

Measurement invariance of the satisfaction with life scale across gender and time

Çiğdem Akin Arıkan 

Ordu University, Faculty of Education, Department of Educational Measurement and Evaluation, Ordu, Turkey, akincgdm@gmail.com

Selen Demirtaş Zorbaz 

Ankara University, Faculty of Educational Sciences, Department of Psychological Counseling and Guidance, Ankara, Turkey, selenpdr@gmail.com



ABSTRACT The aim of this research is to determine whether the Turkish version of Satisfaction with Life Scale (SWLS) has measurement invariance according to gender and longitudinal invariance for Turkish university students. First study's data were collected from 500 university students (366 female, 134 male) whereas 388 students (296 female, 92 male) participated in Study 2 for three times over four-month intervals. Confirmatory factor analysis was used for measurement invariance. According to the results of the Study 1 which was invariance study according to gender, all items on the SWLS provide configural, metric and scalar invariance. In this context, the differences in points obtained on the SWLS by university students according to gender can be used to perform comparison studies. Also, Study 2 revealed that strict longitudinal invariance was found to hold. This means that equality of factor patterns, factor loadings, intercepts and residual variance were determined for university students at four-month time intervals. The SWLS can be used to make longitudinal mean and gender comparisons for Turkish university students.

Keywords: Gender, Longitudinal invariance, Measurement invariance, Satisfaction with life scale

Yaşam doyumu ölçeğinin cinsiyete ve zamana göre ölçme değişmezliğinin incelenmesi

ÖZ Bu araştırmanın amacı, Yaşam Doyumu Ölçeği'nin Türkçe versiyonunun, üniversite öğrenci örnekleminde cinsiyete ve zamana göre ölçme değişmezliğine sahip olup olmadığını belirlemektir. Araştırmanın ilk çalışma grubunu 500 üniversite öğrencisi (366 kadın, 134 erkek), ikinci çalışma grubunu ise dört ay arayla üç kez uygulama yapılan 388 öğrenci (296 kadın, 92 erkek) oluşturmaktadır. Cinsiyete göre ve boylamsal ölçme değişmezliği için doğrulayıcı faktör analizi kullanılmıştır. Cinsiyete göre ölçme değişmezliği sonuçlarına göre, yaşam doyumu ölçeğindeki tüm maddelerin biçimsel, metrik ve ölçek ölçme değişmezliğine sahip olduğu bulunmuştur. Bir başka deyişle kız ve erkek grupların maddelere aynı şekilde yanıt verdiği ve bu gruplardan elde edilen puanların karşılaştırılabilir olduğu bulunmuştur. Zamana göre ise tam ölçme değişmezliğine sahip olduğu bulunmuştur. Bu da üniversite öğrencileri için dört aylık zaman aralıklarında faktör örüntülerinin eşitliği, faktör yükleri, kesişimler ve artık varyansın aynı şekilde belirlendiği anlamına gelir. Yaşam Doyumu Ölçeği Türk üniversite öğrencileri için boylamsal ortalama ve cinsiyet karşılaştırmaları yapmak için kullanılabilir.

Anahtar Sözcükler: Boylamsal değişmezlik, Cinsiyet, Ölçme değişmezliği, Yaşam doyumu ölçeği

Citation: Akin Arıkan, Ç., & Demirtaş Zorbaz, S., (2020). Measurement invariance of the Satisfaction with Life Scale across gender and time. *Turkish Journal of Education*, 9(4), 260-272. DOI: 10.19128/turje.774452

INTRODUCTION

When people are asked about their expectations from life, most respond with something akin to a desire to be happy. Especially in the field of mental health, how to be happy and what makes people happy is a frequently researched topic. More recently, however, the concept of happiness has begun to be replaced with that of subjective well-being. Diener, Suh, Lucas, and Smith (1999) defined subjective well-being as a comprehensive concept, including the individual's emotional responses, contentment, and general judgments related to life satisfaction. Seen as the cognitive dimension of well-being (e.g., Diener, Emmons, Larsen, & Griffin, 1985), life satisfaction can be defined as the individual's general cognitive judgments about their lives (Suh, Diener, Oishi, & Triandis, 1998). As such, which characteristics of an individual influence life satisfaction is a topic frequently researched in the literature. According to Cummins and Nistico (2002), there are three types of beliefs related to life satisfaction: self-value, perceived control, and optimism. These three perceptions are important indicators of whether one feels satisfaction with life or not.

Additionally, having a higher degree of life satisfaction has been shown to be an important protective factor. For example, a study of adolescents divided participants into three groups as having low, moderate, and high life satisfaction and compared the groups. According to the study results, life satisfaction appeared to provide a variety of mental health benefits (Gilman & Huebner, 2006). As a result, it appears that life satisfaction forms beliefs about the individual himself or herself developed after a variety of experiences and that these beliefs affect an individual's life in a multitude of ways.

To accurately measure a concept, such as life satisfaction, that interacts with so many different variables, has great value in terms of the importance of the conclusions reached. While there are different scales measuring life satisfaction, one of the most commonly used scales is the "Satisfaction with Life Scale" (SWLS) developed by Diener et al. (1985). The SWLS has been used to measure life satisfaction in a wide range of cultural and international contexts. There is a strong evidence of the high degree of validity and reliability of the SWLS in many different countries and different languages (Anaby, Jarus, & Zumbo, 2010; Arrindell, Heesink, & Feij, 1999; Bayani, Koocheky, & Goodarzi, 2007; Gouveia, Milfont, Da Fonseca, & de Miranda Coelho, 2009; Sachs, 2003; Yetim, 1993).

Several studies (Emerson, Guhn, & Gadermann, 2017; Jang et al., 2017) have revealed that the structure of life satisfaction scales may change according to culture. This is not surprising since the concept of life satisfaction appears to be affected by the culture that the individual lives in because it requires a comparison of the individual's own perceptions of their lives with that of others within their own group. Notably, in the literature, it appears that life satisfaction may be perceived differently in collectivist and individualistic cultures (e.g., Suh et al., 1998). Moreover, factor analyses of the SWLS in Turkey, a more collectivist culture (e.g., Dağlı & Baysal, 2016; Durak, Senol-Durak, & Gencoz, 2010), appear to prove the validity of the scale, though scale invariance was not examined. In this sense, the present study aimed to reveal invariance of the SWLS measurement in Turkish culture.

