Rasch Measurement Analysis of a 25-Item Version of the Mueller/McCloskey Nurse Job Satisfaction Scale in a Sample of Nurses in Lebanon and Qatar

Michael Clinton¹, Nuhad Yazbik Dumit¹, and Fadi El-Jardali¹

Abstract
The Mueller/McCloskey Nurse Job Satisfaction Scale (MMSS) is widely used, but its psychometric characteristics have not been sufficiently validated for use in Middle Eastern countries. The objective of our methodological study was to determine the psychometric suitability of a 25-item version of the MMSS (MMSS-25) for use in middle-income and high-income Middle Eastern countries. A total of 1,322 registered nurses, 859 in Lebanon and 463 in Qatar, completed the MMSS-25 as part of a cross-sectional multinational investigation of nursing shortages in the region. We used the Rasch rating scale model to investigate the psychometric performance of the MMSS-25. We identified possible item bias among MMSS-25 items. We conducted confirmatory factor analyses (CFA) to compare the fit to our data of five factor structures reported in the literature. We concluded that irrespective of administration in English or Arabic, the MMSS-25 is not sufficiently productive of measurement for use in the region. A core set of 13 items (MMSS-13, Cronbach’s α = .82) loading on five dimensions eliminates redundant MMSS items and is suitable for initial screening of nurses’ satisfaction. Of the five factor structures we examined, the MMSS-13 was the only close fit to our data (comparative fit index = 0.951; Tucker–Lewis index = 0.931; root mean square error of approximation = 0.051; p value = .401). The MMSS-13 has psychometric characteristics superior to MMSS-25, but additional items are required to meet the research-specific objectives of future studies of nurses’ job satisfaction in Middle Eastern countries.

Keywords
Rasch measurement model, nurses’ job satisfaction, Lebanon, Qatar

Introduction
The Mueller/McCloskey Nurse Job Satisfaction Scale (MMSS) is a commonly used measure of nurses’ job satisfaction (Mueller & McCloskey, 1990). McCloskey (1974) developed a precursor to the MMSS to study the influence of incentives and rewards on staff nurse turnover rates. For each of 36 items categorized according to Maslow’s (1954) hierarchy of needs and insights of Burns (1969), staff nurses were asked to indicate which rewards, if available, would have influenced them to remain in the job they left in the previous 4 months. Mueller and McCloskey (1990) conducted exploratory factor analysis (EFA) on 33 of the 36 items and retained the 31 items that have since constituted the MMSS-31 (Table 1).

Mueller and McCloskey (1990) recommended that investigators use the MMSS-31 as a general measure of nurse satisfaction when they have no interest in determining how nurse satisfaction is influenced by particular job or work setting characteristics. When a finer grained understanding of nurse job satisfaction is required, Cronbach’s alpha values ≥.7 suggest that four of the eight sub-dimensions identified by Mueller and McCloskey are sufficiently internally consistent to achieve productive measurement (scheduling satisfaction, six items, Cronbach’s α = .84; interaction, three items, Cronbach’s α = .72; praise/recognition, four items, Cronbach’s α = .80; control/responsibility, five items, Cronbach’s α = .80). Mueller and McCloskey advised investigators to include independent questions on the importance of job satisfaction components if they are interested in which aspects of their work nurses believe are more salient to their job satisfaction. Their rationale was that there is little point in employers...
trying to improve nurses’ job satisfaction unless they understand what nurses themselves regard as important.

Although the MMSS-31 has been used in North America (McCloskey & McCain, 1987; Wilson, Squires, Widger, Cranley, & Tourangeau, 2008), other countries (Gurková et al., 2013; Babiarczyk, Fras, Ulman-Wlodarz, & Jarosova, 2014; Tourangeau, McGillis Hall, Doran, and Petch, 2006) could not replicate the eight-factor model reported by Mueller and McCloskey and suggested an alternative 23-item, seven-factor model (MMSS-23). However, Tourangeau and her colleagues found that three of the eight MMSS-31 sub-dimensions reported by Mueller and McCloskey had unacceptably low Cronbach’s alpha coefficients (extrinsic rewards, Cronbach’s α = .67; satisfaction with coworkers, Pearson’s r = .56; family/work balance, Cronbach’s α = .29). Mueller and McCloskey (1990) reported low reliability coefficients for the same sub-scales.

The studies by Mueller and McCloskey (1990) and Tourangeau et al. (2006) bring into question the internal consistency of MMSS-31 sub-dimensions across nurse populations and the suitability of the scale for use in non-Western countries. In the first study of its kind in the Middle East, Abu Ajamieh, Misener, Haddock, and Gleaton (1996) investigated the cultural appropriateness of the MMSS-31 as part of a study of job satisfaction among Palestinian nurses on the West Bank. Abu Ajamieh et al. (1996) could not extract more than four factors from their West Bank data (Table 1). Five MMSS-31 items did not load on any of Abu Ajamieh et al.’s four factors: Item 3 “Benefits package,” Item 8 “Weekends off per month,” Item 5 “Flexibility in scheduling your hours,” Item 6 “Opportunities to work consecutive days,” and Item 13 “Satisfaction with your immediate supervisor (head nurse or facility manager).” Abu Ajamieh et al. suggested rewording or reconceptualizing these items to improve the validity and reliability of the MMSS-31 for use outside the United States. Furthermore, Abu Ajamieh and his colleagues suggested that the MMSS-31 should be tested with a

Table 1. Item Composition of the MMSS-25 and MMSS-13.

