs, blankets, hospital gowns, and infant dolls to resemble a hospital setting. We filmed in a realistic fashion with minimal video editing. We shared the final vignette videos with First 5 LA for review before showing them to hospital liaisons.

Video Vignette Administration

In December 2015, we administered the video vignettes to 31 hospital liaisons at a computer lab in Los Angeles. We followed the initial administration (time 1 [T1]) with a second administration of the same videos 30 days later (time 2 [T2]). During each administration, liaisons viewed all three vignettes on their computers using headphones and responded to the Modified Bridges questionnaire using the survey tool.

Review of the Data Collection Process

Liaison Characteristics

In November 2015, all participating hospital liaisons ($n = 34$) and hospitals (13) (see Table 2.1) were contacted and asked to participate in the data collection event as a mandatory aspect of continual training. Of the 34 contacted liaisons, 31 agreed to participate (the three nonparticipants declined because of prior conflicts—for example, maternity leave and jury duty). Notably, all 31 liaisons participated in both T1 and T2 data collections; hence, participant response rates were 100 percent at T1 and T2.

Table 2.1. Participating Hospitals and Medical Centers

Hospital or Medical Center	Participants (*n*)
Antelope Valley Hospital	6
Citrus Valley Medical Center	3
White Memorial Medical Center	3
California Hospital Medical Center	2
Miller Children's and Women's Hospital	2
Martin Luther King Jr. Community Hospital	2
Northridge Hospital Medical Center	2
Providence Holy Cross Medical Center	2
St. Mary Medical Center	2
Torrance Memorial Medical Center	2
Valley Presbyterian Hospital	2
Centinela Hospital Medical Center	1
St. Francis Medical Center	1
Missing	1
Total	31

Missing Data Rates

Very little data were missing across the three completed Modified Bridges survey questionnaires at both T1 and T2. Across the 26 Modified Bridges items, the missing data rates ranged from 0 percent to 1.1 percent at both T1 and T2. There was one exception: The liaison did not ask very high-risk mother Clare's age during the recorded interview, and subsequently, for this vignette, we purposely set the question for the mother's age to missing on the Modified Bridges questionnaire for all participating liaisons.

Review of the Results

Table 2.2 provides the means and standard deviations (SDs) of the Modified Bridges and subscales for all completed questionnaires (threshold, high, and very high-risk vignettes) for T1 and T2. As expected, there was very little difference in subscale scores across time, and there were no systematic trends in scores from T1 to T2. Also as expected, the completed questionnaires for the Clare vignette provided the highest risk level, followed by the completed questionnaires for the Brooke and Aimee vignettes.

Table 2.2. The Modified Bridges Total Score and Subscale Means and Standard Deviations Across Time and Vignettes

Subscale	Time	Threshold Risk Vignette Mean (SD)	High-Risk Vignette Mean (SD)	Very High-Risk Vignette Mean (SD)
Medical	T1	31.2 (8.3)	11.0 (3.8)	37.6 (4.6)
	T2	30.4 (7.4)	12.6 (4.8)	38.0 (4.2)
Psychosocial	T1	1.9 (2.8)	53.9 (4.7)	37.0 (8.2)
	T2	2.4 (3.8)	53.1 (4.3)	36.3 (8.1)
Demographics/Basic Needs	T1	29.4 (6.0)	32.1 (3.8)	37.5 (4.5)
	T2	27.5 (4.3)	31.8 (4.3)	40.3 (6.3)
Modified Bridges total score	T1	62.5 (11.5)	97.1 (8.5)	112.1 (12.1)
	T2	60.4 (9.5)	97.5 (7.1)	114.5 (13.8)

Agreement and Inter-Rater Reliability

Inter-rater reliability indicates the degree to which a given hospital liaison will provide the same risk assessment as any other liaison. Using T1 data, we present inter-rater reliability evidence using both kappa coefficients (Cohen, 1960) of agreement and Kendall's W (also known as Kendall's coefficient of concordance) (Kendall, 1948; Kendall and Smith, 1939). Each coefficient provides an indication of the strength of agreement between raters. Table 2.3 presents rules of thumb for interpreting these values (Landis and Koch, 1977; Fleiss, 1981; Gwet, 2010).

Table 2.3. Rules of Thumb for Kappa Coefficients and Kendall's *W* Indices of Inter-Rater Agreement

Magnitude of Agreement	Coefficient Range
Near perfect	>0.80
Substantial	0.60–0.80
Moderate	0.40–0.60
Poor	<0.40

Kappa provides an absolute assessment of agreement (i.e., all misclassification, regardless of severity, is treated equally), while Kendall's coefficient accounts for the magnitude of the misclassification. Because the Modified Bridges uses ordinal response options, we give some preference to Kendall's coefficient, which provides a more realistic estimation of inter-rater reliability, while Kappa, in this instance, might underrepresent the agreement. This is because the Kappa coefficient will treat all disagreement between raters, regardless of how small (e.g., two coders providing a rating of 0 and 1 of a given item) or large (e.g., two coders providing a rating of 0 and 4), as being the most extreme disagreement. Finally, because of the close agreement

8

between raters on individual items (i.e., lack of variability), we also present the item-level percentage of perfect agreement across completed questionnaires for all three vignettes.

Table 2.4 provides the overall assessments of the inter-rater reliability across risk vignettes. The results indicate that overall inter-rater reliability for the Modified Bridges is high (kappa = 0.66 and Kendall = 0.83). As expected, the highest agreement was found for the Demographics/Basic Needs subscale (Kendall = 0.89), followed by the Psychosocial (Kendall = 0.83) and Medical subscales (Kendall = 0.74). Further, there was relatively little variation across the vignettes (Kendall range = 0.76 to 0.86).

Table 2.4. Kappa and Kendall Coefficients of Agreement for the Modified Bridges

Score	Kappa	Kendall
Modified Bridges total score	0.66	0.83
Subscale		
Medical	0.57	0.74
Psychosocial	0.62	0.84
Demographics/Basic Needs	0.73	0.89
Vignette		
Threshold risk	0.68	0.76
High risk	0.68	0.86
Very high risk	0.62	0.85

Inter-rater agreement information can also be computed at the item level; however, because, for many items, the agreement is quite high (in some instances, perfect), there is little variation with which to compute Kendall or kappa coefficients at the item level. As an alternative, Table 2.5 provides, for all risk vignettes, the percentage of perfect agreement for each item—that is, the degree to which all hospital liaisons provide the exact same rating for a given item. For example, an item with 100-percent agreement means that all 31 liaisons provided the exact same rating. Table 2.5 suggests instances in which the agreement is perfect (e.g., mother's use of English, history of domestic violence) and items for which there is less agreement (e.g., strength of maternal bond with infant, average agreement 58 percent; mother's demonstrated awareness of available resources, average agreement 65 percent; infant feeding issues, average agreement 70 percent). Across all risk vignettes, the hospital liaisons demonstrated close agreement with one another. The average percentage of agreement is 84 percent, indicating that liaisons provide identical ratings 84 percent of the time. At the subscale level, the magnitude of the percentage of agreement generally reflects findings from the kappa and Kendall coefficients, with the percentage of agreement highest for the Demographics/Basic Needs subscale.

9

Table 2.5. Percentage of Perfect Agreement Between Liaisons at Time 1 for Modified Bridges Items

Scale or Item	Threshold Risk Vignette	High-Risk Vignette	Very High-Risk Vignette	Average
Medical	76	90	85	84
Mother's health coverage	100	97	100	99
Tobacco smoke in home	100	100	97	99
Adequate and timely PNC	97	81	100	92
Infant health coverage	97	94	94	95
Source of medical care for infant	39	94	81	71
Mother's intent to remain current with well-baby care and immunizations for her infant	52	81	94	75
Infant medical problems	84	74	94	84
Mother's worry about infant's health	48	94	74	72
Infant feeding issues	68	90	52	70
Mother's medical problem	81	100	61	81
Psychosocial	94	86	70	83
Level of family support	77	84	61	74
Strength of maternal bond with infant	84	48	42	58
Mental health problem	100	100	45	82
History of domestic violence	100	100	100	100
History of child abuse or neglect	100	90	100	97
History of excessive alcohol or other drug use by people who will impact infant's well-being	100	94	71	88
Demographics/Basic Needs	91	84	85	87
Mother's age	100	97	Not applicable	98
Mother's use of English	100	100	100	100
Mother's highest education level	97	97	87	94
Mother's marital status	100	100	100	100
Annual household income level	97	100	100	99
Adequate food in the house	77	58	77	71
Mother's demonstrated awareness of available resources	65	65	65	65
Transportation a barrier	100	77	58	78
Other children at home	90	90	94	91
Current housing conditions	87	52	81	73

10

Test–Retest Reliability

Test–retest reliability indicates the degree to which test-takers provide the same rating over time. Test–retest reliability is evaluated by providing the same test to participants at two different occasions and then assessing the correlations between the scores. In this sense, our evaluation of test–retest reliability between T1 and T2 Modified Bridges scores refers to the reliability of the hospital liaisons to provide the same scores at two occasions on the same video vignettes. The Modified Bridges total score correlation between T1 and T2 across all risk vignettes is $r = 0.91$,[2] indicating evidence of high test–retest reliability. Further, the test–retest correlations were highest for the Medical and Psychosocial subscales ($r = 0.96$ and 0.94, respectively) and somewhat lower for the Demographics/Basic Needs subscale ($r = 0.62$), which is more likely an artifact of the properties of the statistics than an actual lack of test–retest reliability.[3]

Because of the close correspondence between T1 and T2, it is useful to evaluate the percentage of perfect agreement between the two assessments (similar to the agreement shown in Table 2.5)—that is, the percentage of T1 ratings that are identical to T2 ratings. For example, if 25 of the 31 liaisons provided the exact same ratings at T1 and T2, the item-level percentage is 81 percent. The findings presented in Table 2.6 provide supporting evidence for the test–retest reliability of the hospital liaisons. At the item level, the findings are similar to the T1 agreement results. The items with highest levels of agreement (i.e., 100 percent) are mother's use of English, mother's marital status, annual household income level, and history of domestic violence. The items with lower agreement are mother's demonstrated awareness of available resources, average agreement 57 percent; strength of maternal bond with infant, average agreement 62 percent; and source of medical care for infant, average agreement 70 percent. Across T1 and T2, the average percentage of agreement is 86 percent, indicating that, 86 percent of the time, liaisons provided identical ratings at T1 and T2.

---

[2] Correlations range from –1.0 to +1.0. A correlation of 1.0 indicates perfect agreement. For example, if an increase in variable A by one unit is always associated by a one-unit increase in variable B, the correlation between variables A and B would be $r = +1.0$.

[3] This phenomenon (lower-than-expected test–retest reliability for the Demographics/Basic Needs subscale) is likely due to the hospital liaisons providing very similar total demographic scores (i.e., an issue of restriction of range). For additional information on range restriction, which is outside the scope of this report, please see Glass and Hopkins, 1996, and Crocker and Algina, 1986.

Table 2.6. Percentage of Perfect Agreement Between Times 1 and 2 for Modified Bridges Items

Scale or Item	Threshold Risk Vignette	High-Risk Vignette	Very High-Risk Vignette	Average
Medical	91	84	87	87
Mother's health coverage	97	100	100	99
Tobacco smoke in home	100	100	97	99
Adequate and timely PNC	68	97	100	88
Infant health coverage	100	100	97	99
Source of medical care for infant	94	48	68	70
Mother's intent to remain current with well-baby care and immunizations for her infant	90	68	90	83
Infant medical problems	77	94	97	89
Mother's worry about infant's health	90	65	81	78
Infant feeding issues	97	77	74	83
Mother's medical problem	97	90	65	84
Psychosocial	83	95	73	83
Level of family support	77	94	48	73
Strength of maternal bond with infant	52	74	61	62
Mental health problem	100	100	52	84
History of domestic violence	100	100	100	100
History of child abuse or neglect	81	100	100	94
History of excessive alcohol or other drug use by people who will impact infant's well-being	87	100	74	87
Demographics/Basic Needs	84	90	83	86
Mother's age	94	94	100	96
Mother's use of English	100	100	100	100
Mother's highest education level	97	97	87	94
Mother's marital status	100	100	100	100
Annual household income level	100	100	100	100
Adequate food in the house	71	77	65	71
Mother's demonstrated awareness of available resources	58	55	58	57
Transportation a barrier	58	100	58	72
Other children at home	87	90	94	90
Current housing conditions	77	90	68	78

Outlier Detection and Predicting Hospital Liaison Performance

In this section, we evaluate the performance of individual hospital liaisons in order to identify liaisons who might be underperforming compared with their colleagues (i.e., hospital

liaison outliers). That is, hospital liaisons who underperform might provide substantially higher or lower ratings on the items in the Modified Bridges. To identify liaisons who might be underperforming, we use specific outlier-detection analyses that compare each liaison's item response against the mean item response (i.e., the average response for an item provided by all liaisons). This assumption is reasonable given the collective experience of the liaisons and the close agreement and inter-rater reliability for most items.

For each liaison, we compute a standardized score (i.e., z-score) for the total Modified Bridges based on the hospital liaisons' ratings. Standardized scores have a mean of 0 and SD of 1. For example, a score of −1.0 indicates a liaison who scores the Modified Bridges one SD below the mean of all the liaisons. For these data, scores plus or minus two SDs indicate potential outliers. Because outliers can be positive or negative, and the goal is to evaluate poor performance in *any* direction, we use the absolute value of the standardized scores in the results presented below. Finally, for these analyses, we make use of just the T1 scores.

Figure 2.1 shows the frequency of the average absolute value of the standardized scores. In other words, the figure represents how similar the hospital liaisons are to one another in terms of scoring the Modified Bridges. Values closer to 0.0 indicate performance that is closer to the mean across hospital liaisons, which identifies liaisons who score the Modified Bridges in ways that are very similar to their peers. Figure 2.1 illustrates that most hospital liaisons are scoring the Modified Bridges in a similar manner—there is relatively little evidence of poor-performing liaisons. Across the 31 liaisons, seven had average (across all risk vignettes) standardized scores greater than 1.0, and two liaisons had average standardized scores greater than 1.5. These two liaisons represent the poorest-performing liaisons (in terms of deviating from other liaisons). Of the two poorest-performing liaisons, additional information available in the survey indicates that one had just started as a hospital liaison (she indicated no prior experience) and the other had slightly more than one year of experience but reported that she did not refer to the Modified Bridges protocol when scoring the vignettes or when scoring mothers in the actual hospital setting.[4] Although these two liaisons perform lower than is desired, their performance was not so low as to cause concern (i.e., their scores were not more than two SDs from the mean). In summary, there is variability in the performance of the hospital liaisons, but there is no evidence that hospital liaisons are consistently underperforming.

---

[4] Hospital liaisons have access to, and are encouraged to use, a protocol that provides details on how to properly score the Modified Bridges and includes example cases.

