VALIDATION OF THE SATISFACTION WITH LIFE SCALE IN A SAMPLE OF HONG KONG UNIVERSITY STUDENTS

John SACHS1)

1) The University of Hong Kong, Hong Kong, SAR

Exploratory and confirmatory analyses were used to examine the Satisfaction With Life Scale in a sample of Hong Kong university students. The exploratory analyses showed that the Satisfaction With Life Scale had good internal consistency reliability and that it appeared to measure a unidimensional construct. Comparison of Hong Kong students’ scores on the Satisfaction With Life Scale with the normative data for American students and mainland Chinese and Korean students showed that Hong Kong students were more similar, on average, to the American students than to the Asian students in their overall satisfaction with life. Further analysis of the scale’s structure using confirmatory factor analysis found that a modified two-factor model defined by present and past life satisfaction items and which allowed the error variances for two items to correlate fit the data best. However, inspection of the fit indices for the modified two-factor model suggested that it over fit the data, while a modified one-factor model, with correlated errors, also showed acceptable fit. Since the modified one-factor model was more parsimonious and agreed with the theoretical development of the scale as a unidimensional measure of general overall life satisfaction, it was retained. It is suggested that because of the Satisfaction With Life Scale’s good reliability and cross-cultural validity along with its brevity and broad-band nature in assessing overall satisfaction with life it could have wide applicability as an ancillary measure in a number of research areas.

Key words: SWLS, reliability, validity, Confirmatory Factor Analysis, Hong Kong University students

Andrews and Withey (1976; cited in Diener, Emmons, Larsen, & Griffin, 1985) defined subjective well-being (SWB) as including both positive and negative affect and cognitive-judgmental aspects. As the term suggests, one’s level of SWB is determined not by some normative comparison but rather by a subjective evaluation of one’s overall satisfaction with life (Diener, 1984; Diener et al., 1985). And while the measurement of the affective aspects of well-being has received considerable attention (Bradburn, 1969; Kammann & Flett, 1983; Kozma & Stones, 1980), an overall general measure of the cognitive-judgmental aspects of SWB was lacking until the development of the Satisfaction With Life Scale (SWLS) by Diener et al. (1985).

The SWLS has a number of desirable features. Firstly, because it measures general life-satisfaction, respondents are free to weight the various aspects of life satisfaction such as health, wealth, love, personal accomplishments etc., in terms of their perceived subjective importance (Diener et al., 1985). Secondly, the SWLS is a very short self-
report inventory composed of only five items. The scale’s brevity makes it easy to administer and incorporate with other individual-difference measures (Pavot & Diener, 1993). Third, the SWLS seems to tap a unidimensional construct (Diener et al., 1985; Lewis, Shevlin, Bunting, & Joseph, 1995; Pavot & Diener, 1993; Shevlin & Bunting, 1994) of general overall life satisfaction so that a simple sum score is sufficient for measuring the construct. Fourth, the SWLS correlates negatively with clinical measures of distress, depression, and anxiety and positively with other measures of well-being (Pavot & Diener, 1993) so that its discriminate and convergent validity are good (also see Pavot, Diener, Colvin, & Sandvik, 1991). And, fifth, besides American data, normative data are available for French-Canadian students, Russian students, mainland Chinese students, Korean University students, and, also, for various adult samples (Pavot & Diener, 1993) making cross-cultural and cross-age group comparisons easier.

Although the undimensionality of the SWLS has generally been accepted, Pavot and Diener (1993) noted that two of the SWLS items (4, 5) seemed to refer to ‘past’ life satisfaction while the other three items (1, 2, and 3) seemed to refer to ‘present’ life satisfaction. Confirmatory factor analyses reported by McDonald (1999) lends some support to this observation since a two-factor congeneric model with ‘present’ life satisfaction items defining one factor and ‘past’ life satisfaction items defining the second factor fit the data better than the one-factor model. But because these factors were rather highly correlated ($r = .86$), McDonald noted that it was difficult to decide whether or not to treat them as truly distinct. A subsequent hierarchical factor analysis by McDonald with one general factor and two uncorrelated group factors (present and past items) showed the loadings on the general factor to be sufficiently high to justify the use of a single sum score with this scale. Although an exploratory factor analysis conducted by the author on a Hong Kong sample showed one dominant eigenvalue thus tentatively supporting a single factor model, there was also some evidence to suggest that a two-factor model should not be ruled out.