| Items | MMSS-25 | MMSS-13 |
|-------|---------|---------|
| 1. Salary | X | X |
| 2. Vacation | X | X |
| 3. Benefits package (insurance, retirement) | X | X |
| 4. Hours that you work | X | X |
| 5. Flexibility in scheduling your hours | X | X |
| 6. Opportunity to work consecutive days | X | X |
| 7. Opportunity for part-time work | X | X |
| 8. Weekends off per month | X | X |
| 9. Flexibility in scheduling time off | X | X |
| 10. Compensation for working weekends | X | X |
| 11. Maternity leave time | X | X |
| 12. Child care for employees’ children at facility | X | X |
| 13. Your head nurse or facility manager | X | X |
| 14. Your nursing peers | X | X |
| 15. The physicians you work with | X | X |
| 16. Care delivery method | X | X |
| 17. Opportunities for social contact at work | X | X |
| 18. Opportunities for social contact after work | X | X |
| 19. Professional interactions with other disciplines | X | X |
| 20. Opportunities to belong to committees (21) | X | X |
| 21. Opportunities for career advancement (23) | X | X |
| 22. Recognition of your work from superiors (24) | X | X |
| 23. Recognition of your work from peers (25) | X | X |
| 24. Encouragement and positive feedback (26) | X | X |
| 25. Participation in organizational decision making (31) | X | X |

Note. Six items included in Mueller and McCloskey’s (1990) 31-item scale were omitted from the El-Jardali, Murray et al. (2013) study because they are not relevant to the scope of nursing practice of most nurses in Lebanon and Qatar. The excluded items are: interaction with faculty, control of practice, research, publication, responsibility, and control of conditions. The italicized items are as worded for the El-Jardali, Murray et al. (2013) study. Numbers in parentheses refer to item numbers in Mueller and McCloskey’s (1990) final set of 31 items. MMSS = Mueller/McCloskey Nurse Job Satisfaction Scale.

*Primary nursing (one nurse is responsible for holistic care of assigned patients); team nursing (nursing team is responsible for the holistic care of assigned patients); functional nursing (nurses are assigned tasks, for example, one nurse is responsible for medications, another for assessing and recording vital signs).
four-category rating scale that excludes the middle neutral category used by Mueller and McCloskey (1990).

MMSS-31 has been used in three studies of nurses in Jordan. Maryyan (2005) used the MMSS-31 to investigate the relationship between nurse satisfaction and patient satisfaction, and in a separate analysis, the relationship between nurses’ job satisfaction and intention to leave (Maryyan, 2006). Maryyan administered the MMSS-31 in English to a sample of nurses on the West Bank and relied on the reliability coefficients originally reported by Mueller and McCloskey (1990).

Abualrub, Omari, and Abu Al Rub (2009) administered the MMSS-31 in Arabic to investigate the effect of social support on stress and satisfaction among hospital nurses in Jordan and reported the Cronbach α coefficient for the scale as .83. However, Abualrub et al. acknowledged that the psychometric characteristics of the Arabic version of the MMSS-31 were not examined prior to data collection. As a result, Abualrub et al. accepted that the results of their study may not relate directly to those of similar studies conducted in other languages. In a related study of support, satisfaction and retention of Jordanian nurses in private and public hospitals (Abualrub et al., 2009) reported Cronbach’s alpha coefficients for the Arabic MMSS-31 as .88 for their private hospital sample and .83 for nurses in public hospitals.

El-Jardali, Dimassi, Dumit, Jamal, and Mouro (2009) used the MMSS-31 to investigate the intention to leave of nurses in Lebanon. The MMSS-31 was administered in Arabic, English, and French. The respondents used four Likert-type scale categories to rate MMSS-31 items (the neutral category in the five-category scale was excluded as recommended by Abu Ajamieh et al., 1996). El-Jardali et al. (2009) reported acceptable Cronbach’s alpha levels for five of Mueller and McCloskey’s (1990) eight sub-dimensions (range = .75 for interaction to .83 for satisfaction with scheduling). Low reliability coefficients were reported for extrinsic rewards (.66); coworkers (.64); and balance of family/work, Pearson’s r = .44). El-Jardali et al. (2009) did not examine whether language of completion biased nurses’ responses to MMSS items.

Al-Enezi, Chowdhury, Shah, and Al-Otabi (2009) used Mueller and McCloskey’s original five Likert-type rating scale categories and 21 MMSS items (MMSS-21) to investigate job satisfaction among nurses with multicultural backgrounds in Kuwait. The Cronbach’s alpha coefficient for the 21-item scale was .90. English and Arabic versions of the MMSS-21 were compiled into a single questionnaire for the convenience of Arabic and non-Arabic speaking respondents. EFA identified five factors with Cronbach’s alpha coefficients ranging from .50 for “extrinsic rewards” to .86 for “professional opportunities.” Al-Enezi et al. (2009) reported that expatriate nurses were dissatisfied with “extrinsic rewards” for two reasons: differences between anticipated and actual earnings in Kuwait; and dissatisfaction with the considerably higher incomes and other benefits enjoyed exclusively by Kuwaiti nurses. Al-Enezi et al. (2009) did not examine the MMSS-21 for possible item bias.

El-Jardali, Alameddine, et al. (2013) used 25 MMSS items (MMSS-25) to investigate the retention of nurses in underserved areas in Lebanon. The six MMSS-31 items excluded (Table 1, footnote) were omitted because they were irrelevant to nurses working in underserved areas. Nurses used four Likert-type rating scale categories to respond to the MMSS-25. The MMSS-25 was administered in Arabic. The study was part of a wider investigation of nurses’ intentions to stay in underserved areas of Yemen, Jordan, Lebanon, and Qatar (El-Jardali, Murray, et al., 2013). In Qatar, the MMSS-25 was administered in English. The MMSS-25 data collected in Lebanon and Qatar provided the data set for our methodological study.