Figure 2.1. Overall Hospital Liaison Performance and Outlier Detection
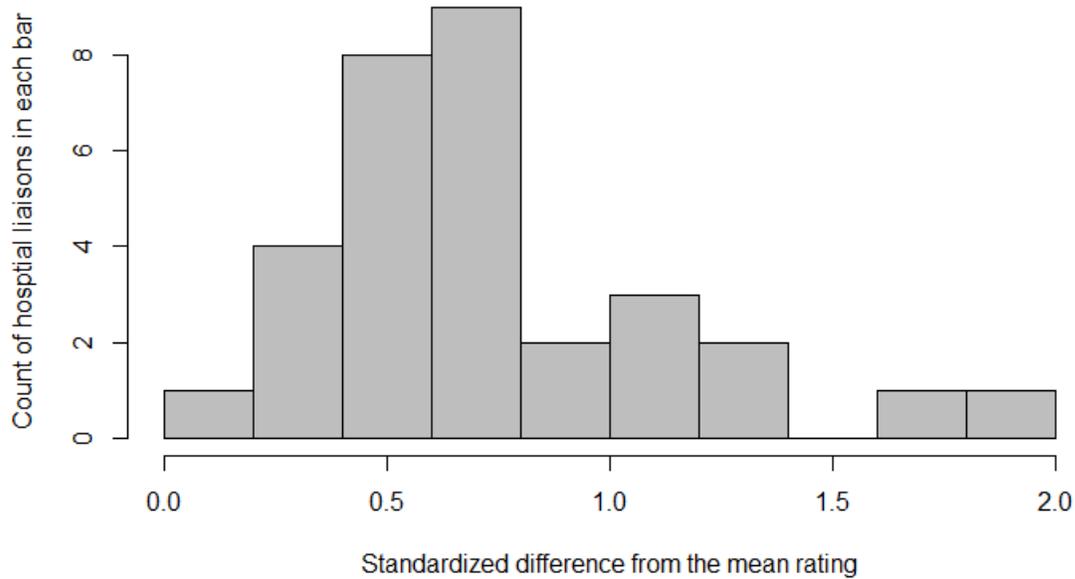
Given the current evidence that the liaisons are consistently scoring the Modified Bridges, we next assessed whether liaisons have a tendency to score mothers higher or lower than their colleagues do. In other words, if a liaison scored risk vignette A above the mean, is that liaison more likely to also score vignettes B and C above the mean? To assess this, we estimated correlations between all three T1 standardized Modified Bridges scores. If there is no trend in rating (i.e., no tendency to score above or below the mean), the correlation between any two risk vignettes would be $r = 0.0$ (because each vignette is independent of the other vignettes),[5] which is the desired correlation. The correlations (r) and associated statistical significance tests (p-values) comparing the three vignettes were $r_{1,2} = 0.14$ ($p = 0.44$), $r_{1,3} = 0.21$ ($p = 0.26$), and $r_{2,3} = 0.54$ ($p = 0.002$). These findings provide mixed evidence that liaisons systematically rate risk levels above or below the mean (i.e., there is statistically significant evidence in only one of the three comparisons). In other words, even though the agreement across liaisons is high, when the liaisons' ratings deviate from the mean, they might do so in a consistent manner—either consistently slightly high or low.

The final analysis compares the (absolute and averaged) standardized scores from the Modified Bridges with the validation items collected during the T1 assessment (i.e., liaison age, years administering the Modified Bridges, and number of times referencing the Bridges protocol during each video). Results suggest that performance is not significantly related to age ($r = 0.03$,

---

[5] Correlations range from –1.0 to +1.0. A correlation of 0.0 indicates no agreement. For example, a correlation of $r = 0.0$ means that scoring high or low on variable A is unrelated to scoring high or low on variable B.

$p = 0.87$) but is marginally related to years of experience administering the Modified Bridges ($r = -0.32$, $p = 0.08$) and significantly related to the number of times that the liaison makes use of the protocol during the videos ($r = -0.54$, $p = 0.002$). In other words, more-experienced liaisons and liaisons who referred to the protocol more frequently tended to provide ratings closer to the mean. Table 2.7 illustrates this finding by showing the average of the standardized scores according to various categories of liaison experience with the Modified Bridges and the number of times the protocol was referenced across the three vignettes. Higher standardized scores indicate greater deviation from the mean rating across liaisons (i.e., 0.0). The best performance was found for liaisons with more than two years of experience and those who referenced the protocols three or more times during the videos.

Table 2.7. Average Absolute Values of Standardized Scores Across Years of Administering the Modified Bridges

Comparison	Average Absolute Value of Standardized Modified Bridges Scores (SD)
Years of experience	
Less than 1 ($N = 10$)	0.86 (0.46)
1–2 ($N = 11$)	0.70 (0.43)
More than 2 ($N = 10$)	0.66 (0.34)
Number of times using the protocol across all three risk vignettes	
0 ($N = 29$)	0.95 (0.69)
1–2 ($N = 42$)	0.68 (0.70)
3–5 ($N = 18$)	0.56 (0.36)
6–9 ($N = 1$)	0.09 (—)
Missing ($N = 3$)	

Limitations

The results presented here provide an accurate representation of the inter-rater reliability of the hospital liaisons; however, at least three key limitations should be considered when evaluating these findings. First, the liaisons are being evaluated only on the degree to which their ratings correspond to each other. That is, this is not an assessment of the liaisons' ability to conduct an interview in a standardized manner. In the hospital setting, liaisons might differ in how they conduct the interviews, which could result in mothers potentially receiving different scores depending on the attending liaison. Because of this, the ratings described here might be biased to provide a more-accurate representation of liaison reliability than is true in the actual hospital setting. Next, we note that the test–retest reliabilities are a reflection of the similarities of the scores that the liaisons provided across a one-month interval. At the second data collection, it is possible that the liaisons could recall the scores they provided at the first data

collection, which would inflate the inter-rater reliability values. However, given the total number of Modified Bridges item ratings provided (78) and the number of possible responses for each Modified Bridges' question (four), we think that it is unlikely that the hospital liaisons would be able to recall their exact responses to the questions. Finally, the actors in the video vignettes are following a narration but occasionally provide responses that are relatively straightforward to code. An example of this is in the sections on history of domestic violence and history of child abuse or neglect, in which the actors' responses were perhaps easier to code than is witnessed during actual mother interviews. That said, we made efforts in each narration to provide some content that would test the liaisons' ability to correctly score the Modified Bridges.

Conclusions

Overall, the inter-rater reliability and test–retest reliability of the Modified Bridges' hospital liaisons is very high, though not perfect. Agreement is highest for the Demographics/Basic Needs subscale and slightly lower for the Medical and Psychosocial subscales. Although agreement was high for many items, a few items could benefit from additional liaison training— namely, strength of maternal bond with infant, source of medical care for infant, mother's demonstrated awareness of available resources, and infant feeding issues. Our findings suggested that all liaisons performed at an acceptable level. Finally, our results suggest that the liaisons who show the highest rates of agreement, or score the Modified Bridges most consistently, are those who refer to the protocol frequently and, to a somewhat lesser extent, are more experienced.

Chapter Three. Does the Modified Bridges for Newborns Screening Tool Identify a Similar Level of Risk Among the Same Women as Other Measures of Psychosocial and/or Medical Risk?

We collected data from liaison–mother interviews to assess the degree to which the Modified Bridges provides risk scores that are similar to those on another widely used measure of medical and psychosocial risk (i.e., the validation instrument). Our process for identifying the validation instrument involved a literature scan, consultation with maternal experts, and discussions with First 5 LA. Once we identified measures for the validation instrument, we conducted a training session with the participating hospital liaisons to administer the validation instrument alongside the Modified Bridges to examine the relationships between the scores. In this chapter, we discuss this process, as well as findings from the data collection.

Measures Reviewed for Inclusion

Identifying Comparison Measures of Risk

To validate the Modified Bridges, we first solicited input from a diverse group of maternal and child health (MCH) public health practitioners and researchers. We reached out to approximately 20 people with MCH master's degrees who collectively have current and prior experience working in federal, state, and local health departments; universities; health care settings; and nonprofit organizations. When asking these MCH informants for guidance, we provided a copy of the Modified Bridges and asked for examples of similar tools that can be used to identify new mothers at risk for poor outcomes. Informants responding to our request recommended primarily tools that the Florida Department of Health uses.

To ensure that we considered all reasonable risk measures, we conducted an Internet search for similar tools that other state and local government agencies and home-visiting programs use. Additionally, we used Google Search and Google Scholar to conduct a broad search for tools that government and nongovernment agencies use and tools mentioned in the academic literature to screen or assess broad risk factors overall and specific to new mothers. We used an approach similar to snowball sampling in that the tools we found continually informed our searches. In other words, in addition to the primary search procedure described below, as we identified new articles, we reviewed their contents (e.g., key words, citations) to identify new journal articles, tools, and search terms that would further improve our search process. We used primarily combinations of the following terms to conduct our search:

- terms describing life stage: *postpartum*, *postnatal*, *pregnant*, *prenatal*, *newborn*, *baby*

- terms describing assessment instruments: *screen*, *assess*, *tool*, *instrument*, *questionnaire*, *survey*
- terms describing focus or content: *risk*, *vulnerable*, *medical*, *psychosocial*, *demographic*, *women*, *family*, *child*.

In Table 3.1, we identify 23 tools and describe the advantages and disadvantages of each as it relates to validating the Modified Bridges. The tools vary in length (anywhere from three items to more than 200 items), intended age range (e.g., child and adult), intended life stage (e.g., prenatal and postpartum), and delivery method (e.g., completed by mother or health care provider).

Table 3.1. Identifying Potential Comparison Measures to Validate the Modified Bridges

Identifier	Instrument	Number of Items	Covers Multiple Domains?	Pros	Cons
1	Florida's Healthy Start Prenatal Risk Screen	21	Yes	- The mother completes the instrument, with the health care provider supplying additional information. - The instrument captures medical, psychosocial, and demographic information. - The instrument is relatively short and has a simple scoring methodology. - The instrument has evidence of validity.	- The survey covers several domains but appears to score only a few. - The instrument does not ask about the newborn's health.
2	Florida's Healthy Start Infant Risk Screen	14	Yes	- The instrument covers infant health, demographics, relationships, and tobacco use. - The instrument is relatively short.	- A health care provider completes the instrument based on information from the birth certificate. - The instrument does not cover domestic violence or substance abuse.

Identifier	Instrument	Number of Items	Covers Multiple Domains?	Pros	Cons
3	Durham Connects Family Support Matrix (North Carolina)	12	Yes	• Home-visiting nurses use the tool to assess risk for negative outcomes and coordinate referral to supportive services. • The instrument is intended to be administered to new mothers in a conversational manner. • The administrator records scores for 12 factors related to four domains of family and child well-being: MCH, caring for the infant, safe home, and parent characteristics.	• Nurses traditionally administer the tool. Hospital liaisons would need extensive training. • It requires a potentially lengthy interview process. • There might be too much overlap in content such that hospital liaisons might rely on the Modified Bridges to score this instrument. • Scoring appears to be specific to each factor rather than cumulative.
4	Kempe Family Stress Checklist (now the Parent Survey in Healthy Families America)	10	Yes	• It is intended to be administered prenatally or within two weeks of the baby's birth. • It is intended to be administered as a conversation. • It covers a variety of domains, including mental health, substance abuse, childhood history, maternal bond, discipline of the child, and stress.	• It is used primarily to predict maltreatment risk. • It does not ask about the baby's or mother's health. • Interviewers would likely need prior training on how to administer and score it.
5	Universal Risk Assessment (Seattle's Thrive by Five Washington)	30	Yes	• It captures a variety of domains, including mental health, substance abuse, housing, economics, education, and domestic abuse. • It is intended to be administered to mothers of children under age 5.	• It is more of a research tool than an assessment tool. • It is long and not very conversational. • It does not ask about the baby's or mother's physical health. • The training session for interviewers is a recommended two hours. • We are uncertain whether there is a scoring method.

19

Identifier	Instrument	Number of Items	Covers Multiple Domains?	Pros	Cons
6	March of Dimes Preconception Screening and Counseling Checklist	74	Yes	• It captures a wide variety of domains, including medical, psychosocial, demographic, lifestyle, and environmental information.	• It is intended to be completed with a health care provider prior to conception. • It does not ask about children. • We are unclear how or whether the instrument is scored.
7	Protective Factors Survey	30	Yes	• It is intended as a self-administered survey that measures protective factors in five areas: family functioning and resiliency, social support, concrete support, nurturing and attachment, and knowledge of parenting and child development. • It takes 10 to 15 minutes to complete.	• It does not ask about the child's or mother's health. • It asks about children but is not specific to new mothers. • It is intended to assess multiple protective factors against child abuse and neglect. • It is used as a pre/post evaluation tool for caregivers receiving child maltreatment–prevention services. • We are unclear how or whether the instrument is scored.
8	Florida Healthy Behaviors Survey	38	Yes	• The survey is specific to women. • It captures information about education, employment, abuse, substance abuse, pregnancy, and physical and psychosocial health.	• The survey is not specific to new mothers and does not ask about children. • We are unclear how or whether the instrument is scored.
9	Georgia's Children 1st Screening and Referral Form	>50	Yes	• It captures a wide variety of domains, including medical, psychosocial, demographic, lifestyle, and environmental information. • It focuses on new mothers and infants.	• It is intended to be completed using vital records or other secondary data. • The form is very long.

Identifier	Instrument	Number of Items	Covers Multiple Domains?	Pros	Cons
10	LSP	43	Yes	• Home visitors use it to assess family functioning. • The instrument covers relationships, education, mental health and substance abuse and other risks, basic essentials, and infant and toddler development. • It has confirmed reliability and validity.	• Cost to obtain this instrument is high. • The LSP is not intended to be administered via interview or parent self-report. • Scoring is based on information obtained through other data sources. • The instrument asks about children but is not specific to new mothers.
11	LSP Modified	8	Yes	• It is intended for use by home visitors. • It is a short instrument that covers multiple domains (relationships, health care, and child care).	• The LSP is not intended to be administered via interview or parent self-report. • Scoring is based on information obtained through other data sources. • The instrument asks about children but is not specific to new mothers. • It does not ask about substance abuse, the child's or mother's health, or the mother's demographics or education.
12	Virginia Behavioral Health Screening Tool for Women of Child Bearing Age adapted from the Institute for Health and Recovery Integrated Screening Tool	8	Yes	• It is specific to women's health. • It captures emotional problems; alcohol, tobacco, and other drug use; and domestic violence. • The instrument is short.	• It does not ask about physical health or demographics. • It does not ask about children.
13	Institute for Health and Recovery Integrated Screening Tool	8	Yes	• It is specific to women's health. • It captures emotional problems; alcohol, tobacco, and other drug use; and domestic violence. • The instrument is short.	• It does not ask about physical health or demographics. • It does not ask about children.