Additionally, as life satisfaction is a concept affected by time, this is a topic that continues to experience debate in the literature. The study by Fujita and Diener (2005) revealed that the life satisfaction of 24% of the participants significantly changes during 17 years of the study period. As the time interval between measurements increases, stability appears to reduce. The order of items on life satisfaction scales may be affected by factors such as one's current mood; however, most times, these factors can be controlled (Diener, Inglehart, & Tay, 2013). Based on all these points, when adapting scales used for the measurement of life satisfaction, researchers should not limit tests to only factor analyses. However, at the same time, it is necessary to perform measurement invariance (MI) analyses.

Measurement Invariance

Invariance is an important psychometric property in addition to the psychometric properties required of a scale tool, like validity and reliability (Brown, 2006; Meredith, 1993). In research, group comparisons using total points or subscale points obtained from scales are frequently encountered. Group comparisons performed without examining measurement invariance (MI) assume the measured structure provided by MI. However, every structure measured may not fit by MI. This brings into question the validity of the results obtained and the interpretations based on these results (Vandenberg & Lance, 2000). MI represents whether a scale measures the same structure without regarding to the group or measurement time (Mellenbergh, 1989). If the MI of a scale is not known, we cannot determine whether the difference observed in points between two groups or two-time measurements is a real difference or merely due to differences between the groups or assessment times in the construct structure (Brown, 2006).

Meredith (1993) and Vandenberg and Lance (2000) note that MI has four stages: configural, metric, scalar, and strict invariance. The configural model is the model with free factor loads, item intercepts, and error variance; the metric model has fixed factor loads, free item intercepts, and error variance; the scalar model has fixed factor loads and item intercepts, while error variance is free. The strict model has fixed factor loads, item intercepts, and error variance. In each stage of measurement invariance, the results obtained in a previous stage are compared, and decisions are made about whether invariance is present or not. If scale invariance is provided in these stages, only the comparison of differences between the means of groups at that time is significant (Millsap & Olivera-Aguilar, 2012). For invariance studies, confirmatory factor analysis (CFA) is one of the commonly used methods. CFA, in addition to determining the invariance of a construct in groups or subgroups, is related to another invariance type of the equivalence in time of the measured construct. Longitudinal MI is an important element in temporal variation research of a construct (Brown, 2006; Geiser, 2012). In longitudinal studies, researchers hope that indicators related to a latent construct do not differ over time. However, if these indicators do differentiate, it means that the variables belonging to a construct obtained at different times do not represent the same construct, or different measurement units of the constructs may be measured with different scales. This difference makes it difficult to compare the construct over time (Geiser, 2012).

SWLS appears to have a variety of MI studies performed in different countries and cultures. For example, Shevlin, Brunson, and Miles (1998) investigated the factor invariance of the scale and the factor loadings, unique variances, and factor variance were found to be invariant according to gender. Another more recent MI study (Jang et al., 2017) found that life satisfaction measured in 26 different countries only provided configural and metric invariance. At the same time, it was revealed that the intercepts of items 2, 4, and 5 were invariable. In addition, a study comparing Argentina, Mexico, and Nicaragua (Dimitrova & del Carmen, 2015) found the SWLS scale had configural, metric, and scalar invariance in different cultures.

In the literature, it appears that the MI of the SWLS in different samples and different cultures according to gender was revealed. For example, based on studies in different countries like Angola (Tomás, Gutiérrez, Sancho, & Romero, 2015), Norway (Clench-Aas, Nes, Dalgard, & Aarø, 2011), Malaysia (Swami, & Chamorro-Premuzic, 2009), Serbia (Jovanović, 2017), Germany (Hinz et al., 2018), and Spain (Checa, Perales, & Espejo, 2019), and in diverse sample groups, like children (Guhn, Ark, Emerson, Schonert-Reichl, & Gadermann, 2018), MI of the SWLS was present. Additionally, some studies (Atienza, Balaguer, & Garcia-Merita, 2003; Sovet, Atitsogbe, Pari, Park, & Villieux, 2016) have found that not all items on the SWLS are invariant. For example, in a study of Norwegian adolescents, Moksnes, Løhre, Byrne, & Haugan (2014) found configural and metric invariance were present; however, scalar and strict invariance did not hold. In a longitudinal MI study by Wu, Chen, and Tsai (2009) with a university student sample, the SWLS scale was found to have partial strict invariance. Also, Kyeunghae Lee, Brekke, Yamada, & Chou (2010) revealed that all five types of longitudinal invariance were supported for individuals with schizophrenia. The invariance study of the Serbian

version of the SWLS (Jovanović, 2017) found that the results of the multi-group CFA supported full scalar invariance across gender and over time. A review study of the MI of the SWLS observed more than half of the studies found that scalar and strict invariance were present for gender (Emerson et al., 2017). Thus, SWLS seems to show sensitivity to gender and time, indicating that gender and time might affect the remarking of SWLS items.

Moving specifically to research in the Turkish context for different sample groups (e.g., Dağlı & Baysal, 2016; Durak, et al., 2010; Köker, 1991; Yetim, 1993), the researchers do not appear to have performed any MI studies (age, gender, longitudinal, etc.) for the life satisfaction scale in the Turkish sample. Additionally, studies using gender comparisons of the SWLS (e.g., Çivitci, 2012; Gündoğar, Gül, Uskun, Demirci, & Keçeci, 2007; Özgür, Gümüş, & Durdu, 2010; Tuzgöl-Dost, 2007) appear to have assumed gender invariance without testing for it. However, the life satisfaction scale should have MI for comparisons between groups and according to time. Therefore, the present study attempts to contribute to the literature in three key ways. First, the factor structure of the Turkish version of SWLS will be replicated. Second, we will conduct a multi-group factor analysis model to test measurement invariance for gender. Last but not least, we will examine changes in the SWLS over time. Therefore, we will reveal the construct of the SWLS differ or not across gender and time in Turkish culture.

METHODOLOGY

Study Group 1

The first study group in the research was comprised of students attending the faculty of education in a state university for the 2018-2019 educational year. The ethical committee approval was obtained from the Social and Humanities Sciences Research Ethics Committee at Ordu University. The data was driven from first time point of the longitudinal data. Data for the first study group were collected from a total of 542 students and 500 students who indicate their gender on the form were included in the analyses. Therefore, study group was comprised 366 female (73.2%) and 134 male (26.8%) at the start of the fall semester in the 2018-2019 educational year. The ages of the students varied from 17 to 22 years, and the mean was 20.58 years.