Our review of studies of nurse satisfaction conducted in the Middle East demonstrates that researchers have used different sets of MMSS-31 items, different rating categories, different languages, and different populations to investigate nurses’ job satisfaction in the region, and that the measurement characteristics of the resulting scales have not been sufficiently examined. Consequently, investigators intending to conduct future studies of nurses’ job satisfaction in Middle Eastern countries need guidance on the following important methodological questions:

1. How many rating scale categories should be used when conducting studies with MMSS items in Middle Eastern countries?
2. Which MMSS items are most productive of measurement when used in the region?
3. Are MMSS items biased when administered to Arabic speaking local nurses or in English to internationally educated nurses in the region?
4. Irrespective of language of administration (Arabic or English) are responses to MMSS items biased by the income level of the country of employment?
5. Within the limitations of the number of MMSS items used in studies of nurse satisfaction in Middle Eastern countries, which reported factor structure for the MMSS is most plausible?
6. What do answers to the first five questions suggest for investigators planning to use MMSS items for future studies of nurse job satisfaction in Middle Eastern countries?

The purpose of our article is to contribute to answering these questions.

Method

Ours is a methodological study that used the Rasch rating scale model (Andrich, 1978) and confirmatory factor analysis (CFA) to examine the psychometric characteristics of the MMSS-25 and item subsets in a secondary analysis of data.
collected from registered nurses in Lebanon and internationally educated nurses in Qatar. Eligibility criteria for the parent study (El-Jardali, Murray, et al., 2013) were as follows: registered nurses of any nationality with at least 12 months in their current employment who were able to complete the study questionnaire in Arabic if working in Lebanon, in English if working in Qatar. The Arabic version of the MMSS-25 was prepared following consultation with a panel of registered nurses in Lebanon and forward and back translation by independent translators.

Participants

The characteristics of the respondents to the MMSS-25 are described in Table 2.

Instrumentation

We used the items listed in Table 1 and the data collected in Lebanon and Qatar (El-Jardali, Murray, et al., 2013) to investigate the measurement performance of the MMSS-25. We examined response category ordering (whether use of rating scale categories increased monotonically as expected), item difficulty (the ordering of items from the least difficult to the most difficult to endorse), dimensionality or fit (whether the items measured a single latent construct), targeting (whether the difficulty of endorsing MMSS items was appropriate for the sample), separation (whether the MMSS-25 was sensitive enough to distinguish between nurses with high and low levels of satisfaction), reliability (whether there is a high probability that respondents with high satisfaction measures were more satisfied than those with lower measures), and differential item functioning (DIF; whether MMSS-25 items were easier to endorse for nurses in Lebanon or in Qatar). We concluded our analysis by using CFA to compare the fit of five factor structures reported in the literature to our data.

The Rasch Measurement Model

The Rasch Measurement Model was developed by the Danish mathematician Georg Rasch (1980) as a special case of the general linear model (GLM). The Rasch model specifies the probability of correct responses to test items and strength of endorsement of rating scale items. The ability of respondents (aptitude for selecting correct answers/tendency to endorse rating scale items) and the difficulty of items (probability of a correct answer or of endorsing a particular category in a rating scale) are modeled on a continuous latent variable measured in logits (additive log-odds units of equal measurement). For Rasch analyses of responses to polytomous items, the data are fit to the following mathematical model: \[
\log e (P_{nij}/P_{n(i-1)}) = B_n - D_i - F_j, \]
where \( \log e \) is the natural logarithm of the probability \( P_{nij} \) of person \( n \) of ability \( B_n \) endorsing category \( j \) in response to a scale item of difficulty \( D_i \), as opposed to the probability \( P_{n(i-1)} \) of the person endorsing the next lowest category \( (j - 1) \). For example, if \( j \) is the endorsement of “4” on a 5-point Likert-type scale, \((j - 1)\) would be the endorsement of “3,” the adjacent lower category. In this formulation, the parameter \( F_j \) defines the same rating scale structure for all items (Linacre, 2012).

Unlike raw scores for a test instrument or questionnaire that have unknown intervals, the Rasch model enables investigators to calibrate item difficulty and measure personal ability using standardized intervals of measurement along a common continuum or latent trait.

More specifically, Rasch analysis examines how well items in a test or rating scale contribute to the useful measurement of an assumed one-dimensional latent variable (Rasch, 1980). Rasch fit statistics indicate how well respondents and their responses fit the response pattern predicted by the Rasch measurement model. Infit and outfit statistics are calculated as chi-square values that range from zero to infinity. Infit and outfit values for an item that perfectly matches the Rasch model have a mean square value (MNSQ) of 1. Items with MNSQ values greater than 1 overfit the model because they lack precision. Items with values less than 1 are too predictable and may not achieve successful measurement. (Linacre, 2012). In Rasch analysis, the dimensionality of rating scales and subscales can

### Table 2. Characteristics of Study Sample (N = 1,322).

| Category                        | n   | %   |
|---------------------------------|-----|-----|
| **Country**                     |     |     |
| Lebanon                         | 859 | 64.9|
| Qatar                           | 463 | 35.0|
| **Nationality**                 |     |     |
| Lebanese                        | 859 | 64.9|
| Qatari/Other Middle Eastern     | 69  | 5.2 |
| Indian                          | 196 | 14.8|
| Filipino                        | 147 | 11.1|
| Other                           | 51  | 3.8 |
| **Gender**                      |     |     |
| Female                          | 1,069 | 80.9|
| Male                            | 233  | 19.1|
| **Age**                         |     |     |
| 18-24                           | 219  | 16.6|
| 25-30                           | 448  | 33.8|
| 31-40                           | 401  | 31.4|
| 41-50                           | 158  | 12.7|
| ≥50                             | 52   | 3.9 |
| **Marital status**              |     |     |
| Married                         | 726  | 54.9|
| Unmarried                       | 596  | 45.1|
| **Highest nursing qualification** |   |     |
| Master’s degree                 | 247  | 18.6|
| Bachelor’s degree               | 197  | 14.9|
| Diploma/other sub-degrees       | 878  | 66.4|
| **Employment**                  |     |     |
| Community nursing               | 98   | 7.4 |
| Hospital service                | 1,224 | 92.6|
be checked by principal component analysis (PCA) of residuals. The Rasch measurement model is the dimension of first comparison. Second and subsequent dimensions with eigenvalues greater than 2 suggest second and higher order dimensions that need investigation (Linacre, 2012).