Identifier	Instrument	Number of Items	Covers Multiple Domains?	Pros	Cons
14	Family Map	>200	Yes	• It is intended to be completed as an interview with early childcare or prenatal educators and families. • The instrument covers employment, education, safety, housing, violence, health, substance abuse, basic needs, and maternal depression. • The instrument identifies both strengths and risks.	• The instrument is long. • The cost to obtain the instrument is high. • The instrument asks about children ages 0 to 5 but is not specific to new mothers. • We are unclear how or whether the instrument is scored.
15	Massachusetts Family Self-Sufficiency Scales and Ladders Assessment Form	>150	Yes	• It is designed for use by social work professionals in the Massachusetts social welfare system. • The instrument covers employment, education, transportation, health, family development, and housing, among other domains.	• The instrument is long. • It asks about children but is not specific to new mothers. • All items appear to be equally weighted.
16	Edinburgh Postnatal Depression Scale	10	No	• It is a short, validated screen for PPD.	• It has limited scope (used to screen for maternal PPD).
17	Center for Epidemiologic Studies Depression Scale Revised	10	No	• This is a short, validated screen for depression.	• It is not specific to new mothers. • It has limited scope (used to screen for depression).
18	Nine-item Patient Health Questionnaire	9	No	• This is a short, validated screen for depression.	• It is not specific to new mothers. • It has limited scope (used to screen for depression).
19	5Ps Prenatal Substance Abuse Screen for Alcohol, Drugs and Tobacco	5	No	• This is a short, validated screen for substance abuse. • It includes questions related to pregnancy.	• It is not specific to new mothers. • It has limited scope (used to screen for substance abuse).
20	Alcohol Use Disorders Identification Test	3	No	• This is a short, validated screen for active alcohol use disorders.	• It is not specific to new mothers. • It has limited scope (used to screen for alcohol use disorders).

Identifier	Instrument	Number of Items	Covers Multiple Domains?	Pros	Cons
21	Women's Experience with Battering Scale Relationship Assessment Tool	10	No	• This is a short, validated screen for intimate-partner violence.	• It is not specific to new mothers. • It has limited scope (used to screen for intimate-partner violence).
22	Ages and Stages Questionnaire, 3rd ed., 2 Month Questionnaire	36	No	• This is a validated, frequently used questionnaire about child development.	• It has limited scope (used to identify developmental delays in children). • It is a proprietary tool with fee for use.
23	Adverse Child Experience screen	10	No	• This captures psychosocial information about the caregiver's upbringing.	• It has limited scope (asks about caregiver's upbringing, not the current situation). • It is not specific to new mothers.

NOTE: LSP = Life Skills Progression. PPD = postpartum depression.

Eight of the tools are used to screen for conditions that are more specific than the general risk assessment that the Modified Bridges provides (e.g., depression, substance abuse, history of domestic violence). Fifteen of the tools cover multiple domains, addressing a range of topics that include demographics, health, psychosocial factors, and substance use. These tools are generally used by state and local agencies to refer families to social services. For example, Florida's Healthy Start Prenatal Risk Screen includes 21 items related to the medical, psychosocial, and demographic characteristics of mothers; has a simple scoring methodology; and is mostly completed by the mother. A limitation of this tool is that it does not capture information about the health of the newborn. However, Florida's Healthy Start Infant Risk Screen includes 14 items related to infant health and the medical, psychosocial, and demographic characteristics of mothers and is intended to be completed by a health care provider using information from the mother's medical record. Both of these tools have scoring systems with established risk cutoffs, with families referred to services if they reach a certain score. Other measures that cover multiple domains have greater complexity in their administration and unknown or more-complex scoring methodologies.

Because the Modified Bridges includes questions about the mother's and newborn's medical, psychosocial, and demographic characteristics, the most-promising tools for this validation study also address multiple domains of risk. Among the many tools reviewed, the most promising, which we ultimately selected, were used by Florida Department of Health to refer pregnant women and new mothers to social services—specifically, Florida's Healthy Start Prenatal Risk

Screen and Healthy Start Infant Risk Screen.[6] We arrived at this selection in partnership with First 5 LA. Specifically, it was important to select a tool that satisfied three criteria: (1) the measure should assess similar risk content as the Modified Bridges (i.e., medical, psychosocial, and demographic characteristics), (2) the measure should be used in a manner similar to that of the Modified Bridges (i.e., to identify mothers and families at risk and in need of additional supportive services), and (3) the measure should allow for administration in an interview format. Florida's risk measures were the only measures that satisfied all these criteria and could be integrated into the hospital liaisons' protocol with minimal burden to their daily routines.

Using the Florida Measures

Prior to administering these tools as validation measures, we made some slight revisions to the measures in an effort to streamline the administration process and to place all the pre- and postnatal items on a single instrument:

1. Because the two Florida measures are intended to be administered at different points in time, respectively (i.e., prenatally and postnatally), some questions become redundant. We removed four such duplicate items (prenatal health care visits, tobacco use, mother's marital status, and age).
2. Because the Florida version of the prenatal screen is intended to be self-reported by the mother, we adapted the scales' items to allow for hospital liaison administration.
3. We also removed the "maternal race black" question from each instrument (two total items).
4. Finally, we removed additional items from the prenatal risk screen instrument that are administered but not scored (seven total items).

This process resulted in an 18-item maternal and infant risk instrument that preserved the content from the original Florida Department of Health measures but revised the format to enable administration by the hospital liaisons concurrently with the Modified Bridges (see Appendix A).

Protocol for Validating the Modified Bridges for Newborns Screening Tool

Data Collection Strategy

To validate the Modified Bridges tool against the Florida risk measures, we partnered with hospital liaisons and managers from Antelope Valley Partners for Health and California Hospital Medical Center, engaging managers from both sites in discussions regarding the coordination process and study protocol beginning in April 2016. We also conducted in-person training with hospital liaisons to review the procedures for administering both the Florida measures (referred to in the hospital liaison protocol as the legacy instrument) and the Modified Bridges and to

---

[6] For additional information, see Florida Department of Health, undated (c), and Florida Department of Health, undated (b).

practice submitting the survey responses electronically. Data collection using the Modified Bridges tool and the Florida measures took place from July 11 to August 5, 2016. During this time, we collected responses from 153 mothers interviewed by nine participating hospital liaisons from the Antelope Valley Hospital ($n = 67$) and California Hospital Medical Center ($n = 85$).

Analytic Strategy

Our analytic process to validate the Modified Bridges used three approaches: (1) evaluating the relationships between the Modified Bridges and Florida measures at various levels, including at the item, subscale, total score, and risk threshold levels; (2) analyses of various Modified Bridges cutoff scores compared with the established Florida cutoff scores; and (3) comparing total scores for Modified Bridges across various factors (i.e., risk groups) known to be associated with increased risk (e.g., age, income, history of drug or alcohol abuse). We discuss these analyses in more detail below.

Although the majority of the validation analyses make use of data collected with the Florida measures, we also report some additional analyses that require larger sample sizes (e.g., risk-group validity analyses). For these analyses, we obtained secondary data downloaded from the Stronger Families database (managed by First 5 LA). We collected these data (to which we refer as the secondary data) between January 2013 and December 2015, and they included responses from 12,165 mothers interviewed by 72 hospital liaisons.

Validity Study Findings

Sample Description

Appendix B shows response frequencies and descriptive statistics for all Modified Bridges items, subscales, and total scores. Nearly 61 percent of the patient sample were 25 years of age or older, the majority spoke English (73 percent fluent in English), 39 percent had not completed high school, 34 percent were married, and 64 percent reported income below $18,000 or were unaware of their income. Most had health insurance (full coverage, 77 percent), consistent food in the house (65 percent), and reported stable and safe housing (71 percent). For one-third of the sample, this was the first pregnancy; 69 percent wanted the pregnancy, and only 15 percent reported any prior poor birth outcomes. The average Modified Bridges total score in this sample was 52.2 (SD = 28.6), with 49 percent of the sample meeting criteria for perinatal risk (at a Modified Bridges threshold of 50).

Relationship Between the Modified Bridges and the Florida Healthy Start Prenatal and Infant Risk Screens

Because the Modified Bridges and Florida measures use different risk thresholds, there is the potential for disagreement in the proportion of mothers who will be identified as being at high risk. For the Healthy Start Prenatal Risk Screen measure, total scores of six (out of 22 possible) or more indicate high risk, while four (out of 20 possible) or more indicate high risk for the Healthy Start Infant Risk Screen measure. Table 3.2 shows the proportion of the validity sample that scored at or above these thresholds indicating high risk. Using these cutoffs, we found only 12 percent and 16 percent of the sample to be at risk according to the Florida pre- and postnatal risk measures, respectively. The differences in these percentages suggest that the Modified Bridges criterion is set at a lower risk level, allowing more mothers to be eligible for follow-up services than would be eligible given the Florida prenatal and infant risk scales' cutoff levels, which essentially require higher levels of risk to be determined high risk.

Table 3.2. Modified Bridges and Florida Prenatal and Infant Risk Frequency and Percentages

Measure	Frequency	Percentage
Modified Bridges (total score ≥50)	75	49.0
Healthy Start Prenatal Risk Screen (total score 6 or more)	19	12.4
Healthy Start Infant Risk Screen (total score 4 or more)	25	16.3

Next, we examined the relationships between the Modified Bridges and Florida instruments using correlational approaches. Table 3.3 provides rules of thumb that might be helpful when interpreting the correlation magnitudes. In most scale assessment contexts, one would expect most correlations between the item and the total scales to be greater than 0.40.

Table 3.3. Rules of Thumb for Correlation Coefficients

Magnitude of Agreement	Coefficient Range
Large	>0.40
Medium	0.20–0.39
Small	<0.19

Table 3.4 provides Pearson correlation coefficients between the individual Modified Bridges items, the total Modified Bridges scores, and both Florida's Healthy Start Prenatal Risk Screen and Healthy Start Infant Risk Screen total scores. The magnitude of the correlation reflects the degree to which the content for a particular item reflects the content that the comparison scale measures.

Table 3.4. Correlations Between Individual Modified Bridges Items and the Modified Bridges Total Score, Healthy Start Prenatal Risk Screen Score, and Healthy Start Infant Risk Screen Score

	Pearson Correlation		
Modified Bridges Item	**Modified Bridges Total Score**	**Healthy Start Prenatal Risk Screen Score**	**Healthy Start Infant Risk Screen Score**
History of child abuse/neglect	0.60***	0.29***	0.32***
Mental health problem	0.56***	0.30***	0.09
History of excessive alcohol or other drug use by people who will impact infant's well-being	0.54***	0.14	0.30***
Level of family support	0.54***	0.41***	0.23**
Adequate and timely PNC	0.53***	0.27**	0.53***
History of domestic violence	0.50***	0.19*	0.09
Infant medical problems	0.49***	0.35***	0.45***
Mother's intent to remain current with care	0.46***	0.20**	0.36***
Mother's medical problem	0.44***	0.32***	0.29***
Transportation a barrier	0.41***	0.22**	0.32***
Current housing conditions	0.39***	0.20**	0.17*
Other children at home	0.37***	0.15	0.20**
Annual household income level	0.36***	0.23**	0.36***
Infant feeding issues	0.36***	0.14	0.16*
Tobacco smoke in home	0.35***	0.23**	0.25**
Awareness of resources	0.34***	0.25**	0.19*
Source of medical care for infant	0.32***	0.13	0.26***
Adequate food in the house	0.32***	0.12	0.14
Mother's worry about infant's health	0.31***	0.18*	0.39***
Strength of maternal bond with infant	0.29***	0.21**	0.18*
Mother's marital status	0.28***	0.38***	0.42***
Mother's highest education	0.27***	0.28***	0.21**
Mother's age	0.22**	0.18*	0.15
Infant health coverage	0.01	–0.15	–0.06
Mother's health coverage	–0.04	–0.01	–0.01
Mother's English	–0.1	–0.06	–0.1
Average item–total correlations	0.35	0.20	0.23
Number of items $r > 0.20$	23 (89%)	15 (58%)	15 (58%)

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Correlations between the Modified Bridges total score and individual items ranged from $r = 0.22$ to $r = 0.60$. The largest correlations occurred for history of child abuse or neglect ($r = 0.60$***), mental health problem ($r = 0.56$***), history of excessive alcohol or other drug use by people who will impact infant's well-being ($r = 0.54$***), and level of family support

($r = 0.54$***). The smallest item–total correlations occurred for mother's age, infant health coverage and mother's health coverage, and mother's use of English, which were not significantly associated with the total Modified Bridges score. Note that these same items were not significantly correlated with either Florida screening measure, which suggests that these risk factors are dissimilar from other aspects of risk that the Modified Bridges and Florida instruments measure and do not contribute to the measures' score reliability.

The items most strongly associated with the Florida measures differed somewhat from those highly correlated with the Modified Bridges measure, indicating slight differences in the content that each measure assesses. In particular, the Modified Bridges items most strongly associated with the Florida prenatal and infant risk measures were mother's marital status ($r = 0.38$ and 0.42), infant medical problems ($r = 0.35$ and 0.45), and adequate and timely PNC ($r = 0.27$ and 0.53).

In summary, the majority of the Modified Bridges items have moderate correlations with the Modified Bridges total score (23 of 26 items, 89 percent, average $r = 0.35$), but fewer are moderately correlated with the Florida prenatal screen (15 of 26, 58 percent, average $r = 0.20$) and the Healthy Start Infant Risk Screen (15 of 26, 58 percent, average $r = 0.23$). These findings suggest that the total Modified Bridges score most reflects content related to severe risk factors, including history of child abuse or neglect, mental health problem, history of excessive alcohol or other drug use by people who will impact infant's well-being, and level of family support (given the high weight of these items).

We also assessed various correlations between the Modified Bridges subscales, total scores, and threshold scores for both the Florida risk measures and Modified Bridges (i.e., using dichotomized scores based on the high versus low/moderate thresholds; see Table 3.5). The Modified Bridges total score is moderately, and similarly, correlated with both Florida's Healthy Start Infant Risk Screen ($r = 0.59$) and Florida's Healthy Start Prenatal Risk Screen ($r = 0.57$). This finding suggests that the Modified Bridges risk content is balanced between prenatal and postnatal risk factors. In addition, the correlation suggests that the content between the separate scales is similar, but not redundant, indicating that the Modified Bridges is measuring content not present on the Florida scales. The correlations based on the Modified Bridges risk threshold and Florida measure risk thresholds are somewhat lower (Florida prenatal risk $r = 0.27$, Florida infant risk $r = 0.45$), providing an initial indication that the Modified Bridges and Florida measures do not provide identical risk classifications. Finally, the correlations between the Modified Bridges subscales and the Florida prenatal and infant risk measures suggest that the Florida measures provide slightly more emphasis on demographic ($r = 0.42$ and 0.42) than psychosocial ($r = 0.39$ and 0.31, respectively) or medical ($r = 0.36$ and 0.36) risk factors.

