
Fan-Hao Chou, Kay C. Avant, Shih-Hsien Kuo, and Han-Fu Cheng
School of Nursing, Kaohsiung Medical University, Kaohsiung, Taiwan, and School of Nursing, The University of Texas at Austin, Austin, Texas, USA.

This study was an initial psychometric test of the Chinese versions of the Index of Nausea, Vomiting, and Retching (INVR) and the Prenatal Self-Evaluation Questionnaire (PSEQ) in pregnant Taiwanese women. Although there already is evidence that the English-language versions of the scales are reliable and valid, it is important to verify the proper psychometric characteristics of the Chinese versions. Forward and backward translation, and a multiphase instrumentation study describing internal consistency, test-retest reliability, and content validity of the translated versions were conducted. A convenience sample was recruited from prenatal clinics in the south of Taiwan. Three measurement instruments were used in this study: the demographic inventory (DI), the INVR, and the PSEQ. Thirty pregnant women participated in the study. Both the internal consistency and stability coefficients of the INVR and PSEQ were satisfactory. The indices of content validity (CVI) for the Chinese versions of these two instruments were both 1.0, indicating that they are acceptable for use among Taiwanese pregnant women. This was the first instrumentation study of the INVR and PSEQ applied to Taiwanese pregnant women. Researchers could use this study as a model for future translation and application of psychometric instrumentation.

Key Words: psychometric equivalence, nausea, vomiting, maternal psychosocial adaptation

Pregnancy can be stressful and often causes uncomfortable physical symptoms. Chief among these are nausea and vomiting, which affect not only the physical and mental condition of pregnant women [1], but also the subsequent relationship between a mother and her newborn. The mother’s adaptation to pregnancy provides a foundation for the mother-newborn relationship to come, and forms one step in achieving the maternal tasks of “seeking and ensuring safe passage through pregnancy and childbirth”, “binding-in to the child”, “acceptance by others”, and “giving of oneself” [2,3].

Nausea and vomiting can increase stress [4,5], which can adversely affect maternal adaptation [6]. Therefore, pregnancy-induced nausea and vomiting may be a significant factor in maternal psychosocial adaptation during pregnancy. Research on this particular issue has been lacking, particularly in Taiwan. Knowledge of the possible effects of nausea and vomiting during pregnancy may help parent-child nurses better understand how to provide appropriate prenatal care, and provide data for comparative research or practical application for other
countries. Nurses can then design and implement suitable comprehensive interventions for these women in Taiwan and elsewhere.

Two instruments, the Index of Nausea, Vomiting, and Retching (INVR) and the Prenatal Self-evaluation Questionnaire (PSEQ), are currently available to study pregnancy-induced nausea and vomiting and maternal psychosocial adaptation. There are, however, no Chinese-language versions of these measurement scales for Taiwanese women, despite the strong empirical evidence for the reliability and validity of the English-language versions. Researchers should consider the issues of cultural and linguistic translation when selecting instruments to be applied in a different country, for example, as with adapting English-language based instruments for use in a Taiwanese population. In adapting an instrument, its reliability and validity are achieved by establishing the equivalence of the original and the translated instrument’s psychometric characteristics, in addition to properly translating the concepts and language [7]. In general, forward and backward translations and expert review satisfy the conceptual issues, while field tests can best establish the proper equivalence of psychometric characteristics [7,8].

Forward and backward translations can be accomplished through direct translation by the researcher, after which another bilingual translator can do the back translation. The two versions of the translation can then be checked by an expert or experts for adequacy [7–9]. In addition to these translation steps to verify and validate the translation of an instrument, decentering, in which “successive iterations of translations are done until both forms are appropriate for their respective cultures” [10], can further enhance the cultural and linguistic translation of an instrument [10,11].

To establish the reliability and validity of the Chinese versions, therefore, is a critical task prior to any study conducted with the INVR and PSEQ in Taiwan. The first aim of this study was to translate the INVR and PSEQ scales into Chinese. The second aim was to conduct initial psychometric testing of the INVR and PSEQ (including evaluating the readability, item clarity and initial psychometric characteristics) in pregnant Taiwanese women before using the scales in an exploratory study with a larger sample [12].

