Translation and Cross-Cultural Validation of Korean Version of the Menstrual Distress Questionnaire

Yaelim Lee1 and Kyeong-Yae Sohng1

Abstract

Given the increase in cross-cultural studies, there is a need for adequately translated and validated study instruments. Using instruments translated into participants’ native language can lower barriers to study participation and increase study validity. The purpose of this study was to describe the translation and validation processes of the Korean version (MDQ-K) of the Menstrual Distress Questionnaire (MDQ). The MDQ was translated into Korean through a forward-and-backward translation process, followed by expert review and pilot testing among 100 bilingual Korean students. The equivalence of MDQ-K to MDQ was tested through bivariate Pearson’s correlations and paired t tests. The psychometric properties of the MDQ-K were evaluated through internal consistency and construct validity (confirmatory factor analysis). The reliability of the questionnaire was good (Cronbach’s $\alpha = .96$). The results of confirmatory factor analysis revealed an acceptable model fit to the data. Overall, the MDQ-K demonstrated acceptable psychometric properties, although paired t tests found significant differences ($p < .05$) between the MDQ and MDQ-K in the means of three items: “restlessness” (Item 22), “bursts of energy and activity” (Item 40), and “blind spots and fuzzy vision” (Item 46). Possible explanations for these discrepancies include the participants’ varying English proficiency levels, issues with understanding medical terminology, and absence of words with the same meanings in different languages. We also discussed possible deletion of questionnaire items through further factor analysis.

Keywords

menstruation, symptom assessment, instruments, psychometrics, Asian

Introduction

Once a month, most reproductive-aged women experience menstruation (Anjum & Jami, 2018). Women often have negative perceptions of menstruation due to multiple unfavorable physical and psychological symptoms as the menstrual period approaches (Gollenberg et al., 2010; Kollipaka et al., 2013). Many studies have examined menstruation-related distress among women with the aims of addressing poor perceptions of menstruation related to a lack of education and improving negative attitudes toward menstruation (Gollenberg et al., 2010; Kollipaka et al., 2013). As women’s cultural circumstances play important roles in shaping menstrual experiences, the types and degrees of menstrual distress vary widely, depending on the specific culture (Anjum & Jami, 2018; Tan et al., 2017). Therefore, studies on menstrual distress in different countries have taken care to consider unique cultural factors and to use their native language when assessing women to better capture the subtle nuances (Chang et al., 1999; Murakami et al., 2008). In conducting such studies, many researchers have utilized pre-existing instruments, translated, and used in their studies.

Some of the commonly used instruments for menstrual symptoms include the Menstrual Distress Questionnaire (MDQ; Moos, 2010), Shortened Premenstrual Assessment Form (SPAF; Allen et al., 1991), and Prospective Record of the Impact and Severity of Menstrual Symptoms (PRISM; Moher et al., 2009).

The MDQ

The MDQ is the most commonly used instrument in the United States to access premenstrual symptoms (Y. Lee & Im, 2016). The purported purpose of the MDQ was to assess menstrual cycle symptomatology (Moos, 2010). The initial evaluation of the MDQ done with 839 women who were...
wives of graduate students in large Western universities (Moos, 2010). It is a 46-item scale focusing on eight somatic factors of reproductive-aged women: (a) pain (Items 1–6), (b) water retention (Items 7–10), (c) autonomic reaction (Items 11–14), (d) negative affect (Items 15–22), (e) impaired concentration (Items 23–30), (f) behavior change (Items 31–35), (g) arousal (Items 36–40), and (h) control (Items 41–46; Moos, 2010). The scale uses 5-point Likert-type scale. A summed total score, total mean score, or scores by eight factors can be used from this scale (Moos, 2010). Although several Korean versions of the modified MDQ exist, it only uses selected items from the original MDQ, which limits its use (e.g., employing either the scoring system or interpretation guideline of the original MDQ and comparing findings of studies that used the original MDQ; Kim, 2004; Moos, 2010). The use of same validated instruments across different cross-cultural and international studies allows ease of comparison of the study findings. However, well-validated translated instruments are often unavailable, and researchers may be faced with a multitude of translated versions of a single instrument (Chang et al., 1999).

**Translation and Validation**

Although most researchers agree on the importance of adequate instrument translation and verification processes, approaches to these processes are inconsistent and vague (Cha et al., 2007; Squires et al., 2013). To retain adequate cross-cultural validity, five criteria should be considered in the translation process: content equivalence, semantic equivalence, technical equivalence, criterion equivalence, and conceptual equivalence (Squires et al., 2013). Adherence to these criteria can be validated by the development of a translation guide, forward-and-back translation, and an expert panel review (Squires et al., 2013; Wild et al., 2005). In the forward-only translation method, an instrument in the original language is translated into the target language (C. C. Lee et al., 2009). In contrast, in forward-and-back translation, the instrument must be translated at least twice by different translators: original-to-target-to-original language (C. C. Lee et al., 2009). The original instrument and back-translated instrument (also in the original language) are then compared with respect to equivalence, and the forward-and-back method is continued until the translators reach a consensus (C. C. Lee et al., 2009). The expert panel review can be done using a scoring process, such as the Likert-type scale (Squires et al., 2013).

