Nursing Research • September/October 2005 • Vol 54, No 5, 313–323

A Psychometric Assessment of the H & H Lactation Scale in a Sample of Thai Mothers Using a Repeated Measurement Design

Busakorn Punthmatharith ▾ Jagdip Singh

- ***Background:*** The H & H Lactation (H & H) scale is a useful instrument for understanding mothers' perceptions of insufficient milk supply, which is the common reason given for premature cessation of breastfeeding. This instrument may be used to intervene in postbirth dropoff in breastfeeding among mothers.
- ***Objectives:*** To examine construct and nomological validity and related psychometric evidence for the translated H & H scale using a repeated measurement design in a sample of Thai mothers.
- ***Method:*** A repeated measurement design with 196 mother-infant dyads. Data were collected at two time periods. At Time 1, data on the H & H and Index of Breastfeeding Status (IBS) scales were obtained between 36 and 48 hr postbirth. For Time 2, the same mothers completed the same instruments at 1 month postbirth. Psychometric properties of the H & H scale were examined using convergent, discriminant, and nomological validity principles within structural equation modeling.
- ***Results:*** Based on judgment/classification task, exploratory/confirmatory factor analysis, and cross-validation using multisample analysis, only nine items of the H & H scale measured at Time 1 and Time 2 were found to be psychometrically sound for measuring insufficient milk supply for Thai mothers. An analysis of the influence of the nine-item of the H & H scale measured at Time 1 and Time 2 on the IBS measured at Time 2 showed that (a) the direct effects of two subscales of the H & H scale measured at Time 2 on the IBS were significant and positive, providing support for the nomological validity of the H & H scale; (b) each H & H factor at Time 1 had a significant influence on its corresponding factor at Time 2, supporting convergent validity; and (c) the crossover effects for each H & H factor for Time 1 and Time 2 data were nonsignificant, indicating discriminant validity.
- ***Discussion:*** Overall, the 20-item H & H scale appears to be unreliable and invalid for measuring the concept of insufficient milk supply in cross-cultural studies. However, the short form of nine items is reliable and valid for Thai mothers. Further testing and development of the H & H scale and the proposed short form in other cultures and contexts is warranted.

- ***Key Words:*** H & H Lactation scale · psychometric assessment · repeated measures · Thai mothers

In the developing world, premature cessation of breastfeeding—cessation of breastfeeding before the infant is 4 months of age due to nonmedical reasons—negatively affects maternal and infant health (Neifert, 1996; Stuart-Macadam, 1995). During breastfeeding, oxytocin is released, reducing maternal blood loss (Lawrence, 1999; Stuart-Macadam), and prolactin levels are elevated, resulting in the onset of milk production (Neifert; Stuart-Macadam). Infant benefits include nutritional, immunological, developmental, and psychological effects (Lawrence), and breastfeeding reduces infant morbidity and mortality worldwide (Armstrong, 1987; Cunningham, Jelliffe, & Jelliffe, 1991). Yet, the proportion of new mothers who rely exclusively on breastfeeding rarely exceeds 50% in the developing countries where breastmilk is usually the only source of nourishment for infants. A major reason given by mothers for premature cessation of breastfeeding is their perception of insufficient milk supply (Durongdej, 1998; Hill, 1992; Hill & Aldag, 1993; Hill & Humenick, 1996). Insufficient milk supply is defined as a state in which a mother has or perceives that she has an inadequate supply of breastmilk to either satiate and/or support adequate weight gain for the infant (Hill & Humenick, 1989). Hill and Humenick have identified several indicators of this state including maternal confidence, maternal satisfaction, and infant satisfaction. Insufficient milk supply leads

Busakorn Punthmatharith, PhD, MSN, MSc, RN, *is an Instructor, Faculty of Nursing, Prince of Songkla University, Songkhla, Thailand.*

Jagdip Singh, PhD, *is a Professor of Marketing, Weatherhead School of Management, Case Western Reserve University, Cleveland, Ohio.*

mothers to feel unsuccessful at breastfeeding and mothering, and thus, to premature cessation of breastfeeding.

To combat the premature cessation of breastfeeding, national and international agencies around the world have implemented intervention programs. For instance, in Thailand, the Baby Friendly Hospital Initiative was outlined to empower all mothers with normal deliveries to practice a 10-step exclusive breastfeeding program with the goal that these infants would be fed exclusively with breastmilk from birth to 4 months of age by at least 30% of mothers by the end of 2006 (Health Promotion Office, 2002). Although the 10-step program has been implemented for all mothers who delivered in Baby Friendly Hospitals since 1991, the rate of exclusive breastfeeding at 4 months stood at only 16.3% by 2001 (Health Promotion Office, 2001). Similarly, in the United States, the *Healthy People 2010* goal for breastfeeding is to increase to at least 75% the proportion of mothers who breastfeed their infants in the early postpartum period and increase to at least 50% the proportion of mothers who continue breastfeeding until their infants are 5 to 6 months old (Hill, 2000). The 1998 Ross Mothers' Survey indicated that 64% of mothers reported breastfeeding during the early postpartum period and 29% at 6 months (Hill, 2000).

Current focus has shifted to the dropoff in breastfeeding rates between postpartum and when the child is 5 to 6 months old. A scientific and meaningful study of dropoff in breastfeeding rests on the availability of sensitive, reliable, and valid instruments for measuring insufficient milk supply. Hill and Humenick (1996) provided the only validated instrument available—referred to as the H & H Lactation (H & H) scale—based on their own conceptual framework for understanding insufficient milk supply (Hill & Humenick, 1989). Unfortunately, the psychometric properties of this scale remain little understood especially in a non-US context. No research was found on adapting the H & H scale in Thai or other populations outside of the United States, nor has any research further examined the psychometric properties of this scale, such as its construct and nomological validity. Using the scale developed for, tested on, and adjusted to fit the norms of one culture may not reflect the scale's ability to be translated or applied to different cultures. Because validity has been viewed as a continuing and accumulative process (Rew, Stuppy, & Becker, 1988), the use of the H & H scale by nurses or nursing researchers to overcome the dropoff in breastfeeding without further validity evidence may foster invalid research and misguided practices.

In this paper, the aim is to provide construct and nomological validity and related psychometric evidence for the H & H scale in a study of Thai mothers using a repeated measurement design. Nomological validity refers to "the empirical support for the relationships between the construct in question and other constructs or variables that are expected based on underlying theory" (Pedhazur & Schmelkin, 1991, p. 72). Specifically, several psychometric properties of a Thai version of the translated H & H scale have been examined including its (a) multidimensionality relative to the original scale developed in the U.S. context, (b) convergent and discriminant validity, (c) invariance of factor structure across time, and (d) nomological validity in predicting self-reported index of breastfeeding status (IBS). These data were collected as part of a repeated measurement design for a randomized controlled trial of the effects of early kangaroo (skin-to-skin) care on maternal feelings, maternal–infant interaction, and breastfeeding success in Thailand (Punthmatharith, 2001).