Reliability and validity of the English versions

The INVR was designed to measure the severity of nausea and vomiting and includes subjective and objective measurements. It is an 8-item instrument that uses a 5-point Likert scale and consists of three subscales: nausea (range, 0–12), vomiting (range, 0–12), and retching (range, 0–8), giving a total range of 0–32. The INVR was developed from the INV-1 and INV-2 [13], both of which had good evidence for reliability and validity [14–18]. For example, the Cronbach’s α of the INV-1 were 0.89–0.97 (n = 25–30) for 12 administrations of a version in oncology patients; the split-half correlations were 0.83–0.99 (n = 25–32) for 11 of the 12 administrations. The concurrent validity of the INV-1 was determined using Spearman’s correlation coefficient, r = 0.87 (n = 18), by comparing the ratings of unhospitalized chemotherapy patients with those of a family member or significant other. The reliability of the INV-2 was 0.98 (Cronbach’s α) (n = 309) for oncology patients [17] and 0.83 (n = 102) to 0.88 (n = 60) in pregnant women [14,18].

In a parallel-form study, Rhodes and McDaniel determined the reliability of the INVR as a replacement for INV-2 [19]. The INVR uses the same symptom components and item-order as the INV-2, but has different descriptions. The INVR is designed with eight introductory statements and for each statement there are five response choices. The INVR is more concise and easier to read than the INV-2. Rhodes and McDaniel compared the responses of subjects to the INV-2 and the INVR. The percentage agreements and Spearman correlations between each of the eight items on the two scales were 79–98% and 0.71–0.95 (n = 159), respectively. The Spearman’s coefficient was 0.87 (n = 159) for the total INVR and the INV-2. The data revealed that the INVR and INV-2 were equivalent in measuring nausea, vomiting, and retching in obstetric (n = 40), oncologic (n = 60), and medical/surgical (n = 59) populations. Rhodes and McDaniel, therefore, recommended use of the INVR in future studies [19].

The PSEQ is used to measure maternal psychosocial adaptation during pregnancy and was developed from the Lederman framework of maternal psychosocial adaptation. It is a 79-item instrument using a four-point Likert scale (1 = not at all, to 4 = very much so) and measures conflict in the maternal developmental tasks of pregnancy. The scale is composed of seven domains: acceptance of pregnancy, identification of a motherhood role, relationship with her mother, relationship with her husband, preparation for labor, fear of helplessness and loss of control in labor, and concern for the well-being of self and baby [2,20–22]. The possible range of summed scores for the PSEQ scale is 79–316. Higher scores indicate greater perception of conflict or lack of adaptation.

This instrument has been used and demonstrated as
valid and reliable through various studies with different populations [21,23,24]. The samples in these studies included primigravidas and multigravidas in different trimesters of pregnancy and in different ethnic groups. For instance, in one study, the Cronbach’s α of the PSEQ in each of the three trimesters of pregnancy among the seven dimensions was, in order, 0.74–0.91 (n = 196–203), 0.71–0.93 (n = 297–302), and 0.73–0.92 (n = 367–377). Previous studies have provided evidence for the predictive and construct validity of the PSEQ. To examine the effects of prenatal maternal adaptation on maternal anxiety in labor, Lederman, for example, performed a longitudinal study (from the third trimester to active-phase labor) in 53 normal multigravidas using the PSEQ and the Labor Anxiety Scale [21]. The results showed that the PSEQ could predict anxiety in labor, supporting the predictive validity of the PSEQ.

In Taiwan, there was a Chinese version of the INV-2, but no reliability and validity data for the INVR or the INV-2 in maternal research have been reported. The researcher, therefore, had to determine the reliability and validity of the INVR Chinese version. There was no Chinese version of the PSEQ. The researcher, therefore, needed to translate the PSEQ and determine the reliability and validity of the Chinese version.

METHODS

Research design and study population
This was a translation and multiphase instrumentation study to determine internal consistency, test-retest reliability, and face and content validity in the translated versions of the INVR and the PSEQ.

The population of concern was Taiwanese pregnant women. Thirty pregnant women were recruited into the study. The criteria for selection were: no previous history or currently diagnosed medical disease and no pregnancy complications; aged from 18 to 44 years old; and speaks, reads, and writes Chinese. The researcher used a convenience sampling technique, collecting data from prenatal clinics in two teaching hospitals in Kaohsiung, Taiwan.

