Urdu translation of Flourishing Scale: Evidence for the validity and measurement invariance across gender

Sadia Niazi[1], Adnan Adil[2]

[1] Lecturer, Department of Psychology, University of Sargodha
[2] Assistant Professor, Department of Psychology, University of Sargodha

livespirit786@yahoo.com

♆ ABSTRACT. Una valutazione completa del benessere comporta la misurazione delle prospettive sia edoniche che eudaimoniche di benessere e la *Flourishing Scale*, composta da 8 item, valuta entrambi questi aspetti del benessere. Pertanto, la ricerca corrente si propone di tradurre la *Flourishing Scale* in urdu e validarla per la popolazione pakistana. La Flr.S è stata tradotta nella lingua urdu attraverso la procedura di back translation standard. Si è esaminata l'affidabilità test-retest (15 giorni) per le versioni urdu-inglese e inglese-urdu in un campione mirato di studenti universitari (N = 60). In un campione casuale separato di 574 studenti universitari pakistani (maschi = 235 e femmine = 339), sono state somministrate le versioni in urdu della Flr.S e della *Core Self-evaluation Scale* per stabilire la validità di costrutto e fattoriale della versione in urdu della Flr.S. L'affidabilità test-retest di due settimane per le versioni urdu-inglese e inglese-urdu era >.90. L'analisi fattoriale confermativa (CFA) della Flr.S ha rivelato una soluzione a fattore singolo, dimostrando l'invarianza configurale, metrica e scalare tra i generi. Le medie latenti dei ragazzi e delle ragazze sulla Flr S erano comparabili: la Flr.S ha dimostrato una validità di costrutto. I nostri risultati suggeriscono che la versione urdu della Flr.S è una misura affidabile e valida del costrutto del fiorire per entrambi i generi nelle popolazioni di lingua urdu.

♆ SUMMARY. *A comprehensive assessment of well-being involves the measurement of both hedonic and eudaimonic perspectives of well-being and the 8-item Flourishing Scale assesses both of these aspects of well-being. Therefore, the current research translated Flr.S into Urdu and validated it for the Pakistani population. Flr.S was translated into the Urdu language through the standard back-translation procedure. Test-retest reliability (15 days) for the Urdu-English and English-Urdu versions was established in a purposive sample of university undergraduates (N = 60). In a separate random sample of 574 Pakistani university undergraduates (males = 235 and females = 339), the Urdu version of Flr.S and Urdu version of Core Self-evaluation Scale were administered to establish the factorial and the construct validity of Urdu Flr.S. The test-retest reliability of two-week for the Urdu-English and English-Urdu versions of Flr.S was >.90. The confirmatory factor analysis (CFA) of Flr.S revealed a single factor solution, which demonstrated evidence for the configural, metric, and scalar invariance across genders. The latent means of males and females on the Flr.S were comparable: the Flr.S demonstrated construct validity. Our findings suggest that the Urdu version of the Flr.S is a reliable and valid measure of flourishing for both genders in the Urdu speaking population.*

Keywords: *Flourishing, Translation, Cross-language validation, Measurement invariance*

INTRODUCTION

Flourishing means to live within an optimal range of human ability. It is a mixture of both functioning and feelings. Flourishing is considered a sign of a high level of mental well-being (Huppert, 2009). However, the concept of flourishing is based on recent theories of social and psychological well-being. One of the most prominent conceptions of psychological well-being is based on the eudaimonic and hedonic paradigms (Woyciekoski, Stenert & Hutz, 2012). Hedonia is subjective happiness that involves seeking pleasure and avoiding pain (Diener, Oishi & Lucas, 2003) and it can be considered as subjective well-being characterized by satisfaction with life and positive affect. On the other hand, eudaimonia refers to the personal experience of self-realization, personal growth, and meaning in life (Ryan & Deci 2001). More precisely, it can be considered as psychological well-being (Waterman et al., 2010). The instruments developed in consonance of this conception of psychological well-being assess people's perception and life evaluation of happiness (Diener et al., 2003).

Previous literature showed that most of the self-report measures for the assessment of well-being were either based on eudaimonic perspective (e.g., Ryff & Keyes, 1995; Waterman et al., 2010) or hedonic perspective (scales assessing negative affect, life satisfaction, and positive affect). Based on the eudaimonic perspective, Ryff and Keyes' (1995) *Psychological Well-being Scale* comprised six theoretically derived dimensions including personal growth, autonomy, the meaning of life, constructive relations with others, mastery, and acceptance of self. The CFA of the *Psychological Well-being Scale* through maximum likelihood estimation in a sample of 1108 adults supported the proposed six-factor model of psychological well-being (χ^2 = 339, df = 120, AGFI = .89, BIC = −167). Moreover, the *Psychological Well-being Scale* was found to be positively related to *Life Satisfaction Index* (Neugarten, Havighurst & Tobin, 1961) and *Affect Balance Scale* (Bradburn, 1969); and negatively related to *Zung Depression Scale* (Zung, 1965), which established evidence for its construct validity (Ryff & Keyes, 1995). Currently, the Ryff's *Psychological Well-being Scale* has three versions. The longest version comprises 84 items (14 for each scale), the mid-length version includes 54 items (9 per scale), and the shortest version comprises 18 items (3 per scale).

Questionnaire of Eudaimonic Well-being (Waterman et al., 2010) is a 21-item scale, which assesses subjective well-being. This questionnaire was based on eudaimonic identity theory (Waterman, 2007) and its unidimensional factorial structure was confirmed in a CFA using the maximum likelihood estimation (χ^2 = 22.59, df = 5, CFI = .99, NNFI = .98, RMSEA = .065, SRMR = .018) in a large sample of college students (N = 1728) of the USA. The factor loading of the items ranged from .60 to .80 and the Cronbach alpha value of scale was .85. The construct validity of *Questionnaire of Eudaimonic Well-being* was established as it was found to be positively correlated with self-esteem (measured through the *Rosenberg Self-Esteem Scale*; Rosenberg, 1965) and internal locus of control (measured through *Locus of Control Scale*; Côté, 1997) and negatively related to general anxiety (measured through *Beck Anxiety Inventory*; Beck, Steer & Carbin, 1988) and social anxiety (measured through *Social Interaction Anxiety Scale*; Habke, Hewitt, Norton & Asmundson, 1997).