The H & H Lactation Scale: Conceptual Foundation and Psychometric Evidence

Hill and Humenick (1989) provided the conceptual foundation for the H & H scale based on their extensive review of the literature relating to the influence factors and etiologies of insufficient milk supply (Allen & Pelto, 1985; Feinstein, Berkelhamer, Gruszka, Wong, & Carey, 1986; Hawkins, Nichols, & Tanner, 1987; Quandt, 1985; Verronen, 1982). Biological, psychological, behavioral, and social etiologies are involved in motivating or impeding breastfeeding behavior (Allen & Pelto, 1985; Hill & Humenick, 1989; Quandt, 1985). Based on these etiologies, several factors contributing to early insufficient milk supply and premature cessation of breastfeeding have been identified including milk maturation rate, early weight gain rate, infant satisfaction, maternal satisfaction and confidence, maternal relaxation, and formula supplementation (Hill & Humenick). Drawing from this literature, Hill and Humenick proposed a framework to capture the potential determinants (e.g., direct and indirect factors) on milk production and indicators of the insufficient milk supply. Using the insufficient milk supply framework, Hill and Humenick focused on these indirect and direct factors that may influence the mother's perception of insufficient milk supply. The H & H scale was used to measure indicators of maternal confidence, maternal satisfaction, and perceived infant satisfaction. From earlier work with fullterm breastfeeding mothers (Dusdieker, Booth, Seals, & Ekwo, 1985; Hill & Aldag, 1993), Hill and Humenick (1996) developed an initial set of 30 items and tested them with two sets of participants: breastfeeding mothers with low-birthweight (LBW) infants (n = 110) and term infants (n = 120). After exploratory factor analysis (EFA), Hill and Humenick refined the H & H scale to 20 items that were posited to measure three distinct factors: (a) Maternal Confidence/Commitment to Breastfeeding (MC) (10 items), (b) Perceived Infant Breastfeeding Satiety (IS) (5 items), and (c) Maternal–Infant Breastfeeding Satisfaction (MS) (5 items). The first three factors in the exploratory analysis accounted for 55.9%, 17.5%, and 8.9% of the variance, respectively, indicating that these factors represent substantive and meaningful aspects of the insufficient milk supply syndrome. To further substantiate these factors, separate EFA was performed with the 20 items in both the original LBW and term samples. The results showed that the original analysis was supported only in the mothers of LBW infants. For the term mothers, all items of the second and third factors performed well, whereas only half the items loading on the first factor performed well.

Other valid (i.e., content, conceptual, concurrent, and predictive) and reliability assessments were reported also by Hill and Humenick (1996). Content validity of the 20-item instrument was established by three experts in the area of lactation. For conceptual validity, the correlation of the Maternal Breastfeeding Evaluation Scale (a similar breastfeeding tool) with sustained breastfeeding was similar to those obtained using the H & H scale (r = .30 to .48). In the LBW and fullterm samples, concurrent validity was supported for the H & H scale with the level of breastfeeding recommended by Labbok and Krasovec (1990). In the LBW sample, the correlation between the level of breastfeeding at Week 8 postpartum and the overall H & H scale was .84, the correlation between the three subscales was .85, .49, and .49. In the fullterm sample, correlation between the level of breastfeeding at Week 6 postpartum and the overall H & H scale was .62, and for the three subscales was .66, .48, and .53. For predictive validity, in the LBW sample, correlation between the MC subscale (of H & H scale) at Week 1 and the level of breastfeeding at Week 8 postpartum was .34. In the fullterm sample, the corresponding correlation was .42. Lastly, the alpha reliability was .96 and .92 for the LBW and fullterm samples, respectively. Few other studies have utilized the complete H & H scale or reported its psychometric properties. For instance, to measure continued commitment, a study used the MS subscale only (Humenick, Hill, & Wilhelm, 1997).

Despite the important contribution of the H & H scale in understanding and diagnosing premature cessation of breastfeeding, validity of the scale has not been fully established or confirmed. Three gaps are noteworthy: (a) validation evidence in non-US samples (Brislin, 1970; Phillip, Hernandez, & Ardon, 1994; Singh, 1995); (b) confirmatory and structural equation modeling procedures to provide stronger psychometric evidence especially in a longitudinal context (DeVellis, 2003; Nunnally & Bernstein, 1994; Pedhazur & Schmelkin, 1991); and (c) development of short forms that balance parsimony and psychometrics. Developing a validated short form appears essential to fully utilize the H & H scale (Punthmatharith, 2001).

Sample

The sampling frame was composed of all eligible breastfeeding healthy mother–fullterm infant dyads at a hospital in southern Thailand. The hospital is a major provincial, 700-bed teaching hospital and one of the Baby Friendly Hospitals where most mothers are encouraged to begin breastfeeding their infants within 30 min after birth for approximately 15–30 min. After providing written informed consent, 196 eligible mother/infant dyads were randomly assigned into two groups by computerized minimization technique (Zeller, Good, Anderson, & Zeller, 1997) between January and August 2000. There were 99 dyads in the control group and 97 in the kangaroo care group who were followed and data collected at two time periods. Specifically, the mothers completed the H & H and IBS scales, between 36 and 48 hr postbirth (Time 1 or T1) in the hospital. For the second data collection, the same mothers completed the H & H and IBS instruments in the privacy of their home 1 month postbirth (Time 2 or T2, n = 170). At both time periods, mothers were asked to respond to the questions based on their current thinking and behaviors. No statistically significant differences at the .05 level were found between the two groups for age, income, occupation, previous breastfeeding experience, infant birthweight, and gender (of infant) (Table 1). Therefore, these two groups were combined together for the psychometric analysis.

