

Further Evaluation of the Psychometric Properties of the Self-Forgiveness Dual Process Scale: Evidence of Measurement Invariance Across Gender

Kendini Affetme İkili Süreç Ölçeğinin Psikometrik Özelliklerinin İleri Düzeyde Değerlendirilmesi: Cinsiyetler Arasında Ölçme Değişmezliğine İlişkin Kanıtlar

Fatih AYDIN¹



Sivas Cumhuriyet University Faculty of
Education Department of Guidance and
Psychological Counseling, Sivas, Türkiye



ABSTRACT

The present study aims to examine the psychometric properties of the Self-Forgiveness Dual Process Scale. Within the scope of the research, it was planned to test the measurement invariance of the Turkish version of the scale across genders. The study sample consists of 261 university students, comprising 168 (64.37%) females and 93 (35.63%) males. The mean age of the participants was calculated to be 21.39 (SD = 2.91). The Personal Information Form and the Turkish version of the Self-Forgiveness Dual Process Scale were used as data collection tools. The results of Confirmatory Factor Analysis (CFA) showed that the scale provided sufficient evidence for the construct validity ($\chi^2/df = 1.654$, CFI = .980, TLI = .973, GFI = .999, RMSEA = .050, SRMR = .067). In addition, multi-group CFA revealed significant findings regarding configural, metric, scalar, and strict invariance across genders. In addition, convergent validity findings for the whole scale were borderline for the Value Reorientation (VRO) Subscale (AVE = .483), while they were sufficient for the Esteem Restoration (ERS) Subscale (AVE = .642). Finally, the analyses regarding the internal consistency reliability (McDonald $\omega = .821$ and Cronbach's $\alpha = .817$ for VRO; McDonald $\omega = .888$ and Cronbach's $\alpha = .886$ for ERS) and split-half reliability ($r = .845$ for VRO; $r = .913$ for ERS) of the scale exhibited good results. The results demonstrated that the Turkish version of the Self-Forgiveness Dual Process Scale is a psychometrically sound measurement tool suitable for use with university students. Specifically, evidence was presented indicating that the scale demonstrated similar measurement performance in both males and females.

Keywords: Self-forgiveness, measurement invariance, gender, validity

Öz

Bu çalışmanın amacı Kendini Affetme İkili Süreç Ölçeğinin psikometrik özelliklerinin incelenmesidir. Araştırma kapsamında bilhassa ölçeğin Türkçe formunun cinsiyetler arasında ölçüm değişmezliğinin test edilmesi planlanmıştır. Araştırmanın örneklemini 168 (%64,37) kadın ve 93 (%35,63) erkek olmak üzere toplam 261 üniversite öğrencisi oluşturmaktadır. Katılımcıların yaş ortalaması 21,39 (SS = 2,91) olarak hesaplanmıştır. Araştırmada veri toplama aracı olarak Kişisel Bilgi Formu ve Kendini Affetme İkili Süreç Ölçeği Türkçe Formu kullanılmıştır. Doğrulamalı Faktör Analizi (DFA) ile elde edilen sonuçlar ölçeğin yapı geçerliğine ilişkin yeterli kanıtlar sunduğunu göstermiştir ($\chi^2/sd = 1,654$, CFI = ,980, TLI = ,973, GFI = ,999, RMSEA = ,050, SRMR = ,067). Bunun yanı sıra, çoklu grup DFA cinsiyetler arasında şekilsel, metrik, skalar ve katı değişmezliğe ilişkin anlamlı bulgular ortaya koymuştur. Ayrıca, ölçeğin tümü için yakınsak geçerlik bulguları Pozitif Değere Yeniden Yönelim Alt Boyutu (PDYY) için sınırdan iken (AVE = ,483), Kişisel Saygının Yenilenmesi Alt Boyutu (KSY) için yeterli düzeydedir (AVE = ,642). Son olarak, ölçeğin iç tutarlık güvenirliği (PDYY için McDonald $\omega = ,821$ ve Cronbach's $\alpha = ,817$; KSY için McDonald $\omega = ,888$ ve Cronbach's $\alpha = ,886$) ve iki yarı güvenirliği (PDYY için $r = ,845$; KSY için $r = ,913$) ile ilgili analizler iyi sonuçlar vermiştir. Sonuçlar Kendini Affetme İkili Süreç Ölçeği Türkçe Formunun üniversite öğrencileriyle kullanılmaya uygun, psikometrik olarak geçerli bir ölçüm aracı olduğunu göstermiştir. Özellikle, ölçeğin kadın ve erkeklerde benzer ölçüm performansı gösterdiğine ilişkin kanıtlar ortaya kovulmuştur.

