TURKISH PSYCHOLOGICAL COUNSELING AND GUIDANCE JOURNAL
2023, VOL. 13, NO. 69, 224 - 239.

Official Journal of
Turkish Counseling and
Guidance Association

ISSN: 1302 – 1370, E-ISSN: 2822 – 6569 https://doi.org/10.17066/tpdrd.1223320mn

Life Satisfaction Scale: A Meta-Analytic Reliability Generalization Study in Turkey Sample

Mehmet Taha ESER[a] Nuri DOĞAN[b]

[a] Aydın Adnan Menderes University, Aydın, Turkey, [b]Hacettepe University, Ankara, Turkey

ARTICLE HISTORY

Received: 23.02.22
Accepted: 29.03.23

KEYWORDS

Reliability Generalization, Meta-Analysis, Life Satisfaction

ABSTRACT

Life satisfaction is the result of comparing one's expectations with the actual situation. The life satisfaction scale developed by Ed Diener and his colleagues is one of the most preferred life satisfaction scales in research. In this study, it was aimed to obtain an approximate estimate of the generalized reliability of the measurement tool through meta-analytic reliability generalization and to determine which sample characteristics of the studies may affect the variability of reliability coefficients. For 24 studies, the generalized reliability coefficient was .84 [.81-.87]. No visual or statistical evidence of publication bias was encountered in the meta-analysis. As a result of the moderator analysis, it was concluded that the year of publication, sample type and the percentage of women in the sample can be shown as sources related to the change in the reliability value. It is thought that it would be useful to report reliability coefficients in accordance with the research conditions and assumptions in future studies.

The concept of life satisfaction has received increasing attention in the last two decades as an important topic in human psychology research. In recent years, researchers have focused on the positive supports and rewards that enable people to form perceptions, be happy and empathize with others, rather than undesirable emotions such as anxiety, depression, sadness and unhappiness. Life satisfaction is among the most widely used concepts to assess subjective well-being. The concept of "life satisfaction", which was first introduced in the literature by Neugarten et al. (1961), has been the subject of many studies. Before defining the concept of life satisfaction, it is first necessary to define the concept of "satisfaction". Satisfaction is the fulfillment of expectations, needs and wishes. "Life satisfaction" is the state or result obtained by comparing a person's expectations (what they want) with what they have (what they have) (Diener, 1984). Life satisfaction can be defined as the cognitive evaluation of one's overall satisfaction with one's current life according to one's own criteria of what a satisfying life means (Diener et al., 1985). Life satisfaction evokes happiness in the mind through a sense of self-worth, peace of mind, satisfaction with work life, tolerance for difficulties, and positive perceptions and attitudes towards oneself, others and society. These factors are important for well-being; they result in perceptions of self-efficacy, self-esteem, good personality, positive emotions and attitudes, and also increase career success.

Numerous studies underscore the positive outcomes associated with high levels of life satisfaction. For instance, research has persistently shown that people who are more satisfied with their lives usually have stronger social bonds, enjoy better social support, and display greater contentment in their marital lives compared to their less satisfied counterparts (Barger, Donoho, & Wayment, 2009; Diener & Seligman, 2002; Pavot & Diener, 2008). From a professional standpoint, life satisfaction proves to be beneficial as well. Employees with a higher degree of life satisfaction tend to perform better at their jobs, express greater

CORRESPONDING AUTHOR Mehmet Taha ESER, taha.eser@hacettepe.edu.tr, ORCID: 0000-0001-7031-1953, Aydın Adnan Menderes University, Aydın, Turkey.

satisfaction with their careers, demonstrate a higher commitment to their organizations, and are less likely to leave their jobs (Erdogan, Bauer, Truxillo, & Mansfield, 2012). Moreover, high life satisfaction is linked to improved health outcomes and longevity. Research suggests that people with high life satisfaction tend to have better overall health and fewer chronic health problems (Siahpush, Spittal, & Singh, 2008). Importantly, high levels of life satisfaction significantly reduce the risk of mortality, thus contributing to a longer lifespan (Lyyra, Törmäkangas, Read, Rantanen, & Berg, 2006; Xu & Roberts, 2010).

