Temporal Stability of the Japanese Versions of the Flourishing Scale and the Scale of Positive and Negative Experience

Katsunori Sumi*
Nagare College, Nagoya Institute of Technology, Japan

Abstract

The Flourishing Scale (FS) and the Scale of Positive and Negative Experience (SPANE) are reliable, valid instruments used to assess aspects of well-being such as psychological flourishing and positive and negative feelings. The Japanese versions of these scales (FS-J and SPANE-J) have been shown to have adequate internal consistency and construct validity. Test-retest reliability of the Japanese versions, however, has not yet been assessed. Therefore, the purpose of this study was to assess the test-retest reliability of the Japanese versions. The temporal stability of the factor structure of the Japanese versions was also evaluated. The FS-J and SPANE-J data were collected from 336 Japanese college students in two sessions conducted one month apart. The participants completed the Japanese versions in both sessions. The results indicated acceptable test-retest reliability for the FS-J (0.87) and SPANE-J (0.57-0.60). Simultaneous confirmatory factor analysis supported the temporal stability of the hypothesized factor structures for the Japanese versions over the one-month interval.

Keywords: Temporal stability; Test-retest reliability; Factor structure; Well-being; Flourishing; Affect; Measure

Introduction

In the past decade, numerous studies have been devoted to understanding well-being [1-3]. To this end, various measurements have been developed to assess this construct [4-6]. Since well-being studies have been proceeding as a matter of course for so long, it would be more productive and more informative for researchers to use a newer well-being measure that can be applied to cumulative research findings. Two well-being measures, the Flourishing Scale (FS) and the Scale of Positive and Negative Experience (SPANE), have been recently introduced by Diener and colleagues [7,8].

The FS and SPANE were developed to complement existing well-being measures; the former is a measure of psychological flourishing or eudaimonic well-being [9,10], and the latter is a measure to assess positive and negative experiences related to feelings of well-being and ill-being. Both scales have already been translated into several languages [11-13], including Japanese (the Japanese versions of the FS and SPANE are the FS-J and the SPANE-J, see Appendix) [13], and they have been described in several books [10,14-17]. One of these advantages of the scales over other well-being measures is that they are brief and easily comprehensible. The FS consists of eight items that describe broad and important aspects of psychological functioning: competence, engagement and interest, meaning and purpose, optimism, self-acceptance, supportive and rewarding relationships, contribution to the well-being of others, and being respected; e.g., “I lead a purposeful and meaningful life” and “I am optimistic about my future.” The SPANE is designed to measure a broad range of positive feelings (SPANE-P), negative feelings (SPANE-N), and the balance between the two (SPANE-B) with a scale that only has 12 items: e.g., “good,” “bad,” and “happy.”

The FS has shown good reliability and validity in college student samples [8]. Cronbach’s alpha coefficients of the FS-J, SPANE-N, and SPANE-B were 0.87, 0.81, and 0.89, respectively. Test-retest reliability coefficients for the three over one month were 0.62, 0.63, and 0.68, respectively. A principal axis factor analysis for SPANE-P and SPANE-N extracted a single-factor solution in both, with eigenvalues of 3.69 and 3.19 that explained 61% and 53% of the total variance, respectively. The factor loadings of SPANE-P and SPANE-N ranged from 0.58 to 0.81 and from 0.49 to 0.78, respectively. The SPANE also was shown to have good reliability and validity in the same samples [8]. Cronbach’s alpha coefficients of the FS-J, SPANE-N, and SPANE-B were 0.87, 0.81, and 0.89, respectively. Test-retest reliability coefficients for the three over one month were 0.62, 0.63, and 0.68, respectively. A principal axis factor analysis for SPANE-P and SPANE-N extracted a single-factor solution in both, with eigenvalues of 3.69 and 3.19 that explained 61% and 53% of the total variance, respectively. The factor loadings of SPANE-P and SPANE-N ranged from 0.58 to 0.81 and from 0.49 to 0.78, respectively. The first and second factors explained 43.4% (eigenvalue of 5.20) and 24.6% (eigenvalue of 2.95) of the variance, respectively. The factor loadings ranged from 0.63 to 0.68 on the first factor, and from 0.42 to 0.54 on the second factor. Confirmatory factor analysis for FS-J and SPANE-J supported the single- and two-factor solutions, respectively. Moreover, acceptable convergent validity of FS-J, SPANE-

*Corresponding author: Katsunori Sumi, Nagare College, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan, Tel: +81-52-735-5399; E-mail: sumi@nitech.ac.jp

Received March 18, 2014; Accepted April 08, 2014; Published April 15, 2014


Copyright: © 2014 Sumi K. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
J-P, SPANE-J-N, and SPANE-J-B was shown by correlations with scores on psychological stress and symptoms measures as well as other well-being measures.