Study Group 2

The second study group in the research comprised students attending the same university and faculty as the first study group for the 2018-2019 educational year. This study group comprised a total of 542 students: 376 female (69.37%) and 166 male (30.36%). Data were collected three times at four-month intervals from six different departments. To match data to students, they were asked to write their student numbers at the start of the measurement sets. Due to missing data between times, analyses were completed with a total of 388 students with 296 females (76.5%) and 91 males (23.5%).

Data Collection Tools

The Satisfaction with Life Scale was developed by Diener et al., (1985), and the Turkish validity and reliability studies for the scale were completed by Köker (1991) and Yetim (1993). The scale comprised five items and one factor related to life satisfaction. The study by Köker (1991) had test-retest validity of $r = .85$, item-test correlations of .71 to .80, while in the study by Yetim (1993) the Cronbach alpha internal consistency coefficient for the scale was .86 with a reliability coefficient of .73 determined with the test-retest method. In this study, the Cronbach alpha for longitudinal data was .795, .744, and .811, while it was .779 for gender.

Data Analysis

Before beginning MI analysis, a model data fit was performed for CFA. CFA analysis used the data set for the 500 individuals in the first study group and investigated assumptions. It was found that skewness and kurtosis were between a -1/+1 interval, and the assumption of normal distribution was obtained. Additionally, Mardia's test was performed for multivariate normality, and it was identified that the data obtained in all three applications did not meet by the multivariate normality assumption ($p = .000$).

For longitudinal invariance data, the 388 individuals present in all three times of the data set were included in the study group. It was found that skewness was between a -1/+1 interval, and the assumption of normal distribution was obtained. Additionally, Mardia's test did not provide by the multivariate normality assumption ($p = 0.00$). All the models were computed by MPlus software (Version 7.4) using a maximum likelihood estimation with robust (MLR) (Muthén & Muthén, 1998-2012). Comparisons related to invariance models used the chi-square difference test based on log likelihood values ($-2\Delta LL$) and scaling correction factors obtained with the MLR estimator. A non-significant chi square (χ^2) difference test means that invariance holds. Since the χ^2 value obtained in the model data fit was sensitive to the sample size, therefore, as the size of the sample increased ($n > 200$), the test was significant (Tabachnick & Fidell, 2007). Therefore, instead of using just χ^2 ; it was recommended to use χ^2/SD (Tabachnick & Fidell, 2007; Kline, 2005). In addition to the χ^2/df value, the fit statistics used in the study were RMSEA, SRMR, TLI, and CFI, which were suggested by Kline (2005).

RESULTS

Study 1

Firstly, the fit statistics for data in the obtained model were investigated. With this aim, the CFA model created with data collected from the first study group is given in Figure 1.

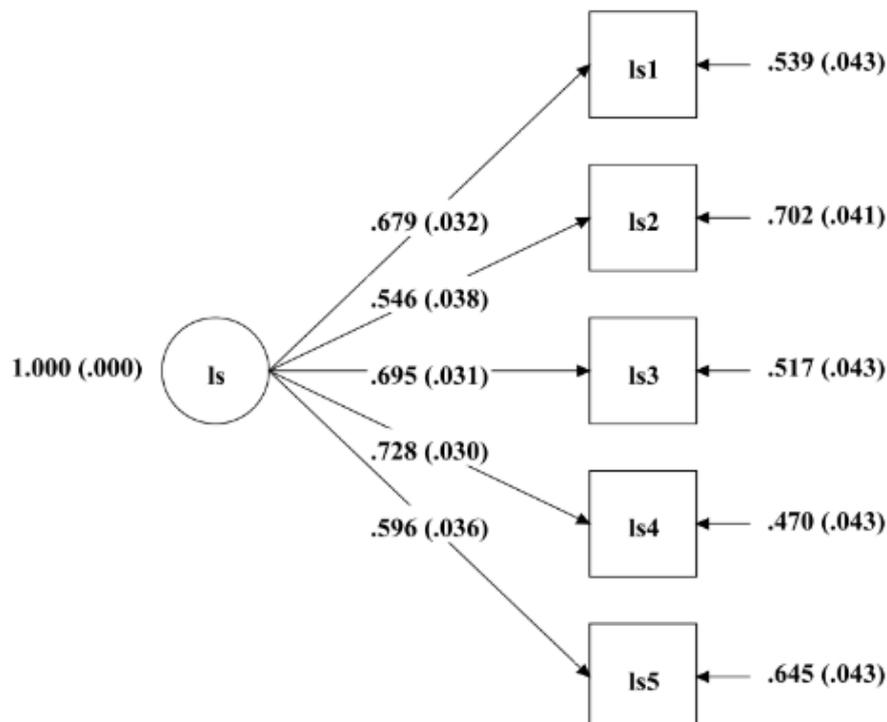


Figure 1. Confirmatory Factor Analysis of SWLS

The fit statistics for the model included in Figure 1 were found to have a perfect level of fit ($df = 5$, $\chi^2 = 10.556$; $\chi^2/df = 2.11$; CFI = 0.991, TLI = 0.982, RMSEA = 0.047; SRMR = 0.019). The standardized factor loads for variables observed in the CFA model, and implicit variables ranged from .546 to .695, with standard errors .041 to .043 and error variance from .470 to .7402 as seen in Figure 1. All intersects were significant at a .001 level. After CFA, MI according to gender was investigated for the life satisfaction scale. The findings obtained related to the MI stages according to gender were given in Table 1.

Table 1.