The reliability and validity of an instrument can be tested using various measures (Mokkink et al., 2010). Its reliability is often based on internal consistency and test–retest correlations, whereas its validity is determined by content validity and construct validity (Che et al., 2017). This process can be done through monolingual or bilingual testing (Sperber, 2004). The bilingual testing method is considered more rigorous and precise, as the translated instruments are tested using individuals who understand both languages and compare the instruments in both languages (Son et al., 2000; Sperber, 2004). The committee approach, in which a number of experts discuss the translated instrument as a team or conduct pilot tests of the translated instrument, is often used in determining the reliability and validity of an instrument ( Maneesriwongul & Dixon, 2004). Within the confines of time and resources, a combination of multiple techniques can help to establish equivalence between the original and translated versions of instruments (Maneesriwongul & Dixon, 2004). The purpose of this study was to describe the translation and validation process of a translated Korean version of the Menstrual Distress Questionnaire (MDQ-K).

**Method**

The development of the MDQ-K involved four steps: (a) obtaining permission to translate, (b) forward-and-back translation, (c) expert review, and (d) pilot testing with bilingual students. The content validity was established through an expert review. Through the pilot tests with bilingual students, the reliability and construct validity of the questionnaire were analyzed.

**Obtaining Permission to Translate**

The research team contacted the copyright holder (Mind Garden) of the MDQ and obtained permission to translate the instrument into Korean. No financial or other conflict of interest was incurred in this process.

**Forward-and-Back Translation**

The two research assistants, Korean–English bilingual female doctoral students in nursing conducted the forward-and-back-translation (Chang et al., 1999). Translator A translated the original version of the instrument (English Version 1) to Korean. Translator B translated the Korean Version 1 to English (English Version 2). Each translation process was blinded (i.e., Translators A and B performed their translation separately and did not communicate). The English Versions 1 and 2 were compared item-by-item, and differences were identified, discussed, and modified in the presence of the two translators and the research project investigator.

**Expert Review**

The agreed-upon translated version was sent to five experts to establish the instrument’s content validity and necessary revisions. The experts were female professors in Korean nursing schools whose research area was menstrual health. The content validity indexing technique was used in the evaluation (Squires et al., 2013). The experts rated each questionnaire item on a 4-point Likert-type scale (1 = inappropriate, 2 = somewhat appropriate, 3 = appropriate, and
4 = very appropriate) (Squires et al., 2013; Yu, 2010). Items that scored less than “2” by two experts were revised for clarity (Yu, 2010).

Pilot Testing With Bilingual Students

Participants. The revised questionnaire from the expert review was named the MDQ-K and pilot tested with 100 bilingual Korean female students studying in the United States. The Institutional Review Board approval was obtained from the primary investigator’s affiliated institution (#820966) before initiating the study. The sample size of the study was calculated to be 90 based on G*Power 3.1.9.2 software, with a power of .80, alpha level of .05, and an effect size of 0.3. The inclusion criteria for the study were the following: (a) female students; (b) aged between 20 and 40 years; (c) South Korean identity; (d) ability to read, write, speak, and understand Korean and English; (e) enrolled in U.S. institutes, with an official school email address; (f) access to the internet; (g) presence of menstruation; and (h) experienced symptomatic changes in the menstrual cycle. The Korean international students studying in the United States were targeted for their bilingual language abilities and them being generally young adults, considering menstrual distress is more often seen in young reproductive-aged women (Meers et al., 2020).

Patient and public involvement. Patient and public were not involved in the design of this study. Korean international students (healthy women; public) were the participants in this study. Once the study has been published, the participants will receive a copy of this article through their provided emails.

Data collection. The study data were collected between September and October of 2014. Participants were recruited using a convenient sampling method through online communities (e.g., Korean international students’ associations, Korean Americans online talk lounges). The study was advertised in online communities for Korean international students studying in the United States. To minimize the possible selection biases, the study advertisement was posted on free bulletin boards where any Korean international students can read without log-ins. The participants were asked to use school emails for the study, so that the researchers can verify them being students.

A total of 100 potential participants were asked to sign electronic informed consent and answer screening questions, which were available through the online study advertisement link. The eligible participants were assigned a study ID and asked to answer questions on sociodemographics and provide baseline information in the original English version of the MDQ and the MDQ-K on the same day. The requirement to complete the questionnaires on the same day was due to inherent natural variability in answers of symptoms on a daily basis due to the menstrual cycle. To minimize possible recall bias, the sequence of items in the translated version of the questionnaire was mixed (Son et al., 2000). For example, when the MDQ was asked from Items 1 to 46, the MDQ-K was asked with mixed order of items. After completing the 10-min-length study questionnaires, each participant received US$15 online gift cards.

Study variables and instruments. Alongside answering questions from the original English version of the MDQ and the MDQ-K, to better interpret the findings of the study, the participants were also asked to answer questions associated with menstrual health (e.g., age, gravidity, age at menarche, and duration and regularity of menstruation) and language/cultural proficiency (e.g., educational status, major, and degree of acculturation; Y. Lee & Im, 2016; Sadler et al., 2010). The participants’ degrees of acculturation was assessed using the 21-item Korean version of Suinn-Lew Asian Self-Identity Acculturation Scale, which includes questions about language, self-identity, community, and cultural preferences (Suinn et al., 1992). The scale has been tested among Korean Americans and demonstrated good construct validity and internal consistency, with a Cronbach’s alpha of .91 (Jackson et al., 2006; Shim & Schwartz, 2008).