Instrument

Insufficiency of Milk Supply Insufficiency of milk supply was measured by the H & H scale (Hill & Humenick, 1996). The H & H scale, a self-report questionnaire, consists of 20 positively or negatively worded items with three subscales. All items are measured on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). For this study, the H & H scale was translated into the Thai language using the iterative translation–backtranslation method wherein the original English and the back-translated English versions were compared for inconsistencies, and the Thai version refined to eliminate such inconsistencies (Brislin, 1970; Jones, 1987; Phillips et al., 1994). Three Thai-born bilingual registered nurse instructors who were enrolled in a graduate program in the United States translated the original English scale into the Thai language independently. Thereafter, the three Thai versions of the instruments were carefully compared and checked for within-version discrepancies as well as for consistency with the original instrument by the lead researcher. Discrepant items were discussed in detail by the lead researcher in a group meeting with the three translators, and revised. The revised Thai version of the instrument was unanimously

TABLE 1. Sample Characteristics

	Control (n = 99)		EKC (n = 97)				
Variables	n	%	n	%	df	χ^2	p
Maternal age (years)					2	0.95	.62
<20	12	12.1	15	15.5			
20–30	65	65.7	65	67.0			
>30	22	22.2	17	17.5			
Income (Baht per month, 42 baht = US$1)					1	0.02	.88
<5,000	52	52.5	52	53.6			
≥5,000	47	47.5	45	46.4			
Occupation					1	0.16	.69
Home	65	65.7	61	62.9			
Outside	34	34.3	36	37.1			
Previous breastfeeding experience					1	0.04	.85
Yes	62	62.6	62	63.9			
No	37	37.4	35	36.1			
Infant birthweight (g)					2	0.85	.66
<2,500	7	7.1	4	4.1			
2,500–3,000	32	32.3	31	32.0			
>3,000	60	60.6	62	63.9			
Infant gender					1	0.08	.78
Male	44	44.4	45	46.4			
Female	55	55.6	52	53.6			

Note. EKC = Early Kangaroo Care.

approved. Next, the revised Thai version was back-translated into English by another bilingual registered nurse who was enrolled also in a U.S. graduate program, and blind to the original English version. Thereafter, an American editor compared the equivalence of the two English versions—the original version and the back-translated version. The Editor identified points of inconsistencies, which were used by the lead researcher to make a final round of revisions in the Thai version. This iterative method was expected to result in equivalence between the original and the backtranslation instruments. After the institutional review board approved the study, additional checks for face validity of the Thai items were performed by three individuals who had expertise in obstetric and pediatric nursing and were familiar with the Thai culture. These experts reviewed and analyzed all items for appropriateness for use in Thai culture. Thereafter, pretesting the Thai instrument with 10 postpartum Thai mothers revealed that most Thai mothers did not experience difficulty in understanding the questions, and were able to easily provide responses. Consequently, the Thai version was administered to the mothers ($N = 196$) in the study. After reversing negatively scored items, the ratings of all items were entered for further measurement analysis.

Level of Breastfeeding The level of breastfeeding was measured by the IBS (Labbok & Krasovec, 1990). The IBS is a single-item, self-report indicator that consists of three major categories of breastfeeding: full (exclusive, almost exclusive), partial (high, medium, low), and token. These categories are ranked from 6 (*for exclusive breastfeeding*) to 1 (*for token*). Face and content validity were established jointly by a group of internationally recognized experts on breastfeeding (Anderson, 1995). Predictive validity was documented by Piwoz, De Kanashiro, De Romana, Black, and Brown (1996). No reports of reliability of this instrument have been found. In this study, the scale was translated using the backtranslation method as noted previously (Brislin, 1970; Jones, 1987; Phillips et al., 1994). The level of breastfeeding status was slightly modified, after testing for clarity and comprehensiveness with 10 postpartum Thai mothers.

Data Analysis Procedures

Testing of psychometric properties of the H & H scale and subscales followed procedures outlined by Bentler (1995) and DeVellis (2003). Exploratory factor analysis and confirmatory factor analysis (CFA) were used to examine the basic psychometric properties of the 20-item H & H scale. The purpose of this analysis was to test the factor structure of the 20 items as proposed by Hill and Humenick (1996). A judgment and classification task was conducted wherein (a) multiple judges (a panel of two Thai and two American nurse experts; all doctoral candidates in Nursing and related areas) were to read each H & H item and classified it into one of three categories corresponding to MC, IS, and MS; (b) interjudge reliability is estimated for each H & H item based on consistency of classification in the judgment task; and (c) items classified in categories consistent with H & H specification were retained for analysis. This judgment and classification task ensured that only items that corresponded conceptually to the underlying dimensions were selected and that empirical regularities such as in a series of EFAs were avoided. The selected items were subjected to EFA and CFA to ensure a consistent factor structure. Cross-validation using multisample analysis was performed to assess whether the factor structure and psychometric evidence obtained for the H & H scales measured at T1 was validated with T2 data. Because high correlation between MC and MS was expected, these dimensions were tested for redundancy and discriminant validity (Singh, 1991). Finally, to take advantage of the longitudinal nature of the data, a nomological model was estimated to examine the affect of IS, MC, and MS on the level of breastfeeding measured by the IBS at T2 (Figure 3).

Results

Exploratory and Confirmatory Factor Analysis of 20 H & H Items

From an EFA of the H & H scale HHT1 ($n = 196$) (Table 2), three factors were extracted that explained about 42.88% of the total variance and were in accord with the underlying theory (Hill & Humenick, 1996). The Maximum Likelihood and Promax methods for extraction and rotation, respectively, were utilized. Based on a report by Hair, Anderson, Tatham, and Black (1998), it was reasoned that an item had a strong and clean loading if: (a) its loading is equal or greater than .30 and (b) the absolute difference between the dominant and the next highest loading was less than .20. Moreover, the item loading must be consistent with theory. For example, items H1 to H10 correspond to MC, while items H11 to H15 correspond to IS, and finally, items H16 to H20 were designed to measure MS. Only 50% of the 20 H & H items meet these criteria (Table 2). Specifically, 15 of the 20 items had a dominant loading that exceeded .30 (corresponding to a communality of 9%) and, of these 15 items, another 5 items did not depict a clean structure where the dominant loading exceeded the next highest loading by .20. For instance, items H18 loaded .36 and .41 on Factors 1 and 2, indicating that its measurement was confounded by two underlying factors. When these two criteria are conjoined with the need for theoretical consistency, only 10 of the 20 items appeared reasonable (Table 2). Overall, these results suggest that the proposed three-dimensional structure does not fit our data on the H & H scale.

A CFA with each item constrained to load on a single factor consistent with theoretical expectations was conducted. This analysis examined statistical significance of the violations observed in the EFA. The results from the Elliptical Reweighted Least Squares (ERLS) estimation procedure in EQS indicate that the overall goodness of fit χ^2 was 253.3 with df of 167, which indicated that the model did not fit the data ($p < .001$). In supporting evidence, the other fit indices were below desirable standards. For instance, the relative fit indices, NFI (Bentler–Bonett Normed Fit Index), CFI (Comparative Fit Index), and AGFI (LISREL Adjusted Goodness of Fit Index) were .85; .94 and .82, respectively—and below the .95 and .90 criteria for CFI and NFI/AGFI for good fitting models (Hair et al., 1998; Tabachnick & Fidel, 2001). In addition,

TABLE 2. Exploratory Factor Analysis of the 20-item H & H Scale (Pattern Matrix and Factor Correlation Matrix)

Pattern Matrix

	Factor		
Items	1 (MS)	2 (MC)	3 (IS)
H1	−.121	.460[a]	.109
H2		.494[a]	
H3		.236	.281
H4	.231	.285	
H5	.226	.420	
H6		.461[a]	−.230
H7	.146	.628[a]	
H8	.292		.181
H9	.210	.369	
H10		.294	
H11	.585		
H12		−.162	.722[a]
H13	−.249	.337	.505
H14			.327[a]
H15		−.117	.824[a]
H16	.319	.310	
H17	.503[a]	.132	
H18	.355	.410	
H19	.587[a]	.102	
H20	.889[a]	−.192	

Factor Correlation Matrix

Factor	1	2	3
1	1.000	.631	.290
2	.631	1.000	.334
3	.290	.334	1.000

Note. Extraction method: Maximum Likelihood; rotation method: Promax with Kaiser normalization; MC = Maternal Confidence/Commitment Breastfeeding subscale = Items H1–H10; IS = Perceived Infant Breastfeeding Satiety subscale = Items H11–H15; MS = Maternal–Infant Breastfeeding Satisfaction subscale = Items H16–H20.
[a]Appear reasonable when two criteria are conjoined.