The purpose of this study, therefore, was to compare the results for the Hong Kong students’ data on the SWLS against the normative scale score data reported in Pavot and Diener (1993), especially for Asian students, and against the earlier exploratory factor analysis results reported by Diener et al. (1985). The second and primary purpose of this study was to test the one- and two-factor models using confirmatory factor analysis to identify a best fitting model for the Hong Kong student data.

**Method**

**Participants:**

A convenience sample of 123 Master of Education (MEd) students from the University of Hong Kong was used. Students’ median age was 32, with an interquartile range of 8 (first quartile=30, third quartile=38). Eighty students were female and 43 were male, resulting in a gender ratio of about two female students for every male student, fairly typical of the student gender ratio in the faculty.

**Measures:**

Besides the SWLS measure, three secondary measures were included to help verify the scale’s validity.
These secondary measures included proactive attitude, general self-efficacy, and internal locus of control.

**Proactive Attitude Scale (PA).** The 8-item Proactive Attitude Scale (Schwarzer, 1999; Schmitz & Schwarzer, 1999, 2000) was used to measure how much responsibility individuals took for their own lives and their belief in their own ability to effect positive life changes. Internal consistency reliability for the scale is good ($\alpha=0.75$) (Schmitz & Schwarzer, 2000), and it has been shown to have moderate positive correlation ($r=0.56$) with the 10-item General Perceived Self-Efficacy Scale (GSE) and moderate negative correlations with measures of Lack of Accomplishment and Depersonalization (Schwarzer, 1999).

Students indicated their responses to each item using a 4-point rating scale: 1 (not at all true of me) to 4 (exactly true of me) so that total scores could range from 8 to 32, with higher scores reflecting a more proactive attitude.

**General Self-Efficacy Scale (GSE).** Based on Bandura’s (1977, 1997) social learning theory, the GSE scale is a short 10-item scale (Jerusalem & Schwarzer, 1992) used to measure belief in one’s ability to deal effectively with stressful life events. Extensive cross-cultural research (Schwarzer, 1993, 1997; Schwarzer & Jerusalem, 1993) has established the scale’s reliability (internal consistency reliability estimates ranged from .75 to .90) and validity (i.e., positive correlations with self-esteem, optimism and negative correlations with anxiety, depression etc.). More recent Hong Kong results confirm these earlier findings (Tang, Au, Schwartzer, & Schmitz, 2001).

Students indicated their responses to each item using a 4-point rating scale: 1 (not at all true of me) to 4 (exactly true of me) so that scale scores could range from a low of 10 to a high of 40, with higher scores reflecting higher self-efficacy.

**Locus of Control (LC).** Locus of control was measured with three items embedded in a larger questionnaire on students’ attitudes toward learning. Two items asked participants whether or not they felt that luck was more important than either hard work or ability in determining their academic performance, while the third item asked them whether they felt that their academic performance (good or bad) was entirely dependent on their own efforts. Students indicated their level of agreement or disagreement with each item using a 7-point rating scale: 1 (strongly disagree) to 7 (strongly agree). Items were appropriately recoded before summing so that higher scores (range: 1 to 21) reflected higher internal locus of control.