Anahtar Kelimeler: Kendini affetme, ölçüm değişmezliği, cinsiyet, geçerlik

Geliş Tarihi/Received 24.06.2025
Kabul Tarihi/Accepted 12.08.2025
Yayın Tarihi/Publication 31.08.2025
Date

Sorumlu Yazar/Corresponding author:
Fatih AYDIN

E-mail: faydin@cumhuriyet.edu.tr

Cite this article: Aydın, F. (2025). Further evaluation of the psychometric properties of the self-forgiveness dual process scale: Evidence of measurement invariance across gender. *Journal of Psychometric Research*, 3(2), 98-106.



Content of this journal is licensed under a
Creative Commons Attribution-
NonCommercial 4.0 International License.

Introduction

Forgiveness is a multifaceted construct that plays a crucial role in individuals' spiritual and psychological well-being (Chen et al., 2019; Noviyanty et al., 2022). It is a pardon people ask from others for their inappropriate behaviors (Yucel & Vaish, 2021). Sometimes it means the mercy that believers ask from God for their offenses, also called *divine forgiveness* (Fincham, 2022). Forgiveness is often conceptualized as a trait-like mechanism, a reasonably stable tendency to respond to the misbehaviors of others toward oneself, regardless of whether the offenders ask for forgiveness or not (Strelan, 2017). Besides, unlike *dispositional forgiveness*, well-established research often adopts forgiveness as a state reaction. Therefore, forgiveness can be described as an act for the peaceful resolution of intrapersonal, or more deeply, internal conflicts. The second, *self-forgiveness*, is identified as a key factor unlocking better flourishing in life (Aydın, 2025).

Self-forgiveness refers to “a willingness to abandon self-resentment in the face of one’s own acknowledged objective wrong while fostering compassion, generosity, and love toward oneself” (Enright, 1996, p. 115). It is a journey to inward world, where individuals confront with their wrongdoings, admit them, take action to repair the damage they have caused, comprehend the circumstances prompt them to offend, accept that they can sometimes do bad things like everyone else, show mercy and love to themselves, gain awareness, learn from their mistakes, find meaning in these experiences, and sincerely commit to not to do such things in the future (internal change) (Webb et al., 2017). Therefore, this process results in an increased amount of knowledge, awareness, and maturation. Hence, people remedy their self-worth and become aware that, as individuals, they are more than the destructive behaviors they displayed earlier (Holmgren, 1998). There are valuable attributes that make any individual special and worthy of love. This type of mindset enables individuals to remove the scars of negative feelings and be free from self-directed hatred (Hall & Fincham, 2005).

Based on individual conceptualizations, researchers have developed numerous psychological assessment tools targeting self-forgiveness. These tools include the Forgiveness of Self and Others Scale (Mauger et al., 1992), State Self-Forgiveness Scale (Wohl et al., 2008), Differentiated Self-Forgiveness Scale (Woodyatt & Wenzel, 2013), Enright Self-Forgiveness Scale (Kim et al., 2022), Forgiveness of Self Scale (Regalia & Pelucchi, 2024), Heartland Forgiveness Scale – Self (Thompson et al., 2005),

Self-Forgiveness Dual-Process Scale (SF-DPS, Griffin et al., 2018), and more. Each scale has a unique item wording, factorial structure, and attributes that are focused on. There is also criticism on some of the scales (e.g., Tangney et al., 2005) that the specific items do not measure self-forgiveness. In Türkiye, several tools are available to assess self-forgiveness. The adapted versions of Heartland Forgiveness Scale – Self (Bugay & Demir, 2010) and State Self-Forgiveness Scale (Aydın & Yerin-Güneri, 2017) are available in Turkish. Additionally, Ersanlı and Vural Batık (2015) developed the Forgiveness Scale, which includes a subscale comprising three items to measure self-forgiveness. More recently, the SF-DPS has been adapted to Turkish by Kaya et al. (2023) in a sample of college students.

The SF-DPS is built on two distinct features of self-forgiveness: value reorientation and esteem restoration. Value reorientation entails acknowledging personal accountability and showing readiness to atone for involvement in events that may have caused moral harm (Griffin et al., 2024). The sample items include “I would take back what I’ve done if I could” and “My actions violated something important to me.” On the other hand, the process of esteem restoration centers on maintaining a sense of personal value and growth potential, regardless of one’s imperfections and failures (Griffin et al., 2018). The sample items include “I still love myself even though I did wrong” and “I feel like a valuable person despite my wrongdoing.” The Turkish version of the scale was reported to exhibit good results in terms of construct, convergent, and discriminant validities (Kaya et al., 2023). Moreover, there was also evidence of excellent reliability. However, the adaptation study lacks psychometric evaluations, including measurement invariance.

Measurement invariance refers to the degree to which a measurement instrument operates equivalently across different groups or conditions (Van de Schoot et al., 2012). Ensuring invariance is a fundamental prerequisite for making valid cross-group comparisons (Jeong & Lee, 2019). The concept is typically examined at multiple levels (Schmitt & Kuljanin, 2008): configural, metric, scalar, and strict. Configural invariance assesses whether the basic factor structure is consistent across groups. Metric invariance tests whether factor loadings are equivalent, suggesting that the items are related to the latent variable in the same way. Scalar invariance further requires equal item intercepts, allowing for a meaningful comparison of latent means. Finally, strict invariance assumes equality of residual variances, which supports direct comparisons of observed scores. Establishing these levels of invariance is

crucial in psychological, educational, and social science research to ensure the validity of inferences drawn from group comparisons.