Standardized Root Mean Squared Residual (SRMR) was .10, again above the expected value of .05 for good fitting models (Hair et al., 1998). Thus, it appeared that the three-dimensional model proposed by H & H scale was an inadequate representation of the H & H scale for the current data both from statistical and substantive standpoints.

Judgment and Classification Task

The judgment and classification task was implemented to identify meaningful correspondence between the H & H items and underlying dimensions. The 20 items were randomly ordered and presented to a panel of two Thai and two American nurse experts; all doctoral candidates in either Maternal–Child Nursing or Community Health Nursing, and related areas. Each expert was requested to independently place each statement into one of three categories, and rate each statement on a 4-point scale for how closely the statement reflects the idea presented by the category definition. In case the statement could not be categorized into one of three categories, the experts were asked to place "NC" (nonapplicable) for that statement including any comments regarding the difficulty experienced in categorizing the NC statement.

The index of interjudge reliability (Fleiss, 1971), interrater reliability (Perreault & Leigh, 1989), and average intensity (summation of rating scores from all raters/numbers of raters) for each item was estimated (Table 3). Overall, a high level of agreement was obtained (κ and *Ir*) for placing an item in the appropriate content domain category for five items (Items 3, 5, 6, 12, and 13), a moderate level for nine items (Items 2, 8, 9, 11, 15, 16, 17, 18, and 19), and a low level of agreement for only one item (Item 7). The rest of them were misplaced (Items 1, 4, 10, 14, and 20). The overall average intensity of rating for correctly placing each item varied from 2.25 to 3.50. Based on this analysis, items that indicated a high or moderate level of agreement were retained for further analysis.

Exploratory and Confirmatory Factor Analysis of Selected H & H Items

The 14 items representing moderate to high level of interjudge agreement were analyzed using EFA. First, the data were examined for multivariate outliers by estimating Mahalanobis distances for each observation and comparing with a critical value of 32 at a .05 level of significance (DeCarlo, 1997). Because multivariate outliers are likely to bias EFA results, observations with Mahalanobis distance values greater than 32 were deleted resulting in 189 remaining cases for T1 and 161 cases for T2. These data were then submitted to EFA using Maximum Likelihood and Promax methods for extraction and rotation, respectively. In agreement with theory, three factors were extracted, and items were examined to ensure that they satisfied the dominant loading (>.30) and cross-loading spread (>.20) criteria (Hair et al., 1998). Five items failed to meet these criteria (i.e., 2, 3, 8, 11, and 18). The remaining nine items were reanalyzed using EFA procedures. Only one item (19) failed to meet one of three criteria, whereas the remaining eight items satisfied all desired criteria. Item 19 loaded along with Items 16 and 17 (as per theory) and yielded a dominant loading exceeding .30 (Table 4). However, it also cross-loaded on another factor and the spread between cross-loadings was less than .20. This item was retained for use in the next step of the analysis.

The MC factor (Items 5, 6, and 9) correlated positively and strongly with the MS factor (Items 16, 17, and 19) with a correlation of .58 (Table 4); however, both MC and MS have relatively lower correlations at .22 and .27, respectively, with the IS factor (Items 12, 13, and 15). This suggested that MC and MS may have in common a higher, second-order construct representing a maternal factor. A two-factor EFA provided a clean and clear factor structure (see Table 5). The interfactor correlation was only .30, which indicated a high level of discriminant validity between the maternal and infant factors.

A CFA with three correlated factors with the nine H & H items was conducted (Figure 1). The hypothesized model of three correlated factors fit the T1 data for nine items reasonably well. Based on the ERLS estimation, the chi-square goodness-of-fit statistic was nonsignificant (χ^2 = 32.2, df = 24, p = .12), which indicated that the discrepancies between the model and data were small (Hair et al., 1998). In addition, the relative and absolute fit statistics exceeded values expected for good fitting models: NFI = .93; NNFI (Nonnormed Fit Index) = .97; CFI = .98; GFI (Goodness of Fit Index) = .96; AGFI = .92; SRMR = .06 with a 90% confidence interval for RMSEA = .00, .08 (Hair et al., 1998; Tabachnick & Fidel, 2001).

All hypothesized loadings are statistically significant ($t \geq 1.96$, $p < .05$) and substantively large (>.30, ranging from .49 to .92) except for item H6 that equaled .28 (Hair et al., 1998). The correlations among three factors for the nine items at T1 were low to high (varying from .22 to .70), especially between MS and MC which showed the highest correlation (equal to .70). To more closely examine the discriminant validity between MS and MC, a redundancy analysis was performed wherein a constrained model was estimated by forcing the correlation between the factors to unity (Singh, 1991). This constrained model was compared with an unconstrained model using a chi-square difference test such that a significant chi-square supported nonredundancy between the factors. Using this procedure, a difference of 12.5 (44.7–32.2) with df = 1 (25–24) was obtained at $p < .001$. This implied that the operational indicators for MS and MC did not converge to a single construct (nonredundancy) and that the discriminant validity between these two factors was empirically tenable.

