

Key Concepts : Korean, POMS, Translation, Reliability, Validity, Cultural Relevancy

## Reliability, Validity, and Cultural Relevancy of the Korean Version of the POMS

*HaeOk Lee<sup>1</sup> · Katherin Lee<sup>2</sup> · Ginger C. V. Kohlman<sup>3</sup>*

### ABSTRACT

The Profile of Mood States (POMS) is the most widely used self-report instrument for the measurement of affect or mood in clinical and nonclinical populations. This paper reports on the translation and testing of a Korean version of the POMS. The translation involved three steps: translation, checking agreement, and panel discussions to arrive at consensus. Then, the Korean version of the POMS was tested with a sample of 47 healthy Koreans who lived in the U.S.; they completed the instrument in the morning, at the beginning of work and in the evening, at the end of work. Internal consistencies for the total scale and subscales were high (alphas = .93 and .94). Face and content validity and the cultural relevance of the Korean version of the POMS were tested through review by five bilingual Korean nursing scholars who were familiar with Korean culture and with the current literature on affect, feeling or mood in both English and Korean. The Korean version of the POMS was then compared with the Haeok Fatigue Behavior Check-list and demonstrated with concurrent validity ( $r = .87, p < .001$ ). Construct validity was established by demonstrating significant differences between the scores on the Fatigue and Vigor subscales ( $p < .001$ ) at the beginning work and at the end of work.

### I. INTRODUCTION

In order to develop appropriate nursing interventions, it is essential not only to measure emotional states physiologically and behaviorally but also to understand subjective aspects of feelings, affect or mood. Increasing attention to mood states and mood changes is evident in the literature on the effects of occupation, exercise, emotional stimulation, psychotherapy

and medication, as well as nursing interventions.

The Profile of Mood States (POMS) was developed to assess transient, distinct mood states and has been used with a broad spectrum of populations from the healthy to the physically and mentally sick, in settings ranging from sports medicine centers to industries (Berglund & Safstrom, 1994; Brunier & Graydon, 1996; Lee, Hicks, & Nino-Murcia, 1991; Lee, Lee, Kohlman, & Schiller, in press;

---

<sup>1</sup> RN, DNSc, Associate Professor, University of Colorado Health Sciences Center, Denver CO, U.S.A.

<sup>2</sup> RN, PhD, Professor, University of California San Francisco, San Francisco CA, U.S.A.

<sup>3</sup> RN, DNSc, Professor, University of California San Francisco, San Francisco CA, U.S.A.

McNair, & Lorr, 1964; McNair, Lorr, & Droppleman, 1971, 1992; Morgan, Costill, Jlynn, Raglin, & O'Connoer, 1988; Pillard & Fisher, 1967). The POMS is now widely used in the health sciences in the USA, Europe, and Asia. However, because it is verbally descriptive, the POMS requires translation for non-English speaking peoples.

During the last decade the POMS has been translated into several languages, including Dutch (Sonnaville et al., 1998), German (Moore, Stanley, & Burrow, 1990), Italian (Carmerino, Cassitto & Gilioli, 1993), Spanish (Escalon, Yanes, Feo & Maizllish, 1994), Swedish (Hassmen & Blomstrand, 1991; Hassmen, Koivular, & Hansson, 1998), Chinese (Liang, Sun, Sun, chen, & Li, 1993), and Japanese (Kuboki et al., 1993; Ohnish, Mori, Kohariyama, Miwata, Murai & Ikeda, 1993). The translated POMS has been reported to have acceptable levels of reliability and validity and has been utilized in broad areas, including sports medicine and occupational medicine, and ill populations. An ability to measure mood states across diverse cultures with the same instrument will increase our understanding of the universal aspects as well as the culturally bounded aspects of human feelings or moods. Using the same tool will also enable researchers to compare their findings with those of previous studies that have included healthy college students, athletes, psychiatric patients, and cancer and cardiovascular patients. However, while the POMS is a standardized instrument that has reliability and validity estimates calculated for a variety of population (McNaire et al., 1971, 1992), validity and reliability are not properties an instrument itself but, rather, properties of an instrument administered to a certain sample under certain conditions for a certain purpose (Polit & Hunger, 1995). Therefore, when measuring a construct in a cultural group other than the one for which an instrument was designed, it is essential to retest

the instrument's validity and reliability. Therefore, in the study reported here, the POMS was translated into Korean, and the reliability and validity of the Korean version were tested.