**Satisfaction With Life Scale (SWLS).** As describe earlier, the SWLS (Diener et al., 1985) is a short 5-item scale developed as an overall general measure of the cognitive-judgmental aspects of subjective well-being (SWB) (Diener, 1984). Numerous studies (Diener et al.,1985; Lewis, Shevlin, Bunting, & Joseph, 1995; Pavot & Diener, 1993; Shevlin & Bunting, 1994) have established the SWLS’s reliability and validity. Reported internal consistency reliability estimates typically range from .79 to .89. Positive validity coefficients have been reported with measures such as self-esteem and self-efficacy, while negative validity coefficients have been reported with measures that reflect dissatisfaction with life such as feelings of lack of accomplishment or feelings of depersonalization. Thus scores on the SWLS for the Hong Kong student data were expected to correlate positively with scores on PA, GSE, and LC, since a lack of general overall well-being or satisfaction with life would be incompatible with a proactive attitude or with believing in one’s ability to deal effectively with stressful life events or with accepting responsibility for one’s successes or failures.

Students indicated their responses to each item by using a 7-point rating scale: 1 (strongly disagree) to 7 (strongly agree), so scores could range from 5 to 35, with higher scores indicating greater satisfaction with life.

**Procedure:** Students completed a 60-item questionnaire, which included the items from the above measures, during the last 20 minutes of a regular educational research module class. The purpose of the questionnaire was explained briefly to all participants before they began along with instructions on how they were to record their responses to each item. Also, participants were reminded that their responses were anonymous so they should respond as truthfully as possible. Most participants were able to complete the entire questionnaire within 15 minutes.

**Analysis:** SPSS 11.0 for Windows (2001) was used for the exploratory descriptive analysis, item analysis, and exploratory factor analysis, while LISREL 8 (Jöreskog & Sörbom, 1996) was used to conduct the confirmatory factor analysis.
RESULTS

Exploratory Analysis

Descriptive Statistics and Reliability Estimates for Measures. Internal-consistency reliability estimates along with scale means and standard deviations are reported in Table 1. All scales showed acceptable internal-consistency reliability estimates, ranging from .70 (LC) to .86 (GSE).

Coefficient alpha was .82 for the SWLS, slightly lower than the alpha of .87 reported by Diener et al. (1985). However, this estimate of internal consistency reliability was still in line with estimates of coefficient alpha for the normative data reported by Pavot and Diener (1993), although at the low end of the range (.79 to .89). Mean responses for each item of the SWLS ranged from 4.32 (item 5) to 5.55 (item 3), similar to that reported by Pavot and Diener (1993).

The mean scale scores for all measures were relatively high indicating that students on average were satisfied with their lives, had a proactive active attitude, believed themselves to be capable of handing stressful life events, and believed that their successes or failures were directly attributable to themselves. The mean SWLS score for the Hong Kong student data ($M=25.29$) was considerably higher than the mean reported for Chinese students ($M=16.1$) or for Korean students ($M=19.77$) but similar to that reported for American college students whose mean SWLS scores ranged from 23.0 to 25.2 (Pavot & Diener, 1993).

Correlations with SWLS. Inter-correlations among the measures including their correlations with age and gender are reported in Table 1, too. As expected, scores on the SWLS showed significant ($p<.01$) but mild positive correlations with scores on GSE, PA, and LC supporting the validity of the scale for the Hong Kong student data. Furthermore, the SWLS scores were not significantly correlated with either age or gender thus agreeing with earlier findings (Pavot & Diener, 1993).

Exploratory Factor Analysis. To facilitate comparisons with Diener et al. (1985), an

Table 1. Inter-correlations Among Measures (Including Age and Gender) Plus Their Means, Standard Deviations, and Reliability Estimates ($n=123$)