Diener et al. [8] reported test-retest reliability— that is, the temporal stability of scale scores [18,19]— for the original FS and SPANE as described above. However, test-retest reliability has not been assessed and currently remains unknown for both the FS-J and the SPANE-J. Test-retest reliability is one of the best known indices of reliability and, along with internal consistency reliability [18-20], is one of the most important features of any new measure [21,22]. Therefore, to assess the stability of scores on the scales over time, it is necessary to evaluate the test-retest reliability of the FS-J and the SPANE-J. The result of the evaluation will provide knowledge on the temporal stability of psychological flourishing and feelings of well-being and ill-being as measured by the FS-J and SPANE-J.

The main purpose of this study was to examine test-retest reliability of the FS-J and SPANE-J in a Japanese sample. The stability of the scales over time was assessed using correlation coefficients between scores obtained from two test sessions, one month apart [18,19]. The time interval between these test sessions was one month, which was used to assess the test-retest reliability of the original FS and SPANE in Diener et al. [7,8]. In addition to test-retest reliability, the temporal stability of the factor structure of both scales was assessed using simultaneous confirmatory factor analysis [23]. Although the temporal stability of factor structure is one indicator of construct validity, it has not yet been examined for either the original FS and SPANE or their Japanese versions. Based on the previous study [7,8,12], it was expected that the FS-J and SPANE-J would have one- and two-factor structures, respectively.

Method

Participants

The participants were 336 college students (139 women, 197 men; ages 18 to 24 years, M=20.75, SD=1.22) from two colleges in large cities in Japan. The researchers obtained informed consent from each participant and assured them that their participation in the two test sessions was strictly voluntary.

Measures

Japanese version of the Flourishing Scale (FS-J): The FS-J consists of 8 items on a 7-point Likert scale ranging from 1 (very rarely or never) to 7 (very often or always). The possible range of total scores is 8 to 56, with higher scores reflecting a higher level of psychological flourishing. These properties are the same as in the original FS.

Japanese version of the Scale of Positive and Negative Experience (SPANE-J): The SPANE-J is composed of 12 items on a 5-point Likert scale ranging from 1 (very rarely or never) to 5 (very often or always). The SPANE-J, sharing the properties of the original SPANE, provides three scores as described above. SPANE-J-P scores are the sum of the scores for the 6 positive feelings items. The scores range from 6 to 30, with higher scores indicating more frequently experienced positive feelings. SPANE-J-N scores are the sum of the scores of the 6 negative feelings items. These scores have the same range as the SPANE-J-P, with higher scores indicating more frequent negative feelings. SPANE-J-B scores are calculated by subtracting SPANE-J -N scores from SPANE-J-P scores. The possible range of total scores is −24 to 24, with higher scores indicating more frequent experiences of positive feelings and less frequent experiences of negative feelings.

Procedure

Two test sessions were held one month apart (Time 1 and Time 2) by teachers after class. In both sessions, the FS-J, SPANE-J, and a demographic questionnaire were administered, and completed by all the participants.

Results

The means, standard deviations, range of scores, and Cronbach’s αs are shown in Table 1 for the scales at Time 1 and Time 2. The FS-J had a Cronbach’s alpha of 0.94, indicating good internal consistency reliability [24,25]. The three subscales of the SPANE-J also showed good internal consistency reliability, with Cronbach’s αs ranging from 0.87 to 0.90.

Table 1 contains the test-retest reliability coefficients for the scales. The FS-J scores yielded a coefficient of 0.87. In comparison with the FS-J, the SPANE-J subscales revealed smaller coefficients, ranging from 0.57 to 0.60.