Translation
The researcher obtained permission to use the INVR and PSEQ from the original authors. The two instruments were then translated from English to Chinese by two bilingual academics, the researcher, and a Taiwanese associate professor with a PhD in a health-care field. A doctoral candidate majoring in foreign language education with a concentration in English as a second language back translated the Chinese versions of the INVR and PSEQ into English. This bilingual doctoral candidate had never seen the English versions of the INVR and PSEQ. The equivalence of the original and back-translated scales was evaluated by a nursing professor who is an expert in parent-child nursing. The instruments were modified as necessary, based on the decentering procedure. Only minor revisions were required.

Data collection
The researcher used a self-report method to collect data. Data collection began after obtaining approval from the Human Subjects Committee of the institution. Participants were informed that all the information gathered from the study would be kept confidential. The researcher first screened potential participants who came to prenatal care clinics, and informed them of the aim of the study and the time needed for self-administration of the instruments. After obtaining informed consent, the researcher gave willing participants the demographic inventory (DI), INVR and PSEQ scales to fill out. When the participants completed the questionnaires at the clinics, the researcher also gave them a second set of the INVR and PSEQ scales in an envelope to take home and fill out. All participants were notified that they had the right to withdraw from the study at any time. They could also refuse to respond to any question that might make them feel uncomfortable. Participants’ names and other identifying information were not on the questionnaires, and a code number was assigned to ensure confidentiality. All the data remained accessible only to the researcher. All participants were asked to provide suggestions about the wording of the instruments, in order to determine face validity.

These steps were continued until 30 participants were obtained. One week later, the researcher received the second set of questionnaires from these participants by return-mail. Each participant received a small gift from the researcher after she had completed the study. In addition, participants could obtain free nursing consultation from the researcher.

Data analysis
Data were analyzed using SPSS, version 9.0 for Windows. The internal consistency and test–retest reliability coefficients were examined for the INVR and the PSEQ scales. The content validity and face validity were also assessed for the Chinese versions of both scales. Data are given as mean ± standard deviation.
RESULTS

The 30 pregnant women in this study ranged in age from 18 to 41 years with a mean age of 27.23 ± 5.14 years. The gestational stage of these women ranged from 7 to 38 weeks with a mean of 25.83 ± 9.90 weeks. Most participants in this sample had a college education (40.0%, n = 12), followed by a senior high school education (36.7%, n = 11). More than half of the participants in this sample were employed (53.3%, n = 16) and more than half (56.7%, n = 17) were primigravida. Most women had nausea and vomiting (73.3%, n = 22).

Content and face validity

According to Streiner and Norman, 3–10 experts are appropriate for a panel of experts [8]. For this study, three experts were chosen from related research and practical fields to estimate the content validity of the Chinese-language versions of the INVR and PSEQ. The first panel member was a School of Nursing professor with a PhD in nursing and more than 20 years of experience researching and working with perinatal women in Taiwan. The second was an obstetric physician with more than 10 years of experience working with pregnant patients with nausea and vomiting. The last was a School of Nursing faculty member who also worked part-time in the obstetric clinics of a hospital. She had several years of research experience with Taiwanese pregnant women.

These three experts were asked to independently rate the Chinese versions of the INVR and PSEQ in terms of the two instruments’ applicability to pregnant women in Taiwan, including their content and cultural relevance and language equivalence to the original instruments. A three-point scale was used to rate each item as not applicable, applicable with some revision, or applicable. Based on the opinions of the three experts, the sixth item of the INVR, which uses “a cup” to describe the amount of vomiting, might need to be further interpreted using concrete numbers, since there are several different sizes of cups in Taiwanese daily life. This item was therefore clarified with a number (“a cup is equal to 250 mL”) after the statement. After the revision in light of the experts’ opinions, the index of content validity (CVI) of both these translated instruments was 1.0 [25]. The CVIs indicated that the INVR and PSEQ Chinese versions were acceptable for use among pregnant women in Taiwan, based on the experts’ evaluation.