On the other hand, instruments based on the hedonic perspective have also been developed and these measures mainly cover three dimensions of hedonic well-being including life satisfaction, negative affect, and positive affect. For instance, *Satisfaction with Life Scale* was developed by Diener, Emmons, Larsen and Griffin (1985). This 5-item scale covered cognitive features of well-being. In an exploratory factor analysis (EFA) using principal axis factoring, Diener et al. (1985) found a unidimensional structure of the scale, which accounted for 66% variance. The item loadings ranged from .60 to .85. The *Satisfaction with Life Scale* showed favorable psychometric properties such as the high value of Cronbach alpha (.87) and test-retest reliability of two months (r = .82). The authors also established the construct validity of *Satisfaction with Life Scale* as it was found to be positively related to Bradburn's *Affect Balance Scale* (1969) and the *Rosenberg Self-Esteem Scale* (1965).

Positive and Negative Affect Schedule (PANAS) was developed by Watson, Clark and Tellegen (1988) for the assessment of the affective feature of well-being. The EFA with principal axis factoring extraction revealed two distinct factors, which accounted for 62.8% variance. The item loadings of all the items remained >.50, which showed that all items of PANAS were good indicators of their corresponding factor. The Cronbach's alpha of positive affect was .86 whereas the same for the negative affect was .87. The PANAS also demonstrated good temporal stability over two months (r = .87).

The 15-item *Subjective Well-being Scale* (Lawrence & Liang, 1988) integrated the *Affect Balance Scale* (Bradburn, 1969) and the *Life Satisfaction Index A* (Neugarten et al.,

1961) into one measure following the subjective well-being model of Liang (1985), which hypothesized subjective well-being to be comprised of four dimensions of including happiness, congruence, positive affect, and negative affect. The CFA of the scale using the maximum likelihood method in a large American sample yielded subjective happiness as a second-order factor with happiness, positive affect, negative affect, and congruence as the first-order factors. The first order loading ranged from -.64 to .89 and the measurement model demonstrated a good fit to the data. Furthermore, the authors established its measurement invariance across age and gender.

Lyubomirsky and Lepper (1999) developed a 4-item measure named *Subjective Happiness Scale* that was capable of measuring the extent to which a person was happy or not. The principal component analysis of the initial 13 items retained four items and suggested a single factor solution in a sample of college students. *Subjective Happiness Scale* was validated on a sample of 2732 adults. Lyubomirsky and Lepper (1999) indicated that this scale had a high Cronbach alpha value (.87) and test-retest reliability of three weeks was .79. The evidence for the convergent validity of the *Subjective Happiness Scale* was established as Lyubomirsky and Lepper (1999) found it to be positively related with self-esteem (measured through *Rosenberg Self-Esteem Scale*; Rosenberg, 1965), and optimism (measured through *Life Orientation Test*; Scheier & Carver, 1985).

The preceding review of the literature suggests that a complete perspective on psychological well-being needs the amalgamation of both hedonic and eudaimonic dimensions into one construct that could be reliably assessed. The construct of flourishing involves both hedonic and eudaimonic dimensions and can be conceptualized as the modern conception of well-being as it refers to a more global view of well-being, which not only covers life satisfaction but also includes personal growth, sense of purpose, self-acceptance, self-esteem, and competence (Keyes, 2006).

Built on the notion for assimilating the viewpoints of earlier assessments of well-being, Diener et al. (2009, 2010) developed the *Flourishing Scale* (*Flr.S*). The Flr.S is a mixture of psychological, emotional, and social well-being that includes the meaning, happiness, purpose in life, engagement, mastery, personal growth, being involved in one's work, being optimistic, and positive social relations with others. The Flr.S assesses core aspects of psychosocial functioning such as the fulfillment of competence and affiliation needs and self-acceptance as well as the ownership of psychological wealth

such as engagement and flow (Diener et al., 2010; Huppert & So, 2013).

The Flr.S comprises of eight items. Diener et al. (2009, 2010) explored the factor structure of the Flr.S through EFA where the factors were extracted through the principal axis factoring in a sample of 689 university students of the USA. The EFA revealed a single factor solution where item loading ranged from .61 to .77. This single factor structure accounted for 53% variance. Cronbach's alpha of the scale was .80. The Flr.S was found to be positively related to *Satisfaction with Life Scale* (Diener et al., 1985), Ryan and Deci's *Basic Need Satisfaction in General Scale* (2000), and Ryff's *Scale of Psychological Well-being* (2008). This established the evidence for the construct validity of the Flr.S.

Flr.S is one of the most widely used measures for assessing psychological well-being around the globe. It has also been translated into different languages i.e. Spanish (Checa, Perales & Espejo, 2018), Italian (Giuntoli, Ceccarini, Sica & Caudek, 2017), French (Villieux, Sovet, Jung & Guilbert, 2016), Chinese (Sumi, 2014), Brazilian (da Fonseca et al., 2015), and Portuguese (Silva & Caetano, 2013). Across all the translations of the Flr.S, the same factor structure has been confirmed as was reported for the original English version. In the present study, the Flr.S was not only translated into Urdu language but it was also validated on a large sample of students. This validation will allow researchers to use Flr.S within Pakistan or in other areas of the world where people speak Urdu.

Numerous studies have examined the gender differences in flourishing, for instance, Keyes and Simoes (2012) found a high level of flourishing in females as compared to males. Hone, Jarden and Schofield (2014) also reported a higher level of flourishing in women as compared to men. Contrary to the aforementioned results, some studies also reported non-significant gender differences in flourishing (Howell & Buro, 2015; Tang, Duan, Wang & Liu, 2016). Howell and Buro (2015) asserted that owing to the inconsistent findings concerning gender differences on the Flr.S, therefore, it needed further examination.