The present study aims to extend the current findings on the Turkish version of the SF-DPS. The scale was selected explicitly since the SF-PDS focuses mainly on self-forgiveness processes, unlike more general scales that capture different types of forgiveness with multiple subscales. Therefore, to fill the gap in evidence regarding measurement invariance, the primary purpose of the present study is to test measurement invariance across gender. Additionally, the construct, convergent, and discriminant validities, as well as the internal consistency and split-half reliabilities, were investigated.

Methods

Participants

By employing the convenience sampling method, the present study recruited 261 college students, 168 females (64.37%) and 93 males (35.63%), who volunteered to participate in the study. The average age was 21.39 with a standard deviation of 2.91. In terms of college grade, the sample comprises 45 freshmen (17.24%), 133 sophomores (50.96%), 42 juniors (16.09%), and 41 seniors (15.71%). In addition, the perceived socio-economic status of the majority of the sample was middle-income ($n = 226$, 86.59%), followed by low-income ($n = 23$, 8.81%) and high-income ($n = 12$, 4.60%).

Measurement Tools

Demographic Information Form

This form was prepared by the author to obtain necessary demographic information. It comprises four questions regarding sex, age, college grade, and perceived socio-economic status.

Self-Forgiveness Dual-Process Scale (SF-DPS)

This scale was developed by Griffin et al. (2018), grounded in Social Cognitive Theory, and aims to capture self-forgiveness based on two distinct components, Value Reorientation (VRO) and Esteem Restoration (ERS). The scale comprises ten items with a seven-point Likert-type scoring system, five items for each subscale. The assessment is based on subscales; hence, there is no total score. Higher scores obtained from the subscales indicate higher self-forgiveness. The scale was adapted into Turkish by Kaya et al. (2023) and exhibited good construct,

convergent, discriminant, and concurrent validities. It was also reported in the Turkish adaptation that the scale performed perfectly in terms of reliability.

Procedures

Ethics committee approval was received for this study from the ethics committee of Sivas Cumhuriyet University (Date: June 19, 2025, Number: E-99711239-050.04-576210). After obtaining the ethics committee approval, the author merged the data collection tools into an online questionnaire and distributed it through social media accounts and instant messaging platforms to reach potential participants. An informed consent form was placed at the very beginning of the online questionnaire. Individuals who did not consent were not permitted to complete the survey. The present study has some inclusion criteria, which are also stated in the informed consent form. Therefore, participants under 18 years of age who were not currently enrolled in an undergraduate program were excluded from the study. Data collection was done in one week. The target sample size for the study was thought to be 200, considering the several cut-offs suggested by widely adopted guidelines (Alavi et al., 2020): minimum of 200 participants, a ratio number of participants to number of model variables to be ≥ 10 , and a ratio of number of participants to number of parameters to be ≥ 5 (Myers et al., 2011). After finishing the data collection, the dataset was prepared for the analysis.

Data Analysis

Before proceeding with the analysis, assumption testing was conducted primarily. The dataset was checked for missing entries first, and no missing data was identified. Next, the item scores were transformed into standardized z-values to investigate significant outliers. Again, no significant outliers have been detected. Lastly, the kurtosis and skewness of the item scores were examined to assess the normality assumption. The kurtosis values ranged from -0.532 to 1.648, and skewness values ranged from -1.277 to -0.025, indicating no departure from normal distribution according to widely adopted criteria of ± 2 (George & Mallery, 2019). Therefore, it was ensured that the dataset was suitable for analysis.

Subsequently, Confirmatory Factor Analysis (CFA) was conducted to test the construct validity of the Turkish version of the SF-DPS. The construct validity of the scale was already examined by Kaya et al. (2023); hence, the present study aims to strengthen the previous findings and provide additional evidence. CFA is a widely used statistical

method that enables researchers to examine the theoretical models – structure of a phenomenon – established upon latent and observed variables (Tabachnick & Fidell, 2014). The Diagonally Weighted Least Squares (DWLS) estimation was selected for the analysis since the data were ordinal (DiStefano & Morgan, 2014). This method is very robust to violations of the normality assumption and is often suggested for conducting CFA with non-normal data (Li, 2016). Additionally, the standard error method was employed because the data met the assumption of normality. At first, parameter estimates were checked. Specifically, factor loadings were examined in line with the widely used criteria. Therefore, factor values can be evaluated as poor ($< .32$), fair ($> .45$), good ($> .55$), very good ($> .63$), and excellent ($> .71$) (Tabachnick & Fidell, 2014). Next, the goodness of fit of the model was investigated. Multiple indices were taken into consideration, including chi-squared divided by degrees of freedom (χ^2/df), comparative fit index (CFI), Tucker-Lewis index (TLI), goodness of fit index (GFI), standardized root mean squared residual (SRMR), and root mean squared error of approximation (RMSEA). Widely recommended criteria were adopted for the indices, namely $\chi^2/df < 3$; CFI, TLI, and GFI $> .95$, and SRMR and RMSEA $< .08$ (Hu & Bentler, 1999; Kline, 2016). For RMSEA, values below .05 were considered indicative of an excellent fit, values ranging from .05 to .08 were deemed acceptable, and values exceeding .10 were viewed as evidence of poor fit (Brown, 2015).