Software

Winsteps 3.80 (Linacre, 2013) was used to conduct the Rasch analyses reported in this article. Version 3.90 (Linacre, 2015) is available from Windows.com.

Results

Response Ordering

We examined response ordering for the six-category rating scale used in the El-Jardali, Murray, et al. (2013) study (0 = no response, 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree, 5 = not important to me). With responses in the “not important to me category” rescored as 0s, response categories for 23 of the 25 MMSS items were disordered. The exceptions were Item 15 “The physicians you work with” and Item 21 “Opportunities for career advancement.” Categories 0 and 1 were disordered for 21 of the 23 disordered items. Categories 1 and 2 were disordered for Item 1 “Salary” and Item 11 “Maternity leave time.” Response category disordering was corrected by excluding null and not important to me responses. All of the following results were obtained using the four-category rating scale: 1 = strongly disagree, 2 = disagree, 3 = agree, and 4 = strongly agree.

Item Difficulty

The measurement characteristics of the MMSS-25 were explored using the corrected four-category rating scale described in the previous paragraph. Respondents’ mean raw satisfaction scores were in the range 2.01 (n = 830) for Item 12 “Child care facilities” to 3.05 (n = 1,162) for Item 14 “Your nursing peers.” When measured with the Rasch dimensioned latent satisfaction variable, the mean measures ranged from −1.63 logits for Item 14 to 1.23 logits for Item 12, with higher logit values indicating more difficulty in endorsing the satisfied and very satisfied rating categories. Respondents expressed their relative satisfaction with Item 3 “Benefit package” (M = 1.19 logits), Item 1 “Salary” (1.08 logits), and Item 10 “Compensation (pay) for weekend work” (M = 1.01 logits). The nurses endorsed higher categories of satisfaction for the items measuring perceptions of their relationships with colleagues: Item 14 “Your nursing peers” (M = −1.63 logits; Item 13 “Satisfaction with your immediate supervisor (head nurse or facility manager)” (M = −1.15 logits); and Item 15 “The physicians you work with” (M = −1.05 logits). With respect to the Mueller and McCloskey (1990) categorization of items, respondents expressed least satisfaction “safety needs” items, somewhat more satisfaction with “psychological needs” items, and most satisfaction with “social needs” items (Figure 1).

Figure 1 shows the relationship between measured nurse satisfaction and difficulty of endorsing MMSS items. The numbers in the first column (bottom to top) from −1 to 3 are measures of nurse satisfaction in logits. The distribution of satisfaction measures for this sample is flatter than a normal distribution and positively skewed. The difficulty of endorsing satisfaction with a MMSS item is shown in the column on the right. Only nurses with satisfaction scores 1.3 standard deviations above the sample mean were satisfied with Item 12 “Child care facilities” and Item 3 “Benefits package.”

Nurses with satisfaction scores 1 standard deviation above the mean measure for the sample expressed satisfaction with Item 10 “Compensation (pay) for weekend work” and Item 11 “Salary.” The easiest item for the nurses to express satisfaction with was Item 14 “Your nursing peers.” Nurses with levels of satisfaction at or above −1.43 standard deviations from the sample mean were able to express satisfaction with this item. The hierarchy of item difficulty can be seen by reading down the column on the right from the most difficult items (Items 12 and 13) to the easiest item (Item 14). More satisfaction with coworkers and interaction opportunities was expressed than with scheduling or with professional opportunities. Least satisfaction was expressed with the extrinsic needs of child care and benefits.

Dimensionality

We examined the dimensionality of the MMSS-25. The measures explained a relatively poor 39.5% of the raw variance (Fisher, 2009). MMSS-25 items explained 12.2% of the raw variance, 1.62 times more than explained by the first contrast. The eigenvalue of 3.1 for the first contrast suggested the presence of a non-Rasch dimension with the strength of approximately three items. Items 22, 21, 24, 25, and 23 loaded ≥.4 (.48-.40) at one pole of the secondary dimension, in contrast to Items 8, 4, 6, 2, and 9 that loaded −.47 to −.41 at the opposite pole.

The first contrast may differentiate responses by types of need, with Items 22, 21, 24, 25, and 23 eliciting responses that reflect psychological needs and Items 8, 4, 6, 2, and 9 reflecting safety needs. We noted that Item 11 “Maternity leave time,” Item 12 “Child care for employees’ children at facility,” and Item 13 “Your head nurse or facility manager” had infit and outfit MNSQ values that exceeded 1.3; our cutoff value for under fitting items. None of the items had MNSQ values less than 0.6, our cutoff value for overly predictable items.

Our analysis indicated the possible existence of a second non-Rasch dimension with a strength of approximately two items (eigenvalue = 2.0) that explained 5.0% of the unexplained variance. Items 24, 22, and 25 loaded ≥.4 (.58-.41) at one pole of this dimension, in contrast to Items 18, 19, and 17 that loaded −.54 to −.45 at the opposite pole. This dimension may distinguish items that measure satisfaction with self-esteem needs from those that measure satisfaction with affiliation needs.
Items that loaded positively and negatively on the contrasts derived from the principal components analysis (PCA) of residuals were used to estimate differential person functioning using independent t tests. Since >5% of the t-test values (9.6% for the first contrast and 10.7% for the second contrast were outside the range ±1.96 multidimensionality was confirmed.