Several research studies have highlighted that life satisfaction also plays a pivotal role in fostering positive educational outcomes. For instance, alongside the absence of psychological distress, a high level of life satisfaction significantly contributes to student engagement and academic success among university students (Antaramian, 2015; Renshaw & Cohen, 2014). Furthermore, university students who report high life satisfaction often express greater contentment with their academic journey (Duffy, Allan, & Bott, 2012; Ojeda, Flores, & Navarro, 2011). Life satisfaction has been linked with optimistic academic expectations, enhanced academic self-confidence, the perception of being closer to achieving personal goals, and reduced academic stress (Ojeda et al., 2011; O'Sullivan, 2011). Additionally, some research indicates a correlation between positive life satisfaction and superior grade point averages (GPAs) among university students, underscoring the significance of life satisfaction in academic performance (Howell, 2009; Rode et al., 2005).

The first scale development study on the concept of life satisfaction was conducted by Diener et al. (1985). Validity and reliability studies were also conducted for different groups (Diener, Pavot, Colvin, & Sandvik, 1991; Diener & Pavot, 1993). In addition to these studies, different scale development studies have taken place in the literature to measure life satisfaction of groups with different characteristics (Andrews & Withey, 1976; Lavallee, Hatch, Michalos, & McKinley, 2007; Lightsey Jr, McGhee, Ervin, Gharghani, Rarey, Daigle, Wright, Powell, 2013). When the literature is examined, life satisfaction scales adapted to Turkish culture are found. It is noteworthy that these are scales created by adapting the life satisfaction scales developed by researchers such as Diener et al. (1985), Lavallee et al. (2007) and Huebner (1991) into Turkish and applying them to different groups (Köse, Çobanoğlu, & Sarı, 2022).

When the literature was examined, it was determined that one of the most preferred life satisfaction scales within the scope of the studies was the life satisfaction scale developed by Ed Diener and his colleagues (Diener et al., 1985; Pavot & Diener, 1993). Life satisfaction is a factor in the more general structure of subjective well-being. Theoretical studies and research in fields other than rehabilitation show that subjective well-being has at least three components: positive emotional appraisal, negative emotional appraisal and life satisfaction (Pavot & Diener, 2008). Life satisfaction differs from emotional appraisal in that it is cognitive rather than emotionally driven. Life satisfaction can be assessed specific to a particular life domain or globally. Life satisfaction scale is considered as a global measure of life satisfaction.

The Life Satisfaction Scale (Diener et al., 1985) consists of 5 items and the items are in a seven-point Likert-type rating form/format. Scores that can be obtained from the scale vary between 5-35. High scores indicate high satisfaction with life. Diener et al. (1985) reported Cronbach's alpha internal consistency coefficient as .87 and test-retest (2 months interval) coefficient as .82. The Life Satisfaction Scale (LSS) has been used in various cultures and populations and translated into various languages. A large body of research has also examined the psychometric properties of the LSS. Exploratory and confirmatory factor analyses supported the unidimensional structure of the LSS as conceptualized by Diener et al. The instrument also has acceptable convergent validity; it is related to, yet distinct from, constructs such as anxiety, depression, happiness, self-esteem, negative and positive affect, as well as psychological distress.

Reliability is one of the most important key concepts in the process of evaluating test scores. Reliability provides information about how close the test or measurement scores are to each other over repeated measures (Traub & Rowley, 1991). Contrary to popular belief, reliability is a dynamic characteristic of test scores rather than a fixed value for measurement results and can vary according to the characteristics of the data (Thompson & Vacha-Haase, 2000). Reliability should therefore be recalculated after each measurement and reported in each study, as reliability can be affected by sample characteristics such as gender, age or language (Thompson, 2002), that is, it can vary across different administrations. Internal consistency, a commonly employed measure of reliability, offers insights into the degree to which the items within a measurement instrument evaluate a

single underlying construct (Semma et al., 2019). There are many methods to assess reliability in terms of internal consistency, such as Cronbach's alpha, dividing the test into two halfes, and Kuder-Richardson. Among these methods, Cronbach's alpha is the most widely used reliability estimator (McNeish, 2018; Dimitrov, 2002) and can be calculated as follows:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} V_i}{V_{total}}\right]$$

The formula mentioned above includes variables such as n, representing the number of items in the scale, Vi, representing the variance of each specific item's score, and Vtotal, representing the variance of all scores (Cronbach, 1951). Psychometric theory points out that reliability is not an inherent characteristic of a test but may vary from one application to another. Considering that reliability can vary from one administration to another and can be greatly affected by sample characteristics, it is necessary to recalculate reliability for different samples. Variables such as gender, age, language and ethnicity can affect reliability estimates (Thompson, 2002). By examining such factors, sources of heterogeneity in reliability values can be examined. Consequently, the best way to meet expectations about the reliability of scores is to quantitatively integrate the various reliability estimates obtained from different administrations of the instrument.