Data analysis. The participants’ answers were automatically coded through the REDCap survey system. The SPSS 22.0 and Mplus 7.3 were used for the data analysis. Two participants missed answering a question on smoking and one participant missed a question on caffeine intake. The missing data of the study was assessed for missing at randomness. Without imputation or deletion for missing data, all analyses were conducted with Full Information Maximum Likelihood (FIML; Enders & Bandalos, 2001).

Descriptive analyses were reported using frequency, percentage, mean, and standard deviation. The equivalence of MDQ-K to MDQ was tested through bivariate Pearson’s correlations and paired t-tests. The reliability of the MDQ-K was calculated by analyzing its internal consistency using Cronbach’s alpha, and the construct validity was assessed using confirmatory factor analysis to confirm the underlying dimensions of an instrument (Son et al., 2000). The model fit was evaluated using chi-square statistics, comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA; Browne & Cudeck, 1993; Elavsky & Gold, 2009; Kilbride et al., 2003; Reid et al., 2015). Generally, the acceptable model is represented by $\chi^2/df$ of below 4, and CFI and TLI more than 0.90 (Browne & Cudeck, 1993; Elavsky & Gold, 2009). The RMSEA below 0.05 represents close model fit and below 0.08 represents reasonable model fit (Browne & Cudeck, 1993; Elavsky & Gold, 2009).

In addition to the purpose of this study, we attempted to screen out relatively insignificant items. Each of the 46 items’ factor loading and each factor’s construct reliability
(CR) and average variance extracted (AVE) were accessed (Pedrosa et al., 2016). Commonly acceptable factor loading is defined as a value more than 0.5, CR as a value more than .7, and AVE as more than 0.5 in previous studies (Chen et al., 2015; Han et al., 2015; Pedrosa et al., 2016).

Results

Expert Review: Content Validity

On the initial version of the translated questionnaire, two items received a score of less than “2” by two experts: “affectionate” (Item 36) and “orderliness” (Item 37). The term “affectionate” was translated as “da-jung-han” by the translators, a word that in Korean is used to describe someone who is kind and pleasant. However, the experts pointed out that the term affectionate has a stronger intonation than “da-jung-han.” Based on the experts’ suggestions, we selected the word “ae-jung-i-num-chi-nun,” which means very affectionate. The word “Orderliness” was translated as “yu-soon-ham,” which is used to describe someone who is pleasant and well-mannered. However, the experts suggested that “Orderliness” in English can mean organized and neat, which needed to be included in the translated version. Thus, we added this meaning by including the expression “jil-so-jung-yeon-ham” in the Korean translated version.

Pilot Testing With Bilingual Students

Sociodemographic data. After the expert review, the questionnaire was named as MDQ-K and tested in a pilot test with 100 bilingual students. The sociodemographic and baseline data on the participants are summarized in Table 1. The mean age of the participants ($N = 100$) was 25.94 years. More than half the participants were enrolled in a graduate program and majoring in more than 20 different majors including applied science (e.g., bioengineering, computer science, environmental science), health professions (e.g., nursing, pharmacology, public health), and business/management/finance/economics majors. The participants’ degree of acculturation score was 2.29, with a score of 1.0 indicating “Asian identified” and a score of 3 denoting “bicultural.” The majority of the participants was single, never pregnant, Christian, had regular menstrual cycles, and had a positive perception of their general health.

Equivalence of MDQ-K to MDQ. The results of the bivariate Pearson’s correlations for the paired items showed that there were significant correlations between each of the 46 items and the total scores of the English version and Korean translated version ($p < .001$), with scores ranging between .572 (“Dizziness and faintness”) and .985 (“Total score”). The results of paired $t$ test used to test equivalence of eight factors showed that factor “Arousal” was significantly higher in Korean translated version ($M = 2.41$, $SD = 3.01$) than English version ($M = 2.68$, $SD = 3.25$); $t(100) = -2.038$, $p < .05$ (Table 2). The results of a paired $t$ test comparing each item between MDQ and MDQ-K showed that the means of three items were significantly different between the English and Korean versions (Table 3). The mean of Item 22 “Restlessness” was significantly higher in the English version ($M = 1.10$, $SD = 1.17$) than the Korean translated version ($M = 0.75$, $SD = 1.03$); $t(100) = 3.697$, $p < .001$. The mean of Item 40 “Bursts of energy and activity” was significantly higher in the Korean translated version