Prior to examining the temporal stability of the factor structure of the FS-J and SPANE-J, the factorial validity of each scale was confirmed using the data from Time 1 and Time 2. Table 2 includes the goodness of fit indices in confirmatory factor analyses. The results indicated that the data from each test session of the FS-J and SPANE-J fit the one- and the two-factor models, respectively. On both scales, goodness of fit indices was quite similar between Time 1 and Time 2. For the FS-J, the Goodness of Fit Index (GFI), Normed Fit Index (NFI), and Comparative Fit Index (CFI) were above 0.95, indicating good fit; the Adjusted Goodness of Fit Index (AGFI) was above 0.90, indicating acceptable fit; and the Root Mean Square Error of Approximation (RMSEA) was below 0.08, showing marginal fit [26,27]. The standardized factor loadings at Time 1 and Time 2 ranged from 0.79 to 0.87 and from 0.82 to 0.87, respectively, and were all statistically significant (p<0.01). For the SPANE-J, the GFI, AGFI, and NFI were 0.90 or higher, indicating acceptable fit; the CFI was above 0.95, indicating good fit; and the RMSEA was below 0.08, showing marginal fit. The standardized factor loadings at Time 1 and Time 2 were all significant (p<0.01), ranging from 0.67 to 0.87 and from 0.60 to 0.91, respectively.

As a result of the simultaneous confirmatory factor analysis, equivalent factor structures were found between Time 1 and Time 2 data for the FS-J and SPANE-J; therefore, the stability of the factor structures across time was supported for both scales. Goodness of fit indices are shown in Table 2. As the main goodness of fit indices regarding the FS-J, the GFI and AGFI were 0.90 or higher, indicating
The results of this study provide support for the stability of the scale scores and factor structures across time for both the FS-J and the SPANE-J. The correlations based on the one-month interval used by the original study [7,8] support reasonable test-retest reliability of the Japanese versions. In fact, the test-retest reliability coefficient of 0.87 for the FS-J is greater than the coefficient of 0.71 obtained for the original FS [8]. Compared with the FS-J, the SPANE-J subscales showed lower test-retest reliability. This tendency is consistent with the data obtained from the study on the original versions [7,8]. The test-retest reliability coefficients for the SPANE-J subscales are generally moderate; however, the test-retest reliability of the subscales is equally as high as the level reported for the Positive and Negative Affect Schedule [28], a well-known measure of affect with good psychometric properties. Moreover, there are very few essential differences in the coefficients between the subscales of the SPANE-J and the original SPANE with adequate test-retest reliability [8]. Because affect or feelings are relatively unstable over time, the test-retest reliability coefficients of these scales cannot reach high values. The SPANE-J subscales, therefore, have acceptable temporal stability of scale scores.

The results of this study confirm the hypothesized factor structures for the Japanese versions that were supported by the previous study [13] and found for both the original versions [8] and the Portuguese versions [9]. The present findings also support the temporal stability of the factor structures for both scales over a one-month period. The one-factor model of the FS-J fits the data from both test sessions, even though they were conducted one month apart. Likewise, the SPANE-J shows the same two-factor structure, indicating positive feelings and negative feelings, on the data from both test sessions.

The present results suggest that psychological flourishing and positive and negative feelings, as measured by the FS-J and SPANE-J, may show the stability of intensity and construct over a short period of time. In addition, it is suggested that the psychological flourishing may exhibit somewhat more stable intensity than do feelings. Further studies using the Japanese versions of these measurements are needed to clarify the detailed nature of both the constructs and their relationships with other constructs (e.g., personality, psychopathology).

In conclusion, this study provides additional information on the psychometric properties of the FS-J and the SPANE-J. However, the time interval used to assess the temporal stability should be mentioned as a limitation to this study. Future studies will need to examine the temporal stability of the scales over longer and varied time periods. Another limitation of this study is the homogeneity of its student sample. Such sample characteristics may have affected the present results, which points to the need for caution in generalizing these results to the broader Japanese populations. Additional research is needed to evaluate the temporal stability in different or heterogeneous groups (e.g., workers, elderly people). This kind of research is also required to more fully generalize the FS-J and SPANE-J. It has often been pointed out that well-being is affected by socio-economic and cultural factors [17,29,30]; thus these factors may cause slight differences between the results of the Japanese versions and the original versions [7,8]. Further studies will need to explore the influences of these factors on temporal stability. Despite these limitations, the findings of this study contribute to a detailed understanding and measurement of psychological flourishing and feelings of well-being and ill-being.