Table 3.5. Pearson Correlations Between the Modified Bridges and Florida Risk Instrument Scores
(*N* = 153)

Score	Modified Bridges Total Score	Modified Bridges Risk (≥50)	Modified Bridges Medical	Modified Bridges Psychosocial	Modified Bridges Demographics/Basic Needs
Florida prenatal total score	0.57	0.42	0.36	0.39	0.42
Florida infant total score	0.59	0.43	0.36	0.31	0.42
Florida prenatal cutoff score (≥6)	0.39	0.27	0.29	0.25	0.23
Florida infant cutoff score (≥4)	0.53	0.45	0.37	0.30	0.26
Modified Bridges total score		0.77	0.57	0.84	0.58
Modified Bridges risk threshold (≥50)			0.49	0.61	0.47
Modified Bridges Medical				0.32	0.14[a]
Modified Bridges Psychosocial					0.24

NOTE: All correlations are *p* < 0.001 unless otherwise noted.
[a] Not significant.

Evaluating the Risk Tool Cutoff Score

In this section, we examine the Modified Bridges risk cutoff scores using a variety of specialized analyses. To assess the utility of a cutoff score of 50, and to evaluate possible alternative cutoffs, we examined sensitivity and specificity, kappa coefficients, receiver operating characteristics (ROCs), and the area under the ROC curve (AUROCC) (see Swets, 1996, and Altman and Bland, 1994). We describe each of these analyses in more detail below.

As a means of understanding differences in the severity of risk that the Modified Bridges and Florida risk measures assess, we provide the tools' means and percentage of the sample who qualify as high risk according to each. Table 3.6 provides the distribution of those who meet criteria for risk according to the range of cutoff scores of the Modified Bridges and the Florida measures. Using the Florida measures' risk thresholds to define high risk, the mean Modified Bridges total score for the prenatal high-risk group was 82.0 (SD = 32.1), whereas the mean score for women without prenatal risk was only 47.9 (SD = 25.6). The mean Modified Bridges total score for the postnatal infant high-risk group was 86.2 (SD = 30.3), whereas the mean score for women without postnatal risk was 45.5 (SD = 23.1). This finding suggests that, given the Florida measures, a comparable Modified Bridges threshold would be about 80 or 85. These

thresholds would also provide a proportion of mothers (16 percent and 12 percent, respectively) that is similar to the proportion that the Florida measures identified.

Table 3.6. Dichotomous Risk Cutoff Frequencies and Percentages

Risk Cutoff Score	Frequency	Percentage
Modified Bridges		
≥50	75	49.02
≥35	102	66.67
≥40	91	59.48
≥45	83	54.25
≥55	61	39.87
≥60	52	33.99
≥65	43	28.10
≥70	33	21.57
≥75	28	18.3
≥80	25	16.34
≥85	18	11.76
≥90	16	10.46
≥95	14	9.15
≥100	10	6.54
Florida prenatal (total score 6 or more)	19	12.42
Florida infant (total score 4 or more)	25	16.34

In this section, we report the sensitivity and specificity of the Modified Bridges. *Sensitivity* refers to the proportion of high-risk families based on the validation tool that the screener is correctly identifying. That is, sensitivity describes how well the Modified Bridges correctly identifies high-risk families. For example, a hypothetical sensitivity value of 0.50 would mean that only 50 percent of mothers identified as high risk according to the primary measure (in this case, the Modified Bridges) would also be identified as having high risk according to the comparison measure (in this case, the Florida risk measures). *Specificity* reflects the proportion of low-/moderate-risk families based on the validation tool that the screener is correctly identifying. That is, specificity describes how well the Modified Bridges correctly identifies low-risk families. For example, a hypothetical specificity value of 0.50 would mean that only 50 percent of mothers identified as low risk according to the primary measure would also be identified as having low risk according to the comparison measure. In a scale development context, it is common to evaluate various cutoff scores in an effort to maximize both the sensitivity and specificity of the measure, occasionally sacrificing one in favor of the other. In this context, the focus of these analyses is on the *sensitivity* of the Modified Bridges because it is

critical that the high-risk families be identified so they are eligible for additional follow-up services. In other words, given that the purpose of the Modified Bridges is to identify mothers and children in need of supportive services, what is most important is that the tool be able to correctly identify the high-risk families (i.e., sensitivity). Further, in focusing on sensitivity, one is more confident in not missing any high-risk families, and the only potential risk is offering additional services to those families who might not have needed it (i.e., low-/moderate-risk families).

Table 3.7 shows the range of sensitivity, specificity, and kappa[7] of different Modified Bridges cutoff thresholds in determining prenatal and postnatal infant risk.

---

[7] The kappa coefficient expresses the agreement between the Florida risk measures and the Modified Bridges. Table 3.7 shows how the agreement between the measures varies as a function of the Modified Bridges cutoff scores.

Table 3.7. Sensitivity and Specificity of Modified Bridges (Thresholds Between 35 and 100)

Florida Risk Measure	Threshold	Sensitivity	Specificity	Kappa
Prenatal (≥6)				
	≥35	17.65	98.04	0.11
	≥40	19.78	98.39	0.15
	≥45	20.48	97.14	0.16
	≥50	21.33	96.15	0.18
	≥55	26.23	96.74	0.26
	≥60	28.85	96.04	0.29
	≥65	34.88	96.36	0.38
	≥70	36.36	94.17	0.36
	≥75	35.71	92.8	0.33
	≥80	36	92.19	0.31
	≥85	33.33	90.37	0.23
	≥90	31.25	89.78	0.19
	≥95	35.71	89.93	0.22
	≥100	50	90.21	0.28
Infant (≥4)				
	≥35	24.51	100	0.18
	≥40	27.27	100	0.23
	≥45	30.12	100	0.28
	≥50	33.33	100	0.34
	≥55	39.34	98.91	0.42
	≥60	42.31	97.03	0.45
	≥65	39.53	92.73	0.37
	≥70	45.45	91.67	0.41
	≥75	46.43	90.4	0.38
	≥80	48	89.84	0.38
	≥85	55.56	88.89	0.38
	≥90	56.25	88.32	0.36
	≥95	50	87.05	0.27

Using a standard cutoff of 50, the Modified Bridges had poorer sensitivity (21.33), higher specificity (96.15), and lower kappa agreement (0.18) than the Florida postnatal risk screen. Put another way, of the mothers whom the Modified Bridges identifies as being high risk, only 21 percent are also identified as being high risk according to the Florida postnatal risk screen. In comparison with the Florida prenatal risk screen, the Modified Bridges performed slightly better, with slightly higher sensitivity (33.33), specificity (100), and kappa agreement (0.34). This means that, of the mothers whom the Modified Bridges identifies as being high risk, only

33 percent (or one out of three) are also identified as being high risk according to the Florida prenatal risk screen. These results show that the current Modified Bridges cutoff of 50, compared with the Florida measures, has poorer sensitivity but higher specificity, suggesting that it is accurate in identifying those who are not at risk but performed worse in accurately identifying those who are high risk.

Given the relatively low sensitivity of the Modified Bridges' current threshold, we examined similar indices based on the other possible risk thresholds ranging from 35 to 100 and generally found low sensitivity (ranging from 17.65 percent to 46.43 percent) but high specificity (ranging from 92.8 percent to 100 percent) for both prenatal and postnatal infant risk. Sensitivity generally improves with a higher risk threshold, while specificity improves with a lower threshold. However, a cutoff score of 70 appears to maximize the sensitivity for prenatal risk, while a cutoff score of 90 maximizes the sensitivity for postnatal infant risk.

Finally, we also considered the ROC curve and the AUROCC. The ROC curve and AUROCC indicate the degree of accuracy of the cutoff point in detecting risk. ROC curves are traditionally used in medical diagnostic settings as a means of assessing the ability of two instruments to correctly classify patients. In the present context, the ROC curves reflect the Modified Bridges' ability to correctly distinguish high-risk mothers from mothers who are not high risk. A measure that has no ability to distinguish between mothers would have an AUROCC of 0.50, and a measure with a perfect classification rate would have an AUROCC of 1.0. There are no published rules of thumb for AUROCC values, but values greater than 0.90 are typically considered excellent, values greater than 0.80 are considered good, and lower values are considered fair or poor.

The ROC curve illustrates that the Modified Bridges performed moderately well in identifying patients with risk. In general, AUROCC values greater than 0.80 are considered good. The AUROCC in this study was 0.81 for prenatal risk (see Figure 3.1) and 0.88 for postnatal infant risk (see Figure 3.2), suggesting that the Modified Bridges has moderately good accuracy in distinguishing between mothers with high risk versus low risk.

Figure 3.1. Receiver Operating Characteristic for Prenatal Risk

ROC Curve for Model
Area Under the Curve = 0.8183
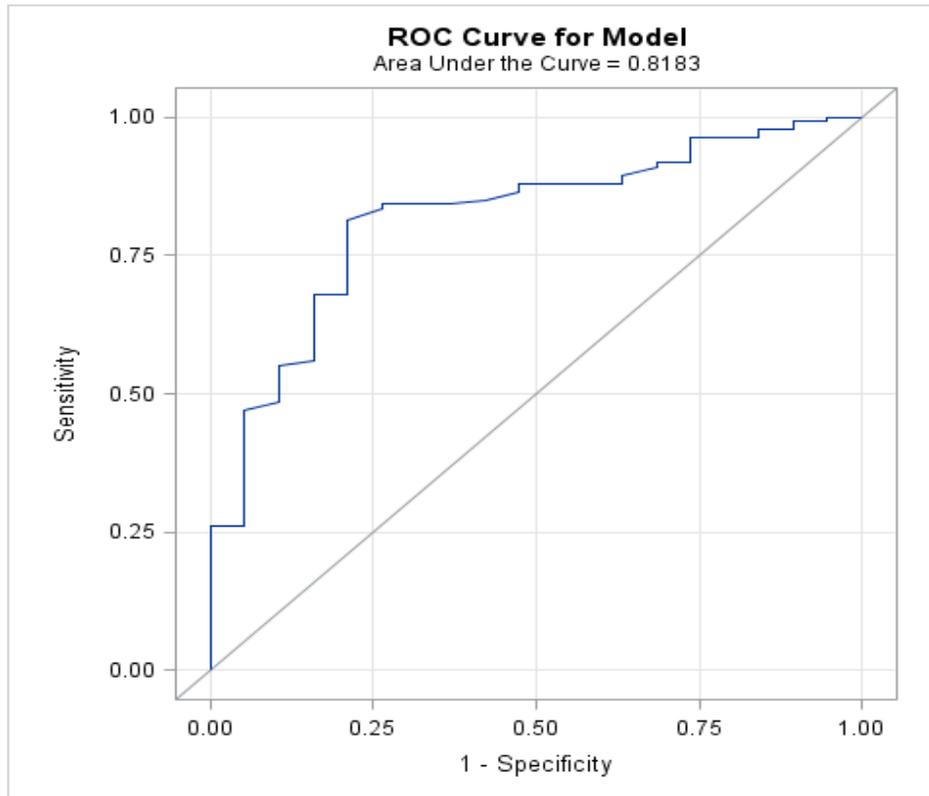
Figure 3.2. Receiver Operating Characteristic for Postnatal Infant Risk
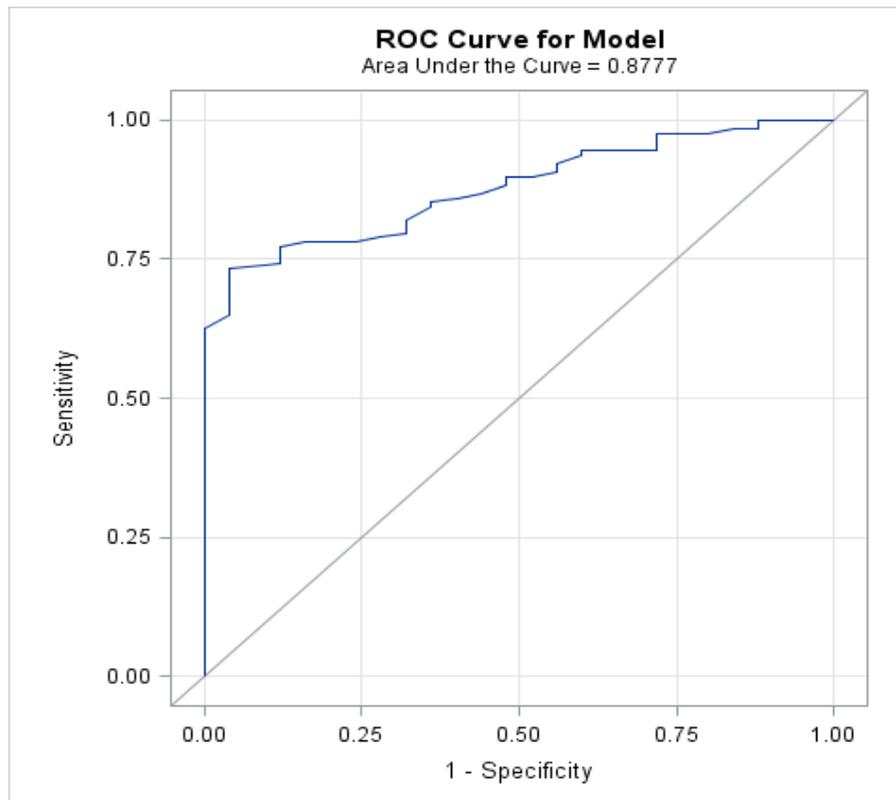
Taken together, these results provide a somewhat mixed set of findings with respect to the Modified Bridges' ability to distinguish between high- and low-risk mothers. The AUROCC, which is relatively good, reflects a combination of *both* the Modified Bridges' strong ability to correctly identify low-risk mothers (i.e., specificity) and its relative weakness at identifying high-risk mothers (i.e., sensitivity). In other words, it lands somewhere in the middle of the Modified Bridges' strength (correctly identifying low risk) and weakness (overidentifying high risk). Readers should consider both the Modified Bridges' sensitivity and specificity when evaluating the AUROCC.

Finally, we note that these findings are dependent on the use of the Florida measures and might not generalize to other risk measures. Thus, even if a higher risk threshold were chosen for the Modified Bridges, such that the Modified Bridges identified a proportion of high-risk mothers similar to what the Florida instruments identified, the two instruments would still select somewhat different mothers as being high risk. This issue is best understood by considering what types of risk factors should lead to a high-risk designation. The Florida measures are somewhat more focused on risk associated with infant health issues, while the Modified Bridges identifies high risk that is related to psychosocial issues pertaining to the mother (e.g., history of child abuse or neglect, mental health problem, drug and alcohol abuse).