Fit statistics for MI according to Gender

Model	χ^2	df	χ^2/df	CFI	TLI	RMSEA	SRMR	-2 Δ LL	Δ df
Configural	13.53	8	1.69	0.99	0.974	0.053	0.022	-	-
Metric	18.54	12	1.544	0.989	0.98	0.047	0.044	5.028	4
Scalar	25.27	16	1.052	0.983	0.978	0.048	0.04	6.8	4
Strict	44.85	21	1.915	0.968	0.966	0.06	0.064	18.092	5
Partial Strict	31.05	19	1.634	0.978	0.976	0.05	0.053	5.712	3

Table 1 presents the results of the model fits. Firstly, considering the goodness of fit statistics for the configural invariance model related to the life satisfaction variable ($\chi^2/df = 1.690$; CFI = 0.990; TLI = 0.974; RMSEA = 0.053; SRMR = 0.022), it appeared that the configural invariance assumption holds. When the chi-square and goodness of fit results for metric invariance were assessed, it showed that the metric model was valid. The -2LL difference test between configural invariance and metric invariance models was not significant ($-2\Delta LL (4) = 5.511$, $p > 0.05$); this result showed that metric invariance holds. Abiding by metric invariance means, the indicators measuring life satisfaction were equivalently across the groups. When the chi-square and goodness of fit results were evaluated for the scalar MI model, it was found the invariance model holds. The -2LL difference test between metric invariance and scalar invariance models were not significant ($-2\Delta LL (4) = 1.755$, $p > 0.05$); this result showed that scalar invariance was present. When the chi-square and goodness of fit results for the strict invariance model were assessed, it was found the invariance model holds. The strict invariance fit was significantly worse than the scalar invariance model, $-2LL (5) = 18.092$, $p < 0.01$, indicating that strict invariance did not hold. Modification indices suggest that freeing residual variances for item 3 and item 4 between groups would help significantly. The partial strict variance invariance model fit was significantly better compared to the strict invariance model. For the partial strict invariance model, when the chi-square and goodness of fit results were evaluated, the invariance model holds. The partial strict invariance model had significantly better fit than the strict invariance model ($-2\Delta LL (2) = 12.996$, $p < 0.01$). The -2LL difference test between scalar invariance and partial strict invariance models were not significant ($-2\Delta LL (5) = 5.712$, $p > 0.05$); this result shows that partial strict invariance was present.

In addition, because the scalar invariance model holds, two additional models (factor variance and factor mean invariance) were tested. In the first model, the factor variance in males was constrained to 1, similar to females. That variance results in a non-significant decrease in fit relative to the residual invariance model ($-2\Delta LL (1) = 0.109$, $p > 0.05$). As a result, males and females appeared to have equivalent amounts of individual differences in life satisfaction. In the second model, the factor mean in men was constrained to 0, similar to women, resulting in a non-significant decrease in fit relative to the factor variance invariance model ($-2\Delta LL (1) = 0.121$, $p > 0.05$). This result indicates that life satisfaction can be compared between males and females.

Study 2

In the second stage, the longitudinal invariance of the SWLS was investigated with data obtained from the second study group. With this aim, firstly, the descriptive statistics for each item on the scale over three applications were calculated and were given in Table 2.

Table 2.

Descriptive Statistics of items across time

Items	N	Mean	Std. Deviation	Skewness	Kurtosis
LS 1T1	388	3.05	.889	-.084	.231
LS 1T2	388	3.03	.900	-.275	.084
LS 1T3	388	2.98	.895	-.112	.050
LS 2T1	388	2.85	.849	-.174	.433
LS 2T2	388	2.85	.827	-.244	.094
LS 2T3	388	2.79	.847	-.081	.299
LS 3T1	388	3.46	.878	-.553	.361
LS 3T2	388	3.42	.857	-.628	.509
LS 3T3	388	3.39	.901	-.216	-.055
LS 4T1	388	3.08	.902	-.094	-.395
LS 4T2	388	3.06	.906	-.075	-.478
LS 4T3	388	3.04	.901	-.108	-.378
LS 5T1	388	2.61	1.152	.276	-.730
LS 5T2	388	2.54	1.098	.249	-.744
LS 5T3	388	2.68	1.067	.062	-.707

LS: Satisfaction with Life Scale; T1: Time 1; T2: Time 2; T3: Time 3; 1: Item 1; 2: Item 2; 3: Item 3; 4: Item 4; 5: Item 5

When the table is examined, it appears the skewness and kurtosis coefficients of the items were in the interval from -1 to +1. The findings related to the longitudinal MI stages tests were given in Table 3.

Table 3.

Fit Statistics related to Longitudinal Measurement Invariance

Model	χ^2	df	χ^2/df	CFI	TLI	RMSEA	SRMR	-2 Δ LL	Δ df
Configural	77.65	63	0.62	0.998	0.997	0.014	0.033	-	-
Metric	86.63	55	1.08	0.998	0.997	0.015	0.041	5.311	8
Scalar	99.13	47	1.13	0.996	0.995	0.018	0.045	6.32	8
Strict	110	37	1.12	0.996	0.995	0.018	0.045	3.241	10

Table 3 presents the results of the model fits. Firstly, considering the goodness of fit statistics ($\chi^2/df = 0.620$; CFI = 0.998; TLI = 0.997; RMSEA = 0.014; SRMR = 0.033) for the configural invariance model related to the life satisfaction variable, the construct had a single factor through time and appeared to fit by the configural invariance assumption. When the chi-square and goodness of fit results for the metric invariance model were evaluated, it showed the metric invariance model holds. The -2LL difference test between configural invariance and metric invariance models was not significant ($-2\Delta LL (8) = 5.311, p > 0.05$); in other words, metric invariance holds. Abiding by metric invariance means that the indicators measuring life satisfaction were equivalently across time. When the chi-square and goodness of fit results for the scalar invariance model were investigated, the invariance model holds. The -2LL difference test between metric invariance and scalar invariance models were not significant ($-2\Delta LL (8) = 2.628, p > 0.05$); this shows that partial scalar invariance is present. Finally, the strict invariance model was investigated. For the strict invariance model, when the chi-square and goodness of fit results are evaluated together, they show the strict invariance model holds. The -2LL difference test between partial scalar invariance and strict invariance models is not significant ($-2\Delta LL (10) = 3.241, p > 0.05$), showed that strict invariance was supported.

After ensuring measurement invariance, the latent factor means that the results across time can be compared. This factor means that T2 and T3 were constrained to be equal, the estimated factor mean at T2 was -0.056, and the factor mean at T3 was -0.058. This results in a non-significant decrease in fit relative to the factor covariance invariance model ($-2\Delta LL (1) = 1.28, p = 0.87$), indicating that the factor means at T3 were insignificantly higher than at T2. The obtained results show the equivalence of the latent factor means through time. Additionally, the stability coefficient across time, in other words the correlation between factors at three time points, was computed using the strict invariance model. To calculate the stability coefficient, the factor variances were fixed to 1. The resulting factor correlation between T1 and T2 is 0.806, the factor correlation between T1 and T3 was 0.645, and the factor

correlation between T2 and T3 was 0.607. Thus, it was found that the SWLS also had moderate stability between T1 and T3 and T2 and T3, but highly stability between T1 and T2.