Removing the three items with infit and outfit MNSQ values greater than 1.3, Items 11 (1.37, 1.38), 12 (1.45, 1.33), and 13 (1.33, 1.39) increased the proportion of raw score variance explained by the Rasch dimension to 39.8%. For all practical purposes, the strengths of the first and second contrasts remained unchanged (eigenvalues = 3.0, 2.0). Therefore, we retained Items 11 to 13.

**Targeting**

The MMSS-25 items were a very good fit for 85% of the sample; person mean discrimination −0.14 logits, model root mean square error (RMSE) 0.36 logits, item mean discrimination 0.00, model RMSE 0.05 logits (Figure 1). The nurses ranged in satisfaction from −6.56 to 7.47 logits. The mean difficulty of the items was slightly greater than the mean satisfaction of the nurses −.14 logits. No items matched the 8.6% of the nurses with the highest levels of satisfaction or the 6.4% of least satisfied nurses.

**Person Separation and Reliability**

The WINSTEPS person separation index indicates the sensitivity of a measure to distinguish between higher and lower measures among a study sample (Linacre, 2012). The MSS-25 was sensitive enough to distinguish approximately four strata of nurse satisfaction, person reliability index 0.87. The person raw score test reliability (estimated traditional test–retest reliability Cronbach’s α) for the MMSS-25 was approximately .95.

**DIF**

DIF occurs when respondents with the same ability level have different probabilities of endorsing response categories.
of the same strength. We examined whether nurses in Qatar \((n = 427)\) and nurses in Lebanon \(n = 786\) with the same Rasch measures for job satisfaction responded differently to MMSS-25 items. None of the items was biased in favor of nurses in either country, confirming that the MMSS items were not biased irrespective of whether nurses responded in Arabic or English, nor were items biased by nurses’ country of employment: middle income (Lebanon) or high income (Qatar).

**Discussion**

**Rating Scale Categories**

Although the MMSS is a well-established and occupation specific measure of nurses’ job satisfaction, its psychometric properties when used with samples outside Western countries have not been critically examined. Until now, no attempt has been made to determine whether MMSS items administered in English or Arabic in countries in the Middle East are biased in favor of local or internationally educated nurses. Furthermore, in the absence of studies conducted in middle-income and high-income countries in the region, there has been a gap in knowledge about whether the MMSS is uniformly productive of measurement when used outside North America. Our results do not support Abu Ajamieh et al.’s (1996) recommendation that a four-category rating scale should be used when investigating nurses’ job satisfaction in the region. Reducing the rating categories to four for our Rasch analyses contributed to the relatively small proportion of variance explained by MMSS-25 items.

We recommend that investigators conducting future studies of nurses’ job satisfaction in Middle Eastern countries rely on the four-category scale (strongly disagree, disagree, agree, and strongly agree) only when they have added items that are easier and more difficult to endorse than those in the MMSS-25, or remove items clustered around the center of the distribution for item difficulty. These strategies are required to increase the proportion of variance explained by MMSS items, irrespective of whether the scale is administered in English or Arabic.

**DIF**

Our results show that responses to MMSS items concerned with salary and benefits are not biased in favor of nurses in either high- or middle-income countries in the Middle East. This finding is counterintuitive because nurses’ salaries in Qatar are significantly higher than the salaries paid to nurses in Lebanon. The benefits available to nurses in Qatar are similarly more generous than those available to nurses in Lebanon. The reason we did not find the MMSS salary and benefit items biased in favor of either Lebanese nurses or internationally educated nurses in Qatar is because we excluded nurses with extreme high and low satisfaction scores from our DIF analyses. When we repeated our DIF analyses with the deleted records reinstated, we found medium to strong evidence of item bias. We found that Item 4 “Hours that you work” was significantly more difficult for nurses in Lebanon to express satisfaction with compared with those in Qatar (effect size = 1.25 logits, \(t = 12.6, df = 966, p < .01\)). A moderate to large DIF value was found for Item 6 “Opportunities to work (consecutive) straight days” (effect size = 0.72 logits, \(t = 7.23, df = 966, p < .001\)), which was also more difficult for nurses in Lebanon to endorse. Conversely, Item 7 “Opportunity for part-time work” (effect size = −0.78 logits, \(t = 7.64, df = 740, p < .001\)) and Item 13 “Satisfaction with your immediate supervisor (head nurse or facility manager)” (effect size = −0.98 logits, \(t = 9.19, df \text{ inf., } p < .001\)) were more difficult for nurses in Qatar to endorse. We suggest that investigators measuring nurse job satisfaction in Middle Eastern countries consider the possibility of bias in MMSS items and consider either dropping seemingly biased items from their data set or including additional satisfaction items to offset the possibility of DIF.

**Dimensionality**

The literature we reviewed indicates that an eight-factor model for the MMSS and a three sub-dimension model (Mueller & McCloskey, 1990), a seven-factor model (Tourangeau et al., 2006), a five-factor model (Al-Enezi et al., 2009), and a four-factor model (Abu Ajamieh et al., 1996) are plausible (Table 3).