## II. THE PROFILE OF MOOD STATES SCALE (POMS)

The Profile of Mood States (POMS) was developed by McNair and Lorr (1964) to measure fluctuating affective states in a wide variety of populations. The original intention was to assess the effects of brief psychotherapy or psychotropic medication on psychiatric patients. The POMS proved to be a sensitive indicator of the responses of these patients to various therapeutic approaches. The POMS consists of 65 adjective items and measures six identifiable mood or affective states: Tension-Anxiety (9 items), Depression-Dejection (15 items), Anger-Hostility (12 items), Vigor-Activity (8 items), Fatigue-Inertia (7 items), and Confusion-Bewilderment (7 items). These subscales have been isolated as dimensions of subjective mood states through factor analysis. While, the complete POMS inventory consists of 65 items, 7 items comprising the Friendliness subscale are rarely reported because of a lack of subscale reliability (McNair et al., 1971).

The 65 adjective checklist uses forced-choice responses on a 5-point Likert scale; respondents are asked to pick the response that describes the degree of their feelings for the past week, ranging from 0 (not at all) to 4 (extremely). Respondents are asked to indicate their mood states during the past week, including the test day or present time. A total mood disturbance score is obtained by adding together the five scale scores of Tension, Depression, Anxiety, Fatigue and Confusion, then subtracting the Vigor subscale score. The higher the score, the greater the mood disturbance. It takes respondents approximately 3-5 minutes to

complete the POMS.

Test-retest reliability was originally assessed with 100 psychiatric patients at a 20-day interval, revealing reliability ranges from .70 to .74 (McNair et al., 1971; 1992). Construct validity of the POMS has been estimated for neurotic psychiatric patients and normal college students. Constructive and predictive validity has been demonstrated through significant correlation of scores with the outcomes of short-term psychotherapy, outpatient drug trials and studies of emotion-inducing conditions. Concurrent validity has been shown by comparing the POMS with other instruments measuring mood state such as the Hopkins Symptom Distress Scale; in male psychiatric patients high correlations were found between the anxiety and depression subscales of the POMS and somatomization ( $r = .60$ ), anxiety ( $r = .77$ ) and depression ( $r = .71$ ) on the Hopkins scale. Also, there was a correlation of  $r = .80$  between tension-anxiety on the POMS and the Taylor Manifest Anxiety Scale. The total Depression score from the Beck Depression Inventory is correlated at .75 with the Depression subscale of the POMS (Jacobs & Bose, 1993). Pearson correlations between the Beck scale and the POMS subscales of Tension, Anger, Fatigue, and Confusion were .67, .66, .67, and .68 respectively. These high relationships between the Beck Depression Inventory and the POMS subscales suggest an underlying convergence of these subscales under the same construct of mood.

### III. THE DEVELOPMENT OF THE KOREAN VERSION OF THE POMS

#### Phase I : Item translation

One long-standing view of the relationship between language and culture proposes that the structure of a language determines the way in which speakers of the language view the world.

Although culture does not determine the structure of a language, it influences how a language is used (Wardhaugh, 1992). Therefore, the objectives in translating the POMS into Korean were to remain faithful to the original English version while at the same time making the tool appropriate and easily understood by the average Korean. In order to avoid the influences of professional translators, the back translation method was not used. Since a language influences how its speakers view the world, purposeful sampling was used to select the widest possible variety of translators. The translation process involved three steps: translation, checking agreement, panel discussions, and consensus.

#### Translation.

Ten bilingual Koreans living in San Francisco were selected to first translate the POMS into Korean: six were first generation immigrants and had lived less than 5 years in the U.S. They included a musician, a homemaker, a waitress, two international graduate students in nursing and an exchange cardiologist.

#### Checking the Agreement.

Forty-five items were translated differently with less than 70% agreement by the translators. Many of these items fell under the subscales of Fatigue, Anger, and Tension. In most of the cases of disagreement, several alternative Korean words or phrases were identified for one English adjective. In all these cases, the Korean words were different but translations were correct.

#### Two Panel discussions.

To reach consensus on a definitive version, two panel discussions were held. Since experts in the translation of an instrument measuring

subjective feelings or moods were necessary, five bilingual Korean nursing scholars with great facility in Korean were selected for the first panel. All these panelists received their basic nursing education in Korea. The second panel discussion was conducted with four bilingual Korean college students who wrote and spoke fluent English and Korean: two were born in Korea but raised in the U.S. in Korean speaking families; two completed elementary school in the U.S. and then received their junior and high school education in Korea.