Comparison of Modified Bridges Total Scores Across Known Risk Groups

Finally, we assessed the convergent validity of Modified Bridges, which refers to the degree to which two measures with hypothetically related content are found through statistical analyses to be related. For example, we might hypothetically assume that the mother's age at pregnancy is related to her overall risk level. Convergent validity simply evaluates the actual relationship between age and risk level: Are these concepts actually related? In this section, we provide an assessment of the convergent validity of each question on the Modified Bridges.

The question we are trying to answer is "how different is the *overall* risk level for a mother who scored in the highest risk category for a particular item and a mother who scored in the lowest risk level for a particular item?" That is, for each Modified Bridges item, we used the highest item response category (i.e., highest risk) and lowest item response category (i.e., lowest risk) to define the high- and low-risk groups. Take, for example, the age risk question on the Modified Bridges. The Modified Bridges uses the following scoring protocol for this question:

Ages 25 and above = zero points (lowest risk)

Ages 20–24 = four points

Ages 17–19 or >39 = eight points

Age <17 = 12 points (highest risk).

For this age risk Modified Bridges question, we want to compare the mean overall risk of all mothers in the lowest risk category (ages 25 and above) and the mean overall risk of all mothers in the highest risk category (age <17). If one found that the very young mothers (<17) have very high levels of overall risk, and the mothers 25 years and above have low levels of risk, this item would have good convergent validity and would be able to distinguish between high- and low-risk mothers. If the overall risk levels do not vary by mothers' age, this question would not do a very good job of distinguishing between high and low risk.[8]

Using the additional statistical power provided by the secondary data set (i.e., data downloaded from the Stronger Families database that contain responses from 12,165 mothers), Table 3.8 reports the differences in the overall mean Modified Bridges scores between mothers who score in the high- and low-risk item response groups for all the Modified Bridges items. In addition to significance tests, we computed effect size estimates using Cohen's *d* to provide a more generalizable interpretation of the magnitude of the group difference. Cohen's *d* values of 0.8, 0.5, and 0.2 represent large, moderate, and small effect sizes, respectively.

---

[8] Note that the overall Modified Bridges score used in this evaluation removes the effect of the particular item's score on the total Modified Bridges score. For example, the Modified Bridges alcohol use item has a mean of 46 for mothers who score in the zero-point item category and 94 for mothers who score in the 18-point category, but the corrected mean of the 18-point category is 76 (i.e., 94 − 18 = 76). This correction allows for an unbiased comparison of total risk between the two groups.

Table 3.8. Comparing the Modified Bridges Total Risk Scores Across Known Risk Groups

Comparison	Modified Bridges Mean (SD)	t-Value (df)	Pr > |t|	Cohen's d
Mother's medical problem: No limitations; fully able to care for child	43.54 (19.62)	−25.36 (6,619)	<0.0001	2.44
Mother's medical problem: Severe limitations; mother is unlikely to be able to care for child	92.20 (42.81)			
Current housing conditions: Stable and safe	40.26 (18.56)	−31.5 (6,648)	<0.0001	2.43
Current housing conditions: Currently homeless or in temporary shelter or car	86.10 (35.39)			
Level of family support: Family and/or partner is supportive, available, and committed to help	42.51 (18.41)	−23.6 (7,282)	<0.0001	2.14
Level of family support: No relatives/friends/partner available or committed; geographically isolated from community services; no phone	82.90 (39.98)			
Adequate food in the house: Consistent	41.28 (19.80)	−31.24 (6,467)	<0.0001	1.96
Adequate food in the house: Chronically inadequate	81.03 (33.99)			
History of child abuse/neglect: No known history of child abuse/neglect	45.90 (18.12)	−52.54 (11,163)	<0.0001	1.82
History of child abuse/neglect: Pending child abuse/neglect investigation; previous abuse/neglect; prior court action, e.g., siblings removed from home; or child abuse suspected or discussed but no system intervention to date	79.48 (27.39)			
Adequate and timely PNC: Accessed PNC within three months; consistent with follow-up appointments	46.75 (19.99)	−26.19 (9,914)	<0.0001	1.79
Adequate and timely PNC: No PNC	83.11 (36.28)			
Mental health problem (mother): No challenges; realistic expectations of child; no history of mental health issues	44.09 (18.42)	−47.59 (8,250)	<0.0001	1.60
Mental health problem (mother): Severe challenges; mother might have difficulty in caring for child	75.00 (32.50)			
History of excessive alcohol or other drug use by people who will impact infant's well-being: No history	46.23 (19.76)	−41.33 (9,249)	<0.0001	1.48
History of excessive alcohol or other drug use by people who will impact infant's well-being: Mother/father/significant other not in drug/alcohol treatment program; individual is in program but attendance is sporadic; (mother) entered program late in pregnancy	76.22 (30.57)			
Infant feeding issues: Mother demonstrates knowledge, confidence regarding infant feeding; sufficient resources available to support a healthy feeding relationship with infant	37.96 (20.86)	−23.37 (1,648)	<0.0001	1.43
Infant feeding issues: Grossly insufficient knowledge regarding infant feeding; lack of interest in improving feeding skills; evidence of or high potential for poor feeding relationship with infant	70.62 (34.80)			
Tobacco smoke in home: No smoking	46.84 (20.25)	−29.19	<0.0001	1.37

Comparison	Modified Bridges Mean (SD)	t-Value (df)	Pr > |t|	Cohen's d
Tobacco smoke in home: Mother smokes	75.40 (33.39)	(10,313)		
Mother's intent to remain current with well-baby care and immunizations for her infant: Very strong	45.50 (20.25)	−26.33 (9,631)	<0.0001	1.35
Mother's intent to remain current with well-baby care and immunizations for her infant: Very weak/no understanding	73.65 (33.84)			
Strength of maternal bond with infant: Strong	45.67 (20.56)	−15.98 (9,401)	<0.0001	1.21
Strength of maternal bond with infant: None	70.67 (36.58)			
History of domestic violence: No known history of domestic violence	45.89 (18.40)	−47.18 (11,162)	<0.0001	1.17
History of domestic violence: Domestic violence investigation; previous domestic violence of serious nature; prior court action	68.01 (27.96)			
Mother's worry about infant's health: Not worried	43.51 (21.08)	−31.21 (7,695)	<0.0001	1.15
Mother's worry about infant's health: Very worried	67.95 (22.86)			
Source of medical care for infant: Regular pediatrician or community clinic	43.81 (21.22)	−25.86 (6,832)	<0.0001	0.98
Source of medical care for infant: No plan for future care	65.56 (30.65)			
Transportation a barrier: Never	41.68 (19.77)	−41 (7,667)	<0.0001	0.96
Transportation a barrier: Frequently	61.66 (23.97)			
Other children at home: None under 5 years old; singleton birth	46.26 (21.56)	−18.53 (3,186)	<0.0001	0.86
Other children at home: Three or more under 5 years old with singleton or multiples, or other child(ren) removed from home or deceased	66.36 (33.64)			
Mother's marital status: Married	43.69 (19.97)	−23.23 (4,263)	<0.0001	0.86
Mother's marital status: Single, living alone	62.13 (30.55)			
Annual household income level: >$60,000	28.51 (16.82)	−17.07 (9,204)	<0.0001	0.78
Annual household income level: Below $18,000 or mother unaware of income	45.26 (21.44)			
Mother's demonstrated awareness of available resources: Full awareness	40.16 (22.18)	−24.24 (4,753)	<0.0001	0.68
Mother's demonstrated awareness of available resources: No awareness	54.93 (20.08)			
Mother's heath coverage: Full	47.32 (23.06)	−8.57 (7,431)	<0.0001	0.59
Mother's health coverage: None	61.01 (35.60)			
Mother's age: 25 and above	46.89 (21.83)	−12.87 (7,185)	<0.0001	0.46
Mother's age: Under 17	56.95 (18.60)			
Infant health coverage: Full	46.72 (22.14)	−8.75 (8,842)	<0.0001	0.40
Infant health coverage: None	55.79 (30.29)			
Mother's highest education level: College graduate	38.27 (20.16)	−20.34	<0.0001	0.38

Comparison	Modified Bridges Mean (SD)	t-Value (df)	Pr > |t|	Cohen's d
Mother's highest education level: Did not complete high school	46.50 (22.18)	(4,978)		
Infant medical problems: No apparent medical or physical problems	44.64 (19.51)	−43.26 (10,035)	<0.0001	0.37
Infant medical problems: Any preterm infant and/or physical or medical problem that significantly impacts vital life functions or physical and intellectual development	52.12 (24.02)			
Mother's use of English: Fluent	49.50 (23.68)	−4.23 (10,245)	<0.0001	−0.03
Mother's use of English: No English	48.89 (18.29)			

NOTE: The *t*-value is the student's *t*-test. df = degrees of freedom.

Results of these risk-group comparisons provide insights into the Modified Bridges content that is most frequently associated with high risk. The Modified Bridges items that most distinguished high- and low-risk groups were level of family support ($d = 2.13$), current housing conditions ($d = 2.43$), and mother's medical problem ($d = 2.13$). The Modified Bridges items that provide the least ability to distinguish between high and low risk are mother's use of English ($d = -0.03$), infant medical problems ($d = 0.37$), and mother's highest education level ($d = 0.38$). The majority of the mothers who scored in the highest risk category for each item had total Modified Bridges scores greater than 60, and, in general, those items at the top of Table 3.8 were associated with the highest risk, while those at the bottom of Table 3.8 were associated with lower risk (or are less likely to distinguish between high- and low-risk mothers).

Limitations

There are several limitations to consider in this section. First, and perhaps most importantly, in research that seeks to compare the performance of two measures against each other, it is most common to select a gold-standard measure for the comparison—that is, a measure that is widely accepted and used in the field. In doing so, the results of the comparison are understood relatively to the gold-standard measure, with which most practitioners would be familiar given its popularity. In the present case, no such gold standard exists, so the results are only relative to the particular measure used (in this case, the Florida risk measures). Although we selected the comparison measure for its comparability to the Modified Bridges (in terms of both content and use), this is a limitation of which readers should be aware. Second, the correlations between the Modified Bridges and Florida risk measures were lower than we expected. These lower correlations in part explain the poor sensitivity results (i.e., the high-risk mothers that the Modified Bridges identified were not the same high-risk mothers that the Florida risk measures identified). The reduced correlations might be explained in part by the third key limitation: The modification in the administration procedures of the Florida risk measures could have affected

responses that the mothers gave. The original format of the Florida risk measures uses a paper-and-pencil self-reported survey to obtain responses from the mothers. It might be that modifying this format to be consistent with the Modified Bridges' use of an interview conducted by the hospital liaisons might have, although unlikely, changed how the mothers answered the questions. That is, they might have provided lower-risk answers to questions, thereby resulting in fewer mothers being identified as being high risk via the Florida risk measures. Taken together, readers should be aware of these limitations, although they are not severe, when interpreting the results presented in this chapter.

Conclusion

Results from the validation exercise provided mixed evidence to support the predictive validity of the Modified Bridges' current threshold of 50. At the item level, we found several items that are strong predictors of risk (e.g., history of child abuse or neglect, level of family support, and history of excessive alcohol or other drug use by people who will impact infant's well-being) along with several items that do not strongly distinguish between risk levels (e.g., mother's age, mother's use of English, and infant health coverage). However, we also found that the Modified Bridges items have only moderate to low correlations with Florida's Healthy Start Prenatal Risk Screen and Healthy Start Infant Risk Screen. This is in part explained by subtle content differences between the Florida and Modified Bridges measures, with the Modified Bridges' content reflecting severe risk factors, including history of child abuse or neglect, mental health problem, history of excessive alcohol or other drug use by people who will impact infant's well-being, and level of family support. Given the modest associations between the Modified Bridges and Florida measures, it is not then surprising that the sensitivity analyses also resulted in lower-than-anticipated values. The Modified Bridges risk threshold is more conservative and allows more mothers the opportunity for follow-up services, while the Florida measures require the presence of a higher level of risk. Our findings suggest that, in order to maximize the sensitivity and specificity of the Modified Bridges, higher thresholds would be needed, perhaps as high as 70 or more, which would, in turn, reduce the number of mothers eligible for SHV.

Chapter Four. Does the Modified Bridges for Newborns Screening Tool Accurately Distinguish Between Low-/Moderate-Risk Mothers and High-Risk Mothers?

A key aspect of assessing the psychometric properties of the Modified Bridges is to understand the degree to which it can precisely distinguish between low-/moderate-risk mothers and high-risk mothers. This question relates to the content validity and reliability of the Modified Bridges. To determine the content validity of the Modified Bridges, we used two separate psychometric approaches: (1) factor analytic techniques to understand the content domains that the Modified Bridges measures and (2) reliability analyses to understand the precision of the Modified Bridges scores and the extent to which the Modified Bridges item weights affect the internal consistency of the total Modified Bridges scores. Factor analyses provide us with an understanding of whether the Modified Bridges' subscales are truly the content that the Modified Bridges questions assess. The reliability analyses describe how much confidence one can have in the scores the Modified Bridges provides. For example, low reliability means that a particular mother's score could vary in unexplained ways if she were assessed repeatedly.

Data Description

Because factor analyses require larger sample sizes than other correlational approaches, we used secondary data from the Stronger Families database. These data were collected from ten Welcome Baby hospitals from January 2013 through December 2015 and consist of 12,166 completed Modified Bridges surveys. We obtained demographic characteristics of mothers from the median responses to the categorical Modified Bridges items: The mothers' median age band was older than 25, the highest educational attainment was high school, the majority were single and living with family, and the median household income was less than $18,000 or unknown.

Using Factor Analysis to Understand the Content of the Modified Bridges

The factor analytic approach we used was intended to identify the main content domains that underlie responses to the Modified Bridges items. That is, because hospital liaisons complete the Modified Bridges items, we expect that the items with similar content will be correlated with one another. When subsets or clusters of items with similar content are correlated, it is possible to extract factors from these content clusters. The factors then provide insight into the content that the Modified Bridges assesses. Typically, the factors obtained from this process are consistent

with one's qualitative understanding of the content measured by the scale; however, this factor analytic process can also uncover factors of content that test developers had not hypothesized.