When measuring gender differences, it is important to determine whether Flr.S measures the same latent structure across gender. The present study is an empirical attempt to bridging this gap as it has explored whether the measurement structure underlying Flr.S is comparable across gender. To the best of our knowledge, it is the first study that has explored the latent structure of Flr.S across gender to establish evidence for its invariance across the two genders. For this

purpose, the CFA was performed to assure whether the original single-factor structure of Flr.S can be replicated in the Pakistani sample (N = 574). Secondly, multi-group CFAs were performed to measure invariance across gender and various models have been tested for establishing configural, metric, and scalar invariance of the scale. Finally, latent mean scores on the Flr.S were compared across both the genders.

To examine the construct validity of the Urdu translated Flr.S, the present study has also examined the pattern of relationships of flourishing with positive and negative core self-evaluation. Core-self-evaluation (CSE) is considered an overall perception of an individual's capability and worth as a human being (Judge et al., 1998) and it comprises of four components including self-esteem (the overall evaluation of one's worth), emotional stability (the ability to feel composed and protected), generalized self-efficacy (assessment of one's ability to effectively accomplish a variety of tasks), and locus of control (the faith that events in the lives of individuals come as a result of their actions or because of powerful others or fate). According to CSE theory, these four faces of personality unite to illuminate an individual's overall judgment of the worth that s/he has as a person (Judge, Locke, Durham & Kluger, 1998).

Positive core self-evaluation (PCSE) may act as a protective factor and contribute positively to psychological well-being. It is a broad latent trait and individual high on the PCSE is likely to be optimistic, composed, efficient, and self-assured. People having positive self-evaluations tend to be more emotionally stable, are more self-efficacious, have a higher degree of self-esteem, and demonstrate an internal locus of control (Judge & Bono, 2001). People's evaluations about their selves determine what they can do and what they can become; this self-evaluation contributes to better psychological and physical health. It promotes healthy functioning such as increasing coping ability, success, and satisfaction with life (Mann, Hosman, Schaalma & De Vries, 2004).

In contrast, people with negative core self-evaluation (NCSE) have poor self-esteem and unstable self-concept. They are emotionally unstable and have an external locus of control, which can play a significant role in the development of an array of mental and social problems (Mann et al., 2004). Judge, Erez, Bono and Thoresen (2003) suggested that a high level of psychological well-being is an outcome of PCSE because people with positive self-evaluations experience more positive emotions.

Zimmerman (2000) asserted that self-esteem and self-efficacy could explain significant variance in psychological well-being and they are key elements of psychological health. Brown (1998) argued that feeling good and mentally healthy is a basic human need that can be fulfilled when one has positive self-evaluation. Roddenberry and Renk (2010) found that people with an external locus of control had poor psychological health than that of their counterparts with an internal locus of control. Additionally, people with a high degree of self-esteem were less likely to be mentally ill as compared to their counterparts having a low degree of low self-esteem. These pieces of research evidence are suggesting that flourishing should have a positive relationship with the PCSE and a negative relationship with the NCSE.

METHOD

The present research comprised of two studies. Study 1 involved the translation of the Flr.S into the Urdu language, its cross-language validation, and assessment of its psychometric properties. Study 2 involved the assessment of factorial validity and measurement invariance of the Flr.S across gender.

STUDY 1

Study 1 comprised of two phases. Phase I involved the Urdu translation of Flr.S whereas phase II aimed at establishing the cross-language validation of the Flr.S.

Phase I: Translation and validation of the Flourishing Scale

This scale is free to use and formal permission for its translation into the Urdu language was sought from the author. In the first phase, the backward translation procedure (Brislin, 1986) was adopted. This procedure was further divided into four steps. The first step was a forward translation, which involved translation from English to the Urdu by following a parallel back-translation procedure. This step was performed to create conceptually equivalent translations for the culture of the target language.

Three bilingual experts (two assistant professors and one lecturer) who were native speakers of the target language and had fluency in English performed forward translations,

emphasizing the conceptual equivalence. While translation, they monitored that the translation and contents associated with semantic features of the original version must be maintained. The second step was the settlement of items for attaining the finest translated items, three independent Urdu translated versions were evaluated through a committee approach for assessing the conjectural consistency among items. They were requested to analyze each item and select one translated version which delivered the best sense of meaning for each item. The committee members made their assessments item-by-item and selected the best translation for each item. This gave us the finalized forward Urdu translations that comprised of best-translated items finalized through the agreement of all bilingual experts. In the next step, the back translation was performed. This step was planned to determine the theoretical equivalence of the finalized forward translation and the original version of the Flr.S Therefore, three bilingual experts who were unfamiliar with the original version of the scale were approached to translate the Urdu translated version of the Flr.S back into English. Thus, three independent English translations of Urdu translated Flr.S were achieved. Finally, the three bilingual expert's committee (two assistant professors and one lecturer) reviewed and compared the back-translated version of Flr.S with the original version. After the agreement on the translated version, to ensure the accuracy of Urdu translation with the original scale, the final version was sent to the original designer (Diener & Biswas-Diener, 2008), who finally approved the Urdu version.

Phase II: Cross-language validation

Cross-language validation of Urdu Flr.S was undertaken by comparing the Urdu translated version with the original English version of the scale. This step aided in establishing the excellence of the Urdu translated version to assess its empirical correspondence with the original English version. For this purpose, four groups of participants were given Flr.S in English to English, English to Urdu, Urdu to Urdu and Urdu to English orders twice with an interval of one week. Firstly, two groups were given Urdu and two groups were given English version of the Flr.S. Subsequently, after one week, all four groups were given Flr.S again but this time two of the groups received the same language version and two received different language version of the scale. The participants'

assignment to these four groups was random. The purpose behind this step was to measure the incongruity or similarity between Urdu and English versions of scales. The sample was distributed in such a way as to control the learning effect that might occur due to the testing of original and translated versions. This empirical equality was calculated by finding the correlations of test and retest phase of a week.

Sample

For cross-language validation, the sample of ($N = 60$) students was randomly selected from the University of Sargodha. The minimum age of students was 20 (range 20-35, $M = 25.3$, $SD = 5.3$) and they all were bilingual in English and Urdu. Further, the sample was grouped into four conditions. Each group comprised of 15 participants in the test and retest phase.

