Validity, Reliability and Readability of Nurse Emotional Labor Scale

Jestoni D. Maniago, Ph.D., DNSc
Department of Nursing, College of Applied Medical Sciences, Majmaah University, 11952, Saudi Arabia

*Corresponding author: Jestoni D. Maniago
DOI:10.21276/sjnhec.2019.2.3.2

Abstract

Emotionally demanding patient interactions, when coupled with organizational directives to focus on exemplary customer service, can prompt nurses to express feelings and emotions that are not genuine and involves a great deal of emotional labor. This study aimed to develop and validate a tool that measures emotional labor of nurses. The measured variables were taken from the underlying structures of nurse emotional labor based on scholarly literature and respondent’s feedback. Based on regular contact with patients and families, it is more difficult for nurses than for other healthcare professionals to escape emotional labor, and recharge their “batteries” for the next encounter.

Keywords: Emotional labor; nursing; readability; reliability; validity.

INTRODUCTION

When considering a career, it’s normal for people to think about what degree of skill(s) a particular job requires but they rarely consider, or sometimes even realize the kind of emotional skills they need in various occupations. In 1983, the Sociologist Arlie Hochschild coined the term emotional labor, in her book “The Managed Heart” [1]. It refers to professions which require individuals to express certain emotions, regardless of how they feel. The point is employees are expected to display the right emotions to properly perform their jobs. The cashier or waitress is expected to smile at you and say, “Have a good day!” even if they are not in a good mood and could not care less how the rest of your day goes.

Every job requires a certain level of emotional labor. For example, a good businessman needs to control his temper during a meeting, a good flight attendant needs to be kind and respectful toward passengers and a good nurse needs to be empathic towards patients. Employees do not often consider the emotional aspects of a job as an integral part of their job description, but this is a very vital element of any organisation that requires will-power and determination to be productive. This can be as draining as physical or mental work.

In 2009, Gray stated that most organizations that involve more contact with humans exact some degree of emotional labor, but it is thought to be most fundamental to nursing [2]. Nurses have contact with patients for prolonged periods of time and are required to deliver compassionate care often under demanding interpersonal circumstances. They are expected to regulate their own emotional reactions while assuaging the fear and distress of patients. Nursing has a set of shared unspoken rules concerning the emotions that should be displayed and those that should not. When performed for altruistic reasons, emotional labor can be a source of satisfaction; if not, it can threaten the well-being of nurses. For example, nurses who have difficulty managing the emotional demands of the job tend to report elevated levels of emotional exhaustion, emotional detachment and general fatigue [3].

Every human is emotional. Human behaviors, thoughts and actions can be affected by how the environment is felt. Emotion is one key value displayed by members of every organization. This perspective is fundamental in health care organizations and plays an important role in helping professions like nursing. Nursing requires intensive emotional, physical and spiritual work, but it is also very rewarding and satisfying. Nurses have constant emotional involvement with patients and their relatives. They have to listen to their complaints and deal with helplessness, suffering and death associated issues. Tension and turmoil come within the territory of being a nurse.

The need for enhanced training and support to help nurses manage the emotional demands of their work has to be recognized. To perform such interventions, there must be an insight into the implications of emotional labor and the factors that...
might protect nurses from the negative effects of the emotional demands they encounter. Although there were several tools [4] to measure emotional labor, there is no specific tool to measure emotional labor of nurses. Thus, the purpose of this study is to develop that tool and take inventory to adequately measure nurse emotional labor and to examine its psychometric properties.

METHODS
Population and Sample of the Study
The target population was selected through Purposive Sampling Technique, in which the researcher decides what needs to be known and deliberately choose the sample. That is, the researcher determines the type of subjects that will participate in the study. Among the 2,330 nurses working in the selected tertiary hospitals in the Philippines, there were 1440(64.6%) females and 790(35.4%) males. The researcher planned on having a sample size of N=2400 for the first drafted questionnaire. However, since two items were deleted during the content validity index (CVI), the final demographic profile of the respondents (N = 2230) responds to the 22-point item of Nurse Emotional Labor Scale (NELS) during field testing. The researcher applied the rule of thumb that there should be at least 10 observations per item [5]. Since NELS has a total of 22 items (questions), the respondents must be equal or more than 220, which was met in the sampling process of this study.

Procedure
Two guidelines [6, 7] provided by different scale developers, which were also widely used by different scale developers, were combined to provide more detailed guidelines in developing and validating the NELS.

The steps of constructing and developing the NELS were as follows:

Phase 1- Background
In this initial step, the objectives of the study were examined. The processes used to select the respondents and determining who may participate in the study was reviewed. A thorough understanding of the problem through literature search and readings was obtained. Good preparation and understanding of Step1 have provided the foundation for initiating Step 2.