We ran a CFA using PCA with varimax rotation to examine the fit of seven of (Mueller & McCloskey, 1990) eight-factor models to our data. The analysis was restricted to data on the 25 items common to the El-Jardali, Murray, et al. (2013) study and the MMSS-31. We ran a second CFA analysis to examine the fit of the three sub-dimensions under which Mueller and McCloskey (1990) subsumed their eight-factor model. We examined the fit of five of Tourangeau et al.’s (2006) seven-factor model. This analysis was restricted to the 13 MMSS items common to the El-Jardali, Murray, et al. (2013) study and the Tourangeau et al. study. We ran additional CFA analyses to examine the fit of the five-factor model proposed by Al-Enezi et al. (2009) and the four-factor model reported by Abu Ajamieh et al. (1996).

The results of our analyses are shown in Table 4. The chi-square values can be ignored due to the large size of our sample. Comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) values indicate that a five-factor model derived from the Tourangeau et al.’s (2006) seven-factor model provided a close fit for 13 MMSS
### Table 3. Comparison of Five Reported Factor Structures for MMSS Items.

| 7/8 factors | 3 factors | 3/4 factors | 4/5 factors | 5/7 factors |
|-------------|-----------|-------------|-------------|-------------|
| Mueller and McCloskey (1990) 25/31 items | McCloskey (1974) Mueller and McCloskey (1990) 25/31 items | Abu Ajamieh, Misener, Haddock, and Gleaton (1996) 19/31 items | Al-Enezi, Chowdhury, Shah, and Al-Otabi (2009) 14/21 items | Tourangeau, McGillis Hall, Doran, and Petch (2006) 13/31 items |
| Control/Responsibility Items: 21, 25 (23, 31) | Psychological needs Items: 20, 21, 25 (21, 23, 31) | Control over work environment No items in MMSS-25 | Professional opportunities No items in MMSS-25 | Professional opportunities: Items: 13, 25 (13, 31) |
| Professional opportunities: Items: 13-19, 23, 24 (13-19, 24, 25, 26) | Social needs Items: 7, 21 (7, 23) | Personal achievement Items: 17, 18, 19, 23 (25) | Praise/Recognition Excluded | Coworkers Items: 14, 15 |
| Interaction Items: 16, 17, 18, 19 | Safety needs Items: 1-12 | Extrinsic rewards Items: 1, 10, 11 | Scheduling satisfaction Items: 2, 4, 5 | Extrinsic rewards Items: 1, 2, 3 |
| Extrinsic rewards Items: 1, 2, 3 | Scheduling satisfaction Items: 4, 5, 6, 8, 9, 10 | Personal achievement Items: 1, 2, 4, 8, 10, 11, 12 | Interaction Items: 14, 15, 16, 17, 18, 19, 20, 22, 23, 24 (21, 24, 25, 26) | Co-workers Items: 15 (15; excluded) |
| Family/Work balance Items: 7, 11, 12 | |

Note. MMSS-25 items only included in the comparison. Numbers in parentheses reference MMSS-31 item numbers (Table 1). MMSS = Mueller/McCloskey Nurse Job Satisfaction Scale. Shaded areas highlight dimensions and items excluded from comparison due to variations in scale composition or dimensionality.

### Table 4. Comparative Fit Indices for Five CFA Models for the MMSS-25.

| 7/8 factors | 3 factors | 3/4 factors | 4/5 factors | 13 items: 5/7 factors |
|-------------|-----------|-------------|-------------|-----------------------|
| Mueller and McCloskey (1990) | McCloskey (1974) Mueller and McCloskey (1990) | Abu Ajamieh, Misener, Haddock, and Gleaton (1996) | Al-Enezi, Chowdhury, Shah, and Al-Otabi (2009) | Tourangeau, McGillis Hall, Doran, and Petch (2006) |
| $\chi^2$ | 1,205.596 | 2,453.256 | 1,548.117 | 982.014 | 230.222 |
| df | 231 | 272 | 149 | 71 | 55 |
| $\chi^2/df$ | 5.21 | 9.02 | 10.39 | 13.83 | 4.18 |
| p value | .001 | .001 | .001 | .001 | .001 |
| CFI | .0896 | .0787 | .0795 | .0832 | .0951 |
| TLI | .0876 | .0765 | .0765 | .0785 | .0931 |
| AIC | 56,881.124 | 59,720.858 | 45,546.244 | 33,425.375 | 32,489.185 |
| BIC | 57,234.043 | 59,991.941 | 45,755.950 | 33,599.277 | 32,673.317 |
| Adjusted BIC | 57,014.869 | 59,823.590 | 45,625.716 | 33,491.278 | 32,558.966 |
| RMSEA | 0.059 | 0.081 | 0.087 | 0.102 | 0.051 |
| 90% CI | [0.053, 0.062] | [0.078, 0.084] | [0.083, 0.091] | [0.097, 0.108] | [0.044, 0.058] |
| p value RMSEA ≤ .05 | .001 | .001 | .001 | .001 | .401 |
| SRMSR | 0.046 | 0.058 | 0.059 | 0.059 | 0.033 |

Note. CFA = confirmatory factor analysis; MMSS = Mueller/McCloskey Nurse Job Satisfaction Scale; CFI = cumulative fit index; TLI = Tucker–Lewis index; AIC = Akaike information criterion; BIC = Bayesian information criterion; Adjusted BIC = sample size adjusted Bayesian information criterion; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval; SRMSR = standardized root mean square residual.

Items (MMSS-13). However, we could not investigate factor structure invariance across samples because we did not have data on all 31 items included in Mueller and McCloskey’s (1990) eight-factor model. Neither could we examine the fit of the full Tourangeau et al. (2006) seven-factor model.
In light of these findings, we reexamined the item-person map for the MMSS-25. Figure 1 shows that as many as 14 MMSS-25 items could be redundant. When we examined item redundancy among MMSS-13 items, we identified five potentially redundant items. The items for possible deletion were Item 3 “Benefits package,” Item 5 “Flexibility in scheduling your hours,” Item 6 “Opportunities to work (consecutive) straight days,” Item 8 “Weekends off per month,” Item 9 “Flexibility in scheduling time-off,” and Item 19 “Professional interactions with other disciplines.” We retained all five items because they did not duplicate the meaning of the items with which they shared similar Rasch item measures. We suggest that investigators conducting studies of nurse job satisfaction in Middle Eastern countries consider using the MMSS-13 as a core set of items with which to screen nurses for work satisfaction.