Instruments

Flourishing Scale was constructed by Diener and Biswas-Diener (2008) and was used to assess the psychological well-being. It is a brief 8-item scale that helps in measuring respondents' self-perceived success in imperative areas of life such as purpose, self-esteem, relationships, and optimism. The items of scales were scored by using a 7-point Likert scale ranging from 1 = strongly agree to 7 = strongly disagree. Diener and Biswas-Diener (2008) reported high reliabilities ($\alpha = .87$, test-retest of a week = .76) and convergence validity (.78) of Flr.S with *Satisfaction with Life Scale* (Diener et al., 1985).

Procedure

This study was designed to assess the validation of translated versions of scales which were administrated twice to the four groups of bilingual Pakistani married people in Urdu-Urdu, English-English, English-Urdu and Urdu-English sequence.

The tests were applied to participants in two settings. In the first trial, two groups were given Urdu versions and two were given English versions of both scales. In the second trial after 7 days, the same 60 participants were contacted to make their responses again, but in this trial, two groups were

given the same language version but the other two groups were given the opposite language version. This procedure was used to identify the point of equivalence or discrepancy between Urdu and English versions of scales. This empirical equivalence was evaluated by finding the correlations of test and retest phase of a week, which were depicted in Table 1.

STUDY 2: FACTORIAL VALIDITY

Since our data were normally distributed, we conducted CFA with maximum likelihood estimation for the confirmation of factor structure and measurement model of the Urdu Flr.S. For that purpose, various fit indices were examined and the fit criteria were established with the help of Root Mean Square Error of Approximation (RMSEA≤.50; the smaller is better); Normed Fit Index (NFI>.90); Comparative Fit Index (CFI>.90) (1990) and Goodness of Fit Index (GFI>.90). Table 3 depicts the model fit indices results and factor loading of CFA for Urdu translated Flr.S.

Sample

The sample (N = 574) of the present study was selected through cluster sampling. Since there was the probability of unidentified internal heterogeneity in the population of university students, therefore, the present study adopted a cluster sampling technique. Out of 36 teaching departments of the University of Sargodha (Main Campus), 18 departments were randomly selected as clusters through draw method. In that way, 50% of the entire population was considered as part of the study to make the sample more representative. In the next step, all students of BS (Honors, Semester VII) programs in various academic disciplines were selected. The sample included 235 males and 339 females with an age range of 18 to 25 years (M = 21, SD = 1.8 years).

Instrument

The 8-item Urdu version of the Flr.S was used to assess flourishing on on a 7-point Likert scale. The test-retest reliability of this scale over one week was .93.

Urdu translated version of the *Core Self-evaluation Scale* (Zia, 2016) was used to measure the core self-evaluations of

the respondents. It comprises of 12 items which are rated on a 5-point Likert scale with the response rate from 1 = strongly disagree to 5 = strongly agree. Item no. 2, 4, 6, 8, 10, and 12 measure NCSE and item no. 1, 3, 5, 7, 9, and 11 measure PCSE. The internal consistency of the scale was satisfactory (α = .84; Zia, 2016).

Procedure

Through random sampling, participants were personally approached in their classrooms. They were directed about the purpose of the study and were briefed about the relevant response format. Participant's queries were resolved and then they were requested to respond to each statement honestly. They were guaranteed the privacy and confidentiality of the information they provided. For scale completion, there was no time limit. Six hundred participants returned the scales while 574 scales were complete in all aspects and suitable for additional data analysis. Participants were appreciated at the end for their help and support.

RESULTS
Cross-Language Validation

Table 1 shows the correlations between all test-retest phases of Flr.S. Results depicted significant correlations among translated versions, which ranged from .93 to .97. Generally, results in Table 1 depicted empirical equivalence of Urdu translated version of Flr.S to its English version. Table 1 also portrays that the highest correlations existed between English-Urdu and Urdu-English versions.

Descriptive statistics of the Urdu Flr.S

Table 2 displays the descriptive statistics of all items of the Flr.S in terms of the entire sample and by the gender. The ratio of skewness coefficients to their standard error and the ratio of kurtosis coefficients to their standard error remained less than 3, which indicated that in the large sample, the distribution of items did not significantly deviate from normality (Brookshier & Boyd, 2016; Tabachnick & Fidell, 2003, 2007).

Table 1 – Correlations between Urdu and English versions of Flr.S (N = 60)

Scales	r
Test-retest Urdu-Urdu	.93***
Test-retest Urdu-English	.95***
Test-retest English-English	.94***
Test-retest English-Urdu	.97***

***$p<.001$.

Table 2 – Descriptive statistics for items of Flr.S (N = 574)

	Total		Males		Females	
Item	M±SD	Sk[a]	M±SD	Sk[b]	M±SD	Sk[c]
1	5.16±1.69	−1.20	5.16±1.7	−1.07	5.16±1.6	−1.31
2	5.52±1.33	−1.59	5.64±1.3	−1.63	5.44±1.3	−1.58
3	5.63±1.32	−1.54	5.61±1.3	−1.43	5.65±1.3	−1.63
4	6.02±1.13	−1.85	5.91±1.2	−1.48	6.10±1.0	−2.19
5	5.84±1.18	−1.60	5.85±1.1	−1.42	5.84±1.0	−1.72
6	5.84±1.27	−1.75	5.67±1.4	−1.41	5.96±1.1	−2.06
7	6.10±1.12	−2.02	6.05±1.0	−1.55	6.13±1.1	−2.33
8	5.92±1.17	−1.75	5.85±1.2	−1.46	5.97±1.1	−2.00

Note. [a] Standard error of skewness = .10; [b] Standard error of skewness = .15; [c] Standard error of skewness = .15.

Internal consistency

The Urdu Flr.S demonstrated a satisfactory level of internal consistency. The Cronbach's alpha of the scale for the whole sample was .80. Across the two genders, the Urdu Flr.S was also found to be quite reliable. For the males' sample, its Cronbach's alpha was .80 whereas the same for the females' sample was .79. Both subscales of the *Core Self-evaluation Scale* were also found to be internally consistent.