Phase 2- Questionnaire Conceptualization and Item Generation
After developing a thorough understanding of the research, the next step was to conceptualize and generate statements/questions for the questionnaire. The researcher adapted the emotional labor scale [4], but since it was not intended for nurses working in the hospital setting, The researcher modified it and were able to develop a new version of Nurse Emotional Labor Scale. The items in this questionnaire are grouped into three: Surface Acting (10 items), Deep Acting (7 items) and Genuine Acting (7 items) with a total of 24 items.

On surface acting, a nurse portrays emotions that are not felt internally, and they mask their true emotions while in the workplace.

On deep acting, a nurse changes internally felt emotions to align with required emotional expressions of the organization.

On genuine acting, a nurse shows natural emotions; in essence, no acting is required.

Phase 3- Format and Data Analysis
In Step 3, the focus was on the selection of questionnaire layout, format, question order, font size, proposed data analysis and appropriate scales of measurement. Scales are devices used to quantify a subject's response to a particular variable. This study used a 5-point Likert Scale (5-Always, 4-Often, 3-Sometimes, 2-Rarely and 1-Never).

Phase 4- Establishing Validity and Reliability
Validity is the extent to which an instrument measures what it is supposed to measure and performs as it is designed to perform. Validation is a process that involves collecting and analyzing data to assess the accuracy of an instrument.

Content validity refers to the appropriateness of the content of an instrument. Content validity was established through the Content Validity Index (CVI) done by the six experts who rated the questionnaire.

The validity of NELS was measured through factor analysis using the Exploratory Factor Analysis (EFA). The aim of factor analysis was attempting to discover the simplest method of interpretation of observed data. Factor analysis of NELS operated on the notion that measurable and observable variables can be reduced to fewer latent variables that share a common variance and are unobservable, is known as reducing dimensionality [8]. EFA was done in NELS by shrinking the data-questions into a smaller data set that is more manageable and more understandable.

EFA consists of seven main steps: reliable measurements, correlation matrix, a factor analysis on principal component analysis, communalities, the number of factors to be retained, factor rotation, use and interpretation of the results.

In the development of NELS, the EFA process first checked whether there were univariate and multivariate normality within the data. The assumptions were next, assuming that the three factors of emotional labor have a relationship on the items. The assumption
is based on the determining factor that there is a linear relationship between the factors and the variables.

Since factor analysis of NELS departs from a correlation matrix, the used variables were first of all measured at (at least) an interval level. Second, the variables were roughly normally distributed which made it possible to generalize the results of the analysis beyond the sample collected [9]. Third, the sample size was taken into consideration, as correlations are not resistant [10], and can hence seriously influence the reliability of the factor analysis [11].

Residuals are computed between observed and reproduced correlations. There are 99 (42.0%) non-redundant residuals greater than absolute value 0.05. It is a rule that a model that has less than 50% of the non-redundant residual is a good fit.

Communality of a variable in NELS represents the proportion of the variance in that variable that was accounted for by all (“common”) extracted factors. Large datasets that consist of several variables were reduced by observing “groups” of variables or factors – that is, factor analysis assembles common variables into descriptive categories. In the factor analysis, the different assumption about the communalities was reflected in a different correlation matrix.

The initial communalities were assumed to be 1 by setting the Eigen value to 1 (values of communality factor loads must be greater than one), which is a criterion for the number of factors to rotate as to Kaiser. It means that an Eigen value less than one implies that the scores on the component would have negative reliability [12].

Most researchers use the Eigen Value Criteria for the number of factors to be extracted. Value of the percentage and Variance Explained method is also used for EFA and the scree test criteria for the selection of factors. In this method, Eigen value is plotted on a graph and factors are selected.

Researcher can determine the statistical power and significance level. For instance, in order to achieve a factor loading of 0.55 with a power of 0.80, a sample of 100 will be needed. In NELS, factor loading was be assessed to determine the strength of the relationships in the interpretation of the factors. The factors was being identified by the largest loadings, but confirmation of the identification of the factors was done by examining the zero and low loadings of the factors. Negatively scored items were being eliminated, and factors load > 0.40 does not exceed and 0.10 will be accepted. The researcher also determined the cut-off for a statistically meaningful rotated factor loading. A general rule to determine the reliability of the factor is to look at the relationship between the individual rotated factor loading and the magnitude of the absolute sample size.

