Journal of Psychology & Psychotherapy

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 coefficient and the test-retest reliability coefficient over one month were 0.87 and 0.71, respectively. A principal axis factor analysis supported a single-factor solution, with eigenvalue of 4.24 that explained 53% of the total variance. The factor loadings ranged from 0.61 to 0.77. The construct validity of the FS was acceptable, based on its moderate to high correlations with scores on several other well-being measures.

The SPANE also was shown to have good reliability and validity in the same samples [8]. Cronbach's alpha coefficients of the SPANE-P, 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 construct validity of the SPANE-P, SPANE-N, and SPANE-B was good, with moderate to very high correlations with scores on several other well-being and affect measures.

The Japanese versions of the FS (FS-J) and the SPANE (SPANE-J), like their original versions, also showed sound psychometric properties when tested with a college student sample [13]. Cronbach's alpha coefficients of the FS-J, SPANE-J-P (positive feelings), SPANE-J-N (negative feelings), SPANE-J-B (balance between the two) were 0.95, 0.91, 0.90, and 0.88, respectively. The results from a principal axis factor analysis for FS-J supported a single-factor solution with an eigenvalue of 5.85 that explained 73.1% of the total variance. The factor loadings ranged from 0.77 to 0.88. As expected, the same factor analysis for the SPANE-J showed a two-factor (positive and negative feelings) solution. 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

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 or 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 (strongly disagree) to 7 (strongly agree). 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 alphas 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 alphas 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

		Mean	Standard Diviation	Range	Cronbach's α	Test-retest reliability[a]
Time 1	FS-J	35.13	8.14	10 to 56	0.94	0.87
	SPANE-J-P	20.04	4.38	6 to 30	0.90	0.60
	SPANE-J-N	17.26	4.68	6 to 30	0.88	0.57
	SPANE-J-B	4.78	7.24	–19 to 23	0.87	0.58
Time2	FS-J	35.81	8.24	10 to 56	0.94	
	SPANE-J-P	20.13	4.30	8 to 30	0.93	
	SPANE-J-N	16.94	4.55	6 to 30	0.86	
	SPANE-J-B	4.59	7.06	–18 to 24	0.88	

[a]Correlation coefficients between scores on the scale for Time1 and Time 2. All the coefficients are significant at p<0.01

Table 1: Means, standard deviations, range of scores, and Cronbach's α coefficients for FS-J, SPANE-J-P, SPANE-J-N, and SPANE-J-B.

Data	$\chi^{2\ a}$	df	GFI	AGFI	NFI	CFI	RMSEA	AIC
FS								
Time 1	50.47	20	0.96	0.93	0.98	0.99	0.07	84.47
Time 2	52.36	20	0.96	0.92	0.98	0.99	0.07	86.36
Time 1 and Time 2[b]	155.62	40	0.94	0.90	0.97	0.98	0.07	219.62
SPANE								
Time 1	144.80	53	0.93	0.90	0.94	0.96	0.07	192.80
Time 2	139.16	53	0.94	0.90	0.94	0.96	0.07	191.16
Time 1 and Time 2[b]	302.12	106	0.93	0.90	0.94	0.96	0.05	402.12

[a]All the values are significant at p<0.01
[b]The results of the simultaneous confirmatory factor analysis of data at both Time1 and Time 2
GFI goodness of fit index, *AGFI* adjusted goodness of fit index, *NFI* normed fit index, *CFI* comparative fit index, *RMSEA* root mean square error of approximation, *AIC* Akaike information criterion

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

acceptable fit; the NFI and CFI were above 0.95, indicating good fit; and the RMSEA was less than 0.08, showing marginal fit. All of the factor loadings were significant (p<0.01), ranging from 0.78 to 0.89. 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 0.05, indicating good fit. All of the factor loadings were significant (p<0.01), ranging from 0.60 to 0.90.

Discussion

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.

References

1. Diener E, Suh EM, Lucas RE, Smith HL (1999) Subjective well-being: Three decades of progress. Psychological Bulletin 125: 276-302.

2. Keyes CLM (2006) Subjective well-being in mental health and human development research worldwide: An introduction. Social Indicators Research 77: 1-10.