Meta-analysis allows researchers to statistically combine multiple reliability coefficients resulting from the application of a particular test to different samples and contexts. Vacha-Haase (1998) used the term Meta-Analytic Generalization of Reliability (MARG) or simply Reliability Generalization (RG) to refer to this type of meta-analysis. MARG, aside from estimating the average reliability of test scores, aims to assess the variability of reliability coefficients reported across different studies that employ the same test. When encountering high variability, another objective is to investigate which characteristics of the studies may be statistically associated with the reliability estimates. This involves examining potential sources of heterogeneity in reliability values using various meta-analytic techniques (Henson & Thompson, 2002; Rodríguez & Maeda, 2006; Sánchez-Meca et al., 2013).

The same measurement result may vary in terms of reliability from case to case depending on the setting and sample, so different study characteristics such as design, format and setting and different participant characteristics such as age, gender and ethnicity may affect the differences in reliability (Vacha-Haase, 1998). When the measurement results of a study are valid for different target groups or different situations, the study is considered to have an acceptable level of generalizability (Onwuegbuzie & Larry, 2000). On the contrary, if the measurement results can only be applied to a subgroup of the sample or to a very specific situation, the generalizability of the relevant study is said to be weak (Vacha-Haase, 1998). Generalizability is one of the three different qualities that researchers use together with validity and reliability to assess the quality of their studies in general (Taylor, 2012). In this case, considering that reliability is the consistency and reproducibility of measurement results determined through a measurement tool applied to a defined sample group at a specific time and under specific conditions (Crocker & Algina, 1986) and the possibility that gender may affect reliability differences, sample type and gender variables were used in moderator analyses within the scope of this reliability generalization study.

The concept of life satisfaction is one of the most widely used well-being measures in the literature (McCulloch, 1992). Considering the fact that the concept of life satisfaction is frequently used as a variable within the scope of research, demographic trends and increasing interest in the concept of healthy aging, it is important that the psychometric properties of the concept of life satisfaction are well understood by researchers. The aim of this study is to apply meta-analytic reliability generalization to empirical studies using the LSS. At this point, it is aimed to obtain an approximate estimate of the overall reliability of the LSS and to determine which sample characteristics of the studies may affect the variability of the reliability coefficients. At the same time, the results to be obtained from this study will inform future research on the range of reliability estimates that can be expected for the LSS.

Method

This section explains the type of the study, the process of data collection, the establishment of coding criteria, the validity and reliability of measurement results, and the method of data analysis.