DISCUSSION AND CONCLUSION

This study investigated the longitudinal MI and the gender MI of the SWLS in a university student sample. First, the findings revealed that the Turkish SWLS has adequate internal consistency (values are between .774 and .811). The results showed that the original single factor structure was replicated in Turkish late adolescent populations. The fit indices show the SWLS has a perfect fit for the late adolescent sample (Kline, 2005; Tabachnick & Fidell, 2007). These findings are in line with previous adaptation studies about the SWLS (Durak et al., 2010; Köker, 1991). As a result, it can be said the SWLS is a valid and reliable scale for life satisfaction among university students.

Based on the results of the invariance study according to gender, all items on the SWLS provide partial strict invariance. In other words, the same construct, factor loadings, and item intercepts were revealed in similar between gender groups. When the literature is examined, it appears that the MI of the SWLS in different samples and different cultures according to gender was revealed (e.g., Checa et al., 2019; Tomás et al., 2015). Additionally, some studies (Atienza et al., 2003; Sovet et al., 2016) have found that not all items on the SWLS are invariant. However, as partial strict MI was present, the SWLS provided similar responses to items in female and male groups, and as such, the points obtained in these groups may be compared. In this context, the differences in points obtained on the SWLS by university students according to gender can be used to perform comparison studies.

Study 2 findings revealed that SWLS provide strict longitudinal invariance. These findings mean that the equality of factor patterns, factor loadings, intercepts, and residual variance were determined similarly for university students at four-month time intervals. Additionally, Study 2 results show that the SWLS has good long-term stability. The stability coefficient for T1 and T2 is 0.806, for T1 and T3 is 0.645, and for T2 and T3 is 0.607. These results show that the construct measured by the SWLS is reliable and stable over time. Stated differently, longitudinal invariance means that the late adolescent understanding of items remains the same throughout the eight-month interval.

There is minimal research demonstrating longitudinal MI of SWLS (Esnaola, Benito, Antonio-Agirre, Axpe, & Lorenzo, 2019; Jovanović, 2017; Kyeunghae Lee et al., 2010; Wu et al., 2009). Studies by Wu et al. (2009) observed that scalar invariance was present in longitudinal invariance studies performed with university students at two-month intervals, while metric invariance was present in a study of adolescents at three-month intervals. Similarly, the study by Jovanović (2017) revealed the SWLS provide scalar invariance. Additionally, like the present study, Kyeunghae Lee et al. (2010) found evidence that all stages of longitudinal invariance were met.

When the literature is examined, it appears that the question whether life satisfaction remains stable or not over time has been investigated. For example, a study by Lucas and Donnellan (2007) revealed the variance in points observed did not vary between 34% and 38%. Additionally, according to the results of the same study, they proposed that life satisfaction may vary to a certain degree linked to contextual situations. Similarly, the study by Koivumaa-Honkanen, Kaprio, Honkanen, Viinamäki, and Koskenvuo (2005) found that the life satisfaction of participants followed for 15 years was moderately stable. In this study, participants were followed for eight months, and it was revealed that life satisfaction had good long-term stability. However, in the literature, studies show that life satisfaction may vary as the time interval increases (Ehrhardt, Saris, & Veenhoven, 2000; Fujita & Diener, 2005). As a result, it should be considered that the longitudinal invariance of life satisfaction may vary depending on the time interval.

As a result, the structure of life satisfaction does measure the same construct for Turkish undergraduate students with varied genders and time. Researchers can compare the life satisfaction concept without worrying about whether the scale measures the same construct according to gender in comparison studies. In the literature, there is no other study encountered showing factor invariance between genders and time for the Turkish version of the SWLS. Also, longitudinal invariance results imply that the construct measured by the SWLS is stable and reliable over time. The SWLS can be used to make longitudinal mean comparisons.

When studies related to the longitudinal invariance (e.g., Jovanović, 2017; Kyeunghae Lee et al., 2010; Wu et al., 2009) and the gender invariance (e.g., Sovet et al., 2016; Moksnes et al., 2014) of the SWLS are investigated, some studies found metric invariance were present, while some studies found full strict invariance. Hence, there is no clarity in the literature related to the gender and longitudinal MI of the SWLS in different cultures, different samples, and at different time intervals. While our findings support its utility, they must be considered in connection with the broader literature base with some conflicting findings. As a result, we would recommend that researchers perform studies to test the longitudinal invariance of the SWLS at different time intervals and in different samples.

There are some limitations of the present study. Firstly, the SWLS scale was adapted to different samples in Turkey, and it appears that comparison studies of different samples were performed. This study only revealed the gender and longitudinal MI for university students of faculty of education. As a result, the results of this study may only be generalized to university students with similar features.

As mentioned previously, the stability of life satisfaction according to time is affected by the time interval and the life events experienced by participants, so the longitudinal invariance study was limited to eight months' duration and three applications. Studies with a longer time interval in longitudinal comparisons may be recommended for repeat longitudinal invariance investigations in future studies. Another limitation of the study is that the life satisfaction concept that the study investigated was dealt with in a single dimension based on Diener et al. (1985). As a result, the results of this study cannot be generalized to conceptual constructs proposing life satisfaction comprises multiple dimensions or scales measuring multidimensional constructs.

To sum up, SWLS is still a widely used measurement all over the world. It is crucial to evaluate its invariance among time and gender to assess mean differences appropriately. Nevertheless, the results of this study add value to the current understanding of life satisfaction stability regarding time and gender.

Acknowledgements

The ethical committee approval of the study was obtained from the Social and Humanities Sciences Research Ethics Committee at Ordu University (Approval Number is 2018-34).