The present study also investigated the convergent validity of the Turkish version of the SF-DPS with average variance extracted (AVE) and composite reliability (CR). It was recommended that the AVE values should be greater than .50, and CR values should be greater than .70 cut-off and AVE values (Fornell & Larcker, 1981). Furthermore, discriminant validity was examined by calculating the square root of AVE. The square root of AVE values for the subscales should be greater than .50, and the correlation between the subscales (Fornell & Larcker, 1981; Hair et al., 2010). Additionally, the scales' internal consistency reliability and split-half reliability were assessed. For internal consistency reliability, McDonald's Omega and Cronbach's Alpha were calculated. The results were evaluated with the criteria of .70 (Hair et al., 2016; Nunnally & Bernstein, 1994). The split-half reliability was examined using Pearson correlation. Again, results above .70 indicate a high relationship between the two halves of the dataset (Mukaka, 2012).

Lastly, for the primary purpose of the present study, the evidence of measurement invariance across gender was

investigated. A multigroup CFA was conducted using the same analysis properties as those employed in the first-order CFA. A four-stage procedure suggested by Meredith (1993) was adopted, where configural invariance is tested initially, followed by metric, scalar, and strict invariance. In each test, the model fit was compared to the earlier one. The difference between each model's CFI values and the previous one (ΔCFI) is recommended to be below .01 to conclude that the model exhibits a good fit (Byrne, 2010; Cheung & Rensvold, 2002). All the analyses described here were conducted using the JASP program version 0.19.2. It is a user-friendly software based on an open-source R package and generally doesn't require users to code.

Results

Structural Validity

The first-order CFA model was analyzed using the DWLS estimator and the standard error method, considering that the data are ordinal and meet the assumption of normality. The results revealed that all parameter estimates, including factor loadings, error terms, and covariance between first-order factors, were significant. After this confirmation, the standardized factor loadings and the goodness of fit indices regarding the first-order CFA model for the Turkish version of the SF-DPS were examined (Figure 1). It was found that the standardized factor loadings ranged from .592 to .889 ($p < .001$). Compared to the suggested criteria, it can be concluded that the model exhibits good results (see Table 1). Furthermore, the analysis demonstrated that the first-order CFA model yielded a good fit to the data without the need for any modifications ($\chi^2/df = 1.654$, CFI = .980, TLI = .973, GFI = .999, RMSEA = .050, SRMR = .067).

Figure 1
Parameter Estimates for the First-Order CFA Model

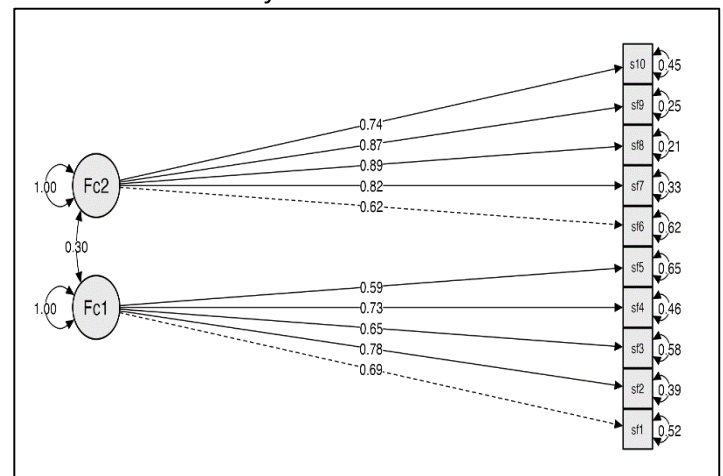


Table 1*Standardized Factor Loadings of the First-Order CFA*

Factor	Item	Estimate	Std. Estimate	Std. Error	z-value	p	95% Confidence Interval	
							Lower	Upper
Value Reorientation	SF1	1.000	.690	0.060	11.468	< .001	0.572	0.808
	SF2	1.095	.780	0.075	10.394	< .001	0.633	0.928
	SF3	0.970	.647	0.060	10.749	< .001	0.529	0.765
	SF4	1.153	.733	0.063	11.599	< .001	0.609	0.857
	SF5	0.925	.592	0.054	10.996	< .001	0.486	0.697
Esteem Restoration	SF6	1.000	.620	0.054	11.528	< .001	0.514	0.725
	SF7	1.205	.820	0.065	12.641	< .001	0.693	0.948
	SF8	1.330	.889	0.071	12.564	< .001	0.750	1.027
	SF9	1.239	.865	0.073	11.904	< .001	0.723	1.007
	SF10	1.124	.738	0.062	11.944	< .001	0.617	0.860