### Table 1. Sociodemographic and Baseline Information of Participants ($N = 100$).

| Variables                              | $M \pm SD$ or $n$ (%) |
|----------------------------------------|-----------------------|
| Age (years old)                        | 25.94 ± 4.10          |
| Educational status                     |                       |
| Undergraduate                          | 35 (35.0)             |
| Graduate                               | 61 (61.0)             |
| Non-degree program                     | 4 (4.0)               |
| Major                                  |                       |
| Applied Science                        | 19 (19.0)             |
| Health Professions                     | 14 (14.0)             |
| Business/Management/Finance/Economics  | 12 (12.0)             |
| Music/Art                              | 10 (10.0)             |
| Media/Communication/Culture            | 8 (8.0)               |
| Education/Human development            | 7 (7.0)               |
| International relations/Policy         | 6 (6.0)               |
| Science                                | 5 (5.0)               |
| Others                                 | 19 (19.0)             |
| Degree of acculturation                | 2.29 ± 0.22           |
| Marital status                         |                       |
| Single                                 | 67 (67.0)             |
| Partnered                              | 21 (21.0)             |
| Married                                | 11 (11.0)             |
| Divorced                               | 1 (1.0)               |
| Gravidity (yes)                        | 4 (4.0)               |
| Religion                               |                       |
| Christian                              | 46 (46.0)             |
| Catholic                               | 13 (13.0)             |
| Buddhism                               | 2 (2.0)               |
| Others                                 | 37 (37.0)             |
| Refused to answer                      | 2 (2.0)               |
| Age at menarche (years old)            | 13.18 ± 1.35          |
| Duration of menstruation (days)        | 5.60 ± 1.22           |
| Regularity of menstruation (irregular) | 38 (38.0)             |
| Perception of general health (negative)| 12 (12.0)             |
| Diagnosed disease (yes)                | 8 (8.0)               |
| Medication intake (yes)                | 8 (8.0)               |
| Smoking                                |                       |
| Previous smoker                        | 3 (3.0)               |
| Current smoker                         | 2 (2.0)               |
| Alcohol consumption (yes)              | 86 (86.0)             |
| Caffeine intake (yes)                  | 96 (96.0)             |
Lee and Sohng

Table 2. Paired t Tests for Eight Factors and Total Menstrual Distress Questionnaire (N = 100).

| Subscales               | English (M, SD) | Korean (M, SD) | t value |
|-------------------------|----------------|----------------|---------|
| Factor 1: Pain          | 5.99 (4.66)    | 6.07 (4.67)    | -0.433  |
| Factor 2: Water Retention | 3.45 (3.17)    | 3.60 (3.10)    | -1.216  |
| Factor 3: Autonomic Reaction | 1.48 (2.48)    | 1.40 (2.45)    | 0.601   |
| Factor 4: Negative Affect | 7.85 (7.25)    | 7.78 (7.14)    | 0.294   |
| Factor 5: Impaired Concentration | 5.32 (6.05)    | 5.34 (5.98)    | -0.120  |
| Factor 6: Behavior Change | 5.70 (4.97)    | 5.96 (5.30)    | -1.763  |
| Factor 7: Arousal       | 2.41 (3.01)    | 2.68 (3.25)    | -2.038* |
| Factor 8: Control       | 1.67 (2.66)    | 1.88 (2.88)    | -1.692  |
| Total                   | 33.87 (26.01)  | 35.71 (26.30)  | -1.796  |

* p < .05.

(M = 0.50, SD = 0.73) than in the English version (M = 0.40, SD = 0.65); t(100) = -2.075, p < .05. The mean of Item 46 “Blind spots and fuzzy vision” was significantly higher in the Korean translated version (M = 0.34, SD = 0.73) than in the English version (M = 0.28, SD = 0.59); t(98) = -2.161, p < .05.

Reliability and construct validity. The reliability of the instrument was determined using Cronbach’s alpha (Table 4). The reliability coefficient alpha of the eight factors of both MDQ and MDQ-K ranged between .76 and .92. The reliability coefficient alpha of both the total MDQ and total MDQ-K was .96. The confirmatory factor analysis of eight factors of MDQ-K was performed to estimate construct validity. The data from MDQ-K represented an acceptable overall fit: χ²/df was 1.546, CFI was .933, TLI was .911, and RMSEA (90% confidence interval) was .074 [.060, .087-->. Based on additional analysis on each item’s factor loading and each factor’s CR and AVE, 18 items with factor loading below 0.5 and factor “control” were removed. Each of the remaining seven factors showed CR above .7 and AVE above 0.5 (Table 5).

Discussion

The overall MDQ-K demonstrated good reliability and construct validity, as shown by the pilot test with the bilingual students. However, the scores for three of the 46 items, “Restlessness” (Item 22), “Bursts of energy and activity” (Item 40), and “Blind spots and fuzzy vision” (Item 46), were significantly different in the MDQ-K versus the original English version. There are several possible reasons for these differences.

First, it is possible that the participants’ English proficiency levels differed and that some of the participants found the meanings of some words difficult to understand. After
Cronbach’s commonly used terminology, such as nervousness or anxiety, had been used. Many Koreans may understand “Restlessness” as having no rest (i.e., “rest” plus “less”) and being tired. Another possibility is that, ironically the translators’ efforts to avoid the use of medical terminology could have contributed to the difference in understanding of certain items between the two languages. For example, many English–Korean dictionaries describe “blind spots” (Item 46) as “meng-jum,” which is a word originated in Chinese characters and more often used in medical settings than in daily life. The translators added a brief explanation of the word in parentheses. This effort could have resulted in the difference in understanding the item between the two languages.

Third, the absence of words with the same meanings in different languages may have generated differences in items between the two languages. When translating a single word into another language, it is almost impossible to find a word with the exact same meaning and nuance (Khalaila, 2013). For example, the Korean translation of “Orderliness” (Item 37) required two concepts “jil-so-jung-yeon-ham” and “yu-soon-ham.” There was no single word in Korean that included both concepts, as described in the “Results” section.

The overall reliability and validity, including results from confirmatory factor analysis, revealed that the MDQ-K can be used as it is. However, additional analysis on each item’s factor loading and each factor’s CR and AVE has suggested research questions for further studies. Although this study was conducted to retain all 46 items from the original MDQ in creating the MDQ-K, it is possible that the 18 items screened out for having weak factor loading or are relatively unimportant items to Korean women. In addition to addressing the aforementioned three items (Items 22, 40, and 46), several refinements of the MDQ-K through exploratory factor analysis and confirmatory factor analysis would be helpful for enhancing the quality of the questionnaire.