The MMSS-13 has acceptable psychometric characteristics: model fit mean-square values range from 0.80 to 1.26, with the exception of Item 13 “Satisfaction with your immediate supervisor (head nurse or facility manager),” outfit MNSQ 1.38, infit MNSQ 1.43; mean non-extreme person measure −.27, model error 0.49; person reliability index .80, estimated Cronbach’s alpha .91; person index 2.03, sufficient to distinguish among three strata of satisfaction (low satisfaction, moderate satisfaction, higher satisfaction). However, the ceiling effect of the MMSS-13 excluded 8.4% of the most satisfied nurses in our sample. The floor effect of the scale excluded a further 16% of our sample. Consequently, our study supports (Mueller & McCloskey, 1990) suggestion that investigators should add items independent of the MMSS-31 according to their specific research objectives.

When we excluded the 12 non-MMSS-13 items, the proportion of raw variance explained by the measures increased from 39.5% to 42.5%, which falls short of the 50% required (Fisher, 2007) to achieve more than a poor quality of measurement. However, 42.5% of explained raw variance in Rasch analysis is sufficient for such practical purpose as initial screening of respondents if the first contrast to the Rasch dimension does not indicate an obvious non-Rasch dimension. The MMSS-13 items explained 12.6% of the raw variance in our data. The proportion of variance explained by the measures can be improved by including participants with a wider range of satisfaction measures, by sampling nurses in a wider range of health care facilities, by conducting multisite studies or by including nurses with more diverse qualifications. In addition to including more items targeted at nurses with lower and higher levels of satisfaction, the proportion of raw variance explained can be increased by including items known to be relevant to nurses in the sample rather than relying on items selected from the (Mueller & McCloskey, 1990) 31-item set.

The empirical usefulness of our shorter scale in Middle Eastern countries needs to be established. Whether investigators should use the MMSS-31 or our shorter version depends on the purpose of their studies. For initial assessment of nurse job satisfaction and for collecting time series data, the shorter scale may be preferred. However for more complex studies, including those that involve causal analyses in which hypotheses are tested about factors that influence nurses’ job satisfaction, the most productive approach would be to combine the MMSS-13 scale with hypothesis-specific items.

More work needs to be done on the dimensionality of the MMSS as a whole and on shorter versions because Maslow’s (1954) hierarchy of human needs does not provide a sufficiently predictive theory for understanding occupation specific job satisfaction. Whereas the hierarchy of needs offers a general framework for understanding human motivation, the assumptions on which it is based, such as the discreteness of levels of human needs, are untenable because it is incapable of capturing situation-specific causal factors. When considering the use of job satisfaction items in samples of local nurses working in Middle Eastern countries, investigators should consider the possibility of item bias among MMSS items. We found that Item 13 “Satisfaction with your immediate supervisor (head nurse or facility manager)” could be biased in favor of nurses working in culturally homogeneous work settings, such as Lebanese nurses working in Lebanon (effect size = −1.16 logits, \( t = 10.6, df = 991, p < .001 \)).