Our process of understanding the content that the Modified Bridges measures consisted of two factor analytic techniques: (1) confirmatory factor analysis (CFA) and (2) exploratory factor analysis (EFA). CFA is a more restrictive psychometric model that imposes a factor structure on the data and tests the extent to which the data support that theory or model. In this context, that means assessing whether the items in each of the Modified Bridges' subscales are actually closely related to their particular subscale. EFA is less restrictive and allows the data to "speak for themselves." That is, no content or factor assumptions are placed on the data—the factors that emerge might be consistent with the scale developer's theories regarding the content of the items or new revelations regarding the clustering of item content might emerge. In this case, that means that we allow the items to *load* on the factor to which they are most closely related, even if that factor is not their original subscale. For examples of these techniques described in further detail, please see Stucky and Edelen, 2015; Stucky, Gottfredson, and Panter, 2012; Stucky, Gottfredson, Panter, et al., 2011; and Irwin et al., 2010. For both techniques, we rely primarily on two sources for our understanding of the factor structure: (1) model fit indices that reflect how closely our model supports the data, including root mean square error of approximation (RMSEA) ≤ 0.08, Tucker–Lewis Index (TLI) ≥ 0.95, and comparative fit index (CFI) ≥ 0.95 (Hu and Bentler, 1999; Browne and Cudeck, 1993), and (2) the magnitude of the factor loadings, which indicate the strength of the relationship between a particular item and the underlying factor (i.e., content domain). In general, the minimum acceptable values are factor loadings greater than 0.50, and more-desirable levels are greater than 0.70. This is because low loadings indicate items that are poorly related to the factor and higher loadings are associated with better (i.e., more-reliable) measures of the intended construct. We evaluated all factor analyses using weighted least squares mean and variance-adjusted estimation that accounts for the categorical nature of the items using interitem polychoric correlations as implemented in the Mplus software (Muthén and Muthén, 2012).

The Factor Structure of the Modified Bridges

We initially fit a three-factor CFA model in which each factor was represented by the Modified Bridges subscales (i.e., Medical, Psychosocial, and Demographics/Basic Needs). The results of this model failed to provide evidence of the hypothesized three content domains. Only 12 of the 26 factor loadings were greater than 0.50, and fit indices suggested that the hypothesized factor structure was not supported by the data (χ^2 [325] = 51,892, $p < 0.001$; RMSEA = 0.080, CFI = 0.553, and TLI = 0.509). Although the magnitude of the factor loadings for the Psychosocial subscale were highest (range = 0.42–0.67), the Medical and Demographics/Basic Needs factor loadings were substantially lower (ranges = 0.24–0.75 and 0.05–0.76, respectively). These findings suggest that the data do not support the hypothesized

Modified Bridges subscales, indicating that the Modified Bridges might assess different risk factors from those defined by the three subscales.

Following the lack of a clear structure from the previous three-factor CFA, we next analyzed an EFA in order to extract content domains not based on any a priori hypothesized domains. The EFA results obtained provided a somewhat complicated view of the Modified Bridges factor structure. The most parsimonious model was a four-factor solution with the extracted factors corresponding to demographics (e.g., age, use of English), resource availability (e.g., having adequate food in the house, annual household income level, housing conditions), psychosocial concerns (e.g., history of child abuse or neglect, history of domestic violence, history of excessive alcohol or other drug use by people who will impact infant's well-being), and infant health (infant medical problems and mother's worry about infant's health). However, these factors had relatively few items with factor loadings greater than 0.50 (e.g., only two to eight items per factor), and nine of the 26 Modified Bridges items had loadings less than 0.50 on all four factors.

Our final factor model used insights gained from the EFA by fitting a four-factor CFA model. The factors corresponded to psychosocial problems, resource availability, health care awareness, and infant health (see Table 4.1). The four-factor model has acceptable fit (χ^2 [130] = 7,332, $p < 0.001$; RMSEA = 0.067, CFI = 0.888, and TLI = 0.868), and the majority of the factor loadings are greater than 0.50; however, nine items did not have estimated factor loadings. The majority of these nine items are from the Demographics/Basic Needs subscale of the Modified Bridges. Taken together, these findings suggest that about one-third of the items on the Modified Bridges measure *unique* content—that is, content that is not related to other items on the scale. Further, given the magnitudes of the correlations in Table 4.1, the dominant Modified Bridges content (i.e., the content that the Modified Bridges most strongly measures) is mother's demonstrated awareness of available resources, psychosocial problems, and health care awareness, while just two items that have high factor loadings assess infant medical problems, indicating that the items are potentially redundant.

Table 4.1. Factor Loadings from a Four-Factor Confirmatory Factor Analysis Model Based on Revised Modified Bridges Subscales

Item Wording	Psychosocial Problems	Mother's Demonstrated Awareness of Available Resources	Health Care Awareness	Infant Medical Problems
History of child abuse or neglect	0.76	—	—	—
History of domestic violence	0.69	—	—	—
Mental health problem (mother; include PPD)	0.67	—	—	—
History of excessive alcohol or other drug use by people who will impact infant's well-being	0.63	—	—	—
Level of family support	0.60	—	—	—
Adequate food in the house	—	0.80	—	—
Current housing conditions	—	0.73	—	—
Transportation a barrier	—	0.64	—	—
Mother's demonstrated awareness of available resources	—	0.61	—	—
Annual household income level	—	0.57	—	—
Mother's intent to remain current with well-baby care and immunizations for her infant	—	—	0.77	—
Infant health coverage	—	—	0.65	—
Source of medical care for infant	—	—	0.64	—
Strength of maternal bond with infant	—	—	0.55	—
Mother's health coverage	—	—	0.43	—
Infant medical problems	—	—	—	0.80
Mother's worry about infant's health	—	—	—	0.80
Stand-alone items				
Mother's age	—	—	—	—
Mother's use of English	—	—	—	—
Mother's highest education level	—	—	—	—
Mother's marital status	—	—	—	—
Other children at home	—	—	—	—
Tobacco smoke in home	—	—	—	—
Adequate and timely PNC	—	—	—	—
Infant feeding issues	—	—	—	—
Mother's medical problem	—	—	—	—

Summary of Factor Analyses

In this section, we provide a brief summary of the factor analytic results for readers who might be unfamiliar with the factor analysis technique. First, we found that the risk factors that the Modified Bridges measures differ somewhat from the original subscale (i.e., Medical, Psychosocial, and Demographics/Basic Needs). The poor fit of the three-factor CFA model and the low factor loadings demonstrated this. Following this result, we turned to the EFA to better understand the content dimensions present on the Modified Bridges. We found that the additional content that the Modified Bridges measures that is not among the original subscales is mother's demonstrated awareness of available resources and, to a lesser extent, infant medical problems. Finally, to identify a parsimonious model for the Modified Bridges, we fit one final four-factor CFA that confirmed the findings from the EFA. That is, in addition to assessing psychosocial problems, the Modified Bridges assesses availability of resources, health care awareness (somewhat similar to the Medical Needs subscale), and infant health (although this factor is represented by only two closely related items). Finally, we found that some *stand-alone* items assess important content, but content that is not found elsewhere on the Modified Bridges.

Reliability of the Modified Bridges

Using the secondary data set described earlier in this report, we evaluated the reliability of the Modified Bridges total and subscale scores using Cronbach's alpha. Unlike earlier chapters in this report that assessed the reliability of the *hospital liaisons*, in this chapter, we are assessing the reliability of the Modified Bridges itself (in other words, the reliability of the scores it produces). We computed two version of Cronbach's alpha: (1) a *standardized* version in which, before computing alpha, each item is standardized to have a mean of 0 and an SD of 1, and (2) a *raw* version in which alpha is computed using the weights of the items. The difference in the reliabilities obtained from the two approaches provides an indication of the impact that the item weights have on the Modified Bridges' reliability. That is, the standardized version is essentially equally weighting the items before computing the reliability, whereas the raw version takes the weights of the items into account when computing reliability. Given that the Modified Bridges is scored using unequal item weights, First 5 LA should consider reporting the raw version of reliability (Cronbach, 1951; Furr, 2011).

Table 4.2 compares the two approaches for computing reliability. The difference in the total Modified Bridges score reliabilities reflects the impact of the weights. The standardized version of reliability is acceptable at 0.74 (a common threshold for minimum reliability is 0.70); however, the raw version is lower than desired at 0.65. Note also that the subscales have standardized reliabilities that are at or less than 0.60, suggesting that it might be inappropriate to report scores at the subscale level.

Table 4.2. Raw and Standardized Cronbach's Alphas for the Modified Bridges

Score	Raw	Standardized
Medical	0.37	0.57
Psychosocial	0.61	0.59
Demographics/Basic Needs	0.56	0.60
Total Modified Bridges score	0.65	0.74

Given that the Modified Bridges is a relatively long instrument at 26 items, the lower-than-desired reliability is a function of having moderate to weak item–total correlations. The poor raw reliability indicates that some items in the Modified Bridges that have high item–total correlations (i.e., those that are more reliable) have *lower* weights, and those items that are less reliable tend to have higher weights. Table 4.3 provides the item correlations for all 26 items arranged in order of magnitude. Note that some items with high item–total correlations have low weights (for example, transportation a barrier is weighted 4) while other items with low item–total correlations have high weights (e.g., infant medical problems is weighted 18).

Table 4.3. Item–Total Correlations Based on Raw and Standardized Scores

Item	Item Weight	Item–Total Correlation	
		Raw	Standardized
Current housing conditions	18	0.41	0.50
Adequate food in the house	9	0.38	0.49
History of child abuse or neglect	18	0.37	0.26
Transportation a barrier	4	0.35	0.42
Level of family support	6	0.35	0.37
Mother's intent to remain current with well-baby care and immunizations for her infant	4	0.35	0.43
Mental health problem (mother; include PPD)	18	0.35	0.27
Mother's worry about infant's health	4	0.34	0.34
History of domestic violence	18	0.30	0.22
History of excessive alcohol or other drug use by people who will impact infant's well-being	18	0.30	0.23
Mother's demonstrated awareness of available resources	4	0.29	0.41
Source of medical care for infant	4	0.29	0.38
Tobacco smoke in home	4	0.28	0.25
Annual household income level	9	0.27	0.34
Adequate and timely PNC	12	0.25	0.24
Strength of maternal bond with infant	15	0.24	0.28
Infant feeding issues	9	0.23	0.26
Mother's medical problem	18	0.22	0.22
Mother's marital status	6	0.20	0.16
Infant health coverage	4	0.19	0.30
Infant medical problems	18	0.14	0.17
Other children at home	4	0.13	0.13
Mother's health coverage	4	0.09	0.18
Mother's highest education level	9	0.08	0.11
Mother's age	12	0.05	0.10
Mother's use of English	4	–0.04	0.00

Considering Potential Revisions to the Modified Bridges to Improve Reliability

In this section, we consider ongoing efforts to improve the reliability of the Modified Bridges by re-weighting and removing select items. This work is intended to provide a demonstration of the improvement in the reliability of the Modified Bridges that might be possible if select items are re-weighted or removed.

The process of identifying items to re-weight involved grouping items based on the magnitudes of their known groups' effect sizes (reported in Chapter Three) and item–total correlations (presented earlier in this chapter). The method of obtaining equally sized groups resulted in item–total correlations in the following groups (from highest to lowest) ($r = 0.50 – 0.34$), ($r = 0.34 – 0.22$), and ($r = 0.22 – 0.00$); this same process resulted in the following equally sized groups of effect sizes (from highest to lowest): ($d = 1.49 – 1.01$), ($d = 0.95 – 0.51$), and ($d = 0.31 – –0.13$). For each Modified Bridges item, we then evaluated both the effect size and item–total correlation groups and compared these groups with the magnitude of the item weight. When there was an inconsistency, we proposed more or less weight for the item.

Table 4.4 provides a comparison of the magnitudes of the effect sizes and item–total correlations, as well as the current and proposed weights. Across the 26 items, we propose changes to 12 items. To summarize the changes, we propose increasing the weight of seven items, decreasing the weight of three items, and removing two items from the scale. For example, the item adequate food in the house has among the highest item–total correlations and the highest effect size estimate but has a weight of only 9; for this item, we propose to increase the weight by six points to a final weight of 15.

Table 4.4. Summary of Proposed Changes to the Modified Bridges Item Weights Based on Item–Total Correlations and Effect Size Estimates

Modified Bridges Item	Current Weight	Proposed Weight (Change)	Average Effect Size	Item–Total Correlation
Items receiving *more* weight				
Adequate food in the house	9	15 (+6)	High	High
Mother's intent to remain current with well-baby care and immunizations	3	9 (+6)	High	High
Level of family support	6	12 (+6)	High	High
Transportation a barrier	3	6 (+3)	Moderate	High
Mother's awareness of available resources	3	6 (+3)	Moderate	High
Source of medical care for infant	3	6 (+3)	Moderate	High
Mother's worry about infant's health	3	6 (+3)	Moderate	High
Items receiving *less* weight				
Infant medical problems	18	Remove or lower	Low	Low
Mother's use of English	3	Remove	Low	Low
Mother's highest education level	9	3 (–6)	Low	Low
Mother's age	12	6 (–6)	Low	Low
Strength of maternal bond with infant	15	12 (–3)	Moderate	Moderate

First 5 LA might want to consider removing two items from the Modified Bridges: infant medical problems and mother's use of English. The infant medical problems item has been

demonstrated to be highly correlated (redundant) with the item mother's worry about infant's health (an item whose weight we suggest increasing). Because the mother's worry about infant's health item outperforms the infant medical problems item in terms of reliability and ability to distinguish between high- and low-risk mothers, we suggest removing the infant medical problems item. Because the items are redundant, including both means that the Modified Bridges is effectively overweighting this particular content. The other item we suggest removing, mother's use of English, has a *negative* effect size estimate (suggesting that the item should be scored in the opposite direction) and a zero item–total correlation. The other items that we suggest receive lower weights—e.g., mother's highest education level and mother's age— similarly have poor psychometric characteristics that are inconsistent with the magnitude of weight they receive.

We implemented the suggested changes in Table 4.4 using the secondary data set ($N = 12,166$), and we scored the Modified Bridges using the revised weights (along with removing the items infant medical problems and mother's use of English). These changes resulted in an overall improvement in the Modified Bridges' total score raw reliability (original = 0.65, revised = 0.73; see Table 4.5). Note that, because the standardized reliabilities are based on standardizing each item prior to calculating reliability, the change in standardized reliability is minor (original = 0.74, revised = 0.75) and is due entirely to the removal of two items. The two subscales receiving the largest improvement in reliability were the Medical (original = 0.37, revised = 0.48) and Demographics/Basic Needs (original = 0.56, revised = 0.65) subscales, with the Psychosocial subscale receiving only a slight improvement in reliability (original = 0.61, revised = 0.62).

Table 4.5. Cronbach's Alphas Using the Raw Scores and Standardized Scores

Score	Revised		Original	
	Raw	Standardized	Raw	Standardized
Medical	0.48	0.55	0.37	0.57
Psychosocial	0.62	0.59	0.61	0.59
Demographics/Basic Needs	0.65	0.62	0.56	0.60
Total	0.73	0.75	0.65	0.74

NOTE: The Cronbach's alpha values reported in this table describe the reliability of the scores that the Modified Bridges produces. In a sense, reliability provides an understanding of how much confidence one can have in the scores that a measure produces. Poor reliability means that, if a mother were repeatedly interviewed, her scores would likely vary across the interviews. High reliability would mean that her scores would be highly stable across interviews.