Convergent Validity

The convergent evidence was investigated by calculating AVE and CR values. The results demonstrated that AVE values were .483 for the Value Reorientation Subscale and .642 for the Esteem Restoration Subscale. Unfortunately, the Value Reorientation Subscale could not meet the criteria of .50; however, it was very close. On the other hand, the Esteem Restoration Subscale exhibited good results. Additionally, the analysis revealed that both subscales performed as desired in terms of composite reliability. The CR values were .821 and .897 for Value Reorientation and Esteem Restoration Subscales, respectively. Therefore, the scale yielded promising results in terms of convergent validity.

Discriminant Validity

The correlations between the subscales and the square root of AVE values were used to investigate discriminant validity. The calculations revealed that the square root of AVE values were .695 for the Value Reorientation Subscale and .801 for the Esteem Restoration Subscale. Since these values were higher than the correlation between the subscales ($r = .300$, $p < .001$), it was confirmed that the scale exhibited favorable results in terms of discriminant validity.

Reliability

The McDonald's Omega and Cronbach's Alpha coefficients were calculated for internal consistency reliability. The findings demonstrated that the subscales yielded satisfactory evidence for internal consistency. The McDonald's Omega and Cronbach's Alpha were .821 and .817 for Value Reorientation Subscale and .888 and .886 for Esteem Restoration Subscale, respectively. The split-half reliability was analyzed with Pearson correlation. Both

subscales performed very well, with values of .845 and .913 for the Value Reorientation and Esteem Restoration Subscales, respectively.

Measurement Invariance Across Gender

A multigroup CFA was employed to investigate measurement invariance across gender (Table 2). The DWLS estimator was selected for the analysis since the data were ordinal. In addition, the standard error method was used as the data met the assumption of normal distribution. First, configural invariance was examined. In this model, the factor loadings and intercepts were estimated freely. The analysis yielded good fit to the data ($\chi^2/df = 1.052$, CFI = .997, SRMR = .071, RMSEA = .020). These findings demonstrate that the factor structure of the Turkish version of the SF-DPS is similar for males and females. Later, metric invariance was tested. The factor loadings were constrained, and intercepts were estimated freely. To evaluate the model fit this time, the criterion of $\Delta CFI < .01$ was employed. Since the ΔCFI for the metric invariance was below .01, it was confirmed that the metric invariance was achieved. These results demonstrated that the factor loadings are similar for males and females. Next, the scalar invariance was examined by constraining both factor loadings and intercepts. Again, the results showed that ΔCFI is below .01, indicating good model fit. This means that both factor loadings and intercepts are equivalent across males and females. Lastly, strict invariance was investigated by constraining error variances in addition to factor loadings and intercepts. This final evaluation also revealed that the ΔCFI was below the .01 threshold, and yet, the strict invariance was achieved. This indicates that both factor loadings, intercepts, and error variances were similar across gender groups.

Table 2*Results for Measurement Invariance Across Gender*

Model	$\chi^2(df)$	χ^2/df	CFI	SRMR	RMSEA	Model comparison		
							$\Delta\chi^2/df$	ΔCFI
1. Configural invariance	71.514(68)	1.052	.997	.071	.020			
2. Metric invariance	90.106(76)	1.186	.988	.082	.038	2 vs 1	18.592(8)	.009
3. Scalar invariance	97.635(84)	1.162	.988	.085	.035	3 vs 2	7.529(8)	.000
4. Strict invariance	103.990(94)	1.106	.991	.089	.029	4 vs 3	6.355(10)	.003

Discussion

The present study extends the findings regarding the psychometric properties of the SF-DPS. First, it was confirmed that the construct validity of the scale is good. Directly comparing the results of goodness-of-fit indices in the original, Turkish adaptation, and the current study may not be a practical approach because all three studies used different estimators. For instance, the original study conducted by Griffin et al. (2018) employed the Maximum Likelihood Robust estimator. In contrast, the Turkish adaptation study preferred the Maximum Likelihood (Standard) method (Kaya et al., 2023). In the current study, the DWLS estimator was selected since the data are ordinal (DiStefano & Morgan, 2014). Regardless of analysis properties, all three studies yielded favorable outcomes. The standardized factor loadings ranged from .592 to .889, which are greater than the criteria of .55 (Tabachnick & Fidell, 2014). Therefore, the factor loadings can be described as a spectrum ranging from good to excellent. For instance, in the current results, χ^2/df was 1.654, significantly below the cut-off value of 3 (Hu & Bentler, 1999; Kline, 2016). The CFI (.980), TLI (.973), and GFI (.999) indices exceed the .95 criterion (Hu & Bentler, 1999; Kline, 2016) and approach a value of 1.00. The RMSEA indices of .050 seemed to be borderline compared to the perfect fit criteria, but acceptable according to the more flexible criteria of .080 (Brown, 2015). Similarly, the SRMR indices were below the criteria of .080. In the original study (Griffin et al., 2018), the RMSEA and SRMR indices appear to be more favorable than the current results, .015 and .055, respectively. However, by applying two modifications, the Turkish adaptation study yielded similar results to the present study, with RMSEA = .045 and SRMR = .063.