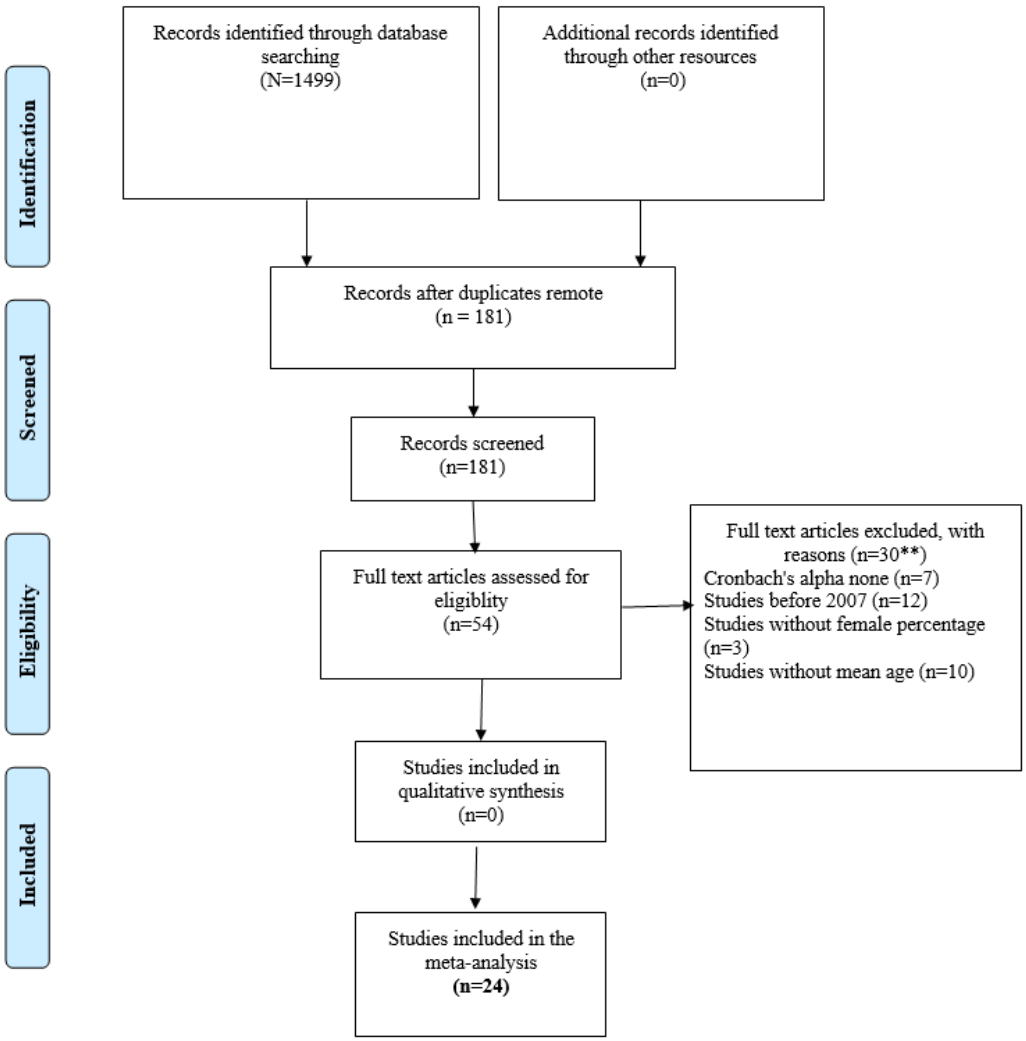
Type of Research

This research falls under the category of RG research. RG, a meta-analytical method, investigates the origins of error variance when multiple studies employing a specific instrument or a set of instruments to measure the same construct are taken into account (Vacha-Haase, Henson, &Caruso, 2002).

Data Collection Process

In the global academic community, it is advised to incorporate the flow diagram as outlined in the PRISMA statement and adhere to the PRISMA guidelines when conducting systematic review and meta-analysis studies. The purpose of following these guidelines is to ensure the accurate execution and enhance the presentation and reporting of systematic review and meta-analysis studies (The PRISMA Group, 2009). In this context, two researchers independently scanned the studies published in TR-Dizin and Adnan Menderes University library database between 2007 and 2022 in which the LSS was used. The databases were searched with the keywords "Life satisfaction" and "Yaşam doyum" and a total of 200 studies were analyzed according to their titles and abstracts. Then, the same studies were removed and the full texts of the remaining studies were analyzed.

Figure 1. Flow Diagram



Data Analysis

After the identical studies were removed, the full texts of the remaining studies were analyzed. Before proceeding to the analysis phase, coding reliability was calculated.

Coding of the Studies

At this stage, inclusion criteria were determined. The criteria determined are; i) ULAKBIM, Google Scholar and Adnan Menderes University library database, ii) Cronbach's Alpha coefficient is reported or calculable, iii) the sample group, sample size, mean age of the sample, scale form or number of items are included in the study, iv) the participants in the study were individuals of Turkish origin, and v) the language used in the research was either English or Turkish. Two authors independently coded the studies based on the inclusion criteria, and the inter-coder agreement percentage was found to be 90%, indicating a high level of agreement. The Krippendorff Alpha coefficient, which measures inter-coder reliability, was calculated to be .88, further supporting the high level of agreement between the coders. Ultimately, 24 studies met the inclusion criteria and were included in the meta-analysis, ensuring a robust dataset for analysis.

Statistical Analyses

All analyses of the study were conducted using JAMOVI and R software.The jamovi program was used to create funnel plots and forest graphs with high resolution. Other calculations related to the study were carried out through the "meta" package available within the R software.The distributions of Cronbach's Alpha coefficients are not normally distributed. Therefore, the distribution of reliability coefficients was normalized by using Bonett (2002) transformation. The reliability generalization study was carried out using the coefficients obtained as a result of the transformation, and the comments were made by converting the last value obtained into Cronbach's Alpha coefficient. In determining whether the studies evaluated within the scope of meta-analysis show a heterogeneous distribution, that is, in determining the sources of variability of reliability values within the scope of individual studies, the Q statistic, the I^2 statistic, which is a function of the Q statistic, and the lower and upper confidence interval for the generalized reliability value were used. Due to the fact that Bonett's VC model is fundamentally a random effects model according to Holland (2015), and the random effects model is considered to be a more accurate reflection of real-world scenarios as stated by Field (2003b), the research opted for the random effects model (REM). For the estimation of between-study variance under REM, the Sidik-Jonkman estimator, which has better features and produces better results than other estimators, was preferred.