Factorial validity

To determine whether the original factor structure of the Flr.S can be replicated in the sample of the current study, the CFA was performed using the maximum likelihood estimation because our data were normally distributed. The results showed that the proposed model showed a very good model fit. The results of CFA indicated that the chi square to *df* ratio was 2.3 (χ^2 = 33, *df* = 14, *p*<.05), and other indices of model fit also showed a good model fit between the data and the model. The values of CFI, GFI, and NFI were all above .90 and hence met the criteria of fit indices. The value of RMSEA was .04 with a non-significant p-value (*p* = .48) with 95% CI (LL = .02, UL = .07) and the standardized RMR was well below the cut-off point of .05. The factor loading ranged from .35 to .74.

The values in Table 3 shows the standardized factor loading of all items of Flr.S in the whole sample, the sample of males only, and the sample of females only. The factor loadings of all items on a single latent factor were ≥.30, which supported the unidimensional structure of the Flr.S in the whole sample as well as in the separate groups of males and females. Furthermore, the fit indices indicated that the data fitted well to the model in the whole sample as well as its subgroups.

Table 4 displays the correlations and covariance of Flr.S computed in the entire sample, which suggests that all items of the Flr.S are positively correlated with one another.

Measurement invariance

Table 5 depicts the invariance analyses across gender for the CFA of the Flr.S through maximum likelihood estimation since our data were normally distributed. For the assessment of configural invariance, the factor structure of the scale was kept the same across both the gender i.e., the same number of factors with the same indicators were specified for both men and women. Results revealed that the data fitted well to both the genders, which provided the evidence that the measure was configuraly invariant across the two genders. For the assessment of metric invariance, the factor loadings of the measure were constrained to be

Table 3 – Standardized factor loadings and fit indices (N = 574)

	Standardized factor loading of indicators										Fit indices of models				
	1	2	3	4	5	6	7	8	χ^2	df	GFI	CFI	NFI	RMSEA	St.RMR
Total	.35	.45	.50	.53	.60	.72	.74	.53	33**	14	.98	.98	.97	.04	.02
Males	.30	.41	.50	.56	.62	.80	.79	.44	37***	14	.97	.97	.95	.07	.03
Females	.46	.52	.48	.47	.58	.65	.70	.63	20	12	.97	.98	.95	.04	.03

Legenda. df = degree of freedom; GFI = Goodness of Fit Index; CFI = Comparative Fit Index; NFI = Normed Fit Index; RMSEA = Root Mean Square Error of Approximation; RMR = Root Mean Residual.
**p<.01, *** p<.001.

Table 4 – Covariance and correlation matrices of Urdu version of Flr.S (N = 574)

	Flr.S1	Flr.S2	Flr.S3	Flr.S4	Flr.S5	Flr.S6	Flr.S7	Flr.S8
Flr.S1	–	.38***	.30***	.20***	.22***	.24***	.26***	.18***
Flr.S2	.87	–	.44***	.26***	.43***	.26***	.37***	.34***
Flr.S3	.69	.78	–	.36***	.43***	.30***	.38***	.22***
Flr.S4	.39	.39	.55	–	.52***	.39***	.36***	.26***
Flr.S5	.45	.52	.68	.71	–	.43***	.42***	.31***
Flr.S6	.53	.45	.51	.57	.66	–	.55***	.41***
Flr.S7	.50	.55	.58	.46	.57	.78	–	.39***
Flr.S8	.37	.54	.35	.35	.43	.61	.52	–

Note. Below the diagonal is the covariance matrix.
***$p<.001$.

Table 5 – Invariance tests for the Flr.S across gender (N = 574)

Model	χ^2	df	CFI	RMSEA	Model Comparison	$\Delta\chi^2$	Δdf	ΔCFI	ΔRMSEA
1. M_1	57.99	28	.976	.043					
2. M_2	69.61	35	.967	.046	2 vs 1	11.62	7	.009	.003
3. M_3	80.80	43	.957	.047	3 vs 1	22.81	15	.019	.004

Legenda. df = degree of freedom; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; M_1 = invariant form model (configural invariance), M_2 = invariant loading model (metric invariance), M_3 = invariant intercept model (scalar invariance).

equal across both genders. Results suggested that the data again fitted well to the model, which provided the evidence for metric invariance. Furthermore, the comparison of configuraly invariant and metrically invariant models revealed nonsignificant chi square difference tests and negligible differences in other measures of fit indices, which were well below the critical value suggested by Chen (2007). This indicated that data fitted to both models 1 and 2 equally well. Finally, the scalar invariance was estimated by constraining the factor loadings as well as intercepts of the measure to be equal across the two genders. This scalarly invariant model also demonstrated excellent fit to the data. Moreover, the comparison of scalarly invariant models with the configuraly invariant models suggested

non-significant chi square difference tests with negligible differences in CFI and RMSEA values, which were quite small as compared to the critical values suggested by Chen. This demonstrated full scalar invariance of the Flr.S across both the gender and revealed that both scalarly invariant models and the configuraly invariant models fitted to the data equally well.

Latent mean differences

To estimate the differences in the latent means of males and females on the Flr.S, females were selected as the reference group and its latent mean was fixed to zero. However, the latent mean of the males was estimated. The results revealed that the latent means of males and females did not differ significantly from each other (CR = 1.45, p = .15).

Construct validity

Table 6 shows the means, standard deviations, and reliability coefficients of Flr.S and CSE scales. There is no evidence of the restricted range in the scores. The ratio between the values of skewness and its standard error is less than 2, which suggests that the variables were symmetrical in their distribution. Table 6 provides the evidence for the convergent validity of the Flr.S as it indicates a significant

positive correlation between Flr.S and PCSE and a significant negative correlation of Flr.S with NCSE. Furthermore, the relationship between flourishing and positive core self-evaluation was stronger than the relationship between flourishing and NCSE. These results provide evidence for the construct validity of Urdu translated Flr.S.

DISCUSSION

The primary purpose of the present study was to translate and validate Flr.S. for the Pakistani population. It also intended to establish the empirical evidence for the measurement invariance of the Urdu Flr.S across gender. Overall, results showed a single-factor structure of Flr.S with a reasonable level of internal consistency and temporal stability.