REFERENCES

- Anaby, D., Jarus, T., & Zumbo, B. D. (2010). Psychometric evaluation of the Hebrew language version of the Satisfaction with Life Scale. *Social Indicators Research*, 96(2), 267-274. DOI: 10.1007/s11205-009-9476-z
- Arrindell, W. A., Heesink, J., & Feij, J. A. (1999). The satisfaction with life scale (SWLS): Appraisal with 1700 healthy young adults in The Netherlands. *Personality and Individual Differences*, 26(5), 815-826. DOI: 10.1016/S0191-8869(98)00180-9
- Atienza, F. L., Balaguer, I., & García-Merita, M. L. (2003). Satisfaction with life scale: Analysis of factorial invariance across sexes. *Personality and Individual Differences*, 35(6), 1255-1260. DOI: 10.1016/S0191-8869(02)00332
- Bayani, A. A., Koocheky, A. M., & Goodarzi, H. (2007). The reliability and validity of the satisfaction with life scale. *Journal of Iranian psychologists*, 3(11), 259-60.

- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press Publications.
- Checa, I., Perales, J., & Espejo, B. (2019). Measurement invariance of the Satisfaction with Life Scale by gender, age, marital status and educational level. *Quality of Life Research*, 28(4), 963-968. DOI: 10.1007/s11136-018-2066-2
- Çivitçi, A. (2012). Üniversite öğrencilerinde genel yaşam doyumu ve psikolojik ihtiyaçlar arasındaki ilişkiler [The relationships between global life satisfaction and psychological needs in university students]. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 21(2), 321-336.
- Clench-Aas, J., Nes, R. B., Dalgard, O. S., & Aarø, L. E. (2011). Dimensionality and measurement invariance in the Satisfaction with Life Scale in Norway. *Quality of Life Research*, 20(8), 1307-1317. DOI: 10.1007/s11136-011-9859-x
- Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness Studies*, 3(1), 37-69. DOI: 10.1023/A:1015678915305
- Dağlı, A., & Baysal, N. (2016). Yaşam doyumu ölçeğinin Türkçe'ye uyarlanması: geçerlik ve güvenilirlik çalışması [Adaptation of the satisfaction with life scale into Turkish: The study of validity and reliability]. *Electronic Journal of Social Sciences*, 15(59), 1250-1262. DOI:10.17755/esosder.75955
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71-75. DOI: 10.1207/s15327752jpa4901_13
- Diener, E., Inglehart, R., & Tay, L. (2013). Theory and validity of life satisfaction scales. *Social Indicators Research*, 112(3), 497-527. DOI: 10.1007/s11205-012-0076-y
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125(2), 276-302.
- Dimitrova, R., & del Carmen, D. E. A. (2015). Measurement invariance of the satisfaction with life scale in Argentina, Mexico and Nicaragua. *Social Inquiry into Well-being*, 1, (1) 32-39. DOI: 10.13165/SIIW-15-1-1-04
- Durak, M., Senol-Durak, E., & Gencoz, T. (2010). Psychometric properties of the satisfaction with life scale among Turkish university students, correctional officers, and elderly adults. *Social Indicators Research*, 99(3), 413-429. DOI: 10.1007/s11205-010-9589-4
- Ehrhardt, J. J., Saris, W. E., & Veenhoven, R. (2000). Stability of life-satisfaction over time. *Journal of Happiness Studies*, 1(2), 177-205.
- Emerson, S. D., Guhn, M., & Gadermann, A. M. (2017). Measurement invariance of the Satisfaction with Life Scale: reviewing three decades of research. *Quality of Life Research*, 26(9), 2251-2264. DOI: 10.1007/s11136-017-1552-2
- Esnaola, I., Benito, M., Antonio-Agirre, I., Axpe, I., & Lorenzo, M. (2019). Longitudinal measurement invariance of the Satisfaction with Life Scale in adolescence. *Quality of Life Research*, 1-7. DOI: 10.1007/s11136-019-02224-7
- Fujita, F., & Diener, E. (2005). Life satisfaction set point: stability and change. *Journal of Personality and Social Psychology*, 88(1), 158-164. DOI: 10.1037/0022-3514.88.1.158
- Geiser, C. (2012). *Data Analysis with Mplus (Methodology in the Social Sciences)*. New York, NY: Guilford Press.
- Gilman, R., & Huebner, E. S. (2006). Characteristics of adolescents who report very high life satisfaction. *Journal of Youth and Adolescence*, 35(3), 293-301. DOI: 10.1007/s10964-006-9036-7
- Gouveia, V. V., Milfont, T. L., Da Fonseca, P. N., & de Miranda Coelho, J. A. P. (2009). Life satisfaction in Brazil: Testing the psychometric properties of the satisfaction with life scale (SWLS) in five Brazilian samples. *Social Indicators Research*, 90(2), 267-277. DOI: 10.1007/s11205-008-9257-0
- Guhn, M., Ark, T. K., Emerson, S. D., Schonert-Reichl, K. A., & Gadermann, A. M. (2018). The Satisfaction with Life Scale adapted for Children: Measurement invariance across gender and over time. *Psychological Assessment*, 30(9), 1261. DOI: 10.1037/pas0000598
- Gündoğar, D., Gül, S. S., Uskun, E., Demirci, S., & Keçeci, D. (2007). Üniversite öğrencilerinde yaşam doyumunu yordayan etkenlerin incelenmesi [Investigation of the predictors of life satisfaction in university students]. *Klinik Psikiyatri*, 10(1), 14-27.
- Hinz, A., Conrad, I., Schroeter, M. L., Glaesmer, H., Brähler, E., Zenger, M., ... & Herzberg, P. Y. (2018). Psychometric properties of the Satisfaction with Life Scale (SWLS), derived from a large German community sample. *Quality of Life Research*, 27(6), 1661-1670. DOI: 10.1007/s11136-018-1844-1
- Jang, S., Kim, E. S., Cao, C., Allen, T. D., Cooper, C. L., Lapiere, L. M., ... & Abarca, N. (2017). Measurement invariance of the satisfaction with life scale across 26 countries. *Journal of Cross-Cultural Psychology*, 48(4), 560-576. DOI: 10.1177/0022022117697844
- Jovanović, V. (2017). Measurement invariance of the Serbian version of the satisfaction with life scale across age, gender, and time. *European Journal of Psychological Assessment*, 35, 555-563. DOI: 10.1027/1015-5759/a000410