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Coefficient Alpha</th>
<th>$M$</th>
<th>(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SWLS</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.82</td>
<td>25.29 (5.123)</td>
<td></td>
</tr>
<tr>
<td>2. PA</td>
<td>0.313**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77</td>
<td>25.56 (3.108)</td>
<td></td>
</tr>
<tr>
<td>3. GSE</td>
<td>0.247**</td>
<td>0.516**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td>0.86</td>
<td>30.60 (4.225)</td>
<td></td>
</tr>
<tr>
<td>4. LC</td>
<td>0.246**</td>
<td>0.531**</td>
<td>0.399**</td>
<td>1.000</td>
<td></td>
<td></td>
<td>0.70</td>
<td>15.31 (3.242)</td>
<td></td>
</tr>
<tr>
<td>5. Age</td>
<td>0.128</td>
<td>0.036</td>
<td>0.234**</td>
<td>0.015</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Gender</td>
<td>-0.035</td>
<td>-0.027</td>
<td>-0.086</td>
<td>-0.166</td>
<td>-0.117</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Correlation is significant and the 0.01 level (2-tailed)**
exploratory factor analysis was conducted using principal axis factor analysis. A dominant eigenvalue of 3.04 for the first factor accounted for 61% of the total variance. This was slightly lower than the 66% of accounted variance for the single factor reported in Diener et al. (1985). Table 2 presents the factor loadings and corrected item-total correlations for the single-factor model for the Hong Kong student data along with data from Diener et al. (1985). Factor loadings and corrected item-total correlations for the Hong Kong data were lower on three items (1, 4, 5) and higher on two items (2, 3). However, a comparison of the SWLS factor loadings and corrected item-total correlations results for the Hong Kong student data with results for the normative data reported by Pavot and Diener (1993) shows that the Hong Kong data results were well within the range of these earlier results.

Confirmatory Factor Analysis

Two basic models were tested in the confirmatory factor analysis — the one-factor model and the two-factor model. For the two-factor model, the SWLS items 1 to 3 defined the factor dealing with ‘present’ life satisfaction while items 4 and 5 defined the factor dealing with ‘past’ life satisfaction. Based on the initial confirmatory factor analysis results, other models were considered.

Need for confirmatory factor analysis. While the results of the exploratory analysis of the SWLS seem to agree fairly well with earlier findings, it cannot provide sufficient evidence for the scale’s undimensionality and, therefore, its construct validity. Confirmatory analyses that can test theoretical factor models in which the number of factors, pattern of loadings, and inter-factor correlations are explicitly specified are required (McDonald, 1985, 1999).

Sample Size. Although the sample size used here is relatively small, it is still considered acceptable (Marsh, Balla, & McDonald, 1988). Furthermore, with only five variables, there were nearly 25 observations per variable.

Table 2. Exploratory Factor Analysis SWLS Factor Loadings and Corrected Item-total Correlations for the Hong Kong Student Data and Diener et al.’s (1985) American Undergraduate data

<table>
<thead>
<tr>
<th>Item</th>
<th>Hong Kong Student Data n=123</th>
<th>Diener et al. Student Data n=176</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor Loadings</td>
<td>Corrected Item-Total</td>
</tr>
<tr>
<td>1. In most ways my life is close to my ideal.</td>
<td>.64</td>
<td>.54</td>
</tr>
<tr>
<td>2. The conditions of my life are excellent.</td>
<td>.86</td>
<td>.74</td>
</tr>
<tr>
<td>3. I am satisfied with my life.</td>
<td>.86</td>
<td>.76</td>
</tr>
<tr>
<td>4. So far I have gotten the important things I want in life.</td>
<td>.62</td>
<td>.58</td>
</tr>
<tr>
<td>5. If I could live my life over, I would change almost nothing.</td>
<td>.58</td>
<td>.54</td>
</tr>
</tbody>
</table>
**Model fit indices used.** Confirmatory factor analysis results are presented for the one- and two-factor models and for two modified models. All analyses were conducted on the covariance matrix with maximum likelihood estimation. Absolute model fit was assessed with the chi-square ($\chi^2$) statistic, the goodness-of-fit (GFI) statistic, where fit is good for values of GFI greater than .95 and acceptable for values greater than .90 (McDonald, 1999), and the root mean square residual (RMR), where fit is acceptable if RMR is less than .10 (McDonald, 1999). Incremental fit (compares the fitted model to a baseline model, usually the null model) was assessed with the Tucker-Lewis index (TLI also known as the non-normed fit index) (Tucker & Lewis, 1973), where values of .90 or greater indicate acceptable fit (Bentler, 1990) and the comparative fit index (CFI) (Bentler, 1990), where values of .90 or greater indicating acceptable fit (Bentler). Parsimonious fit (adjusts for the degrees of freedom) was assessed with the normed $\chi^2$ statistic (Jöreskog, & Sörbom, 1979), where values between 1 and 3 are generally taken as indicative of acceptable fit (Carmines & McIver, 1981), and the adjusted goodness-of-fit index (AGFI), where fit is considered good for values of AGFI greater than .90 (Loehlin, 1987). The Root Mean Square Error Approximation (RMSEA) was used to indicate how well the model fit in the population, with values less then .10 indicating good fit and values less than .05 indicating very good fit (Steiger, 1988). Nested models were compared using the chi-square difference or delta chi-square ($\Delta \chi^2$) test (Bollen, 1989; Loehlin, 1987).