The findings of the present study indicate that the Urdu translated version of Flr.S has an adequate level of reliability and construct validity for the Pakistani population. The CFA of the Urdu Flr.S indicated that the single factor solution demonstrated a very good fit to the data, which is consistent with the pertinent literature. The original English version of the Flr.S and its various translations have confirmed the single factor structure of the scale and the present research was no exception. Thus, we have established the factorial validity of the Urdu Flr.S and empirically demonstrated its factorial equivalence with the original English version as well as the

Table 6 – Descriptive statistics, alpha reliability, and correlations of the scales of the present study (N = 574)

Variables	M	SD	α	2	3	Range		Ska
						Actual	Potential	
1. Flr.S	46	6.6	.80	.39***	−.18***	14-56	8-56	−.05
2. PCSE	22	3.5	.83	–	−.28***	9-30	6-30	−.06
3. NCSE	18	3.5	.79	–	–	8-30	6-30	−.03

Legenda. Flr.S = Flourishing Scale; PCSE = Positive Core self-evaluation subscale of Core Self-evaluation Scale; NCSE = Negative Core self-evaluation subscale of Core Self-evaluation Scale.

Note. [a] Standard error of skewness = .10; ***p<.001.

other translated versions of the scale. The brevity and simple factor structure make the Flr.S one of the most frequently used instruments for the measurement of well-being, specifically when investigators have a limited period or they want to administer a battery of instruments in a single study.

Albeit the Flr.S has been used in varied cultural and linguistic contexts and with numerous scale versions, none of the studies deliberated on the assessment of measurement invariance of the Flr.S across the gender. Factor uniformity requires to be confirmed statistically because it is crucial for the comparison and clarification of psychological constructs in dissimilar groups such as across the two genders.

We followed Vandenberg and Lance's (2000) proposal to investigate the measurement invariance with successively restrictive phases. In step 1, the configural invariance test was performed to induct a baseline model across groups. The configural invariance test permits factor loadings, intercepts, and residuals were assessed freely. The establishment of a configural invariance test infers that the conceptual framework is the same across groups. If the data is not in the favor of configural invariance test, the measurement invariance test should be dismissed. Our data delivered strong empirical support for the configural invariance as the equivalent factorial structure of the Flr.S was fitted well to both males' and females' sample, therefore we advanced to step 2 in which we tested the metric invariance model. In this model, all factor loadings were embarrassed the same. The metric invariance is a weak invariance test and the establishment of this test means that different groups reacted to the indicators in the same way. Our results supported the metric invariance of the Flr.S. since the model with constrained loadings and the model with freely estimated loadings did not significantly vary in terms of their fit to the data.

In step 3, we performed the scalar invariance model in which the factor loading and indicator intercepts were constrained to be identical across groups. The scalar invariance test is a strong invariance test, which advocates that the measurement model has alike scale across the different groups. Our results provided support for the scalar invariance of the Flr.S as the model with constrained factor loadings and intercepts and the model where these parameters were freely estimated did not significantly differ in terms of their fit to the data. The metric and scalar invariances tests were inspected by measuring the change in the χ^2, CFI, and RMSEA values. The establishment of this test is a prerequisite before the latent means can be compared across groups.

Finally, we did the comparison of latent mean differences between gender. Explicitly, a full scalar invariance model was used as the baseline. To compare latent mean between genders, we constrained the females' group latent mean to 0 and the latent means of the males' group was free to estimate. We used the value of the critical ratio (CR) to calculate latent mean differences. CR is calculated by parameter estimate divided by its standard error, which tests whether the coefficient is significantly different from 0. A CR value larger than 1.96 indicates statistically significant differences in the latent means (Byrne, 2013). A positive CR suggests that the comparison group has a greater latent mean than the reference group. Equally, a negative CR submits that the comparison group's latent mean is lesser than the reference group. In our case, the CR was non-significant, which advocated that the latent means of both males and females on the Flr.S were equivalent.

We also established the evidence for the construct validity of the Urdu Flr.S. The FLr.S was found to be correlated with the PCSE and the NCSE in the expected direction, which provided the evidence for the convergent validity. More specifically, our results suggested that flourishing was positively related to positive core self-evaluation and it was negatively related to the negative core self-evaluation. Furthermore, the Flr.S was more strongly related to the positive core self-evaluation than the negative core self-evaluation, which yielded evidence for the discriminant validity of the Urdu Flr.S.

Pertinent literature supports this pattern of relationships between flourishing and core self-evaluations. For instance, Valkenburg, Peter and Schouten (2006) claimed that evaluations regarding self, such as self-esteem and self-efficacy might influence one's degree of well-being. Bornstein, Davidson, Keyes and Moore (2003) defined psychological well-being as the successful state of performance throughout life, which can lead to the integration of cognitive, physical, and socio-emotional functions that results in productive activities. There are many sources of psychological well-being among which PCSE is an important factor involving self-esteem, self-efficacy, internal locus of control, and emotional stability. Diener and Suh (2000) claimed that these sources can be individual's perceptions about his/her self, appraisals of one's happiness, life happiness, emotional state, personal worthiness, personal value, positive evaluation of one's success, and happiness. All these psychological factors have a positive influence on the psychological well-being of individuals (Bornstein et al., 2003). According to

Kelly (2017), well-being is considered as a positive sense of personal efficacy, strong self-esteem, and emotional stability, which are positively associated with better health outcomes; contrarily, low self-esteem is strongly associated with poor health outcomes.

CONCLUSION

The current research provided details of the Urdu translations of the Flr.S and established the evidence for its factorial structure, internal consistency, temporal stability, and construct validity. Our results on the measurement invariance of the Urdu Flr.S demonstrated the construct of flourishing remains the same across both genders in the Pakistani population. Our findings suggest that the Urdu Flr.S is a psychometrically sound measure that can be reliably used to assess well-being. It has demonstrated a good degree of temporal stability and internal consistency. Furthermore, its single-factor structure was comparable to that of the original English version, which makes it equivalent to the original scale in terms of the factor structure. It also demonstrated construct validity as its positive relationship with the positive core self-evaluations was stronger than its negative correlation with the negative core self-evaluations albeit the statistical significance of both correlations. Our findings on the measurement invariance indicate that the Urdu Flr.S has

equivalent structure, similar responses, and comparable latent scores for both males and females, therefore, the construct of flourishing seems to be the same for both the genders.