- Kline, R. B. (2005). *Principles and practice of structural equation modeling: Methodology in the social sciences*. The Guilford Press: New York.
- Koivumaa-Honkanen, H., Kaprio, J., Honkanen, R. J., Viinamäki, H., & Koskenvuo, M. (2005). The stability of life satisfaction in a 15-year follow-up of adult Finns healthy at baseline. *BMC Psychiatry*, 5(1), 4. DOI: 10.1186/1471-244X-5-4
- Köker, S. (1991). *Normal ve sorunlu ergenlerin yaşam doyumu düzeylerinin karşılaştırılması [A comparison of life satisfaction of normal and delinquent adolescence]* (Unpublished master's thesis) Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Kyeunghae Lee, K., Brekke, J. S., Yamada, A. M., & Chou, C. P. (2010). Longitudinal invariance of the satisfaction with life scale for individuals with schizophrenia. *Research on Social Work Practice*, 20(2), 234-241. DOI: 10.1177/1049731509347860
- Lucas, R. E., & Donnellan, M. B. (2007). How stable is happiness? Using the STARTS model to estimate the stability of life satisfaction. *Journal of Research in Personality*, 41(5), 1091-1098. DOI: 10.1016/j.jrp.2006.11.005
- Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, 13(2), 127-143. DOI: 10.1016/0883-0355(89)90002-5
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525-543.
- Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 380–392). New York, NY: Guilford Press.
- Moksnes, U. K., Løhre, A., Byrne, D. G., & Haugan, G. (2014). Satisfaction with life scale in adolescents: Evaluation of factor structure and gender invariance in a Norwegian sample. *Social Indicators Research*, 118(2), 657-671. DOI: 10.1007/s11205-013-0451-3
- Muthén, L.K., & Muthén, B.O. (1998-2012). Mplus user's guide. Seventh Edition. Los Angeles, CA: Author.
- Özgür, G., Gümüş, A. B., & Durdu, B. (2010). Evde ve yurttan kalan üniversite öğrencilerinde yaşam doyumu [Life satisfaction of university students living at home or in the dormitory]. *Psikiyatri Hemşireliği Dergisi*, 1(1), 25-32.
- Sachs, J. (2003). Validation of the satisfaction with life scale in a sample of Hong Kong university students. *Psychologia*, 46(4), 225-234. DOI: 0.2117/psysoc.2003.225
- Shevlin, M., Brunson, V., & Miles, J. N. V. (1998). Satisfaction with life scale: Analysis of factorial invariance, mean structures and reliability. *Personality and Individual Differences*, 25(5), 911-916. DOI: 10.1016/S0191-8869(98)00088-9
- Sovet, L., Atitsoğbe, K. A., Pari, P., Park, M. S. A., & Villieux, A. (2016). Psychometric evaluation of the satisfaction with life scale in Togo: A three-step approach. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 66(5), 243-250. DOI: 10.1016/j.erap.2016.06.002
- Suh, E., Diener, E., Oishi, S., & Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology*, 74(2), 482. DOI: 10.1037/0022-3514.74.2.482
- Swami, V., & Chamorro-Premuzic, T. (2009). Psychometric evaluation of the Malay satisfaction with life scale. *Social Indicators Research*, 92(1), 25. DOI: 10.1007/s11205-008-9295-7
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (Vol. 5). Boston, MA: Pearson.
- Tomás, J. M., Gutiérrez, M., Sancho, P., & Romero, I. (2015). Measurement invariance of the Satisfaction with Life Scale (SWLS) by gender and age in Angola. *Personality and Individual Differences*, 85, 182-186. DOI: 10.1016/j.paid.2015.05.008
- Tuzgöl-Dost, M. (2007). Üniversite öğrencilerinin yaşam doyumunun bazı değişkenlere göre incelenmesi [Examining life satisfaction levels of university students in terms of some variables]. *Pamukkale University Journal of Education*, 2(22), 132-143.
- 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: 10.1177/109442810031002
- Wu, C. H., Chen, L. H., & Tsai, Y. M. (2009). Longitudinal invariance analysis of the satisfaction with life scale. *Personality and Individual Differences*, 46(4), 396-401. DOI: 10.1016/j.paid.2008.11.002
- Yetim, Ü. (1993). Life satisfaction: A study based on the organization of personal projects. *Social Indicators Research*, 29(3), 277-289.

TÜRKÇE GENİŞLETİLMİŞ ÖZET

Yaşam doyumu kavramı, bireylerin yaşamlarına yönelik kendi algılarını ve bireyin kendini grup içinde karşılaştırmasını gerektirdiği için bireyin içinde bulunduğu kültürden etkilenebilmektedir. Literatüre bakıldığında yaşam doyumunun toplulukçu ve bireyci kültürlerde farklı algılanabildiği görülmektedir (Suh, Diener, Oishi & Triandis, 1998). Çeşitli araştırmalar (Emerson, Guhn & Gadermann, 2017; Jang et al., 2017) yaşam doyumu ölçeklerinin yapısının kültüre göre değişebildiğini ortaya koymuşlardır. Ayrıca Yaşam Doyumu Ölçeği'nin (YDÖ) toplulukçu kültüre daha yakın olan Türkiye'de yapılan faktör analizleri (Dağlı & Baysal, 2016; Durak, Senol-Durak & Gencoz, 2010) ölçeğin geçerliğine bir kanıt olarak yorumlansa da ölçme değişmezliğine bakılmadığı görülmektedir. Bununla birlikte yaşam doyumunun zamandan etkilenen bir kavram olduğu da literatürde tartışma konusudur. Fujita ve Diener (2005) çalışmalarındaki katılımcıların %24'ünün yaşam doyumunun ilk beş yıl ve son beş yıl arasında anlamlı olarak değiştiğini ortaya koymuşlardır. Yine ölçümler arasındaki zaman aralıkları arttıkça kararlılığının azaldığı görülmüştür. Ayrıca, yaşam doyumu ölçeklerinin maddelerin sırası, o anki ruh hali gibi faktörlerden etkilenebileceği ancak çoğu zaman bu etkenlerin kontrol edilebileceği vurgulanmaktadır (Diener, Inglehart & Tay, 2013). Tüm bunlardan hareketle yaşam doyumunun ölçülmesinde kullanılan ölçeklerin kültüre uyarlanırken yalnızca faktör analizi ile sınırlı kalmaması aynı zamanda ölçme değişmezliği çalışmalarının da yapılması gerektiği söylenebilir.