According to the responses from the expert panel and the participants in the study, the instrument was modified as necessary to build face validity. Face validity “refers only to the appearance of the instrument to the layman” and “does not provide evidence for validity” [25]. If an instrument has face validity, however, response rates may increase [25]. Therefore, the face validity of the INVR and PSEQ Chinese-language versions was estimated by a bilingual pregnant woman who had earned a bachelor’s degree in Taiwan, had lived in the USA for more than 2 years and, at the time of the study, was 33 weeks pregnant. She could respond easily to all items on both scales. She proposed that only one item of the PSEQ needed to be slightly reworded in Chinese. In the item, “It will be hard for me to balance child care with my other commitments and activities,” the “my” in the Chinese-language version seemed redundant. The word “my” was therefore omitted.

Reliability

The INVR total scores ranged from 0 to 29 (mean, 8.0 ± 8.8) for the 30 participants. For the three subscales of the INVR, the means for nausea, vomiting, and retching were 3.7 ± 3.8, 2.1 ± 3.2, and 2.2 ± 2.6, respectively.

In this study, the reliability coefficient for the INVR scale using Cronbach’s α was 0.94. The 1-week test-retest reliability coefficient determined by intraclass correlation coefficient (ICC) was 0.97. Both the internal consistency and stability coefficients of the INVR scale were high and satisfactory.

The PSEQ total scores ranged from 84 to 204 (mean, 144.0 ± 27.9). The scores on the seven subscales were:

| Table 1. Mean and standard deviation of the three subscales of the Index of Nausea, Vomiting, and Retching compared with two different studies using the Index of Nausea and Vomiting—2 |
|-------------------------------|----------------|----------------|----------------|----------------|
| Subscale                      | Present study (n = 30) | Zhou et al (n = 103) [18] | Belluomini et al (n = 60) [14] |
|                               |                     | Pre-treatment | Control group |
| Nausea                        | 3.7 ± 3.8           | 3.4 ± 0.7     | 8.4 ± 2.2      | 8.0 ± 2.5     |
| Vomiting                      | 2.1 ± 3.2           | 1.5 ± 0.6     | 2.1 ± 2.5      | 1.8 ± 2.7     |
| Retching                      | 2.2 ± 2.6           | 1.9 ± 0.8     | –              | –             |

Kaohsiung J Med Sci July 2005 • Vol 21 • No 7
acceptance of pregnancy, 25.9 ± 6.8; identification of a motherhood role, 25.2 ± 6.6; relationship with her mother, 16.0 ± 4.8; relationship with her husband, 15.3 ± 4.1; preparation for labor, 19.5 ± 5.5; fear of helplessness and loss of control in labor, 19.1 ± 5.6; and concern for the well-being of self and baby, 23.1 ± 6.0.

The reliability coefficient for the entire PSEQ scale using Cronbach’s α was 0.93. The ICC test-retest reliability coefficient at 1 week was 0.95. The Cronbach’s α coefficients of the seven subscales were: acceptance of pregnancy, 0.82; identification of a motherhood role, 0.81; relationship with her mother, 0.77; relationship with her husband, 0.68; preparation for labor, 0.80; fear of helplessness and loss of control in labor, 0.79; and concern for well-being of self and baby, 0.81. The 1-week test-retest reliability coefficients of these subscales were 0.91, 0.85, 0.89, 0.93, 0.96, 0.94, and 0.83, respectively. Both the internal consistency and stability coefficients of the PSEQ scale were satisfactory.

**DISCUSSION**

These results show that the Chinese-language versions of the INVR and PSEQ scales had satisfactory internal consistency and stability coefficients. The CVI for both these Chinese versions was 1.0, indicating that they are acceptable for use among Taiwanese pregnant women.

There are no related studies using the INVR. However, as US studies reveal that the INVR and INV-2 are equivalent in measuring nausea, vomiting, and retching in obstetric populations, the INVR scores in this study were compared with INV-2 scores from previous studies. In this study, the reliability of the INVR scale was 0.94. This was a little lower than Rhodes et al reported when using the INV-2 in oncology patients [17], but higher than Belluomini et al and Zhou et al reported with the INV-2 in pregnant women [14,18]. The reliability of the PSEQ scale was 0.93, which was consistent with previous Western studies [21,23,24].