Cross-Validation Using Data

To cross-validate the obtained results, the model of the T1 data (n = 189) was fitted (Figure 2) to the corresponding T2 data (n = 161). Based on the overall goodness-of-fit statistics (Hair et al., 1998; Tabachnick & Fidel, 2001), the hypothesized model provided acceptable fit to the data (χ^2 = 72.5, df = 60, p = .13; NFI = .90; NNFI = .98;

TABLE 3. Index of Interjudge Reliability, Estimated Reliability, and Average Intensity of Each Item of the 20-item H & H Lactation Scale

	Categories		
	1 = MC	2 = IS	3 = MS
Items (H)	κ/Ir/AVI	κ/Ir/AVI	κ/Ir/AVI
1. I feel breastfeeding is providing my baby with an ideal food.			.25/.50/2.5
2. I made the right decision when I decided I would breastfeed my baby.	.63/.79/3.0		
*3. Even though I can breastfeed I would rather not be breastfeeding.	1.0/1.0/3.5		
4. Breastfeeding is a special way to console my baby.			1.0/1.0/3.5
5. My baby would only get a bottle if I am not available for breastfeeding.	1.0/1.0/2.75		
6. I believe I can solve any breastfeeding problems which come along.	1.0/1.0/3.0		
7. I feel a sense of pride from watching my baby grow from my breastmilk.	.25/.50/3.0		
*8. I am so upset about breastfeeding problems that I become upset at the thought of breastfeeding.	.63/.79/2.33		
9. I arrange my life so that breastmilk is almost the only thing my baby gets.	.63/.79/3.0		
10. Overall, I would describe breastfeeding as a relaxing activity.			.63/.79/2.67
11. My baby was satisfied with the amount of breastmilk received.		.63/.79/3.0	
*12. My baby would be hungry if I did not use formula along with breastfeeding.		1.0/1.0/3.25	
*13. I believe that following breastfeeding with a bottle is how to find out if baby got enough.		1.0/1.0/2.25	
*14. I would describe my baby as being fussy after breastfeeding.			.63/.79/3.0
*15. I feel I have to give formula after breastfeeding to satisfy my baby.		.63/.79/3.0	
16. In general, I believe my baby was satisfied with breastfeeding.			.63/.79/3.0
17. In general, I was satisfied with breastfeeding.			.63/.79/3.0
18. I became more relaxed as I sat and breastfed.			.63/.79/3.33
19. My baby appeared to enjoy breastfeeding.			.63/.79/3.0
20. In general, I feel successful at breastfeeding my baby.	.63/.79/2.67		

Note. MC = Maternal Confidence/Commitment Breastfeeding subscale = Items 1–10; IS = Perceived Infant Breastfeeding Satiety subscale = Items 11–15; MS = Maternal–Infant Breastfeeding Satisfaction subscale = Items 16–20; κ = kappa; Ir = Estimated reliability; AVI = average intensity; *scoring reversed.

TABLE 4. Exploratory Factor Analysis of Selected H & H Items (Pattern Matrix and Factor Correlation Matrix)

Pattern Matrix

	Factor		
Items	1 (IS)	2 (MS)	3 (MC)
H5			.517
H6	−.186		.330
H9	.110		.662
H12	.746		−.111
H13	.476	−.126	.292
H15	.793		
H16		.732	.176
H17		.849	
H19		.375	.303

Factor Correlation Matrix

Factor	1	2	3
1	1.000	.273	.219
2	.273	1.000	.576
3	.219	.576	1.000

Note. Extraction method: Principal Axis Factoring; rotation method: Promax with Kaiser normalization; MC = Maternal Confidence/Commitment Breastfeeding subscale = Items H5, H6, H9; IS = Perceived Infant Breastfeeding Satiety subscale = Items H12, H13, H15; MS = Maternal–Infant Breastfeeding Satisfaction subscale = Items H16, H17, H19.

CFI = .98; GFI = .96; AGFI = .93; SRMR = .08; 90% CI of RMSEA = .00, .04). To provide a stronger test of validity, factor loadings and covariances were constrained for the corresponding items and factors of T1 and T2 data. The Lagrange Multiplier Test for releasing constraints (multivariate test statistics) showed that there was no statistically significant difference in the estimated parameters between the two groups at a .05 level. The evidence of identical parameters and covariance matrices in T1 and T2 data gave support to the proposed underlying structure of the H & H scale and confirmed that the measurement structure for the nine H & H items was stable and consistent across time.

Structural Analysis of the Influence of T1 and T2 Data on IBS

A model was developed to account for repeated measurements to examine the structural influence of MC, IS, and MS at T1 on their subsequent values at T2 (MCt2, ISt2, and MSt2) and on the level of breastfeeding measured by the IBS at T2 (Figure 3). In accord with time-dependent phenomenon, the T2 variables of MCt2, ISt2, and MSt2 were hypothesized to mediate the influence of T1 variables on IBS. This model was estimated using EQS after allowing for appropriate constraints for repeated measurement data. Overall, the fit for the model (Figure 3) to the data was acceptable (χ^2 = 143.88, df = 141, p = .42; NFI = .83; NNFI = 1.00; CFI = 1.00; GFI = .91; AGFI = .88; SRMR = .06; 90% CI of RMSEA = .00, .04) (Hair et al., 1998; Tabachnick & Fidel, 2001). When considering the standardized solution, all loadings were significant ($t \geq 1.96$, $p < .05$) and large (>.3, ranging from .30 to .91) for the nine items at T1 and T2 except for Item H6 measured at T1. In addition, the direct effects of T1 variables (MCt1, ISt1, and MSt1) on IBS were nonsignificant (all $p > .10$), whereas their direct effects on the corresponding T2 variables (MCt2, ISt2, and MSt2) achieved significance and provided support for the hypothesized mediation effect. This suggested that the posited model was useful for understanding the time dependent effects of the H & H scale. The estimated coefficients are displayed in Figure 3.

Discussion

Several aspects of these results are noteworthy. First, each of the H & H dimensions at T2 had a significant influence on the breastfeeding status. Specifically, the MC and IS had a positive and significant effect on the breastfeeding status of the infant (B = .24 and .22 respectively, $p < .01$). The mother's satisfaction with infant breastfeeding was negatively associated with the breastfeeding status, although the magnitude of this effect was relatively marginal (B = −.08). Whether this effect represents complacency on

TABLE 5. Exploratory Factor Analysis of Selected H & H Items Using a Two-Factor Solution (Pattern Matrix and Factor Correlation Matrix)

Pattern Matrix

	Factor	
Items	1	2
H5	.505	
H6	.362	−.194
H9	.477	
H12	−.104	.759
H13	.126	.446
H15		.810
H16	.817	
H17	.637	
H19	.621	

Factor Correlation Matrix

Factor	1	2
1	1.000	.302
2	.302	1.000

Note. Extraction method: Principal Axis Factoring; rotation method: Promax with Kaiser normalization.