Literatür incelendiğinde farklı örneklem gruplarında Türkçe'ye uyarlanan (Dağlı & Baysal, 2016; Durak, Senol-Durak & Gencoz, 2010; Köker, 1991; Yetim, 1993) yaşam doyumu ölçeğinin Türkiye örneğinde herhangi bir ölçme değişmezliği çalışmalarının (yaş, cinsiyet, medeni durum gibi...) yapılmadığı görülmüştür. Bununla beraber ölçeğin cinsiyetler arası karşılaştırmalarda kullanıldığı çalışmalar (Çivitci, 2012; Gündoğar, Gül, Uskun, Demirci & Keçeci, 2007; Özgür, Gümüş & Durdu, 2010; Tuzgöl-Dost, 2007) cinsiyet değişmezliğinin olduğu varsayılarak yapılmıştır. Ancak yaşam doyumu ölçeğinin gruplar arası veya zamana göre karşılaştırılabilirliği için ölçme değişmezliğine sahip olması istenir. Bu nedenle bu araştırmanın amacı; yaşam doyumu ölçeğinin cinsiyete ve zamana göre ölçme değişmezliğine sahip olup olmadığını belirlemektir.

Bu araştırmanın çalışma grubunu 2018-2019 eğitim-öğretim yılında bir devlet üniversitesinin eğitim fakültesinde öğrenim gören öğrenciler oluşturmaktadır. Birinci çalışma grubunda 2018-2019 eğitim-öğretim yılının güz dönemi başında 366'sı (%73,2) kadın ve 134'ü (%26,8) erkek olmak üzere toplam 500 öğrenciden veri toplanmıştır. Öğrencilerin yaşları 17 ile 22 arasında değişmekte ve yaş ortalaması 20,58'dir. İkinci çalışma grubunu 376'sı (%69,37) kadın ve 166'sı (%30,63) erkek olmak üzere toplam 542 üniversite öğrencisi oluşturmaktadır. Çalışmada altı farklı bölümden dört ay arayla üç kez veri toplanmıştır. Verileri eşleştirmek için öğrencilerden ölçek setlerinin başına öğrenci numaralarını yazmaları istenmiştir. İkinci ve üçüncü uygulamalarda veri kayıpları olduğundan analizler 297'si (%76,5) kadın ve 91'i (%23,5) erkek olmak üzere toplam 388 öğrenciden toplanan veri üzerinde yapılmıştır.

Araştırmada Diener, Emmons, Laresen ve Griffin (1985) tarafından geliştirilen ve Türkçe geçerlik, güvenilirlik çalışması Köker (1991) ve Yetim (1993) tarafından gerçekleştirilen Yaşam Doyumu Ölçeği kullanılmıştır. Ölçek, yaşam doyumuna ilişkin beş maddeden oluşmaktadır. Araştırma kapsamında öncelikle doğrulayıcı faktör analizi gerçekleştirilmiştir. Daha sonra ele alınan yaşam doyumu ölçeğinin cinsiyet ve boylamsal değişmezliğini test etmek için ölçme değişmezliği modelleri test edilmiştir. Bütün analizler Mplus programı (Version 7.4) ile yapılmıştır (Muthén & Muthén, 1998-2012). Değişmezlik modellerine ilişkin karşılaştırmalarda log olabilirlik değerlerine dayalı olan Ki-Kare fark testi kullanılmıştır. Bu çalışmada biçimsel model, faktör yükleri, madde kesişimleri ve hata varyansları serbest olan modeli; metrik model faktör yükleri sabit modeli, madde kesişimleri ve hata varyansları serbest modeli; ölçek değişmezliği modeli, faktör yükleri ve madde kesişimleri sabit, hata varyansları serbest modeli ve katı değişmezlik modeli ise faktör yükleri, madde kesişimleri ve hata varyansları sabit modeli temsil etmektedir.

Değişmezlik, bir ölçme aracının sahip olması gereken önemli bir psikometrik özelliktir (Meredith, 1993; Brown, 2006). Ölçme değişmezliğinin sağlanması gruplar arasında ve farklı zamanlarda anlamlı karşılaştırmaların yapılabilmesi için bir ön koşuldur. Eğer değişmezlik sağlanmazsa puan ve grup farklılıkları ile ilgili yanıltıcı yorumlar yapılmasına neden olabilir. Cinsiyete göre değişmezlik sonuçlarına göre, Yaşam Doyumu Ölçeği'nin tüm maddelerinde yapısal, metrik ve ölçek değişmezliğini sağlandığı gözlenmiştir. Bir başka deyişle, cinsiyet fark etmeksizin bireylerin maddelere aynı şekilde yanıt verdiği ve ölçekten elde edilen puanların karşılaştırılabilir olduğu bulunmuştur. Boylamsal ölçme değişmezliğinin sonuçlarına göre, dört aylık zaman dilimine göre Yaşam Doyumu Ölçeği'nin tüm maddelerinde yapısal, metrik, ölçek ve katı değişmezliği sağlandığı bulunmuştur. Elde edilen bu sonuçlar YDÖ ile ölçülen yapının zamana göre güvenilir ve stabil olduğunu göstermektedir.

Yaşam Doyumu Ölçeği'nin boylamsal ve cinsiyete göre değişmezliğini ortaya koyan bu çalışmanın bazı sınırlılıkları bulunmaktadır. Öncelikle, Yaşam Doyumu Ölçeğinin Türkiye'de farklı örneklemelere uyarlandığı ve farklı örneklemelerde karşılaştırma çalışmaları yapıldığı görülmektedir. Bu çalışmada sadece üniversite öğrencilerinde cinsiyet ve boylamsal ölçme değişmezliği ortaya konmuştur. Dolayısıyla bu çalışmanın sonuçları sadece benzer niteliklere sahip üniversite öğrencilerine genellenebilir. Yaşam Doyumu Ölçeği'nin boylamsal değişmezliği ve cinsiyet değişmezliği ile ilgili çalışmalar incelendiğinde bazı çalışmaların metrik değişmezlik bazı çalışmaların ise katı değişmezliği sağladığı görülmüştür. Dolayısıyla yaşam doyumu ölçeği'nin farklı kültürlerde, farklı örneklem ve farklı zaman aralıklarındaki cinsiyet ve boylamsal ölçme değişmezliği ile ilgili literatürde bir netlik yoktur. Bu nedenle araştırmacılara Yaşam Doyumu Ölçeği'nin farklı zaman aralıklarında ve farklı örneklemelerinde boylamsal değişmezliğini test eden çalışmalar yapmaları önerilebilir.