During the coding of the studies selected according to the criteria for inclusion in the meta-analysis, the following study characteristics were considered: (i) study name, (ii) author(s) name, (iii) year of publication, (iv) language of publication, (v) reliability coefficient, (vi) type of reliability, (vii) sample size, (viii) sample type. The studies were coded by two researchers according to the specified characteristics, and the percentage of inter-coder agreement was 93% and the Krippendorff Alpha coefficient was .91, and these results were considered as an indicator of high inter-coder reliability.

This meta-analysis study has some limitations like other meta-analysis studies. The first one is the limitations of the meta-analysis method itself. Within the scope of the study, only the studies in which Cronbach's alpha value was calculated and only the studies conducted within the scope of Turkey sample were evaluated. At the same time, only articles were reviewed within the scope of the study. Although all of the postgraduate studies of individual researchers were reached, only those that were published and open to the access of readers could be reached due to the different dates of acceptance and publication of the articles in accordance with the publication policies of peer-reviewed scientific journals. Although a rich keyword pool was created for individual studies obtained with the help of search engines and databases, the fact that studies that were not shown or could not be reached as a result of the search could not be included in the list is seen as another important limitation. In addition, the study is limited to the analysis of coded moderator variables.

Within the scope of the study, funnel diagram, Egger's regression test and Kendall's tau were used to examine publication bias. In addition, the fail-safe N method was also used to obtain information about how many studies with an effect size value of zero should be conducted in order to eliminate the significance of the meta-analysis result. Within the scope of the study, the year of publication (Between 2007 and 2018/2018 and beyond, sample type (university student/non-university student), percentage of female students (more than half female students/more than half male students), mean age (continuous) were considered as moderator variables.

Meta-regression and Analog ANOVA were used to determine the effect of these moderator variables on the variability of reliability estimation.

Results

In this section, publication bias findings are presented first, followed by heterogeneity, effect size and moderator analysis findings.

Table 1 presents the descriptive statistics of the studies analyzed within the scope of the research.

Table 1. Characteristics of the Studies Included in the Meta-Analysis

Authors	Year	Sample Size	Female Populaion	Mean Age	Target Population
Bulut, M. B., & Yıldız M.	2020	602	324	20.86	University students
Recepoğlu, E., & Tümlü Ülker, G.	2015	94	31	45.02	Others
Sincar et al.	2020	1180	592	20.33	University students
Odacı et al.	2021	235	133	21.23	University students
Akbıyık, M.	2020	5208	2742	22.00	University students
Aydıner Boylu, A., & Güner, G.	2018	516	165	66.00	Others
Tepeli Temiz, S., & Ulusoy Gökçek, M.	2020	206	89	22.38	University students
Şahin, G., & Balcı Akpınar, R.	2016	194	57	37.24	Others
Balaban, T. et al.	2021	987	439	32.76	Others
Meryem Kara, F. et al.	2018	336	165	20.45	University students
Kolbaşı, E., & Bağcı, Z.	2019	131	28	50.00	Others
Atasoy, I., & Turan, Z.	2019	214	191	31.10	Others
Amanak, K., & Sevil, Ü.	2020	112	112	37.50	Others
Gençay, S., & Akkoyunlu, Y.	2012	233	108	22.14	University students
Yıldırım, J. C. et al.	2021	403	266	23.38	Others
Uğurlu, O.	2013	246	95	21.09	University students
Çattık, M., & Aksoy, M.	2018	225	139	35.80	Others
Erdoğdu, M. Y.	2021	435	215	16.21	Others
Akyüz, H.	2020	146	58	21.92	University students
Çırpan, H. et al.	2019	284	119	31.31	Others
Kahyaoğlu Süt, H.	2019	455	394	20.50	University students
Dil, K. et al.	2020	230	104	45.80	Others
Bozoğlan, B.	2015	444	236	54.12	Others
Şeker, B. D., & Sirkeci, İ.	2014	125	125	35.06	Others

When Table 1 is examined, it is noticeable that there are 24 studies evaluated within the scope of the research. When Table 1 is re-examined, it is seen that the publication year of the studies varies between 2013 and 2021, and the study with the largest sample is Sincar et a. (2020). In addition, the study with the largest female population is Akbıyık et al. (2020).