During the first session, panelists were asked to examine items which received less than 70% agreement by the translators. The percent of agreement and the different translations were provided to the panelists, who then compared the Korean version of the POMS to the English version. The first panel focused on the cultural relevance of adjective descriptors of feelings or mood states and on the appropriateness of particular words for describing subjective feeling or mood states, as well as the relevance and applicability of these words for average Koreans. In the second panel discussion, the focus was on the clarity and accuracy of the Korean words. The two panels reached consensus on the final Korean version of the POMS.

Face and content validity and the cultural relevance of the Korean version of the POMS were also assured by review by the five bilingual Korean nursing scholars, who were familiar with Korean culture and had knowledge of the current literature on affect, feeling or mood in English and in Korean.

#### Phase 2 : Assessment of the Korean version of the POMS

To assess the reliability and validity of the Korean version of the POMS, the translated POMS was administered to 47 healthy Koreans living in the San Francisco Bay Area. Subjects had lived in the U.S. less than 5 years and

spoke Korean at home and at work. Using convenient sampling, the subjects were recruited from Korean restaurants, Korean grocery markets, and a Korean catholic church in the SF bay area. The subjects ranged in age from 24 to 69 years, they included 26 women and 21 men: all were employed full-time. Data were collected twice, once in the morning, at the beginning of work and once in the evening at the end of work. The subjects were asked to rate how they felt "right now" on each item, on a 5-point Likert scale. At the end of the questionnaire, the subjects were asked to describe how they liked this tool.

#### Internal Consistency.

The internal consistency of the Korean version of the POMS was high. Cronbach's alpha coefficient for the total scale was .93 in the morning at the beginning of work and .94 in the evening at the end of work (Table 1). All the reliabilities of subscales were highly satisfactory except for the Confusion subscale at the beginning of the work (alpha = .57): however the internal consistency for this subscale was considerably higher at the end of work (alpha = .75) than at the beginning of the work.

<Table 1> Internal Consistency and Test-ReTest Reliabilities of the POMS

Subscales	Beginning of Work	End of Work
	Alpha	Alpha
Fatigue	.80	.78
Vigor	.94	.95
Tension-Anxiety	.94	.84
Depression-Dejection	.93	.94
Anger-Hostility	.87	.86
Confusion-Bewilderment	.59	.75
Global	.93	.94

### Test-retest reliability.

The test-retest method requires administration of an instrument on two separate occasions to the same group of subjects. However, emotions or mood states can change quickly in individuals, depending upon a variety of circumstances. Moreover, in the sample studied here, it should be noted that working at least six hours was probably in itself associated with some change in mood states. Therefore, we did not examine the test-retest reliability of the instrument in this study, since the results might not reflect the stability of the instrument.

### Validity.

One common approach to testing construct validity is the known group technique. In this procedure groups of people with certain "known" characteristics are administered a scale, and differences in scores and the directions of these differences are predicted. To examine the construct validity of the Korean version POMS, in this study, it was hypothesized that working more than six hours is to a fatigue inducing condition and therefore, at the end of work fatigue scores showed increase while vigor scores decreased. Construct validity was measured by testing the hypothesis that the mean fatigue score would be higher and the mean vigor score lower at the end of work than at the beginning of work (Table 2). Paired t tests were used to test for the differences between the mean scores. As indicated, Fatigue scores were significantly higher at the end of work than at the beginning of the work ( $p < .01$ ), while Vigor scores were significantly lower ( $p < .01$ ).

Concurrent validity was measured by examining the relationship between the POMS Fatigue and Vigor subscales and the HaeOk Fatigue Behavior scale. Pearson correlations were used to establish concurrent validity,

correlations  $> .30$  ( $p < .01$ ) were considered significant. The Fatigue and Vigor subscales were moderately-to-highly correlated with HaeOk Fatigue Behavior scale at the beginning of work and at the end of work.

Means and standard deviations of the POMS subscale scores of these healthy Korean subjects at the beginning of work were similar to the baseline mean scores on the POMS subscales reported for college students by Pillard and others (1967).

**<Table 2> Difference between Beginning and End of Work with POMS subscales**

Subscales	Beginning of Work	End of Work
	<u>M</u> (SD)	<u>M</u> (SD)
Fatigue	4.8 (5.1)	11.8 (6.1)*
Vigor	15.0 (7.1)	9.2 (7.5)*
Tension-Anxiety	7.8 (4.3)	8.4 (11.5)*
Depression-Dejection	6.3 (8.7)	11.3 (5.3)*
Anger-Hostility	5.1 (7.1)	10.5 (9.6)*
Confusion-Bewilderment	5.2 (3.9)	8.5 (4.7)*
Global	14.6 (29.2)	-26.1 (29.7)*