This study has several limitations. First, we did not assess the participants’ English proficiency level prior to the study. We could have achieved more reliable study findings had we included participants with a language level matched to that of the participants in the original MDQ. A second limitation is that the study consisted of only Korean international students enrolled in U.S. institutes. Although we intentionally limited the study participants to this population because their Korean proficiency levels were inferred to be suitable for comparing questionnaires in two languages, the homogeneous characteristics of the study population (e.g., the participants’ educational level) limit the generalizability of the study findings. Third, culturally specific menstrual symptoms or factors affecting the symptoms were not explored in this study. Future qualitative studies can explore these points to further test the adequacy of MDQ-K to Korean women. Finally, this study is limited for conducting factor analysis by its small sample size. Adequate sample size for factor analysis is determined by the ratio of the number of items to the number of samples (i.e., 1:3 means three samples are needed to measure one item; Bujang et al., 2012). For factor analysis in the medical field, where

| Table 4. Reliability Analyses of Menstrual Distress Questionnaire (English and Korean Versions; 46 Items). |
|----------------------------------|-----------------|-----------------|
| Value                            | English version | Korean version  |
| Factor 1: Pain                   |                 |                 |
| Item mean (SD)                   | 1.01 (0.19)     | 1.01 (0.23)     |
| Cronbach’s α                     | .84             | .84             |
| Factor 2: Water Retention        |                 |                 |
| Item mean (SD)                   | 0.86 (0.11)     | 0.90 (0.11)     |
| Cronbach’s α                     | .78             | .76             |
| Factor 3: Autonomic Reactions    |                 |                 |
| Item mean (SD)                   | 0.37 (0.03)     | 0.35 (0.01)     |
| Cronbach’s α                     | .84             | .83             |
| Factor 4: Negative Affect        |                 |                 |
| Item mean (SD)                   | 0.97 (0.02)     | 0.97 (0.03)     |
| Cronbach’s α                     | .92             | .92             |
| Factor 5: Impaired Concentration |                 |                 |
| Item mean (SD)                   | 0.66 (0.09)     | 0.67 (0.09)     |
| Cronbach’s α                     | .90             | .89             |
| Factor 6: Behavior Change        |                 |                 |
| Item mean (SD)                   | 1.16 (0.03)     | 1.19 (0.03)     |
| Cronbach’s α                     | .89             | .91             |
| Factor 7: Arousal                |                 |                 |
| Item mean (SD)                   | 0.49 (0.01)     | 0.54 (0.02)     |
| Cronbach’s α                     | .88             | .90             |
| Factor 8: Control                |                 |                 |
| Item mean (SD)                   | 0.29 (0.01)     | 0.31 (0.01)     |
| Cronbach’s α                     | .76             | .76             |
| Total                            |                 |                 |
| Item mean (SD)                   | 0.76 (0.15)     | 0.76 (0.15)     |
| Inter-item correlation mean (SD) | 0.32 (0.03)     | 0.33 (0.02)     |
| Cronbach’s α                     | .96             | .96             |

completion of the questionnaires, some of the participants stated that they found the English version of the instrument slightly difficult to understand. Although the inclusion criteria for the study required that the participants could read, write, speak, and understand Korean and English and that they were enrolled in U.S. institutes, these criteria do not necessarily mean that all the participants were fluent in English. It is possible that the participants found the expression “Bursts of energy and activity” (Item 40) relatively difficult to understand, as many of the other items were single words. Previous studies have emphasized the importance of considering study participants’ reading comprehension level when administering a pre-existing survey questionnaire, as the level does not always meet the targeted literacy level of the questionnaire (Flory & Emanuel, 2004; Suka et al., 2014). Assessing the readability grades of MDQ and MDQ-K would have been helpful.

Second, some of the medical terminologies could have confused the participants. For example, the word “Restlessness” (Item 22) may have been more easily understood if more commonly used terminology, such as nervousness or anxiety, had been used. Many Koreans may understand “Restlessness” as having no rest (i.e., “rest” plus “less”) and being tired. Another possibility is that, ironically the translators’ efforts to avoid the use of medical terminology could have contributed to the difference in understanding of certain items between the two languages. For example, the word “Restlessness” would have been helpful.

For example, the Korean translation of “Orderliness” (Item 37) required two concepts “jil-so-jung-yeon-ham” and “yu-soon-ham.” There was no single word in Korean that included both concepts, as described in the “Results” section.

The overall reliability and validity, including results from confirmatory factor analysis, revealed that the MDQ-K can be used as it is. However, additional analysis on each item’s factor loading and each factor’s CR and AVE has suggested research questions for further studies. Although this study was conducted to retain all 46 items from the original MDQ in creating the MDQ-K, it is possible that the 18 items screened out for having weak factor loading or are relatively unimportant items to Korean women. In addition to addressing the aforementioned three items (Items 22, 40, and 46), several refinements of the MDQ-K through exploratory factor analysis and confirmatory factor analysis would be helpful for enhancing the quality of the questionnaire.