The scree test of 22-point item scale will be used in conjunction with the Eigen values to determine the amount of number of factors to be retained. The number of factors to be retained is the data points that are above the break or the point of inflexion. The scree plot is typically interpreted as, the number of factors appropriate for a particular analysis, is the number of factors before the plotted line turns sharply to the right.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett’s test of sphericity were examined to decide if the data were performing the factor analysis appropriately. In this study, the criterion that was used to determine the number of factors to retain is Kaiser’s criterion rule of thumb. This criterion according to Kaiser suggests all retaining factors that are above the eigenvalue of 1 is accepted [12].

After validating the content of the instrument, the reliability of NELS was then measured. Reliability can be thought of as consistency. As to Carlson and Winquist, a measure is said to have a high reliability if it produces similar results under consistent conditions. In this study reliability of the instrument will be tested with the aid of Cronbach’s Alpha, coefficient (homogeneity reliability) and stability coefficient [13]. Cronbach’s Alpha scores were calculated to test the internal consistency of each subscale and each overall scale of the adapted instrument. The Alpha coefficients served as indicators for the quality of the instrument [14].

Phase 5- Final Version of NELS and its Readability

After all the steps above were taken, the NELS was then drafted for its final version. Its question numbers were rearranged according to the EFA done in Phase 4. Items that were extracted in the process of validity and reliability were removed. When the final arrangement was made, a test of readability was done to check the reading ease of the newly developed Nurse Emotional Labor Scale tool. The readability was computed by the use of Flesch Reading Ease formula.

The Flesch Reading Ease was used to test the readability of the questionnaire by measuring textual difficulty which indicates how easy a text is to read. According to Flesch, reading-ease test, higher scores indicate material that is easier to read and lower numbers mark passages that are more difficult to read. With the reading ease formula, the interpretation would be based on the reading ease assessment scale of: 90 - 100 (very easy); 80-89 (easy); 70-79 (fairly easy); 60-69 (standard); 50 -59 (fairly difficult); 30-49 (difficult) and 0-29 very confusing. The Flesch–Kincaid readability tests are readability tests designed to
indicate how difficult a passage in English is to understand and was used to better ensure reliability and validity of the NELS.

**Ethical Considerations**

The institutional review board signed the ethical clearance form to ensure that this study did not involve any form of invasion of respondents’ integrity. Prior to data gathering, the respondents received written information about informed consent, purpose of the study(including its voluntary nature), study procedures, benefits of the participation, voluntary participation (including the right to withdraw anytime), and a guarantee of the maintenance of anonymity, and confidentiality of the data.

**RESULTS**

**Validity and Reliability of NELS**

The initial draft of NELS was answered by 2,230 Filipino nurses with at least two years experience. By assumption, according to Gorsuch, the three factors of emotional labor have a linear relationship between the factors and the variables [15]. Since factor analysis of NELS departs from a correlation matrix, the used variables were first of all measured at (at least) an interval level. Secondly, the variables were roughly normally distributed; this makes it possible to “generalize the results of the analysis beyond the sample collected” [9]. Thirdly, the sample size was taken into consideration, as correlations are not resistant [10], and can seriously influence the reliability of the factor analysis [11]. Results showed that NELS has a 42.5% of non-redundant residuals. Residuals are computed between observed and reproduced correlations. There are 99 (42.0 %) non-redundant residuals greater than absolute value 0.05. It is a rule that a model has less than 50% of the non-redundant residual is a good fit. Therefore, since NELS has less than 50% non-redundant residuals. The model is a good fit.

Large datasets that consist of several variables were reduced by observing groups of variables or factors. Factor analysis assembles common variables into descriptive categories. In the factor analysis the different assumption with regard to the communalities was reflected in a different correlation matrix.

After running the communality factor, the remaining 15 items are 2, 3, 4, 5, 6, 9, 10, 12, 13, 14, 17, 18, 19, 20, and 21. The initial communalities were assumed to be 1 by setting the Eigen value to 1 (values of communality factor loads must be greater than one) which is a criterion for the number of factors to rotate [12]. All the remaining questions have Eigen value of 1. This implies that all the score on the component have positive reliability. Therefore, NELS model is considered valid in its communalities.

The scree test of 22-point item scale was used in conjunction with the Eigen values to determine the amount of number of factors to retain, where the number of factors to be retained is the data points that are above the break or the point of inflexion. The scree plot is typically interpreted as follows; the number of factors appropriate for a particular analysis is the number of factors before the plotted line turns sharply right. The elbow in the scree plot occurs at Factor Number 3, indicating that there are three factor solutions for the 22 items.