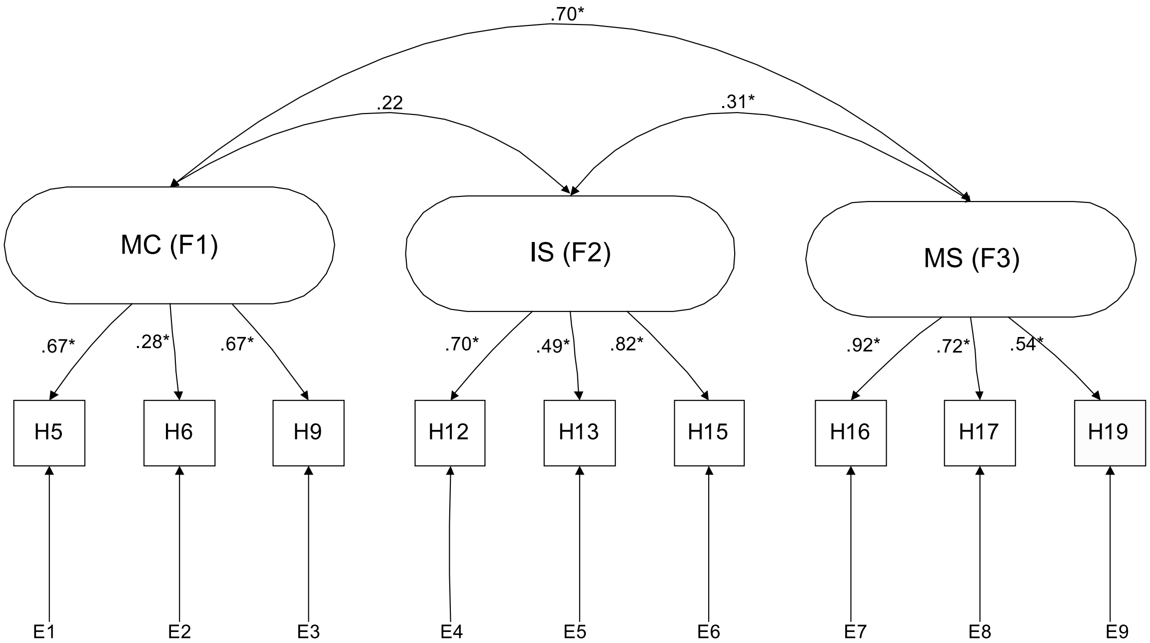

FIGURE 1. Estimated loadings and factor intercorrelations from CFA of the nine-item HHT1 based on the ERLS estimation procedure in EQS. Note: MC = Maternal Confidence/Commitment Breastfeeding subscale; IS = Perceived Infant Breastfeeding Satiety subscale; MS = Maternal–Infant Breastfeeding Satisfaction subscale. Latent factors are indicated by F, and measurement error by E. Individual items from the H & H scale are represented by H5 to H19. Straight arrows indicate factor loadings and double-headed curved arrows represent intercorrelations. $^*p < .05$.

the part of the mothers or a statistical artifact of the collinear relationship between MC and MS is difficult to disentangle. Second, convergent validity was evident because each H & H factor at T1 had significant influence on its corresponding factor at T2. For instance, the maternal confidence factor at T1 had a coefficient of .99 ($p < .01$) for its effect on maternal confidence at T2. Likewise, similar evidence was available for IS and MS (B = .71 and .32, respectively). Third, there was a strong evidence of discriminant validity because the crossover effect for each H & H factor was marginal and nonsignificant, without exception. For instance, the effects of MCt1 on ISt2 and

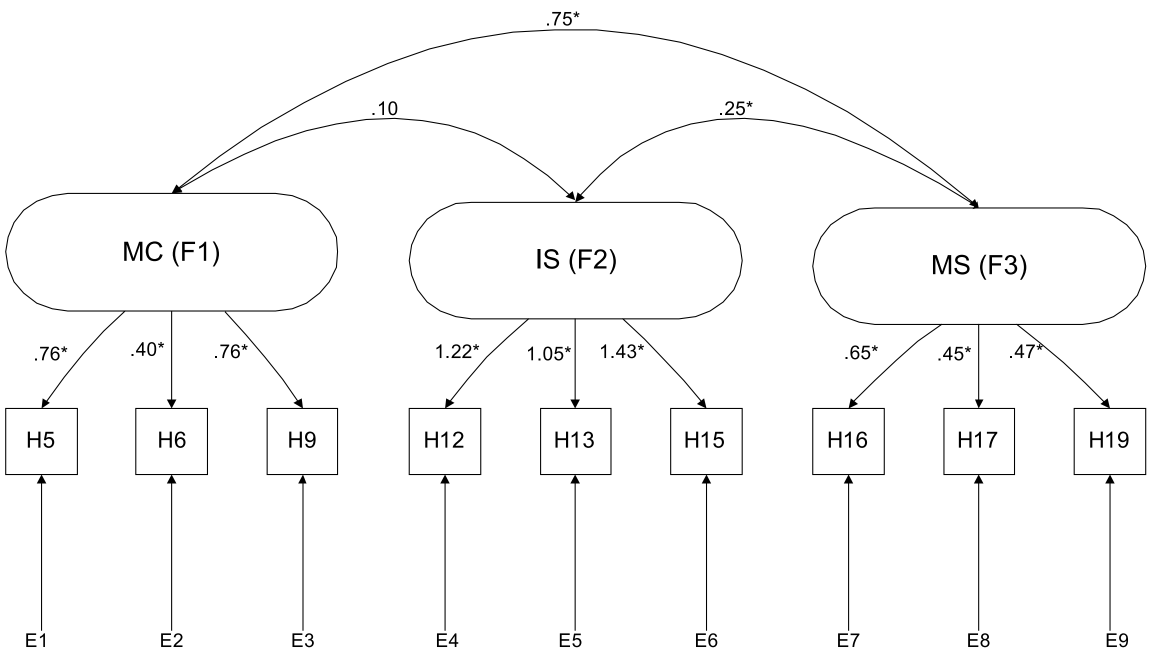

FIGURE 2. Estimated loadings and factor intercorrelations from a multisample CFA (cross-validation) of the nine-item HHT1 and the nine-item HHT2 using EQS. Note: MC = Maternal Confidence/Commitment Breastfeeding subscale; IS = Perceived Infant Breastfeeding Satiety subscale; MS = Maternal–Infant Breastfeeding Satisfaction subscale. Latent factors are indicated by F, and measurement error by E. Individual items from the H & H scale are represented by H5 to H19. Straight arrows indicate factor loadings and curved arrows represent intercorrelations. $^*p < .05$.