This study has several limitations. First, we did not assess the participants’ English proficiency level prior to the study. We could have achieved more reliable study findings had we included participants with a language level matched to that of the participants in the original MDQ. A second limitation is that the study consisted of only Korean international students enrolled in U.S. institutes. Although we intentionally limited the study participants to this population because their Korean proficiency levels were inferred to be suitable for comparing questionnaires in two languages, the homogeneous characteristics of the study population (e.g., the participants’ educational level) limit the generalizability of the study findings. Third, culturally specific menstrual symptoms or factors affecting the symptoms were not explored in this study. Future qualitative studies can explore these points to further test the adequacy of MDQ-K to Korean women. Finally, this study is limited for conducting factor analysis by its small sample size. Adequate sample size for factor analysis is determined by the ratio of the number of items to the number of samples (i.e., 1:3 means three samples are needed to measure one item; Bujang et al., 2012). For factor analysis in the medical field, where
recruitment of participants may be difficult, statisticians recommend a ratio of 1:2 (Bujang et al., 2012). Therefore, our sample size of 100 participants for a 46 item questionnaire may be considered adequate. However, this is the minimum acceptable sample size. Especially for the confirmatory factor analysis, when there are six to 12 factors in a measure, a sample size of approximately 500 should be used whenever possible (Koran, 2016). As inadequate sample size can risk the power of the study findings, we recommend further studies with a larger sample size to confirm the results of factor analysis of the MDQ-K. Moreover, there is no single analytic test for testing hypothesis. We recommend future studies to test MDQ-K using various analytic methods (e.g., congruence coefficient analyses, concurrent validity tests, content validity tests from women with menstrual distress; Mokkink et al., 2010).

**Conclusion**

Many of the reproductive-aged women experience multiple menstrual symptoms. A good scale can support the adequate assessment of these symptoms. This study introduced a rigorous four-stage process for the translation and validation of a pre-existing instrument (the MDQ) into a new language, Korean. Overall, the MDQ-K demonstrated acceptable psychometric properties. However, some of the items may be open to further modification (Items 22, 40, and 46). Possible explanations for discrepancies between some of the items in the MDQ and MDQ-K and possible deletion of items through further factor analysis were discussed.

This study has a number of implications for future research and practice. First, when administering a translated questionnaire, it is important to determine whether the equivalence of the translated version to the original has been established before administering the instrument to the study participants. Second, the translation process of a questionnaire into a different language needs to be further explored. The rigorous four-stage questionnaire translation process introduced in this study was based on the literature (C. C. Lee et al., 2009; Squires et al., 2013). Although the MDQ-K carefully followed the steps described in the literature, some items may require further modification. We suggest repeating some of the steps to improve the equivalence of the translated version of the

| Path | Standardized path coefficient | Conceptual reliability | Average variance extracted |
|------|--------------------------------|------------------------|----------------------------|
| Q1 → Pain | .686 | .102 | 6.746*** | .831 | .555 |
| Q3 | .650 | .066 | 9.796*** |
| Q4 | .897 | .047 | 19.072*** |
| Q6 | .724 | .060 | 12.166*** |
| Q9 → Water retention | .778 | .074 | 10.526*** | .841 | .728 |
| Q10 | .922 | .069 | 13.390*** |
| Q11 → Autonomic reaction | .786 | .074 | 10.545*** | .895 | .742 |
| Q12 | .818 | .065 | 12.653*** |
| Q13 | .969 | .070 | 13.828*** |
| Q16 → Negative affect | .836 | .035 | 24.054*** | .910 | .630 |
| Q17 | .731 | .057 | 12.908*** |
| Q19 | .709 | .060 | 11.793*** |
| Q20 | .894 | .030 | 29.937*** |
| Q21 | .776 | .051 | 15.221*** |
| Q22 | .803 | .041 | 19.583*** |
| Q25 → Impaired concentration | .752 | .046 | 16.438*** | .914 | .728 |
| Q26 | .774 | .045 | 17.096*** |
| Q27 | .930 | .029 | 32.362*** |
| Q28 | .940 | .027 | 34.716*** |
| Q31 → Behavior change | .974 | .032 | 30.790*** | .886 | .616 |
| Q32 | .642 | .063 | 10.216*** |
| Q33 | .715 | .052 | 13.730*** |
| Q34 | .620 | .069 | 9.004*** |
| Q35 | .907 | .042 | 21.409*** |
| Q36 → Arousal | .871 | .035 | 25.240*** | .892 | .674 |
| Q37 | .841 | .037 | 22.765*** |
| Q39 | .789 | .044 | 17.994*** |
| Q40 | .780 | .046 | 17.007*** |

* ***p < .001.
questionnaire to the original. Third, more studies are necessary to ensure a comprehensive validation of the MDQ-K. In addition to calculating the internal consistency of the instrument and conducting paired *t* tests or confirmatory factor analysis, other validation techniques, including validation against other instruments or assessment methods (e.g., biomarkers), could be useful. Moreover, as the original MDQ is based on symptoms experienced by Western women, it would be useful to conduct studies to screen out MDQ items irrelevant to non-Westerners, newly group the items through exploratory factor analysis, or explore potential ethnic-specific symptoms that could be added.

As we are living today in an era of supraterritoriality, where physical place is becoming less and less meaningful, the importance of cross-cultural and international studies will continue to grow. We expect this study to be of help to researchers who intend to conduct studies using either English or previously translated questionnaire to non-native English speakers, or who plan to personally translate an English questionnaire for use in their study.