Items are arranged from largest result to smallest result by each factor. Item 3 has the largest absolute value of 0.716. Item 5 and 21 has an absolute value of 0.711. Item 9 has 0.660, item 2 has 0.648 and item 13 has 0.625. Item 19 has 0.575, item 20 has 0.569, item 18 has 0.556, item 17 has 0.533 and item 12 has 0.522. Item 14 has 0.495, item 4 has 0.418, item 10 has 0.417 and item 6 has 0.453. Out of 22-point item initially developed questions, only 15 were left. The items eliminated were item numbers: 1, 7, 8, 11, 15, 16, and 22. These items received an absolute value >0.40, therefore they were extracted.

To determine the strength of the relationships in the NELS in the interpretation of the factors, the loadings were then assessed. The factors were identified by the largest loadings, but confirmation of the identification of the factors was also done by examining the zero and low loadings of the factors. Negatively scored items were eliminated, where factors load >0.40 does not exceed and 0.10 was accepted. The researcher also determined the cut-off for a statistically meaningful rotated factor loading. A general rule to determine the reliability of the factor is to look at the relationship between the individual rotated factor loading and the magnitude of the absolute sample size.

In this study, the criterion used to determine the number of factors to retain is Kaiser’s criterion which is a rule of thumb. This criterion according to Kaiser suggests retaining all factors that are above the Eigenvalue of 1. The KMO score is 0.853, It is a rule that, the sample is adequate if the value of KMO is greater than 0.5. Since we have 0.853 as the KMO score, the sample is then concluded to be adequate [12]. The Bartlett’s test of sphericity was significant at a level of 0.000. Both tests indicated that the data was appropriate for factor analysis.

In the number of factors to be retained, it is important to know that the Eigen values and scree test are used in the NELS to determine how many factors to retain. Before the extraction, the Eigen value was set at a level greater than 1. Factor 1 (Surface Acting) received an Eigen value of 6.5, factor 2 (deep acting) earned an Eigen value of 3.2 and factor 3 got an Eigen value of 2.3. Since all three have Eigen values of greater than 1, it can then be concluded that all factors...
have positive internal consistency reliability. Three factors explain 63.60% of the variance in the 22-point items. This means that the scale is very useful to measure the emotional labor in terms of the three constructs of surface acting, deep acting and genuine acting.

The Cronbach’s reliability that was used to measure the internal consistency of the NELS test score showed a computed value of 0.87 for SA, and in internal consistency, it is describe in the Cronbach’s alpha rule as “good” because it falls on the alpha range of (0.9 > α ≥ 0.8); 0.81 for DA, is interpreted as “good” (0.9 > α ≥ 0.8) and 0.78 for GA is interpreted as ‘acceptable’ because it falls on the alpha range of 0.8 > α ≥ 0.7, which indicated that the factors for scale have an acceptable internal consistency.

Readability of NELS
The final NELS scored 87 in the Flesch Reading Ease which can be concluded that it is easy to read. As indicated, its reading grade level is 4, meaning that even individuals in the fourth grade can read and understand the developed Nurse Emotional Labor Scale.

<table>
<thead>
<tr>
<th>Surface Acting</th>
<th>1 Never</th>
<th>2 Rarely</th>
<th>3 Sometimes</th>
<th>4 Often</th>
<th>5 Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feign emotions I don’t really feel when telling patients bad news.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I pretend to be sad when my patient tells me a bad experience.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I put on an act in order to deal with patients so in an appropriate way.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have to cover up my true feelings when dealing with patients.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I behave in a way that differs from how I really feel.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Deep Acting</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I work hard to feel the emotions that I need to show when dealing with my patients.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>When dealing with difficult patients, I imagine them being little kids so that I can be more patient with them.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I try to show appropriate emotion at the right time when dealing with patients.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I try to actually experience the emotions that I must show when interacting with patients.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have to concentrate more on my behavior when I display an emotion that I don’t actually feel.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Genuine Acting</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I show the same feelings that I feel inside.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I actually feel the emotions that I need to show to do my job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The smiles I share with my patients are sincere.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I genuinely care for my patients and their well-being.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The emotions I show to patients match what I truly feel.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DISCUSSION**

Scoring high on one subscale does not preclude the fact that you can also score high on the rest of the subscales. This tool measures which emotional strategies nurses are using in their everyday encounters with their patients the most. Individual differences do not significantly affect the way nurses perform emotional labor and an individual’s affective style or empathic level do not significantly associate with the type of acting one adopts to perform emotional labor. This tool is intended for hospital administrators or nursing managers to give to nurses working in hospital contexts to ensure its consistency. The primary purpose of this tool is to adequately measure emotional labor among nurses.
settings. It can be given during their year-end review to assess and measure where nurses may stand in terms of emotional labor. From there, the hospital administrators or nurse managers can take stock and re-evaluate these test scores, along with an individual’s characteristics, and performance to see if any adjustments are needed to be made.