We readily acknowledge the limitations of our study. Our Rasch and CFA analyses were conducted as a secondary analysis of data on only 25 of Mueller and McCloskey’s (1990) 31 MMSS items. Furthermore, the response categories we used were different from those used by Mueller & McCloskey (1990). Despite these limitations, we have established that the MMSS-31 is not a suitable measure for investigating nurses’ job satisfaction in Middle Eastern countries. The scale explains a relatively low proportion of raw score variance in nurses’ job satisfaction scores, and up to 45% of MMSS items may be redundant depending on the setting and context. Whereas the MMSS is clearly not unidimensional, further studies are required to establish its factor structure. Four of five factor models tested were plausible but not a close fit to our data. We have reported on a close fitting factor structure for MMSS-13 scale items, which are suitable for administering in Arabic or English to nurses in the region. At the same time, we have confirmed that the MMSS-13 provides a psychometrically acceptable core set of items to initially screen nurses in the Middle Eastern countries for levels of job satisfaction, irrespective of whether they work in middle- or high-income countries in the region.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The primary study was funded by the Alliance for Health Policy and Systems Research. The author(s) received no financial support for the research and/or authorship of this article.
References
Abu Ajamieh, A. R., Misener, T., Haddock, K. S., & Gleaton, J. U. (1996). Job satisfaction correlates among Palestinian nurses in the west bank. *International Journal of Nursing Studies, 33*, 422-432.
Abualrub, R. F., Omari, F. H., & Abu Al Rub, A. F. (2009). The moderating effect of social support on the stress-satisfaction relationship among Jordanian hospital nurses. *Journal of Nursing Management, 17*, 870-878.
Al-Enazi, N., Chowdhury, R. I., Shah, M. A., & Al-Otabi, M. (2009). Job satisfaction of nurses with multicultural backgrounds: A questionnaire survey in Kuwait. *Applied Nursing Research*, 22, 94-100.
Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573.
Babiarczyk, B., Fras, M., Ulman-Wlodarz, I., & Jarosova, D. (2014). The level of job satisfaction and its relation to midwives’ subjective quality of life. *Medycyna Pracy, 65*, 99-108.
Burns, T. (1969). Motivation and human relations. *American Journal of Medical Technology, 35*, 12-23.
El-Jardali, F., Alameddine, M., Jamal, D., Dimassi, H., Dumit, N. Y., McEwen, M. K., . . . Murray, S. F. (2013). A national study on nurses’ retention in healthcare facilities in underserved areas in Lebanon. *Human Resources for Health, 11*, Article 49.
El-Jardali, F., Dimassi, H., Dumit, N., Jamal, D., & Mourou, G. (2009). A national cross-sectional study on nurses’ intent to leave and job satisfaction in Lebanon: Implications for policy and practice. *BMC Nursing, 8*, Article 3.
El-Jardali, F., Murray, S. F., Dimassi, H., Jamal, D., AbuAlRub, R., Al-Surimi, K., . . . Dumit, N. Y. (2013). Intention to stay of nurses in current posts in difficult-to-staff areas of Yemen, Jordan, Lebanon and Qatar: A cross-sectional study. *International Journal of Nursing Studies, 50*, 1481-1494.
Fisher, W. (2007) Rating scale instrument quality criteria. *Rasch Measurement Transactions, 21*(1), 1095.
Gurková, E., Haroková, S., Džuka, J., & Žiaková, K. (2014). Job satisfaction and subjective well-being among Czech nurses. *International Journal of Nursing Practice, 20*, 194-203.
Gurková, E., Soosová, M. S., Haroková, S., Žiaková, K., Šerfelová, R., & Zamboriová, M. (2013). Job satisfaction and leaving intentions of Slovak and Czech nurses. *International Nursing Review, 60*, 112-121.
Hickson, J. (2013). New nurses’ perceptions of hostility and job satisfaction: Magnet® versus non-magnet. *Journal of Nursing Administration, 43*, 293-301.
Linacre, J. M. (2012). *Winsteps® Rasch measurement computer program user’s guide*. Beaverton, OR: Winsteps.
Linacre, J. M. (2013). *Winsteps® (Version 3.80.0) [Computer software]*. Beaverton, OR: Winsteps.com.
Linacre, J. M. (2015). *Winsteps® (Version 3.90.0) [Computer software]*. Beaverton, OR: Winsteps.com.
Maslow, A. H. (1954). *Motivation and personality*. New York, NY: Harper.
McCloskey, J. (1974). Influence of rewards and incentives on staff nurse turnover rate. *Nursing Research, 23*, 239-247.
McCloskey, J. C., & McCain, B. E. (1987). Satisfaction, commitment and professionalism of newly employed nurses. *Journal of Nursing Scholarship, 19*, 20-24.
McGilton, K. S., Tourangeau, A., Kavec, C., & Wodchis, W. P. (2013). Determinants of regulated nurses’ intention to stay in long-term care homes. *Journal of Nursing Management, 21*, 771-781.
Mrayan, M. T. (2005). Nurse job satisfaction and retention: Comparing public to private hospitals in Jordan. *Journal of Nursing Management, 13*, 40-50.
Mrayan, M. T. (2006). Jordanian nurses’ job satisfaction, patients’ satisfaction and quality of nursing care. *International Nursing Review, 53*, 224-230.
Mueller, C. W., & McCloskey, J. C. (1990). Nurses’ job satisfaction: A proposed measure. *Nursing Research, 39*, 113-117.
O’Brien-Pallas, L., Murphy, G. T., Shamain, J., Xiaoqiang, L., & Hayes, L. J. (2010). Impact and determinants of nurse turnover: A pan-Canadian study. *Journal of Nursing Management, 18*, 1073-1086.
Rasch, G. A. (1980). Probabilistic models for some intelligence and attainment tests. Chicago, IL: University of Chicago. (Original work published 1960)
Tourangeau, A. E., McGillis Hall, L., Doran, D. M., & Petch, T. (2006). Measurement of nurse job satisfaction using the McCloskey/Mueller Satisfaction Scale. *Nursing Research, 55*, 128-136.
Wilson, B., Squires, M., Widerg, K., Cranley, L., & Tourangeau, A. (2008). Job satisfaction among a multigenerational nursing workforce. *Journal of Nursing Management, 16*, 716-723.

Author Biographies
Michael Clinton RN, BA(Hons), MSc, PhD is professor of Nursing at the American University of Beirut. He has extensive experience in academic leadership and health care administration. He is vice-chair of the Social and Behavioral Sciences Institutional Review Board. He has a strong research track record in mental health, mental health workforce development, and ambulatory care. His current research interest is the conceptualization and measurement of phenomena relevant to nursing and ethical nursing practice.

Nuhad Yazbik Dumit RN, MA, PhD has more than 30 years’ experience as a university teacher at the Hariri School of Nursing, American University of Beirut, where she is convener of the Bachelor of Science in Nursing Program, a member of the AUB Board of Graduate Studies and course coordinator for courses in health service development and quality management in teams in the Master of Science in Nursing Program. Her research work focuses on nursing workforce issues, including: nurse migration; nurse workforce recruitment and retention; sickness and absenteeism rates; and, assessment and continuing education of bedside nurses.

Fadi El-Jardali MPH, PhD is associate professor and director of the Knowledge to Policy (K2P) Center, Co-Director of the Center for Systematic Reviews in Health Policy and Systems Research (SPARK), and co-director of the Nodal Institute at the American University of Beirut. He is a research program director for the Research, Advocacy and Public Policy (RAPP) Program in the Arab World at Issam Fares Institute for Public Policy and International Affairs at AUB; and associate professor at the Department of Clinical Epidemiology & Biostatistics, McMaster University in Canada. His research activities and publications are focused on health policy and systems including policymaking, knowledge translation, human resources for health and quality of care. He has led major programs and initiatives including national, regional and international studies. He is the author of 50+ original scientific publications.