Note. \*  $p < .01$

## DISCUSSION

In our study, the reliability and validity of the Korean version POMS were assessed in order to determine whether this instrument for assessing mood states can be recommended for Koreans. When the POMS was translated into Korean, there was some initial skepticism about the usefulness of the translation because of the high rate of disagreement among the translators, who offered several alternative words and phrases for one English adjective. Only 20 items out of 65 were given the same translation by 7 out of the 10 translators. Also, unlike English, single words were often not available to translate the meaning of items into Korean. In the Korean version, therefore, phrases or clauses were

sometimes used instead of single words: this is similar to the POMS translated into Chinese, Italian, and Japanese. However, two panel discussions reached consensus on the translated POMS and the instrument was then administered to healthy Koreans. Yet some subjects still complained about its ambiguity, saying that many adjectives seemed to have almost the same meaning. This is also a common problem in administering the POMS in the U.S. (Eichman, 1978; Albrecht & Ewing, 1989).

Another problem was found in translating the ordinal measurement of a Likert scale: this limitation, however, involves more than the meaning of words translated from English to Korean. When the POMS is administered and subjects read the words "not at all", "a little", "moderately", "quite", and "extremely", as part of the 5-point Likert scale, the points are assumed by researchers to be equidistant from each other. In current research applications, ordinal measurements are analyzed as though they were interval measurements. However, in the translated Korean version, it was more difficult to assume equidistance. Quite a few Korean subjects asked if they could put an extra point between the Korean translations of "not at all" and "a little." This means that for Koreans, some elements possess more or less of some characteristic than other elements, but it is not possible to say exactly how much more or less they possess. That supports Lee and colleagues' concern that Likert-type scales force subjects to respond according to a scale with arbitrary intervals.

The internal consistency of the Korean version of the POMS was high in this group for all subscales except Confusion subscale at the beginning of work (.57). Nunnally (1978) suggested that an instrument reliability of .70 or higher will suffice in preliminary research, thus the reliability of the Confusion subscale at the end of work (coefficient of .75) was within the

acceptable range. The lower reliability of the Confusion subscale at the beginning of work might be attributable, in part, to the fact that the Korean subjects did not understand the first time the meaning of a particular phrase appearing on the subscale of Confusion. It would be useful to provide standardized alternative Korean phrases and to conduct additional testing to see if this would enhance this measure. However, overall, the high values of Cronbach's alphas for the total scores and other subscales indicate high internal consistency of the scale.

The means and standard deviations of the average scores on the POMS subscales of Vigor and Fatigue obtained at the beginning of work and at the end of work in this study showed good agreement with those of Lee and colleagues (1991) from healthy adult American subjects (Fatigue at the beginning was 4.8 vs 5.2; at the end of work, 11.8 vs 12.1; Vigor at the beginning of the work was 15.0 vs 16.8; at the end of work, 9.2 vs 10.7).

## CONCLUSIONS

In our study, the reliability and validity of the Korean version of the POMS were assessed in order to determine whether this translated instrument for assessing mood states should be recommended for use in cross-cultural studies of affect or mood. Even with complaints from subjects and skepticism at the beginning stage, the translated Korean version of the POMS showed high reliability and validity, and was culturally acceptable and relevant to measure fluctuating mood states. Translation of the POMS into a different language and validation of this instrument is important for international understanding of an important psychological experience that influences an individual's health. This work provides considerable confidence in the translated POMS for cross-cultural comparisons, particularly for Koreans. However,

further study is needed to provide more information with respect to factor analysis for construct validity.

Some caution, of course should be used in the interpretation of these findings because sampling procedures were non-random and the sample size was only made up of 47 health Koreans who lived in the U.S.

Nevertheless, since the original POMS and other translated POMS have been used extensively with both clinical and nonclinical populations, the Korean version now needs to be used to compare mood states and test nursing interventions in clinical and nonclinical populations within Korea and with other ethnic populations.