Second, the convergent validity was examined with AVE and CR values. The original study does not provide such information. However, the Turkish adaptation study and the current study demonstrated similar results. For AVE values, both studies revealed that the Value Reorientation subscale yielded unfavorable results, performing below the

criterion of .50 (Fornell & Larcker, 1981). In the adaptation study, the AVE for the VRO and ERS subscales were .41 and .61, respectively.

In comparison, the AVE for the VRO and ERS subscales was calculated to be .483 and .642, respectively, in the present study. Therefore, the current results seem to be slightly better. However, the low AVE value in both studies may indicate a problem with the wording of the items. Griffin et al. (2024) modified the existing scale by changing some wording in the items. This might solve the issue, but they did not provide any information on the AVE values of the subscales. Despite the unfavorable result for the VRO subscale, CR values seem to meet the necessary criteria. The Turkish adaptation study yielded promising results, with CR values of .78 and .88 for the VRO and ERS subscales, respectively. The current findings were slightly better, with CRs of .821 for VRO and .897 for ERS subscales.

Third, the analysis for discriminant validity exhibited good results. Both the adaptation study and the present study yielded similar results. The results exceeded both the .50 cut-off and the correlation between the subscales (Fornell & Larcker, 1981; Hair et al., 2010). In the adaptation study, the correlation between the subscales was $r = .34$. The square root of AVE for the VRO and ERS subscales was .68 and .78, respectively. In parallel, the present study revealed that the correlation between the subscales was $r = .300$, and the square root of AVE of VRO and ERS subscales was .695 and .801, respectively. Again, slightly better results have been achieved. This finding supports the notion that the two subscales are distinct entities and not similar.

Furthermore, the present findings support that the internal consistency and split-half reliabilities of the scale were excellent. Both adaptation and the present study performed similarly, exhibiting values greater than the criterion of .70 on both internal consistency and split-half reliability (Hair et al., 2016; Mukaka, 2012; Nunnally & Bernstein, 1994). The Cronbach's Alpha was calculated as .76 for VRO and .87 for ERS subscales in the adaptation study. The present findings echoed these results and revealed that the McDonald's Omega and Cronbach's Alpha

for both subscales were greater than .80, which is slightly better than the results of previous research. Additionally, the split-half reliability was somewhat better in the present study, with $r = .845$ for VRO and $r = .913$ for the ERS subscales. At the same time, the adaptation study yielded results of .71 and .83 for the VRO and ERS subscales, respectively.

Lastly, the present study contributes to the literature by demonstrating evidence of measurement invariance across gender in the Turkish version of the SF-DPS, using a sample of college students. The original and Turkish adaptation studies did not investigate this property of the scale. Therefore, the present findings are vital. Measurement invariance was examined at four levels: configural, metric, scalar, and strict (Schmitt & Kuljanin, 2008). The configural model exhibited good results in terms of goodness of fit. The significance of the model fit of the latter models was evaluated based on ΔCFI . Since the ΔCFI was smaller than the criteria of .01 (Byrne, 2010; Cheung & Rensvold, 2002). In all the latter models, it was concluded that the metric, scalar, and strict models exhibited a good fit to the data. This means that the factorial structure, items' loadings to the factors, latent means, and residual variances are similar for male and female college students. The scale performs similarly in both male and female groups. Therefore, future research can utilize the Turkish version of the SF-DPS as a valid and reliable tool and make cross-group comparisons based on gender.

Limitations and Future Directions

The present study revealed valuable information for researchers in the field. However, several limitations should be considered when interpreting the current results. First, the SF-DPS has been recently modified by Griffin et al. (2024). Therefore, the current findings do not cover the modified items and their psychometric properties in the Turkish sample. Griffin et al. (2024) reported that the modified items performed better than the older ones. Hence, it is recommended to adapt the latest version of the scale and compare the performance of both scales in the Turkish context. Second, the AVE of the Value Reorientation subscale did not meet expectations, despite the good CR values. This might stem from the wording of the items; therefore, the modified version might not have this issue. Third, the data was collected online, which can be problematic in terms of under-representation of the target population and ethical restrictions, including confidentiality, if the security protocols were not set at the beginning of the study. Hence, to eliminate such risks, a

highly trusted platform was selected to create the online questionnaire. No personal information was collected to ensure the confidentiality of participants in the event of potential security breaches. The dataset was deleted from the platform after being downloaded to the computer and stored on an encrypted device, and the online questionnaire was distributed across different faculties to ensure maximum diversity. Lastly, the present study provides evidence for measurement invariance across gender but does not promise further information based on other grouping variables to conduct cross-group comparisons. For instance, measurement invariance can be investigated based on factors such as religiosity, age, educational status, perceived socio-economic status, employment status, and others. It may be a subject of future studies to examine the measurement invariance of the Turkish version of the SF-DPS, based on various grouping variables.