Limitations and suggestions

The present study was not free of weaknesses. This study was based on a sample of university undergraduate students that could not be a true representative of the general population. Forthcoming research with longitudinal design and the larger, more diverse sample will be more valuable to continue the research of the measurement invariance of the 8-item version of the Flr.S. In the present study the invariance across gender was determined, the same procedure should be employed to determine variance across other demographics. Notwithstanding these restrictions, the findings provided support for the use of the Flr.S as a reliable and valid measurement tool for the assessment of well-being across both the genders. In conclusion, even if the current research displayed the factor invariance of the 8-item version of the Flr.S across gender, further cross-cultural studies are essential to well understand the structure and the nature of this construct. This will lead to the growth of knowledge and warrant the generalizability of the Flr.S as well as its applicability in different educational and cultural contexts.

References

BECK, A.T., STEER, R.A. & CARBIN, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8* (1), 77-100. doi. org/10.1016/0272-7358(88)90050-5

BORNSTEIN, M.H., DAVIDSON, L., KEYES, C.L. & MOORE, K.A. (Eds.) (2003). *Well-being: Positive development across the life course*. Psychology Press.

BRADBURN, N.M. (1969). T*he structure of psychological wellbeing*. Chicago: Aldine.

BRISLIN, R.W. (1986). Research instruments. *Field Methods in Cross-*

cultural Research, *8*, 137-164.

BROOKSHIER, A. & BOYD, R. (2016). *SPSS data screening*. Retrieved May 30, 2020 from https://nau.edu/wp-content/ uploads/sites/30/COERC-SPSS-Data-Screening-Tutorial.pdf

BYRNE, B.M. (2013). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Psychology Press.

CHECA, I., PERALES, J. & ESPEJO, B. (2018). Spanish validation of the Flourishing Scale in the general population. *Current Psychology*, *37* (1), 949-956. doi.org/10.1007/s12144-017-9581-0

CHEN, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14* (3), 464-504. doi.org/10.1080/10705510701301834

CÔTÉ, J.E. (1997). An empirical test of the identity capital model. *Journal of Adolescence*, *20* (5), 577-597. doi.org/10.1006/jado.1997.0111

da FONSECA, P.N., da SILVA NASCIMENTO, B., BARBOSA, M., GOMES, L.H., VIONE, K.C. & VELOSO GOUVEIA, V. (2015). Flourishing scale: Evidence of its suitability to the Brazilian context. *Social Inquiry into Well-Being*, 1 (2), 33-40.

DIENER, E. & BISWAS-DIENER, R. (2008). *Happiness: Unlocking the mysteries of psychological wealth*. Malden, MA: Blackwell Publishing.

DIENER, E., EMMONS, R.A., LARSEN, R.J. & GRIFFIN, S. (1985). The satisfaction with Life Scale. *Journal of Personality Assessment*, *49* (1), 71-75. doi.org/10.1207/s15327752jpa4901_13

DIENER, E., OISHI, S. & LUCAS, R.E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual Review of Psychology*, *54* (1), 403-425. doi.org/10.1146/annurev.psych.54.101601.145056

DIENER, E. & SUH, E.M. (2000). Measuring subjective well-being to compare the quality of life of cultures. In E. Diener & E.M. Suh (Eds.), *Culture and subjective well-being*. Cambridge: MIT Press.

DIENER, E., WIRTZ, D., BISWAS-DIENER, R., TOV, W., KIM-PRIETO, C., CHOI, D. et al. (2009). New measures of well-being. In E. Diener (Ed.), *Assessing well-being: The collected works of Ed Diener*. New York: Russell Sage Foundation. doi: 10.1007/978-90-481-2354-4 12

DIENER, E., WIRTZ, D., TOV, W., KIM-PRIETO, C., CHOI, D., OISHI, S. et al. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97* (2), 143-156. doi:10.1007/s11205-009-9493-y

GIUNTOLI, L., CECCARINI, F., SICA, C. & CAUDEK, C. (2017). Validation of the Italian versions of the Flourishing Scale and the Scale of Positive and Negative Experience. *Sage Open*, *7* (1). doi.org/10.1177/2158244016682293

HABKE, A.M., HEWITT, P.L., NORTON, G.R. & ASMUNDSON, G. (1997). The Social Phobia and Social Interaction Anxiety Scales: An exploration of the dimensions of social anxiety and sex differences in structure and relations with pathology. *Journal of Psychopathology and Behavioral Assessment*, *19* (1), 21-39. doi.org/10.1007/BF02263227

HERRICK, C.A. & BROWN, H.N. (1998). Underutilization of mental health services by Asian Americans residing in the United States. *Issues in Mental Health Nursing*, *19* (3), 225-240. doi.org/10.1080/016128498249042

HONE, L., JARDEN, A. & SCHOFIELD, G. (2014). Psychometric properties of the Flourishing Scale in a New Zealand sample. *Social Indicators Research*, *119* (2), 1031-1045.

HOWELL, A.J. & BURO, K. (2015). Measuring and predicting student well-being: Further evidence in support of the flourishing scale and the scale of positive and negative experiences. *Social Indicators Research*, *121* (3), 903-915.

HUPPERT, F.A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being*, *1* (2), 137-164. doi.org/10.1111/j.1758-0854.2009.01008.x

HUPPERT, F.A. & SO, T.T. (2013). Flourishing across Europe: Application of a new conceptual framework for defining well-being. *Social Indicators Research*, *110* (3), 837-861. doi.org/10.1007/s11205-012-0030-z

JUDGE, T.A. & BONO, J.E. (2001). Relationship of core self-evaluations traits - self-esteem, generalized self-efficacy, locus of control, and emotional stability - with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, *86* (1), 80-92.

JUDGE, T.A., EREZ, A., BONO, J.E. & THORESEN, C.J. (2003). The Core Self-evaluation Scale: Development of a measure. *Personnel Psychology*, *56* (2), 303-331. doi.org/10.1111/j.1744-6570.2003.tb00152.x

JUDGE, T.A., LOCKE, E.A., DURHAM, C.C. & KLUGER, A.N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, *83* (1), 17-34.