3. Ryan RM, Deci EL (2001) On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. Annu Rev Psychol 52: 141-166.

4. Keyes CLM, Magyar-Moe JL (2003) The measurement and utility of adult subjective well-being. In: Lopez SJ, Snyder CR (Eds.), Positive psychological assessment: A handbook of models and measures. American Psychological Association, New York, 411-425.

5. Norrish JM, Vella-Brodrick DA (2008) Is the Study of Happiness a Worthy Scientific Pursuit? Social indicators research 87: 393-407.

6. Veenhoven R (2014) World database of happiness: Archive of research findings on subjective enjoyment of life. Erasmus University Rotterdam.

7. Diener E, Wirtz D, Biswas-Diener R, Tov W, Kim-Prieto C et al. (2009) New measures of well-being. In: Diener E (Edr.), Assessing well-being: The collected works of Ed Diener. Russell Sage Foundation, New York, NY, 247-266.

8. Diener E, Wirtz D, Tov W, Kim-Prieto C, Choi D et al. (2010) New well-being measures: Short scales to assess flourishing and positive and negative feelings. Social Indicators Research 97: 143-156.

9. Silva AJ, Caetano A (2011) Validation of the Flourishing Scale and Scale of Positive and Negative Experience in Portugal. Social Indicators Research 110: 469-478.

10. OECD (2013) OECD guidelines on measuring subjective well-being, OECD Publishing, Paris, France.

11. Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2009). New measures of well-being: Flourishing and positive and negative feelings. Social Indicators Research 39: 247-266.

12. Diener E, Wirtz D, Tov W, Kim-Prieto C, Choi DR, et al. (2009). New measures of well-being: Flourishing and positive and negative feelings. Social Indicators Research 39: 247-266.

13. Sumi K (2013) Reliability and validity of Japanese versions of the Flourishing Scale and the Scale of Positive and Negative Experience. Social Indicators Research, Advance online publication.

14. Hefferon K, Mutrie N (2012) Physical activity as a "stellar" positive psychology intervention. In: Acevedo EO (Edr.), The Oxford handbook of exercise psychology. Oxford University Press, New York, NY 117-131.

15. Hefferon K, Boniwell I (2011) Positive psychology: Theory, research and applications. Open University Press, New York, NY.

16. Simmons CA, Lehmann P (2012) Tools for strengths-based assessment and evaluation. Springer, New York, NY.

17. Sirgy MJ (2012) The psychology of quality of life: Hedonic well-being, life satisfaction, and eudaimonia (2nd Edn.), Springer, New York, NY.

18. DeVellis RF (2003) Scale development: Theory and applications (2nd Edn.), Sage, Thousand Oaks, CA.

19. Gregory RJ (2004) Psychological testing: History, principles, and applications (3rd Edn.), Allyn & Bacon, Boston, MA.

20. Widaman KF, Little TD, Preacher KJ, Sawalani GM (2010) On creating and using short forms of scales. In: Trzesniewski KH, Donnellan MB, Lucas RE (Eds.), Secondary data analysis: An introduction for psychologists. American Psychological Association, Washington DC, 39-61.

21. Anastasi A, Urbina S (2009) Psychological testing (7th Edn.), Upper Saddle River, Prentice Hall, NJ.

22. Kaplan S (1996) Critical evaluation of standardized tests. Occup Ther Health Care 10: 3-14.

23. Arbuckle JL (2012) IBM SPSS Amos 21 user's guide. IBM, Chicago: IL.

24. Henson RK (2001) Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. Measurement and Evaluation in Counseling and Development 34: 177-189.

25. Nunnally JC, Bernstein IH (1994) Psychometric theory (3rd Edn.), McGraw-Hill, New York, NY.

26. Bowen NK, Guo S (2011) Structural equation modeling. Oxford University Press, New York, NY.

27. Schumacker RE, Lomax RG (2010) A beginner's guide to structural equation modeling (3rd Edn.), Routledge, New York, NY.

28. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology 54: 1063-1070.

29. Diener E (2000) Subjective well-being. The science of happiness and a proposal for a national index. Am Psychol 55: 34-43.

30. Eckersley R (2001) Culture, health and well-being. In: Eckersley R, Dixon J, Douglas B (Eds.), The social origins of health and well-being. Cambridge University Press, Cambridge, UK, 51-70.

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
私には大切な活動（勉強,仕事等）のための能力や実力がある
私はよい人間であり，よい生活をおくっている
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