The INVR total score in this study (mean, 8.0 ± 8.8) was lower than that reported by Belluomini et al, using the INV-2 in American women within the first 10 weeks of pregnancy who received pretreatment (mean, 12.6 ± 5.7; n = 30) or no pretreatment (mean, 11.5 ± 4.9; n = 30) [14]. Table 1 compares the three subscales of the INVR in this study with the two subscales of the INV-2 in two other studies [14,18]. The findings in this study are roughly equivalent to those of previous studies except that the score on the nausea subscale was lower than Belluomini et al reported for pretreated pregnant women (mean, 8.4 ± 2.2) and the control group
(mean, 8.0 ± 2.5) [14]. One explanation for the difference may be the different sampling criteria used: Belluomini et al. included women with symptoms of nausea and vomiting during pregnancy, while this study included women with or without nausea and vomiting during pregnancy. Another possible reason for the higher scores for nausea reported by Belluomini et al. may be due to the absence of a retching subscale and inclusion of retching in the nausea subscale.

Table 2 compares the seven subscales of the PSEQ in this study with those from three other studies [21,23,24]. The results of this study are similar to those of previous studies, where, for example, the mean score of the “relationship with her mother” subscale was similar to those Halman et al. reported for third-trimester fertile and infertile pregnant women [23]. However, it was lower than the means reported by Stark and Lederman for third-trimester pregnant women [24,21]. The mean score of the “relationship with her husband” subscale was higher than the scores reported by Halman et al for third-trimester fertile and infertile pregnant women [23], but lower than the scores reported by Stark [24] and Lederman [21] for third-trimester pregnant women. The mean scores of “acceptance of pregnancy,” “identification of a motherhood role,” and “concern for well-being of self and baby,” however, were higher than the mean scores in the three previous studies. One possible reason is that this study included all three trimesters and the three other studies focused only on the third trimester. Because the third trimester is closer to labor and delivery, the women may have been more accepting of their pregnancy, had more identification with a motherhood role, and had more concern for the well-being of themselves and their babies.

**Methodological issues**

Two instruments, the INVR and PSEQ, that did not have Chinese-language versions were subjected to direct and back translations. The psychometric properties of the INVR and PSEQ Chinese versions indicated satisfactory validity and reliability with indices of content validity of 1.0 and of reliability above 0.93. To increase the external validity and the theoretical basis for these two scales in Taiwan, further research should evaluate these Chinese versions in different Taiwanese populations, such as a larger sample in similar populations or other pregnant populations and settings in Taiwan.

During the process of this study, many pregnant women proposed that the PSEQ scale had too many items (79 items). Therefore, a short form may be worth developing for both cost-effectiveness and the benefit of subjects.

**References**


本研究目的為：將噁心、嘔吐及乾嘔量表 (INVR) 與孕期自我評量表 (PSEQ) 中文版並進行初步的心理計量測試。儘管此二量表的英文版已有良好的信度和效度，但適當之中文版心理計量特徵的確認亦非常重要。本研究採雙向翻譯與多面向之工具性研究設計來描述中文版之內在一致性、再測信度及內容效度。以方便取樣的方式，於台灣南部地區的產科門診進行資料的收集。本研究使用三種測量工具：人口學基本資料 (DI)、INVR及PSEQ，並以 SPSS 之 Windows 第 9 版進行資料的分析。共有 30 位懷孕婦女全程參與本研究，結果顯示，INVR 與 PSEQ 的內在一致性與穩定係數皆屬良好；此二量表的中文版之內容效度係數 (CVI) 皆為 1.0，顯示它們可被使用於台灣的懷孕婦女群體。這是 INVR 與 PSEQ 應用於台灣孕婦的初步工具性研究，研究者可利用本研究作為未來心理計量工具之翻譯及其應用的參考典範。

關鍵詞：心理計量對等性，噁心，嘔吐，孕期心理社會適應
(高雄醫學會 2005;21:314–21)

收文日期：94年 11月 24日
接受刊載：94年 5月 27日
通訊作者：周汎忠博士
高雄醫學大學護理學系
高雄市 807 十全一路 100 號

Kaohsiung J Med Sci July 2005 • Vol 21 • No 7
321