## REFERENCES

- Albrecht, R. R., & Ewing, S. S. (1989). Standardizing the administration of the Profile of Mood States (POMS): development of alternative word lists. Journal of Personality Assessment, 53, 31-39.
- Berglund, B., & Safstrom, H. (1994). Psychological monitoring and modulating of training load of world-class canoeists. Medicine and Sciences in Sports and Exercise, 26, 1036-1040.
- Brunier, G., & Graydon, J. (1996). A comparison of two methods of measuring fatigue in patients on chronic fatigue in patients on chronic haemodialysis: visual analogue vs Likert scale. International Journal of Nursing Study, 33, 338-348.
- Carmarino, D., Cassitto, M.G., & Gilioli, R. (1993). Prevalence of abnormal neurobehavioral scores in populations exposed to different industrial chemicals. Environmental Research, 61, 252-257.
- Eichman, W. J. (1978). Review of the Profile of Mood States. In O. K. Buros (Ed.), Eighth mental measurements yearbook(vol. 1, pp. 1016-1018). Highland Park, NJ: Gryphon.
- Escalona, E., Yanes, L., Feo, O., & Maizllish, N. (1994). Neurobehavioral evaluation of Venezuelan workers exposed to organic solvent mixtures. American Journal of Industrial Medicine, 27, 15-27.
- Hassmen, P., & Blomstrand, E. (1991). Mood change and marathon running: a pilot study using a Swedish version of the POMS test. Scandinavian Journal of Psychology, 32, 225-232. (Sweden)
- Hassmen, P., Koivular, N., & Hansson, T. (1998). Precompetitive mood states and performance of elite male golfers: do trait characteristics make a difference? Perceptual & Motor Skills, 86, 1443-1457.
- Heller, J., Edelmann, R. J. (1991). Compliance with a low calorie diet for two weeks and concurrent and subsequent mood changes. Appetite, 17, 23-28.
- Hooper, S.L., Mackinnon, L. T., & Ginn, E. M. (1998). Effects of three tapering techniques on the performance, forces and psychometric measures of competitive swimmers. European Journal of Applied Physiology, 78, 258-263.
- Jacobs, K. W., & Boze, M. M. (1993). Correlations among scales of the Beck Depression Inventory and the Profile of Mood States. Psychological Reports, 73, 431-434.
- Kuboki, F., Noura, S., Wada, M., Akabayashi A., Nagataki, M., Suematsu, H., Yokiyama, K., & Araki, S. (1993). Multidimensional assessment of mental state in occupational health care-combined application of three questionnaires: Tokyo University Egogram (TEG), Time Structuring Scale (TSS), and Profile of Mood States (POMS). Environmental Research, 61, 285-298.
- Lee, K. A., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. Psychiatry Research, 36, 291-298.
- Lee, H., Lee, K. A., Kohlman, G. C., &

- Schiller, N. B. (in press). Fatigue, mood, and hemodynamic patterns after myocardial infarction. Applied Nursing Research.
- Liang, Y. S., Sun, R.K., Sun, Y., Chen, Z., & Li, L. H. (1993). Psychological effects of low exposure to mercury vapor: application of a computer-administered neurobehavioral evaluation system. Environmental Research, 60, 320-327.
- McNair, D. M., & Lorr, M. (1964). An analysis of mood in neurotics. Journal of Abnormal and Social Psychology, 69, 620-627.
- McNair, D. M., Lorr, M., & Dropplemen, L. F. (1971). Profile of Mood States Manual, San Diego, California. EdITS/Educational and Industrial Testing Service.
- McNair, D. M., Lorr, M., & Dropplemen, L. F. (1992). Profile of Mood States Manual, San Diego, California. EdITS/Educational and Industrial Testing Service.
- Moore, K., Stanley, R., & Burrow, G. (1990). Profile of Mood States: Australian normative data. Psychological Reports, 66, 509-510.
- Morgan, W. P., Costill, D. L., Jlynn, M.G., Raglin, J. S., O' onnoer, P. J. (1988). Mood disturbance following increased training in swimmers. Medical Science Sports Exercise, 20, 408-414.
- Nunnally, J. C. (1978). Psychometric Theory. New York: Harper.
- Ohnishi, A., Mori, K., Kohariyama, K., Miyata, M., Murai Y., & Ikeda, M. (1993). Japanese translation of Profile of Mood States (POMS)- interpretation of the sixty five items of Mood factors and their application in a manufacturing automotive parts factory. Japanese University of Environmental Health, 2, 147-154.
- Pillard, R. C., & Fisher, S. (1967). Effect of chlordiazepoxide and secobarbital on film-induced anxiety. Psychopharmacologia (Berl.), 12, 18-23.
- Polit, D. F., & Hunger, B. P. (1995). Assessing Data Quality: Reliability of Measuring Instruments. Nursing Research: Principles and methods. Philadelphia, J.P. Lippincott.
- Sonnville, F. J., Deville, W., Snoek, F., Wijkel, D., colly, L. P., & Heine, R. J. (1998). Well-being and symptoms in relation to insulin therapy in type 2 diabetes. Diabetes Care, 21, 919-924.
- Wardhaugh, R. (1992). Sociolinguistics. Cambridge: Blackwell.