Limitations

In this chapter, we present results demonstrating the improvement in the score reliability of the Modified Bridges if some items were removed and if others were re-weighted. When interpreting this, readers should be aware that these findings are based on results that assume that the revisions to the Modified Bridges and scoring protocol have been implemented. That is, the revised psychometric properties of the Modified Bridges presented here assume that changing the Modified Bridges would not have other unintended consequences. Put more simply, to ensure that the reliabilities of the revised Modified Bridges are accurate, according to the changes proposed in this chapter, additional data would need to be collected using the revised Modified Bridges and revised item weights.

Conclusions

In this chapter, we assessed the Modified Bridges' ability to distinguish between low-/moderate-risk mothers and high-risk mothers by evaluating the Modified Bridges' content validity and score reliability. According to our factor analytic review, the content that the Modified Bridges assesses does not closely correspond with the original subscale (i.e., Medical, Psychosocial, or Demographics/Basic Needs) but is more closely aligned with psychosocial problems, availability of resources, and health care awareness. In addition, we found the Modified Bridges' reliability, estimated using the raw item weights, to be relatively low (0.65). In practical terms, this means that there is a reasonable degree of uncertainty associated with the risk scores that mothers receive following their interviews with the hospital liaisons.

Given the relatively low reliability of the Modified Bridges, we conducted a hypothetical re-weighting of certain items and removed others, in an effort to improve the reliability of the Modified Bridges (i.e., to improve its ability to distinguish between high- and low-risk mothers). The potential improvements in reliability were large enough to warrant further consideration. Notably, in high-stakes testing environments (e.g., standardized educational assessments), reliabilities of 0.70 are considered the minimum acceptable value, which the weighting revisions reported here achieve, although, prior to these changes, the Modified Bridges' total scale reliability failed to reach this criterion. Further, with these revisions, the raw reliability (0.73) is closer in magnitude to the standardized value (0.75); this is an improvement from the disparity between the original raw (0.65) and standardized (0.74) reliabilities. In other words, if these revisions were adopted, mothers' scores on the Modified Bridges would be more precise and First 5 LA would have greater confidence that the Modified Bridges is able to correctly identify at-risk mothers.

Chapter Five. Conclusion

The purpose of this report is to assess the psychometric properties of the Modified Bridges. This report provides results relating to both the reliability and the validity of the Modified Bridges and indicates the ability of the hospital liaisons to accurately and precisely score the risk level of new mothers. Below, we arrange the results found in this report according to their respective research questions:

1. Do hospital liaisons assign similar scores when using the Modified Bridges? In general, our evaluation of the Modified Bridges found strong evidence that hospital liaisons provide similar risk assessments. The overall inter-rater reliability of the Modified Bridges is high (Kendall's $W = 0.83$). The highest-performing hospital liaisons are those with more experience and those who refer to the Modified Bridges protocol more frequently.

2. Does the Modified Bridges identify a similar level of risk among the same women as other measures of psychosocial and/or medical risk? We found limited evidence that the Modified Bridges can identify a similar level of risk as two risk instruments sponsored by the Florida Department of Health. The Modified Bridges total scores were only moderately related to both Florida risk measures. In particular, at the standard risk cutoff of 50, the Modified Bridges has low sensitivity and high specificity in detecting prenatal and postnatal risk compared to the Florida risk thresholds. Higher Modified Bridges cutoff scores (e.g., 75 to 90) would align better with Florida prenatal and infant risk scales and would improve sensitivity; however, such a shift would reduce the number of mothers eligible for supportive services.

3. Does the Modified Bridges accurately distinguish between low-/moderate-risk mothers and high-risk mothers? Finally, the extent to which the Modified Bridges can accurately distinguish between risk levels was lower than desired. Using a reliability based on the item weights, the Modified Bridges had a suboptimal score reliability (0.65). In particular, mother's use of English, mother's age, and mother's highest education level are the items that most contribute to lower reliability for the Modified Bridges. Our revisions suggest increasing the weight of seven items, decreasing the weight of three items, and removing two items: infant medical problems, which is redundant with mother's worry about infant's health, and mother's use of English, which does not distinguish between maternal risk levels. Adopting these changes would increase the score reliability of the Modified Bridges from 0.65 to 0.73, an increase that would achieve the minimum desired reliability of 0.70.

51

Appendix A. Legacy Instrument Protocol

In this appendix, we provide the protocol for the legacy instrument verbatim, unedited.

Purpose

The legacy measure risk questions were developed for a state-wide program in Florida with similar goals as Welcome Baby—to identify families at greatest risk and need and to link those families to supportive services. The legacy measure questions assess prenatal and infant risk factors that are similar to those assessed by the Modified Bridges. The legacy items are administered in order to provide a comparison with the Modified Bridges. At the close of the study, RAND Corporation will compare the risk scores produced by the legacy instrument to those provided by the Modified Bridges, and will use this information to better understand how the risk is categorized by the Modified Bridges.

Administering the Legacy Instrument

The 14 legacy measure risk questions are very similar to those on the Modified Bridges and should only be asked *after* you have completed the Bridges. In some instances the responses to the legacy questions may be informed by questions on the Bridges; however, you should ask the legacy question if you are at all uncertain. In other instances, new content is being asked and will require you to initiate a discussion with the mother. Unlike the Bridges, the legacy measure questions were *not* developed to be asked by an interviewer. Instead, they were developed to appear on a paper-and-pencil survey that the mother would complete herself. Because of this, *hospital liaisons should, when possible, ask the question directly as it appears on the questionnaire*.

Administering the Legacy Instrument

You may find that easiest way to complete and upload the instrument is to use a paper-and-pencil version to score the questions during the interview, which you would then refer to when uploading the responses electronically.

You are required to upload *both* the Modified Bridges questions and the legacy questions.

Uploading the Legacy Instrument to the Web

    Using the link provided to you, navigate the electronic survey website. Once there, you will see the following screens and will be guided through uploading the Modified Bridges and legacy instrument responses:

F5LA Legacy Measure Survey
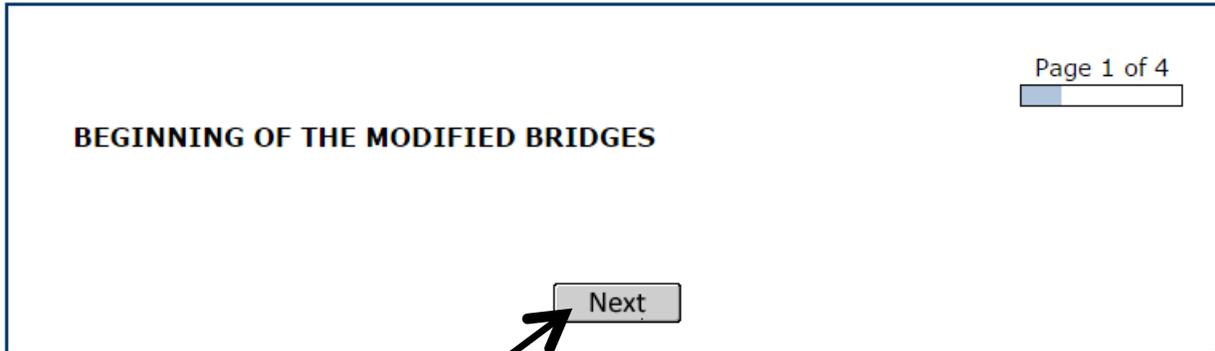
Page 1 of 4

BEGINNING OF THE MODIFIED BRIDGES

Next

1. Click "Next" to begin the Bridges

F5LA Legacy Measure Survey
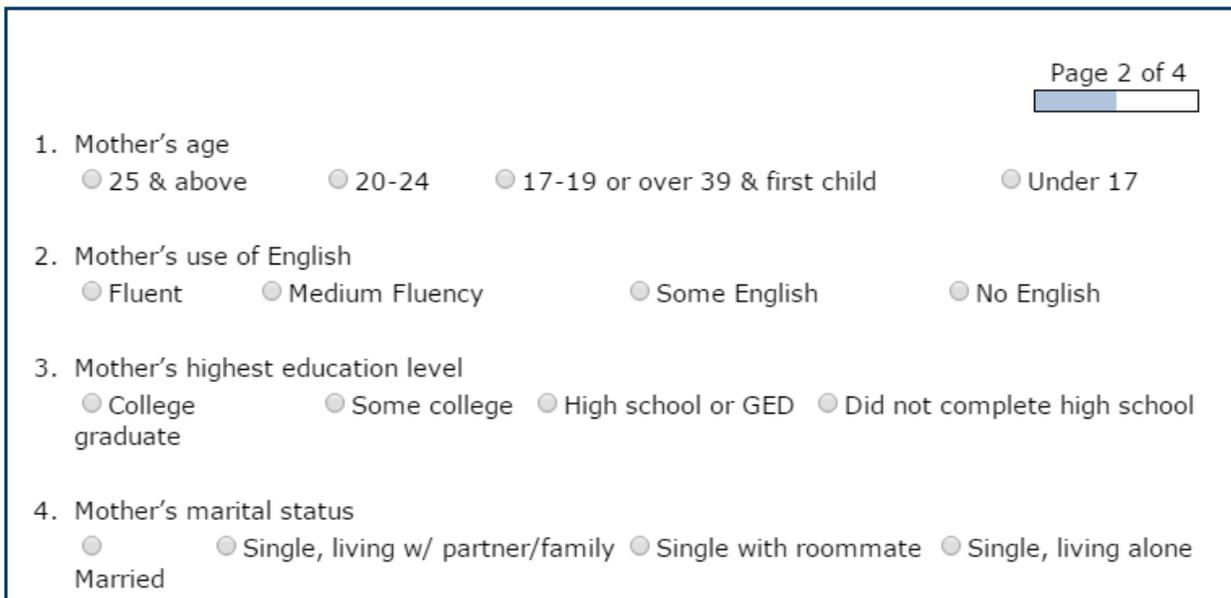
Page 2 of 4

1. Mother's age
   ○ 25 & above   ○ 20-24   ○ 17-19 or over 39 & first child   ○ Under 17

2. Mother's use of English
   ○ Fluent   ○ Medium Fluency   ○ Some English   ○ No English

3. Mother's highest education level
   ○ College   ○ Some college   ○ High school or GED   ○ Did not complete high school
   graduate

4. Mother's marital status
   ○   ○ Single, living w/ partner/family   ○ Single with roommate   ○ Single, living alone
   Married

2. Enter the Modified Bridges responses

25. History of child abuse/neglect

Include parents' own childhood history of abuse or any non-spousal assault

○ No known history of child abuse/neglect

○ No low level for this category

○ Prior protective services provided to siblings, with that episode resolved and case closed; mother retains custody of siblings; parent of child was a victim of childhood abuse.

○ Pending child abuse/ neglect investigation; previous abuse/neglect; prior court action, e.g., siblings removed from home; or child abuse suspected or discussed but no system intervention to date

26. History of excessive alcohol or other drug use by people who will impact infant's wellbeing

○ No History

○ Some history but not currently using any drugs/alcohol

○ Mother/Father/Significant Other is receiving drug/alcohol treatment, remains in program, and is considered compliant

○ Mother/Father/Significant Other not in drug/alcohol treatment program; individual is in program but attendance is sporadic; (mother) entered program late in pregnancy

[Back] [Save] [Next]

3. Finish entering the responses and click "Save" (you may have to wait a few seconds)

4. To upload the Bridges, click "Next" (you may have to wait a few seconds)

F5LA Legacy Measure Survey

Page 3 of 4

BEGINNING Of THE LEGACY MEASURE

[Back] [Save] [Next]

5. Click "Next" to begin the Legacy

27. Maternal age is less than 18
- ○ Yes
- ○ No

28. In the last month, how many alcoholic drinks did you have per week?
- ○ Did not drink
- ○ At least one drink

29. Since you have been pregnant, how many cigarettes did you smoke a day (a pack has 20 cigarettes)?
- ○ Did not smoke   ○ 1-10   ○ 11-20   ○ 20+

---

6. Begin entering the responses

40. Is the mother's principal source of insurance Medical?
- ○ Yes
- ○ No

41. Hospital Liaison Name:*

42. Date (Month/day/year):

43. Hospital name:
--None--

[Back] [Save] [Done]

---

7. Finish entering the responses and click "Save" (you may have to wait a few seconds)

8. To upload the Legacy, click "DONE" (you may have to wait a few seconds)

Note that you *are required* to provide your name.

Do not close the window until you see this screen:

F5LA Legacy Measure Survey

Completed

Thank you for taking the Legacy Measure Survey!

Enter Next Response

You have successfully completed the upload. Clicking "Enter Next Response" will open a new survey for another mother.

Legacy Instrument Protocol and Response Instructions

27. Maternal age is less than 18

   Yes

   No

   NOTE: You may have obtained information for this question from Modified Bridges question 1

28. In the last month, how many alcoholic drinks did you have per week?

   Did not drink = Mother most have had not even one drink

   At least one drink = Even having half a drink counts as having one drink

   NOTE: You may have obtained information for this question from Modified Bridges question 26

29. Since you have been pregnant, how many cigarettes did you smoke a day (a pack has 20 cigarettes)?

   Did not smoke = Mother did not smoke at all during pregnancy

   1–10 = Select if mother smoked even just one cigarette

11–20 = More than a half a pack, but less than a full pack per day

20+ = At least a pack of cigarettes a day

NOTE: You may have obtained information for this question from Modified Bridges question 13

30. Thinking back to just before you got pregnant, did you want to be . . .

Pregnant = The mother planned to become pregnant.

Pregnant later = The mother wanted to become pregnant soon (within the next 1–2 years), just not right now.

Not pregnant = The mother did not want to become pregnant and did not have a timeframe in mind when she wanted to become pregnant

31. Is this your first pregnancy?

Yes = This is the mother's first time being pregnant

No = The mother has been pregnant before

32. Have any of the following happened to you?

Had a baby that was not born alive (does not include abortions or miscarriages)

NOTE: *Only* endorse this item if the mother has had a stillbirth

Had a baby born 3 weeks or more before due date *(stillbirth or live birth)*

Had a baby that weighed less than 5 pounds, 8 ounces

None of the above = If the mother has never been pregnant before, or if none of the above have happened, select "None of the above"

NOTE: Select all that apply

This might require asking each question individually.

33. Before being pregnant, what was the mother's weight (in pounds)?

Enter weight number in lbs (e.g., "130")

NOTE: If the mother can't recall, they should provide their best guess. Note this information might be available from the patient chart.

34. What is the mother's current height (in inches)?

Inches equivalences are included in the selection box below.

Select the height in feet or total inches

NOTE: Note this information might be available from the patient chart.

35. Thinking about the time between your last pregnancy and your most recent pregnancy. Was the time interval between pregnancies less than 18 months?

(If this is the mother's first pregnancy, mark "Not pregnant before".)