Conclusion and Recommendations

The Turkish version of the SF-DPS demonstrated good construct, convergent, and discriminant validity in a sample of college students. The analysis also revealed that the scale demonstrated good internal consistency and split-half reliability. The results suggest that the Turkish SF-DPS exhibits measurement invariance across genders at the configural, metric, scalar, and strict levels of measurement invariance.

Etik Komite Onayı: Bu çalışma için etik komite onayı Sivas Cumhuriyet Üniversitesi'nden (Tarih: 19 Haziran 2025, Sayı: E-99711239-050.04-576210) alınmıştır.

Katılımcı Onamı: Yazılı onam bu çalışmaya katılan tüm katılımcılardan alınmıştır.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar, çıkar çatışması olmadığını beyan etmiştir.

Finansal Destek: Yazar, bu çalışma için finansal destek almadığını beyan etmiştir.

Ethics Committee Approval: Ethics committee approval was received for this study from the ethics committee of Sivas Cumhuriyet University (Date: June 19, 2025, Number: E-99711239-050.04-576210).

Informed Consent: Written informed consent was obtained from patients who participated in this study.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflicts of interest to declare.

Financial Disclosure: The author declared that this study has received no financial support.

References

- Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Chi-square for model fit in confirmatory factor analysis. *Journal of Advanced Nursing*, 76(9), 2209-2211. <https://doi.org/10.1111/jan.14399>
- Aydın, F. (2025). Does attachment to God matter? Role of spiritual attachment in mental health through self-forgiveness: lessons from Turkish college sample. *Frontiers in Psychology*, 16, Article 1603654. <https://doi.org/10.3389/fpsyg.2025.1603654>
- Aydın, G., & Yerin-Güneri, O. (2017, May). *Testing factor structure and reliability of the Turkish version of State Self Forgiveness Scale (SSFS)* [Conference presentation]. EJER 4th International Eurasian Educational Research Congress. Denizli, Turkey.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Bugay, A., & Demir, A. (2010). A Turkish version of heartland forgiveness scale. *Procedia Social and Behavioral Sciences*, 5, 1927-1931. <https://doi.org/10.1016/j.sbspro.2010.07.390>
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.
- Chen, Y., Harris, S. K., Worthington Jr, E. L., & VanderWeele, T. J. (2019). Religiously or spiritually-motivated forgiveness and subsequent health and well-being among young adults: An outcome-wide analysis. *The Journal of Positive Psychology*, 14(5), 649-658. <https://doi.org/10.1080/17439760.2018.1519591>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 425-438. <https://doi.org/10.1080/10705511.2014.915373>
- Enright, R. D. (1996). Counseling within the forgiveness triad: On forgiving, receiving forgiveness, and self-forgiveness. *Counseling and Values*, 40(2), 107-126. <https://doi.org/10.1002/j.2161-007X.1996.tb00844.x>
- Ersanlı, K., & Vural Batık, M. (2015). Development of the forgiveness scale: A study of reliability and validity. *Turkish Studies-International Periodical for the Languages, Literature and History of Turkish or Turkic*, 10(7), 19-32. <http://doi.org/10.7827/TurkishStudies.8201>
- Fincham, F. D. (2022). Towards a psychology of divine forgiveness. *Psychology of Religion and Spirituality*, 14(4), 451-461. <https://doi.org/10.1037/rel0000323>
- Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), 382-388. <https://doi.org/10.1177/002224378101800313>
- George, D., & Mallery, P. (2019). *IBM SPSS statistics 25 step by step: A simple guide and reference* (15th ed.). Routledge.
- Griffin, B. J., Norman, S. B., Weber, M. C., Hinkson Jr, K. D., Jendro, A. M., Pyne, J. M., Worthington Jr., E. L., & Maguen, S. (2024). Properties of the modified self-forgiveness dual-process scale in populations at risk for moral injury. *Stress and Health*, 40(5), Article e3413. <https://doi.org/10.1002/smi.3413>
- Griffin, B. J., Worthington Jr, E. L., Davis, D. E., Hook, J. N., & Maguen, S. (2018). Development of the self-forgiveness dual-process scale. *Journal of Counseling Psychology*, 65(6), 715-726. <https://doi.org/10.1037/t70152-000>
- Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- Hair, J. F., Jr., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Sage Publications.
- Hall, J. H., & Fincham, F. D. (2005). Self-forgiveness: The stepchild of forgiveness research. *Journal of Social and Clinical Psychology*, 24(5), 621-637. <https://doi.org/10.1521/jscp.2005.24.5.621>
- Holmgren, M. R. (1998). Self-forgiveness and responsible moral agency. *Journal of Value Inquiry*, 32, 75-91. <https://doi.org/10.1023/A:1004260824156>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jeong, S., & Lee, Y. (2019). Consequences of not conducting measurement invariance tests in cross-cultural studies: A review of current research practices and recommendations. *Advances in Developing Human Resources*, 21(4), 466-483. <https://doi.org/10.1177/1523422319870726>