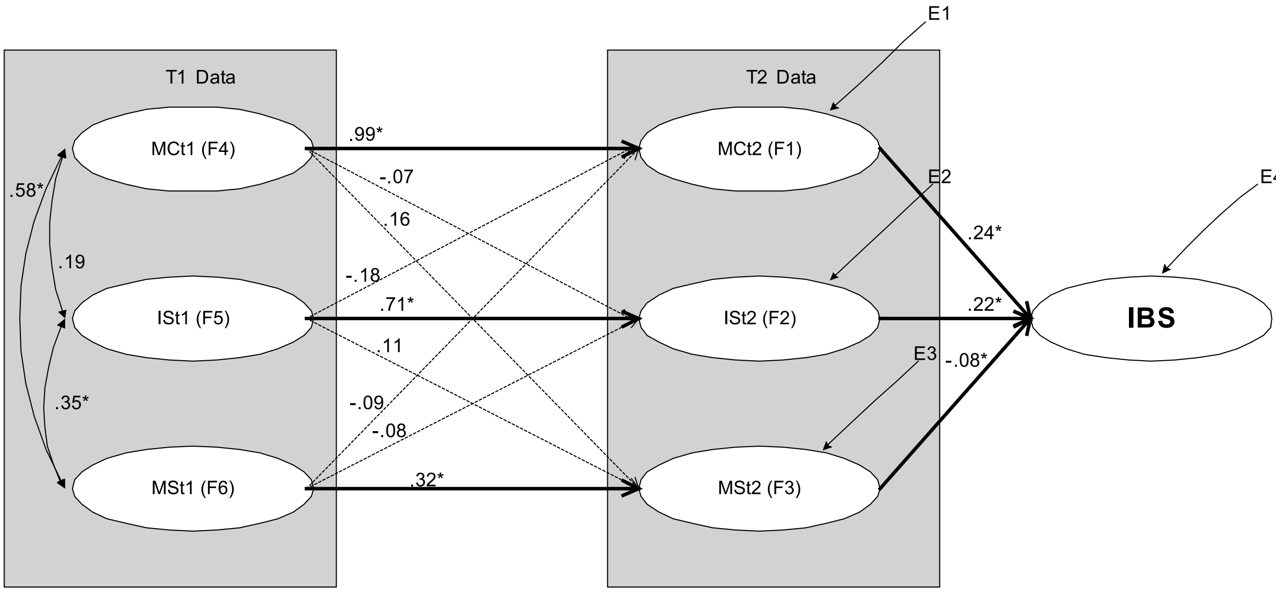

FIGURE 3. Estimated coefficients from a structural equation model involving the influence of the nine-item HHT1 (Time 1) and HHT2 (Time 2) on IBS (Time 2) using EQS. Note. MC = Maternal Confidence/Commitment Breastfeeding; IS = Perceived Infant Breastfeeding Satiety; MS = Maternal–Infant Breastfeeding Satisfaction; IBS = Index of Breastfeeding Status measured at Time 2. Measures obtained in the first time period are indexed by T1 or t1, and those obtained in the second time period by T2 or t2. Latent factors are indicated by F, and structural error in equations by E. Significant paths at $p = .05$ level are in bold. $^*p < .05$.

MSt2 were both marginal and nonsignificant ($B = -.07$ and .16, respectively, $p > .05$). Likewise, the corresponding crossover effects for ISt1 ($B = .18$ and .11, $p > .05$) and for MSt1 ($B = -.09$ and $-.08$, $p > .05$) followed a similar pattern of nonsignificance. Combined with the evidence of significant convergent validity, the nonsignificance of crossover effects suggested that each H & H dimension was distinct. Finally, in agreement with the theoretical structure of H & H, the intercorrelations among the three dimensions were positive and significant, with one exception. The correlation between ISt1 and MSt1 had an intercorrelation of .35 ($p < .05$). The only exception was the intercorrelation between MCt1 and ISt1 that was positive but only marginally significant ($B = .19$, $p < .10$).

Nine H & H items, identified by the current study, demonstrated sound, consistent, and acceptable psychometric properties for use with Thai women. Using both EFA and CFA procedures, confirming support was not found for the original set of items (Hill & Humenick, 1996). Only 50% of the original set of items qualified. The judgment and classification task identified promising items, which were subjected to EFA and CFA analyses. This process narrowed the potential items to nine, with three items each for the three dimensions. Further analysis provided clear and compelling evidence in support of the nine H & H items, although one of the items was relatively weak (Item 19). Cross-validation yielded identical parameters and covariance matrices for a nine-item scale. The use of the original 20-item H & H scale in cultures other than the United States is not without risks and its use is not warranted unless psychometric evidence for the culture in question is available. In the case of Thailand, the authors are confident in recommending the nine-item H & H scale instead of the original 20-item scale. For Thai mothers, all nine items appear easy to understand, and respond to in a relatively short period (5–10 min). Although more evidence is needed, the results indicate that the proposed nine-item scale may be used in other contexts, including the United States.

These results also confirm the multidimensionality of the lactation scale. Questions have been raised about the lack of discriminant validity between the MC and MS dimensions. Both dimensions pertain to material disposition toward breastfeeding—how satisfied they are with it and how committed they feel toward breastfeeding. Although satisfaction and commitment are related concepts, the evidence from this study supports their discriminant validity. For both time periods, the correlations among the three dimensions meet standard criteria for discriminant validity. Structural analysis provided additional evidence of discriminant validity. Specifically, each H & H factor at T1 had significant influence *only* on its corresponding factor at T2, and the crossover effects for each H & H factor are marginal and nonsignificant. Finally, each H & H factor at T2 has a significant and independent influence on the infant breastfeeding status. Yet, the intercorrelation between the MC and MS dimensions exceeds .70, suggesting that further improvements in their discrimination are required.

The usefulness of the H & H scale depends on its potential to predict infant breastfeeding status over time. The maternal commitment and infant satiety dimensions of the lactation scale independently and significantly influence the infant breastfeeding status at any given point in time. Moreover, whether the mother was committed to, and confident about breastfeeding at the time of delivery, is less important to postpartum breastfeeding status. What is important is how the mothers feel about breastfeeding after they leave the hospital. Hence, if the goal is to stem the dropoff in breastfeeding rates, it is critical to monitor and enhance mother's commitment to, and confidence in, continued breastfeeding. Likewise, providing continued information to mothers to help them accurately evaluate their infant's satiety with breastmilk appears essential to maintaining breastfeeding rates. Continued monitoring of perceived infant satiety would be needed for this purpose. The nine-item H & H scale can offer a sound foundation for such monitoring and maintenance programs.

Several limitations of the current work should be considered for a balanced assessment of its contributions. First, because the study is based on a sample of Thai women from one hospital in Southern Thailand, the generalizability of our findings is constrained by the sampling design. Second, these results are based on data collected over two time periods that are separated by 1 month. Although the repeated measurement design allows testing for the consistency and stability of psychometric properties and mitigates the problem with common method variance from cross-sectional survey data, future testing with longitudinal designs is needed. Third, the translation of instruments may confound the results. The iterative translation–backtranslation procedures recommended by cross-cultural researchers were utilized to address this concern. The CFA, EFA, and multisample factor analysis results confirm that the nine-item H & H scale is not confounded because they measure one and only one dimension, and this measurement is consistent across time. More work is needed to rule out cultural and translation factors in the psychometric property of H & H items.

Much of the work on the H & H scale is based in the U.S. context. Culturally grounded studies in non-U.S. contexts may reveal concerns that are not captured in current H & H scale dimensions or items. The essential test for a viable and valid measure of insufficient milk supply phenomenon is in its ability to predict breastfeeding status. Without this quality, the power to change nursing practice and influence mother's choice to address premature cessation of breastfeeding is diminished. For nursing research and practice, the effectiveness of efforts to understand and combat premature cessation of breastfeeding rests on this line of inquiry.