**One-Factor Model.** This model did not fit the data well. All fit indices were outside the range considered acceptable: $\chi^2 = 38.926$ ($p = .000$); GFI = 0.875; RMR = 0.146; TLI = 0.747; CFI = 0.874; normed $\chi^2 = 7.785$; AGFI = 0.624; RMSEA = 0.252. A large modification index (36.123) was observed for the error variances between items 1 and 2. The model was modified accordingly and refitted.

**Modified One-Factor Model.** Allowing the error variances for items 1 and 2 to correlate resulted in a substantial improvement in fit: $\chi^2 = 9.770$ ($p = .045$); GFI = 0.971; RMR = 0.084; TLI = 0.946; CFI = 0.979; normed $\chi^2 = 2.442$. Fit as measured by the AGFI and RMSEA indices were borderline but still acceptable: AGFI = 0.890; RMSEA = 0.103.

**Two-Factor Model.** Items 1, 2, and 3 defined the factor for ‘present’ life satisfaction, while items 4 and 5 defined the factor for satisfaction with life in the ‘past’. Because an additional parameter was estimated, this model showed improved fit over the one-factor model: $\chi^2 = 27.207$ ($p = .000$); GFI = 0.918; RMR = 0.066; TLI = 0.784; CFI = 0.88; normed $\chi^2 = 6.802$; AGFI = 0.692; RMSEA = 0.218. Since the one-factor model was nested within the two-factor model, the difference in fit between these two models was directly compared with the delta chi-square, which showed that the two-factor model fit significantly better than the one-factor model: $\Delta \chi^2 = 11.719$ ($p < .001$). The inter-factor correlation of 0.72 ($p < .01$) is somewhat smaller than the 0.86 correlation reported by McDonald (1999), but still large enough to question whether these factors should be considered as truly distinct constructs. Although this two-factor model did fit significantly better than the one-factor model, the modification index between the error variances of items 1 and 2 was still large (24.704).

**Modified Two-Factor Model.** The modified two-factor model allowed the errors for
items 1 and 2 to correlate. This modified model fit the data exceptionally well: $\chi^2=2.543$ ($p=.468$); GFI=0.992; RMR=0.027; TLI=1.006; CFI=1.000; normed $\chi^2=0.848$; AGFI=0.992; RMSEA=0.000. The factors were also significantly and positively correlated, $r=0.81$ ($p<.01$). Compared to the modified one-factor model, the modified two-factor model showed a significant improvement of fit: $\Delta\chi^2=7.227$ ($p<.01$).

**Which Model?**

In terms of fit statistics, the modified one- and two-factor models fit the data best. However, the substantial improvement in fit associated with the modified models over the non-modified models resulted from allowing the error variances of items 1 and 2 to correlate. One justification for allowing these errors to correlate is the possible redundant item content (Byrne & Shavelson, 1996) associated with these two items. Item 1 states: “in most ways my life is close to my ideal” while Item 2 states: “the conditions of my life are excellent” suggesting that the difference between an “ideal life” and an “excellent life” may not have been perceived as particularly distinct to the Hong Kong students. In fact, the mean response for each item was just over 5, and their correlation was fairly high (.70).