- Kaya, F., Odacı, H., & Aydın, F. (2023). Kendini Affetme İkili Süreç Ölçeği'nin geçerlik ve güvenirliğinin test edilmesi [Testing the validity and reliability of the self-forgiveness dual process scale]. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 43(1), 1-29. <https://doi.org/10.17152/gefad.1110236>
- Kim, J. J., Volk, F., & Enright, R. D. (2022). Validating the Enright self-forgiveness inventory (ESFI). *Current Psychology*, 41(11), 7604-7617. <https://doi.org/10.1007/s12144-020-01248-4>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Li, C. H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods*, 21(3), 369-387. <https://doi.org/10.1037/met0000093>
- Mauger, P. A., Perry, J. E., Freeman, T., & Grove, D. C. (1992). The measurement of forgiveness: Preliminary research. *Journal of Psychology and Christianity*, 11(2), 170-180. <https://psycnet.apa.org/record/1992-41508-001>
- Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58(4), 525-543. <https://doi.org/10.1007/BF02294825>
- Mukaka M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal: The Journal of Medical Association of Malawi*, 24(3), 69-71. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3576830/>
- Myers, N. D., Ahn, S., & Jin, Y. (2011). Sample size and power estimates for a confirmatory factor analytic model in exercise and sport: A Monte Carlo approach. *Research Quarterly for Exercise and Sport*, 82(3), 412-423. <https://doi.org/10.1080/02701367.2011.10599773>
- Noviyanty, H., Ismail, Z., Hamjah, S. B. H., & Mohamad, A. D. (2022, February). *Spiritual psychotherapy and mental health: The forgiveness therapy in achieving spiritual well-being of drug addicts with depression disorders* [Conference presentation abstract]. Second International Conference on Public Policy, Social Computing and Development (ICOPOSDEV 2021) (pp. 7-14). Atlantis Press.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Regalia, C., & Pelucchi, S. (2024). Forgiveness of self scale. In A. C. Michalos (Ed.), *Encyclopedia of quality of life and well-being research* (pp. 2576-2578). Springer International Publishing.
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18(4), 210-222. <http://doi.org/10.1016/j.hrmr.2008.03.003>
- Strelan, P. (2017). The measurement of dispositional self-forgiveness. In L. Woodyatt, E. Worthington, Jr., M. Wenzel, B. Griffin (Eds.), *Handbook of the psychology of self-forgiveness* (pp. 75-87). Springer. https://doi.org/10.1007/978-3-319-60573-9_6
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Pearson Education Limited.
- Tangney, J. P., Boone, A. L., & Dearing, R. (2005). Forgiving the self: Conceptual issues and empirical findings. In E. L. Worthington, Jr. (Ed.), *Handbook of forgiveness* (pp. 143-158). Routledge.
- Thompson, L. Y., Snyder, C. R., Hoffman, L., Michael, S. T., Rasmussen, H. N., Billings, L. S., Heinze, L., Neufeld, J. E., Shorey, H. S., Roberts, J. C., & Roberts, D. E. (2005). Dispositional forgiveness of self, others, and situations. *Journal of Personality*, 73(2), 313-360. <https://doi.org/10.1111/j.1467-6494.2005.00311.x>
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492. <https://doi.org/10.1080/17405629.2012.686740>
- Webb, J. R., Bumgarner, D. J., Conway-Williams, E., Dangel, T., & Hall, B. B. (2017). A consensus definition of self-forgiveness: Implications for assessment and treatment. *Spirituality in Clinical Practice*, 4(3), 216-227. <https://doi.org/10.1037/scp0000138>
- Wohl, M. J., DeShea, L., & Wahkinney, R. L. (2008). Looking within: Measuring state self-forgiveness and its relationship to psychological well-being. *Canadian Journal of Behavioural Science*, 40(1), 1-10. <https://doi.org/10.1037/0008-400x.40.1.1.1>
- Woodyatt, L., & Wenzel, M. (2013). Self-forgiveness and restoration of an offender following an interpersonal transgression. *Journal of Social and Clinical Psychology*, 32, 221-254. <https://doi.org/10.1521/jscp.2013.32.2.225>
- Yucel, M., & Vaish, A. (2021). Eliciting forgiveness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 12(6), Article e1572. <https://doi.org/10.1002/wcs.1572>