**Authors’ Note**

The research materials (e.g., data sets) used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Consent for Publication**

Research Edition Translation TA-442 Menstrual Distress Questionnaire (MDQ)—Forms C & T (All 93 items) performed by Yaelim Lee on this date November 17, 2014. Translated into Korean and reproduced by special permission of the Publisher Mind Garden, Inc. www.mindgarden.com from Menstrual Distress Questionnaire by Rudolf H. Moos. Copyright © 1968, 1989, 1990, 1991, 1999, 2000, 2010, by Rudolf H. Moos. All rights reserved in all media. Further reproduction is prohibited without the Publisher’s written consent. Published by Mind Garden, Inc. www.mindgarden.com

**Declaration of Conflicting Interests**

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The translation of this instrument was done as a preparation process for conducting a study exploring Korean women’s menstrual symptoms. The original license holder (Mind Garden, Inc.) retains ownership of translated Korean version of MDQ. The authors purchased the original MDQ and a manual, and conducted this study for an academic purpose. The license holder played no role in the study funding, design of the study, data collection, analysis, interpretation, and writing the manuscript.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the Sigma Theta Tau Xi Chapter & National Research Foundation of Korea in 2018–2019 (2017R1C1B5075221). The funding body plays no role in the design of the study, data collection, analysis, interpretation, and writing the manuscript.

**ORCID iDs**

Yaelim Lee [https://orcid.org/0000-0003-4200-8011](https://orcid.org/0000-0003-4200-8011)
Kyeong-Yae Sohng [https://orcid.org/0000-0002-4391-4481](https://orcid.org/0000-0002-4391-4481)

**References**

Allen, S. S., McBride, C. M., & Pirie, P. L. (1991). The Shortened Premenstrual Assessment Form. *Journal of Reproductive Medicine*, 36(11), 769–772.

Anjum, A., & Jami, H. (2018). Attitude towards menstruation, social adjustment, and mood states during menstruation among young women. *Pakistan Journal of Psychological Research*, 33(2), 591–606.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Bujang, M. A., Ghani, P. A., Soelar, S. A., & Zulkifli, N. A. (2012, September 10–12). *Sample size guideline for exploratory factor analysis when using small sample: Taking into considerations of different measurement scales* [Paper presentation]. The ICSSBE 2012—Proceedings, 2012 International Conference on Statistics in Science, Business and Engineering: “Empowering Decision Making with Statistical Sciences,” Langkawi, Kedah, Malaysia.

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: Issues and techniques. *Journal of Advanced Nursing*, 58(4), 386–395.

Chang, A. M., Chau, J. P. C., & Holroyd, E. (1999). Translation of questionnaires and issues of equivalence. *Journal of Advanced Nursing*, 29(2), 316–322.

Che, C. C., Hairi, N. N., & Chong, M. C. (2017). A systematic review of psychometric testing of instruments that measure intention to work with older people. *Journal of Advanced Nursing*, 73, 2049–2064. [https://doi.org/10.1111/jan.13265](https://doi.org/10.1111/jan.13265)

Chen, T. F., Chou, K. R., Liao, Y. M., Ho, C. H., & Chung, M. H. (2015). Construct validity and reliability of the Chinese version of the Disaster Preparedness Evaluation Tool in Taiwan. *Journal of Clinical Nursing*, 24(7–8), 1132–1143. [https://doi.org/10.1111/jocn.12721](https://doi.org/10.1111/jocn.12721)

Elavsky, S., & Gold, C. H. (2009). Depressed mood but not fatigue mediate the relationship between physical activity and perceived stress in middle-aged women. *Maturitas*, 64(4), 235–240. [https://doi.org/10.1016/j.maturitas.2009.09.007](https://doi.org/10.1016/j.maturitas.2009.09.007)

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, 8(3), 430–457.

Flory, J., & Emanuel, E. (2004). Interventions to improve research participants’ understanding in informed consent for research: A systematic review. *Journal of the American Medical Association*, 292(13), 1593–1601. [https://doi.org/10.1016/j.jama.2009.13.1593](https://doi.org/10.1016/j.jama.2009.13.1593)

Gollenberg, A. L., Hediger, M. L., Mumford, S. L., Whitcomb, B. W., Hovey, K. M., Wactawski-Wende, J., & Schisterman, E. F. (2010). Perceived stress and severity of perimenstrual symptoms: The biocyte study. *Journal of Women’s Health*, 19(5), 959–967.
Han, S. S., Han, J. W., & Lee, J. M. (2015). Development of an instrument for assessment of Korean nurses’ attitudes toward obese patients. *Japan Journal of Nursing Science, 12*(3), 249–257. https://doi.org/10.1111/jjns.12064

Jackson, S. C., Keel, P. K., & Ho Lee, Y. (2006). Trans-cultural comparison of disordered eating in Korean women. *International Journal of Eating Disorders, 39*(6), 498–502. https://doi.org/10.1002/eat.20270

Khailala, R. (2013). Translation of questionnaires into Arabic in cross-cultural research: Techniques and equivalence issues. *Journal of Transcultural Nursing, 24*(4), 363–370. https://doi.org/10.1177/1043659613493440

Kilbride, H. W., Powers, R., Wirtschafter, D. D., Sheehan, M. B., Charsha, D. S., LaCorte, M., . . . Goldmann, D. A. (2003). Evaluation and development of potentially better practices to prevent neonatal nosocomial bacteremia. *Pediatrics, 111*(4, Pt. 2), e504–e518.