Accepted for publication February 1, 2005.

Partial funding from Sigma Theta Tau International and the Frances Payne Bolton School of Nursing Alumni Association.

The authors thank Mary Anthony, Associate Professor of Nursing, Case Western Reserve University for detailed comments and suggestions on earlier drafts. The authors also acknowledge the critical and helpful comments of two anonymous reviewers.

Corresponding author: Busakorn Punthmatharith, PhD, MSN, MSc, RN, Faculty of Nursing, Prince of Songkla University, Songkhla, Thailand (e-mail: busakorn.p@psu.ac.th).

References

Allen, L., & Pelto, G. (1985). Research on determinant of breastfeeding duration: Biocultural studies. *Medical Anthropology*, *9*, 97–105.

Anderson, G. C. (1995). *Self-regulatory preterm infant care: Adaptation postbirth*. Grant application, Case Western Reserve University, Cleveland, Ohio. (funded by the National Institute of Nursing Research, NIH, 1996–2000).

Armstrong, H. C. (1987). Breastfeeding low birthweight babies: Advances in Kenya. *Journal of Human Lactation*, *3*(2), 34–37.

Bentler, P. M. (1995). *EQS: Structural equations program manual*. Los Angeles, CA: BMDP Statistical Software.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216.

Cunningham, A. S., Jelliffe, D. B., & Jelliffe, E. F. (1991). Breastfeeding and health in the 1980s: A global epidemiological review. *Journal of Pediatrics*, *118*, 1–8.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*, 292–307.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Durongdej, S. (1998). *Evaluation of the sustainability of the Baby Friendly Hospital Initiatives and its impact upon breastfeeding practices in urban communities*. Bangkok, Thailand: Department of Nutrition Faculty of Public Health, Mahidol University.

Dusdieker, L., Booth, B., Seals, B., & Ekwo, E. (1985). Investigation of a model for the initiation of breastfeeding in primigravidae women. *Social Science and Medicine*, *20*, 695–703.

Feinstein, J., Berkelhamer, J., Gruszka, M., Wong, C., & Carey, A. (1986). Factors related to early termination of breast-feeding in an urban population. *Pediatrics*, *78*, 210–215.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*, 378–382.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Hawkins, L., Nichols, F., & Tanner, J. (1987). Predictors of duration of breastfeeding in low-income women. *Birth*, *14*, 204–209.

Health Promotion Office, Department of Health, Ministry of Public Health. (2001). *Health promotion evaluation in the 8th National Economic and Social Development Plan*. Bangkok, Thailand: Ministry of Public Health.

Health Promotion Office, Department of Health, Ministry of Public Health. (2002). *Health promotion indicators in the 9th National Economic and Social Development Plan (Public Health, 2002–2006)*. Bangkok, Thailand: Ministry of Public Health.

Hill, P. D. (1992). Insufficient milk supply syndrome. *NAACOG's Clinical Issues in Perinatal and Women's Health Nursing*, *3*, 605–612.

Hill, P. D. (2000). Update on breastfeeding: Healthy people 2010 objectives. *American Journal of Maternal Child Nursing (MCN)*, *25*, 248–251.

Hill, P. D., & Aldag, J. C. (1993). Insufficient milk supply among black and white breast-feeding mothers. *Research in Nursing & Health*, *16*, 203–211.

Hill, P. D., & Humenick, S. S. (1989). Insufficient milk supply. *IMAGE: Journal of Nursing Scholarship*, *21*(3), 145–148.

Hill, P. D., & Humenick, S. S. (1996). Development of the H & H lactation scale. *Nursing Research*, *45*, 136–140.

Humenick, S. S., Hill, P. D., & Wilhelm, S. (1997). Postnatal factors encouraging sustained breastfeeding among primiparas and multiparas. *The Journal of Perinatal Education*, *6*(3), 33–45.

Jones, E. (1987). Translation of quantitative measures for use in cross-cultural research. *Nursing Research*, *36*, 324–327.

Labbok, M., & Krasovec, K. (1990). Toward consistency in breastfeeding definitions. *Studies in Family Planning*, *21*(4), 226–230.

Lawrence, R. A. (1999). *Breastfeeding: A guide for the medical profession* (5th ed.). St. Louis, MO: Mosby.

Neifert, M. (1996). Early assessment of the breastfeeding infant. *Contemporary Pediatrics*, *13*(10), 142–166.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York, NY: McGraw-Hill.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Perreault, W. D., & Leigh, L. E. (1989). Reliability of nominal data based on qualitative judgements. *Journal of Marketing Research*, *26*, 135–148.

Phillips, L. R., Hernandez, I. L. D., & Ardon, E. T. D. (1994). Focus on psychometrics: Strategies for achieving cultural equivalence. *Research in Nursing & Health*, *17*, 149–154.

Piwoz, E. G., De Kanashiro, H. C., De Romano, G. L., Black, R., & Brown, K. H. (1996). Feeding practice and growth among low-income Peruvian infants: A comparison of internationally-recommended definitions. *International Journal of Epidemiology*, *25*, 103–106.

Punthmatharith, B. (2001). *Randomized controlled trial of early kangaroo (skin-to-skin) care: Effects on maternal feelings, maternal–infant interaction and breastfeeding success in Thailand*. Unpublished doctoral dissertation, Case Western Reserve University, Cleveland, Ohio.

Quandt, S. (1985). Biological and behavioral predictors of exclusive breastfeeding duration. *Medical Anthropology*, *9*, 139–151.

Rew, L., Stuppy, D., & Becker, H. (1988). Construct validity in instrument development: A vital link between nursing practice, research, and theory. *Advances in Nursing Science*, *10*, 10–22.

Singh, J. (1991). Redundancy in constructs: Problem, assessment, and an illustrative example. *Journal of Business Research*, *22*, 255–280.

Singh, J. (1995). Measurement issues in cross-national research. *Journal of International Business Studies, third quarter*, 597–619.

Stuart-Macadam, P. (1995). Biocultural perspectives on breastfeeding. In P. Stuart-Macadam and K. A. Dettwyler (Eds.), *Breastfeeding: Biocultural perspectives* (pp. 1–37). New York, NY: Aldine De Gruyter.

Tabachnick, B. G., & Fidel, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

Verronen, P. (1982). Breast feeding: Reasons for giving up and transient lactation crisis. *Acta Paediatrica Scandinavica*, *71*, 447–450.

Zeller, R., Good, M., Anderson, G. C., & Zeller, D. L. (1997). Strengthening experimental design by balancing potentially confounding variables across treatment groups. *Nursing Research*, *46*, 345–349.