**Table 2: Goodness of fit indices for tests of factorial validity of FS-J and SPANE-J.**

<table>
<thead>
<tr>
<th>Data</th>
<th>χ² *</th>
<th>df</th>
<th>GFI</th>
<th>AGFI</th>
<th>NFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>50.47</td>
<td>20</td>
<td>0.96</td>
<td>0.93</td>
<td>0.98</td>
<td>0.99</td>
<td>0.07</td>
<td>84.47</td>
</tr>
<tr>
<td>Time 2</td>
<td>62.36</td>
<td>20</td>
<td>0.98</td>
<td>0.92</td>
<td>0.98</td>
<td>0.99</td>
<td>0.07</td>
<td>86.36</td>
</tr>
<tr>
<td>Time 1 and</td>
<td>155.62</td>
<td>40</td>
<td>0.94</td>
<td>0.90</td>
<td>0.97</td>
<td>0.98</td>
<td>0.07</td>
<td>219.62</td>
</tr>
<tr>
<td>Time 2b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPANE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>144.80</td>
<td>53</td>
<td>0.93</td>
<td>0.90</td>
<td>0.94</td>
<td>0.96</td>
<td>0.07</td>
<td>192.80</td>
</tr>
<tr>
<td>Time 2</td>
<td>139.16</td>
<td>53</td>
<td>0.94</td>
<td>0.90</td>
<td>0.94</td>
<td>0.96</td>
<td>0.07</td>
<td>191.16</td>
</tr>
<tr>
<td>Time 1 and</td>
<td>302.12</td>
<td>106</td>
<td>0.93</td>
<td>0.90</td>
<td>0.94</td>
<td>0.96</td>
<td>0.05</td>
<td>402.12</td>
</tr>
</tbody>
</table>

*All the values are significant at p<0.01

The results of this study provide support for the stability of the scale scores and factor structures across time for both the FS-J and the SPANE-J. The correlations based on the one-month interval used by the original study [7,8] support reasonable test-retest reliability of the Japanese versions. In fact, the test-retest reliability coefficient of 0.87 for the FS-J is greater than the coefficient of 0.71 obtained for the original FS [8]. Compared with the FS-J, the SPANE-J subscales showed lower test-retest reliability. This tendency is consistent with the data obtained from the study on the original versions [7,8]. The test-retest reliability coefficients for the SPANE-J subscales are generally moderate; however, the test-retest reliability of the subscales is equally as high as the level reported for the Positive and Negative Affect Schedule [28], a well-known measure of affect with good psychometric properties. Moreover, there are very few essential differences in the coefficients between the subscales of the SPANE-J and the original SPANE with adequate test-retest reliability [8]. Because affect or feelings are relatively unstable over time, the test-retest reliability coefficients of these scales cannot reach high values. The SPANE-J subscales, therefore, have acceptable temporal stability of scale scores.

**References**


Appendix

The Japanese version of the Flourishing Scale (FS-J)
© Copyright by Ed Diener and Robert Biswas-Diener, January 2009.

あなたは次の8つの文にどの程度同意しますか？
それぞれの文について，1から7の選択肢から，当てはまるものを選んで下さい。

7. まったくそうだ
6. そうだ
5. ややそうだ
4. どちらでもない
3. ややちがう
2. ちがう
1. まったくちがう

私は目的や意味のある生活をおくっている
私の人間関係は私にとって支えであり，価値あるものだ
私は日々の活動に興味をもって取り組んでいる
私は他者の幸福や健康に，積極的に貢献している
私は大切な活動（勉強，仕事等）のための能力や実力がある
私は将来について楽観的だ
人は私を尊重してくれる

The Japanese version of the Scale of Positive and Negative Experience (SPANE-J)
© Copyright by Ed Diener and Robert Biswas-Diener, January 2009.

ここ4週間に，あなたは次の気持ちをどのくらい経験しましたか？
それぞれの気持ちについて，1から5の選択肢から，当てはまるものを選んで下さい。

1. ほとんどなかった
2. あまりなかった
3. ときどきあった
4. しばしばあった
5. ほぼいつもあった

前向きの気持ち
後ろ向きの気持ち
良い気持ち
悪い気持ち
快適な気持ち
不快な気持ち
幸せな気持ち
悲しい気持ち
心配な気持ち
うれしい気持ち
怒りの気持ち
満ち足りた気持ち