Kim, H. (2004). Perimenstrual symptoms of Korean women living in the USA: Applicability of the WDHD (Women’s Daily Health Diary) on prospective report. *Taehan Kanho Hakhoe Chi, 34*(8), 1395–1401.

Kollipaka, R., Arounassalame, B., & Lakshminarayanan, S. (2013). Does psychosocial stress influence menstrual abnormalities in medical students? *Journal of Obstetrics and Gynaecology, 33*(5), 489–493.

Koran, J. (2016). Preliminary proactive sample size determination for confirmatory factor analysis models. *Measurement and Evaluation in Counseling and Development, 49*(4), 296–308. https://doi.org/10.1177/0748175616664012

Lee, C. C., Li, D., Arai, S., & Puntilllo, K. (2009). Ensuring cross-cultural equivalence in translation of research consents and clinical documents: A systematic process for translating English to Chinese. *Journal of Transcultural Nursing, 20*(1), 77–82. https://doi.org/10.1177/1043659608325852

Lee, Y., & Im, E. O. (2016). Stress and premenstrual symptoms in reproductive-aged women. *Health Care for Women International, 37*(6), 646–670. https://doi.org/10.1007/s10439-015-0493-2

Maneesrivongul, W., & Dixon, J. K. (2004). Instrument translation process: A methods review. *Journal of Advanced Nursing, 48*(2), 175–186.

Meers, J. M., Bower, J. L., & Alfano, C. A. (2020). Poor sleep and emotion dysregulation mediate the association between depressive and premenstrual symptoms in young adult women. *Archives of Women’s Mental Health, 23*, 351–359. https://doi.org/10.1007/s00737-019-00984-2

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed1000097

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Moos, R. H. (2010). Menstrual distress questionnaire manual, instrument, and scoring guide. http://www.mindgarden.com/products/mdq.htm

Murakami, K., Sasaki, S., Takahashi, Y., Uenishi, K., Watanabe, T., Kohri, T., . . . Suzuki, J. (2008). Dietary glycemic index is associated with decreased premenstrual symptoms in young Japanese women. *Nutrition, 24*(6), 554–561. https://doi.org/10.1016/j.nut.2008.02.003

Pedrosa, R. B. S., Rodrigues, R. C. M., Oliveira, H. C., & Alexandre, N. M. C. (2016). Construct validity of the Brazilian version of the self-efficacy for appropriate medication adherence scale. *Journal of Nursing Measurement, 24*(1), E18–E31. https://doi.org/10.1891/1061-3749.24.1.E18

Reid, C., Courtney, M., Anderson, D., & Hurst, C. (2015). Testing the psychometric properties of the Brisbane Practice Environment Measure using Exploratory Factor Analysis and Confirmatory Factor Analysis in an Australian registered nurse population. *International Journal of Nursing Practice, 21*(1), 94–101. https://doi.org/10.1111/ijn.12225

Sadler, C., Smith, H., Hammond, J., Bayly, R., Borland, S., Panay, N., . . . Inskip, H. (2010). Lifestyle factors, hormonal contraception, and premenstrual symptoms: The United Kingdom Southampton women’s survey. *Journal of Women’s Health, 19*(3), 391–396.

Shim, Y. R., & Schwartz, R. C. (2008). Degree of acculturation and adherence to Asian values as correlates of psychological distress among Korean immigrants. *Journal of Mental Health, 17*(6), 607–617.

Son, G. R., Zauszniewski, J. A., Wykle, M. L., & Picot, S. J. F. (2000). Translation and validation of caregiving satisfaction scale into Korean. *Western Journal of Nursing Research, 22*(5), 609–622. https://doi.org/10.1177/01939450022044629

Sperber, A. D. (2004). Translation and validation of study instruments for cross-cultural research. *Gastroenterology, 126*(1), S124–S128.

Squires, A., Aiken, L. H., van den Heede, K., Sermeus, W., Bruyneel, L., Lindqvist, R., . . . Matthews, A. (2013). A systematic survey instrument translation process for multi-country, comparative health workforce studies. *International Journal of Nursing Studies, 50*(2), 264–273. https://doi.org/10.1016/j.ijnurstu.2012.02.015

Suinn, R. M., Ahuna, C., & Khoom, G. (1992). The Suinn-Lew Asian Self-Identity Acculturation Scale: Concurrent and factorial validation. *Educational and Psychological Measurement, 52*(4), 1041–1046.

Suka, M., Odajima, T., Okamoto, M., Sumitani, M., Nakayama, T., & Sugimori, H. (2014). Reading comprehension of health checkup reports and health literacy in Japanese women. *Environmental Health and Preventive Medicine, 19*(4), 295–306. https://doi.org/10.1007/s12199-014-0392-8

Tan, D. A., Hathihtouwa, R., & Fraser, I. S. (2017). Cultural aspects and mythologies surrounding menstruation and abnormal uterine bleeding. *Best Practice & Research: Clinical Obstetrics and Gynaecology, 40*, 121–133. https://doi.org/10.1016/j.bpobgyn.2016.09.015

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for Patient-Reported Outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health, 8*(2), 94–104. https://doi.org/10.1111/j.1524-4733.2005.00454.x

Yu, D. S. F. (2010). Insomnia Severity Index: Psychometric properties with Chinese community-dwelling older people. *Journal of Advanced Nursing, 66*(10), 2350–2359. https://doi.org/10.1111/j.1365-2648.2010.05394.x