KELLY, S. (2017). *The effects of gender role orientation, self-esteem, and self-efficacy on help-seeking behaviors and wellbeing*. [Higher Diploma Final Year Project, Dublin Business School]. hdl.handle.net/10788/3307

KEYES, C.L. (2006). Subjective well-being in mental health and human development research worldwide: An introduction. *Social Indicators Research*, 77 (1), 1-10. doi.org/10.1007/s11205-005-5550-3

KEYES, C.L. & SIMOES, E.J. (2012). To flourish or not: Positive mental health and all-cause mortality. *American Journal of Public Health*, 102 (11), 2164-2172. doi: 10.2105/AJPH.2012.300918

LAWRENCE, R.H. & LIANG, J. (1988). Structural integration of the Affect Balance Scale and the Life Satisfaction Index A: Race, sex, and age differences. *Psychology and Aging, 3* (4), 375-384. doi.org/10.1037/0882-7974.3.4.375

LIANG, J. (1985). A structural integration of the Affect Balance Scale and the Life Satisfaction Index A. *Journal of Gerontology*, *40*, 552-561.

LYUBOMIRSKY, S. & LEPPER, H.S. (1999). A measure of

subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46* (2), 137-155. doi: 10.1023/A:1006824100041

MANN, M.M., HOSMAN, C.M., SCHAALMA, H.P. & De VRIES, N.K. (2004). Self-esteem in a broad-spectrum approach for mental health promotion. *Health Education Research*, *19* (4), 357-372. doi.org/10.1093/her/cyg041.

NEUGARTEN, B.L., HAVIGHURST, R.J. & TOBIN, S.S. (1961). The measurement of life satisfaction. *Journal of Gerontology*, *16*, 134-143. doi.org/10.1093/geronj/16.2.134

RODDENBERRY, A. & RENK, K. (2010). Locus of control and self-efficacy: Potential mediators of stress, illness, and utilization of health services in college students. *Child Psychiatry and Human Development*, *41*(4), 353-370. doi.org/10.1007/s10578-010-0173-6

ROSENBERG, M. (1965). Rosenberg's Self-esteem Scale (RSE). *Acceptance and Commitment Therapy. Measures Package*, *61* (52), 18-60.

RYAN, R.M. & DECI, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*, 68-78. doi:10.1037/0003-066X. 55.1.68

RYAN, R.M. & DECI, E.L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52* (1), 141-166. doi.org/10.1146/annurev.psych.52.1.141

RYFF, C.D. (2008). *Scales of psychological well-being*. The University of Wisconsin, Institute on Aging: The Author.

RYFF, C.D. & KEYES, C.L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69* (4), 719-727. doi.org/10.1037/0022-3514.69.4.719

SCHEIER, M.F. & CARVER, C.S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, *4* (3), 219-247. doi.org/10.1037/0278-6133.4.3.219

SILVA, A.J. & CAETANO, A. (2013). Validation of the Flourishing Scale and scale of positive and negative experience in Portugal. *Social Indicators Research*, *110* (2), 469-478. doi.org/10.1007/s11205-011-9938-y

SUMI, K. (2014). Reliability and validity of Japanese versions of the Flourishing Scale and the scale of positive and negative experience. *Social Indicators Research*, *118* (2), 601-615. doi.org/10.1007/s11205-013-0432-6

TABACHNICK, B.G. & FIDELL, L.S. (2003, May). *Preparatory data analyses*. Paper presented at the annual meeting of the Western Psychological Association, Vancouver, BC, Canada.

TABACHNICK, B.G. & FIDELL, L.S. (2007). *Using multivariate statistics (5th ed.)*. Boston, MA: Allyn & Bacon.

TANG, X., DUAN, W., WANG, Z. & LIU, T. (2016). Psychometric evaluation of the simplified Chinese version of the Flourishing Scale. *Research on Social Work Practice*, *26* (5), 591-599. doi.org/10.1177/1049731514557832

VALKENBURG, P.M., PETER, J. & SCHOUTEN, A.P. (2006). Friend networking sites and their relationship to adolescents' well-being and social self-esteem. *CyberPsychology & Behavior*, *9* (5), 584-590. doi.org/10.1089/cpb.2006.9.584

VANDENBERG, R.J. & LANCE, C.E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3* (1), 4-70. doi.org/10.1177/109442810031002

VILLIEUX, A., SOVET, L., JUNG, S.C. & GUILBERT, L. (2016). Psychological flourishing: Validation of the French version of the Flourishing Scale and exploration of its relationships with personality traits. *Personality and Individual Differences*, *88*, 1-5. doi.org/10.1016/j.paid.2015.08.027

WATERMAN, A.S. (2007). Doing well: The relationship of identity status to three conceptions of well-being. *Identity: An International Journal of Theory and Research*, *7* (4), 289-307. doi.org/10.1080/15283480701600769

WATERMAN, A.S., SCHWARTZ, S.J., ZAMBOANGA, B.L., RAVERT, R.D., WILLIAMS, M.K., BEDE AGOCHA, V. & BRENT DONNELLAN, M. (2010). The Questionnaire for Eudaimonic Well-Being: Psychometric properties, demographic comparisons, and evidence of validity. *The Journal of Positive Psychology*, *5* (1), 41-61. doi.org/10.1080/17439760903435208

WATSON, D., CLARK, L.A. & TELLEGEN, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54* (6), 1063-1070. doi:10.1037/0022-3514.54.6.1063

WOYCIEKOSKI, C., STENERT, F. & HUTZ, C.S. (2012). Determinantes do bem-estar subjetivo. [Determinants of subjective well-being]. *Psico*, *43* (3), 280-288.

ZIA, S. (2016). *Impact of parental acceptance-rejection and core self-evaluations on the pride and mental health of university students*. Unpublished BS Thesis. Sargodha: University of Sargodha, Pakistan.

ZIMMERMAN, B.J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25* (1), 82-91. doi:10.1006/ceps.1999.1016

ZUNG, W.W. (1965). A self-rating depression scale. *Archives of General Psychiatry*, *12* (1), 63-70. doi:10.1001/archpsyc.1965.01720310065008