Although the modified two-factor model fit significantly better than the modified one-factor model, examination of the fit indices for the modified two-factor model suggests possible over fitting: (1) normed $\chi^2$ less than one, (2) RMSEA equal to zero, and (3) other fit indices near or at their maximum values. Another problem with the modified two-factor model included the relatively high correlation (.81) between the factors, which makes treating these factors as distinct questionable, and the fact that with two factors and
only five items one factor is defined by three items while the other factor is defined by only two items. Clearly a doublet factor is to be avoided, if possible, as defining a construct (McDonald, 1985). Therefore, the modified one-factor model — which showed acceptable fit, was more parsimonious, and fit within the theoretical framework associated with the scale’s development as a unidimensional measure of general overall life satisfaction — was retained for the Hong Kong student data (see Fig. 1).

**DISCUSSION**

Treating the SWLS as a undimensional measure of general overall life satisfaction was argued to be appropriate for the Hong Kong student data, justifying the use of a single sum score for the scale. More work, however, is needed to better understand why the mean SWLS score for Hong Kong student data was more similar to the normative data for American university students than it was to the normative data for mainland Chinese and Korean university students. One possible answer is that it reflects a greater degree of westernization in the Hong Kong university student population. While on the surface such an argument might seem credible, in reality it has little support. Even though Hong Kong has many trappings of a Western city especially in the commercial and economic sense, it is still a uniquely Chinese city where traditional Chinese values and culture hold sway for the majority of its inhabitants. Thus cross-cultural differences in how Hong Kong students respond to Western instruments even after translation are often observed. A case in point is the elevated response level on some MMPI scales noted by Cheung and Song (1989) and shown to reflect cultural differences rather than being indicative of an elevation in psychopathology. Another plausible explanation is that the Chinese and Korean studies reported in Pavot and Diener (1993) were both conducted in the early 1990s, so the situation may be considerably different in these student populations now. And since both studies used a translated instrument, this might also account for some of the large mean score differences.

However, some support for treating the general concept of life satisfaction or subjective well-being as measured by the SWLS as relatively culture free comes from a study conducted by Zhang and Norvilitis (2002) who found that the concept of psychological well-being could be validly assessed in Chinese culture with Western-developed instruments.

Therefore, higher SWLS scores for the Hong Kong student data might merely reflect their elite group status as graduate students. Intuitively, such an argument would seem to make sense. Graduate students are indeed an elite group, but it does not logically follow that they will experience a greater sense of general overall life satisfaction given that they are faced with constant competition amongst themselves for academic standing and financial resources. Couple this with heavy workloads, deadline pressures, and high expectations to perform well and one has the makings for a very stressful life situation suggesting that such a group of people might actually show less overall satisfaction with life. In fact, however, as already mentioned, the average overall life satisfaction for the
participants in this study and the variability of their scores were similar for the most part with results for American college students and for various adult groups in the general population as reported in Pavot and Diener (1993). The lowest average overall life satisfaction scores were reported for groups such as inmates and hospital inpatients. So differences do exist between groups, but participants in this study do not appear to be particularly more satisfied with their life than other groups in the general population as reported in Pavot and Diener (1993). Nevertheless, further research is warranted.

In summary, the good reliability and validity of the SWLS across cultures and even across time along with the scale’s brevity makes it ideally suited for inclusion within a battery of scales where an overall measure of life satisfaction might be of ancillary use such as in the teacher burnout research or in research interested in gauging the impact of changes associated with major life events such as divorce, job change, or job loss, say. While many instruments exist to assess specific constructs within these research areas, their specificity may in fact miss a general sense of overall well-being or lack thereof which a broad-band instrument like the SWLS is designed to capture.

REFERENCES