Yes = The mother has been pregnant at least once before in the past 18 months (not including current pregnancy)

No = The mother has not been pregnant in the past 18 months *(but was pregnant before then)*

Not applicable ("Not pregnant before) = The mother has never before been pregnant

36. Has your baby had any of the following services or conditions: Assisted Ventilation (30 min. or more), Assisted Ventilation (6 hrs. or more), NICU [neonatal intensive care unit]

admission, newborn given Surfactant Replacement Therapy, Hyaline Membrane Disease/RDS [respiratory distress syndrome], or seizure or serious neurological dysfunction?

Yes = Only select 'yes' if one of the above conditions apply

No = If the mother reports other less severe conditions, but none of the conditions listed above, select "no."

37. At birth, did your baby weigh less than 4 pounds, 7 ounces?

Yes

No

NOTE: This information should be available from the patient chart or the baby card. Confirm with the mother.

38. Was your baby transferred to a different facility for additional medical care within 24 hours of delivery?

Yes

No

NOTE: This might apply in situations in which the baby is being transferred for medical interventions (e.g., cardiac, neurological, or GI [gastrointestinal] surgeries).

39. Does the mother know the father's name?

Yes

No

NOTE: The purpose of this question is to understand if the mother knows who is the child's father. Mothers may be uncomfortable telling you the name of father. What is most important is that the mother knows who fathered the baby.

If the mother won't tell you the father's name, but she knows who the father is, select "Yes."

40. Is the mother's principal source of insurance Medical?

Yes

No

41. Hospital Liaison name_____

42. Date (month/day/year)_____

43. Hospital name _____

After you have entered the responses for the Modified Bridges and Legacy Measure, click "DONE."

Frequently Asked Questions

Q1: How do I find the link to the website?

The link to the survey's website can be accessed in the email that you have received from RAND. You can also add the link to your "favorite websites" and you can access the website directly from your browser.

Q2: Do I need to have a password or username?

You do not need to have a password or username to access the website.

Q3: What happens if I have to close the website before I've finished entering the responses?

If you did not click the **SAVE** button to save your answers in each section, you will need to enter the answers again.

Q4: How long can I take to enter a mother's responses?

The window will close after approximately 4-hours of non-use

Q5: Can I upload more than one mother at a time?

Yes, you can, but it is preferable not to do it.

Q6: Can I skip questions?

You can skip questions but you are encouraged to make every attempt to enter a response. The only question that is required is that you enter your name; the survey cannot be uploaded with the name field missing.

Q7: Can I go back . . . that is, can I change my answers if I've already selected the *NEXT* button?

Yes, you can. Push the **Back** button to change your answers.

Q8: What do I click to submit the responses?

To submit the responses, you will need to click the **DONE** button.

Q9: How do I know my responses have been submitted?

The survey is completed when you will receive the screen "Completed and thank you message"

Q10: What does the final screen mean that says "*ENTER NEXT RESPONSE*"?

By clicking this button, you will be taken to a new window to enter the responses for the next mother.

Q11: Who should I contact for technical issues or any additional questions?

Brian Stucky: bstucky@rand.org, Gabriela Castro: gabriela@rand.org

Appendix B. Modified Bridges Frequencies and Descriptive Statistics from the Validity Study

Table B.1. Modified Bridges Frequencies and Descriptive Statistics from the Validity Study

Modified Bridges Item	Count	Percentage	Mean	SD	Range
1. Mother's age			1.97	2.80	0–12
25 and above	93	60.78			
20–24	45	29.41			
17–19 or over 39 and first child	12	7.84			
Under 17	2	1.31			
2. Mother's use of English			0.61	1.10	0–3
Fluent	112	73.2			
Medium fluency	5	3.27			
Some English	15	9.8			
No English	19	12.42			
3. Mother's highest education level			6.04	2.88	0–9
College graduate	10	6.54			
Some college	38	24.84			
High school or equivalent	43	28.1			
Did not complete high school	60	39.22			
4. Mother's marital status			1.73	1.69	0–6
Married	52	33.99			
Single, living with partner or family	84	54.9			
Single with roommate	3	1.96			
Single, living alone	14	9.15			
5. Annual household income level			7.46	2.42	0–9
>$60,000	6	3.92			
$60,000–36,000	12	7.84			
$36,000–18,000	36	23.53			
<$18,000 or mother unaware of income	98	64.05			
6. Mother's health coverage			0.48	0.89	0–3
Full	118	77.12			
Temporary	32	20.92			
None	3	1.96			

Modified Bridges Item	Count	Percentage	Mean	SD	Range
7. Adequate food in the house			1.37	2.09	0–9
Consistent	100	65.36			
Adequate	37	24.18			
Inconsistent	15	9.8			
Chronically inadequate	1	0.65			
8. Mother's demonstrated awareness of available resources			1.08	0.90	0–3
Full awareness	48	31.37			
Moderate awareness	50	32.68			
Some awareness	46	30.07			
No awareness	7	4.58			
9. Transportation a barrier			0.74	1.03	0–3
Never	91	59.48			
Rarely	22	14.38			
Sometimes	26	16.99			
Frequently	13	8.5			
10. Level of family support			1.29	1.57	0–4
Family or partner is supportive, available and committed to help	83	54.25			
Inconsistent or limited family or partner support	40	26.14			
Family supportive but not in geographic area; some support from friends and neighbors; limited community services available	29	18.95			
11. Other children at home			1.21	0.85	0–3
None under age 5; singleton birth	35	22.88			
One or two under age 5; singleton birth	54	35.29			
First-time mother or zero to two under age 5 with twin or multiple birth	55	35.95			
Three or more under age 5 with singleton, or multiples, or other child removed from home or deceased	6	3.92			
12. Current housing conditions			2.05	3.54	0–18
Stable and safe	108	70.59			
Adequate	37	24.18			
Rents a motel, garage, or portion of a living space; frequent migration; staying with friends	6	3.92			
Currently homeless or in temporary shelter or car	1	0.65			
13. Tobacco smoke in home			0.43	0.89	0–3
No smoking	120	78.43			
Visitors smoke	4	2.61			
Household members smoke	20	13.07			
Mother smokes	7	4.58			

Modified Bridges Item	Count	Percentage	Mean	SD	Range
14. Adequate and timely PNC			0.81	2.23	0–12
Accessed PNC within 3 months; consistent with follow-up appointments	128	83.66			
Accessed PNC between 3 and 6 months; consistent with follow-up appointments	14	9.15			
Did not access PNC until after 6 months or inconsistent with follow-up appointments	5	3.27			
No PNC	2	1.31			
15. Infant health coverage			0.12	0.45	0–2
Full	142	92.81			
Temporary	8	5.23			
16. Source of medical care for infant			0.34	0.67	0–3
Regular pediatrician or community clinic	110	71.9			
Attending pediatrician	31	20.26			
Emergency room	4	2.61			
No plan for future care	4	2.61			
17. Mother's intent to remain current with well-baby care and immunizations for her infant			0.10	0.44	0–3
Very strong	143	93.46			
Moderately strong	4	2.61			
Weak	2	1.31			
Very weak or no understanding	2	1.31			
18. Infant medical problems			4.17	6.86	0–18
No apparent medical or physical problems	104	67.97			
Minor medical or physical problems that do not significantly affect infant's vital functions or physical and intellectual development	13	8.5			
Medical or physical problems that moderately affect infant's vital functions or physical and intellectual development	10	6.54			
Any preterm infant or physical or medical problem that significantly impacts vital life functions or physical and intellectual development	24	15.69			
19. Mother's worry about infant's health			0.51	0.84	0–3
Not worried	98	64.05			
Minor worries	35	22.88			
Moderate worries	7	4.58			
Very worried	9	5.88			

Modified Bridges Item	Count	Percentage	Mean	SD	Range
20. Infant feeding issues			3.02	2.25	0–9
Mother demonstrates knowledge, confidence regarding infant feeding; sufficient resources available to support a healthy feeding relationship with infant	38	24.84			
Mother requires some education regarding infant feeding; adequate resources available to support a healthy feeding relationship	79	51.63			
Great uncertainty or lack of experience or knowledge with infant feeding or limited resources available to support a healthy feeding relationship	31	20.26			
Grossly insufficient knowledge regarding infant feeding; lack of interest in improving feeding skills; evidence of or high potential for poor feeding relationship with infant	4	2.61			
21. Strength of maternal bond with infant			1.43	3.24	0–15
Strong	121	79.08			
Moderate	17	11.11			
Weak	10	6.54			
None	2	1.31			
22. Mother's medical problem			3.72	4.32	0–18
No limitations; fully able to care for child	76	49.67			
Mild limitations that might affect ability to care for child	57	37.25			
Moderate limitations that could significantly affect ability to care for child	15	9.8			
Severe limitations; mother is unlikely to be able to care for child	2	1.31			
23. Mental health problem (mother; include PPD)			3.63	6.07	0–18
No challenges; realistic expectations of child; no history of mental health issues	103	67.32			
Mild challenges that might impact ability to care for child	22	14.38			
Moderate challenges that could significantly affect ability to care for child	11	7.19			
Severe challenges; mother might have difficulty in caring for child	16	10.46			
24. History of domestic violence			2.64	5.79	0–18
No known history of domestic violence	123	80.39			
Partner currently in treatment for domestic violence	15	9.8			
Domestic violence investigation; previous domestic violence of serious nature; prior court action	12	7.84			
25. History of child abuse or neglect			2.03	5.18	0–18
No known history of child abuse or neglect	128	83.66			
No low level for this category	2	1.31			
Prior protective services provided to siblings, with that episode resolved and case closed; mother retains custody of siblings; parent of child was a victim of childhood abuse	12	7.84			
Pending child abuse or neglect investigation; previous abuse or neglect; prior court action, e.g., siblings removed from home; or child abuse suspected or discussed but no system intervention to date	9	5.88			

Modified Bridges Item	Count	Percentage	Mean	SD	Range
26. History of excessive alcohol or other drug use by people who will impact infant's well-being			2.17	5.01	0–18
No history	121	79.08			
Some history but not currently using any drugs or alcohol	18	11.76			
Mother, father, or significant other is receiving drug or alcohol treatment, remains in program, and is considered compliant	2	1.31			
Mother, father, or significant other not in drug or alcohol treatment program; individual is in program but attendance is sporadic; (mother) entered program late in pregnancy	11	7.19			
Total score			52.15	28.63	0–161
Medical subscale			7.51	5.36	0–30
Psychosocial subscale			13.13	16.59	0–74
Demographics/Basic Needs subscale			25.12	10.41	0–57

NOTE: Some item percentages do not equal to 100 due to missing observations

References

Altman, Douglas G., and J. Martin Bland, "Diagnostic Tests 1: Sensitivity and Specificity," *British Medical Journal*, Vol. 308, No. 6943, June 11, 1994, p. 1552.

Browne, Michael W., and Robert Cudeck, "Bootstrapping Goodness-of-Fit Measures in Structural Equation Models," in Kenneth A. Bollen and J. Scott Long, eds., *Testing Structural Equation Models*, Newbury Park, Calif.: Sage Publications, 1993, pp. 111–135.

Cohen, Jacob, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, Vol. 20, 1960, pp. 37–46.

Crocker, Linda M., and James Algina, *Introduction to Classical and Modern Test Theory*, New York: Holt, Rinehart, and Winston, 1986.

Cronbach, Lee J., "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, Vol. 16, No. 3, September 1951, pp. 297–334.

Fleiss, Joseph L., *Statistical Methods for Rates and Proportions*, 2nd ed., New York: Wiley, 1981.

Florida Department of Health, "Healthy Start," undated (a). As of January 4, 2017: http://www.floridahealth.gov/programs-and-services/childrens-health/healthy-start/

———, "Help Your Baby Have a Healthy Start in Life!" undated (b). As of January 5, 2017: http://www.floridahealth.gov/programs-and-services/childrens-health/healthy-start/_documents/english-prenatal-screen.pdf

———, "Infant Risk Screen," undated (c). As of January 5, 2017: http://www.floridahealth.gov/programs-and-services/childrens-health/healthy-start/_documents/final-dh-3135-infant-risk-screen-01-15%20English.pdf

Furr, R. Michael, *Scale Construction and Psychometrics for Social and Personality Psychology*, Los Angeles, Calif.: SAGE, 2011.

Glass, Gene V., and Kenneth D. Hopkins, *Statistical Methods in Education and Psychology*, 3rd ed., Boston, Mass.: Allyn and Bacon, 1996.

Gwet, Kilem Li, *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*, 2nd ed., Gaithersburg, Md.: Advanced Analytics, 2010.

Hu, Li-tze, and Peter M. Bentler, "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, Vol. 6, No. 1, 1999, pp. 1–55.

Irwin, Debra E., Brian Stucky, Michelle M. Langer, David Thissen, Esi Morgan DeWitt, Jin-Shei Lai, James W. Varni, Karin Yeatts, and Darrin A. DeWalt, "An Item Response Analysis of the Pediatric PROMIS Anxiety and Depressive Symptoms Scales," *Quality of Life Research*, Vol. 19, No. 4, May 2010, pp. 595–607.

Kendall, Maurice G., *Rank Correlation Methods*, London: C. Griffin, 1948.

Kendall, M. G., and B. Babington Smith, "The Problem of *m* Rankings," *Annals of Mathematical Statistics*, Vol. 10, No. 3, 1939, pp. 275–287.

Landis, J. Richard, and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, Vol. 33, No. 1, March 1977, pp. 159–174.

Muthén, Linda K., and Bengt O. Muthén, *Mplus User's Guide*, 7th ed., Los Angeles, Calif.: Muthén and Muthén, September 2012. As of January 4, 2017: https://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf

Stucky, Brian D., and Maria Orlando Edelen, "Using Hierarchical IRT Models to Create Unidimensional Measures from Multidimensional Data," in Steven Paul Reise and Dennis A. Revicki, eds., *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment*, Part II: *Classic and Emerging IRT Modeling Approaches*, New York: Routledge, Taylor and Francis Group, 2015, pp. 183–206.

Stucky, Brian D., Nisha C. Gottfredson, and A. T. Panter, "Item-Level Factor Analysis," in Harris Cooper, Paul M. Camic, Debra L. Long, A. T. Panter, David Rindskopf, and Kenneth J. Sher, eds., *APA Handbook of Research Methods in Psychology*, Vol. 1: *Foundations, Planning, Measures and Psychometrics*, Washington, D.C.: American Psychological Association, 2012, pp. 683–697.

Stucky, Brian D., Nisha C. Gottfredson, A. T. Panter, Charles E. Daye, Walter R. Allen, and Linda F. Wightman, "An Item Factor Analysis and Item Response Theory–Based Revision of the Everyday Discrimination Scale," *Cultural Diversity and Ethnic Minority Psychology*, Vol. 17, No. 2, 2011, pp. 175–185.

Swets, John A., *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*, Mahwah, N.J.: L. Erlbaum Associates, 1996.