If nurse scores high in the Surface Acting category, the management must reassess its policies or develop retraining programs to teach necessary skills to perform emotional labor. If the hospital administration could provide appropriate training that would help nurses learn to deep-act emotional labor, nurses could provide more sincere hospitality. They would also gain a sense of satisfaction from their ability to act out emotional labor. Nurses need to be courteous to clients. However, clients have no obligation to return empathy or even courtesy. In some situations where clients exercise the privilege of “customers are always right,” nurses face real challenges suppressing their true feelings. It is critical for both nurses and organizations to learn how to deal with such situations. Openly discussing the frustration on their jobs is a cure to heal nurses’ wounds when they are hurt or insulted by clients.

This type of training has several effects. First, it informs nurses that the organization is aware of and acknowledges the emotional contribution that nurses put into jobs. This positive feedback can motivate nurses to increase their productivity and be more committed to their jobs and organizations. Second, it provides an opportunity to ventilate nurses’ negative emotions caused by their jobs. Third, by implementing this type of training, organizations can develop in their nurses the ability to suppress anger or avoid frustration. If companies do not provide formal or informal ways to ventilate employees’ anger and frustration, sooner or later, employees will express anger to their clients.

For many within the nursing profession, the job description involves a great deal of emotional work or emotional labor. By virtue of their near-constant contact with patients and families, it is more difficult for nurses than for other healthcare professionals to escape emotional labor, and recharge their batteries for the next encounter. Unfortunately, emotional labor is not something that nurses are prepared for. They aren’t taught in nursing school to put on a brave face so that they can exude confidence around patients and coworkers, even when they are falling apart inside. Nurses just intuitively know that they have to do this, and with experience, they perfect the art. This doesn’t mean that nurses should just accept emotional labor as an inevitable part of the job and get on with their work. Emotional labor might be integral to the practice of nursing, but it impairs psychological health. Furthermore, depression, absenteeism, burnout, and turnover intention are among the consequences of unrelieved emotional labor.

CONCLUSION

The measured variables can be taken from the underlying structures of nurse emotional labor based on scholarly literatures, respondent interviews, expert feedback consist of 15 components, thus making the NELS final version a 15-point item tool, which were initially 22 items during the initial conceptualization. The latent constructs that underlie the battery of measured variables are surface acting, deep acting and genuine acting.

The NELS was validated using the CVI and EFA. In CVI, 22 out of 24 items were retained which the six experts validated by giving relevance ratings of 3 and 4. In EFA, the KMO test score was greater than 0.5, which concludes that the sampling size was adequate. Bartlett’s test of sphericity was significant at a level of 0.000. This indicates that the data was appropriate for factor analysis. All three factors have an Eigen value higher than one, indicating that the factors are stable and valid. The reliability of the Nurse Emotional Labor scale as to its factors is that surface acting and deep acting have good reliability and genuine acting has an acceptable reliability. Overall, the reliability of all subscales of the scale had higher internal consistency than a suggested value of 0.70 [16], which can be interpreted as adequate. In Flesch Reading Ease test, NELS scored 87, which concludes that it is easy to read. Its reading grade level is 4, meaning that even individuals in the fourth grade can read and understand the developed Nurse Emotional Labor Scale. These findings measure the emotional labor of nurses.

RECOMMENDATIONS

In depth investigation of other constructs to include in the study of NELS should be further investigated. Emotional Labor may result from the overlap between task-related and patient-related job characteristics. For example, nurses who are expected to provide a good service to patients (patient-related) and at the same time do charting and endorsement (workload), may experience role conflict that can lead to emotional dissonance. Using the development guidelines [6, 7] would still help prove useful when doing further in-depth research in Emotional Labor. Moreover, confirmatory factor analysis should be done to further validate the NELS. Further applications of other reliability test (test-re-test) should be conducted to better test the consistency of the scale. Further tests should also not be limited to tertiary government hospitals, but also include private hospitals, primary and secondary hospitals. All nurses should also be included in the study, regardless of how long they have worked.
Consequently, interventions are needed to inform nurses on the detrimental influences of surface acting and train them in the daily use of deep acting as the healthier emotional regulation strategy. Interventions targeted at the enhancement of the individual capacity for self-control would be beneficial to nurses' health. More research is needed to seek solutions to the problem, as well as to better understand the impact of emotional labor on nurses and nursing care.

Conflicts of Interest
The authors declared there is no conflicts of interest on this manuscript.

REFERENCES