Below are the results of the publication bias analysis conducted within the scope of the study.

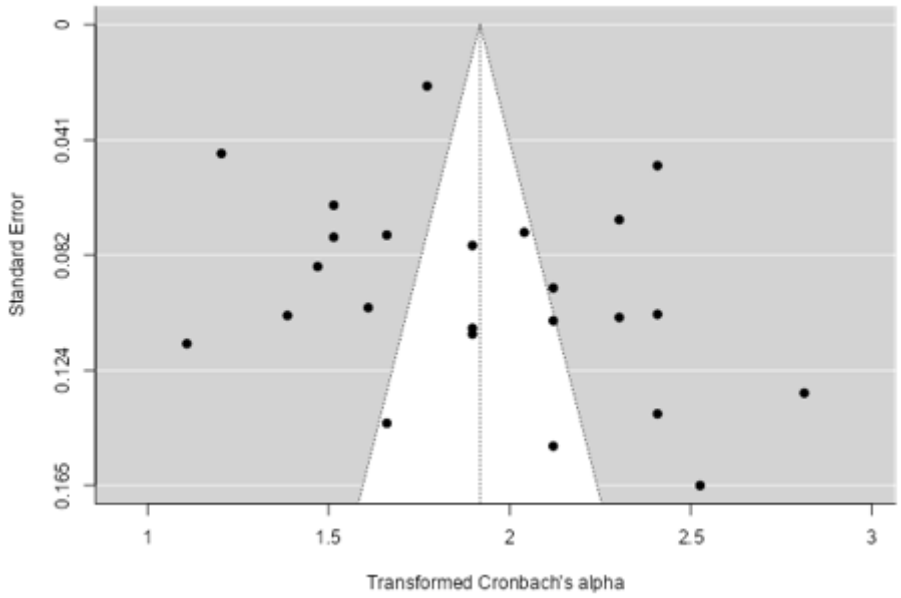
Results on Publication Bias

Within the scope of the research, before calculating the average effect size of the meta-analysis, evidence for the detection of publication bias in the studies examined within the scope of the analysis was sought. For this purpose, Mullen, Muellerleile, and Bryant's (2001) formula, funnel plot, Egger's linear regression test, Kendall's tau value, and Duval and Tweedie clipping and filling method were used.

Mullen, Muellerleile, and Bryant (2001) pointed out that the resilience of the results of meta-analysis studies against future studies can only be realized when the value calculated using the formula N/(5k+10) is greater than 1. Using the relevant formula, it was determined that the value obtained for the total population of 13,241 people was greater than 1. The fact that the calculation results for these values are greater than 1 can be interpreted as the publication bias of this meta-analysis study is very low.

One of the most widely used methods for collecting visual evidence of publication bias is the funnel plot. Figure 2 shows the funnel plot.

Figure 2. Funnel Graph



In a funnel plot, studies are expected to be symmetrically distributed around the uncertainty line (the line that cuts the overall magnitude of reliability). Although Figure 2 shows that individual studies are approximately symmetrically distributed to the right and left of the overall effect size for Cronbach's alpha, this interpretation is subjective (Borenstein, 2019). For a more objective interpretation, Egger's regression test should be utilized. The fact that the result of Egger's Linear Regression test (EggerValue= 1.70, p=0.08>0.05) is not statistically significant is another indication that there is no evidence of publication bias. In addition, the Begg and Mazumdar rank correlations statistic was examined for contributing/not contributing to the lack of asymmetry in the funnel plot. When the Begg and Mazumdar rank correlations statistic is analyzed, it is concluded that there is no asymmetry in the funnel plot (Kendall's Tau=0.13, p=0.16>0.05). Finally, the Duval and Tweedie trimming and filling test result was examined and it was concluded that there was no difference between the observed and actual effect sizes. In conclusion, as a result of the tests on the asymmetry of the funnel diagram, it was concluded that the funnel diagram was symmetrical and there was lack of publication bias.

Within the scope of the research, the numerical output of the meta-analysis regarding the Fail-Safe N, which is a way of defining the p value, was also analyzed. The fact that the p-value for the Fail-Safe N is smaller than the alpha value (p<0.001) indicates that the study is a strong study with low reliability. Within the scope of the study, it was determined that the p value for the Number of Error Protections was less than the alpha value of 0.05 (FSN=117.03, p<.001). According to this result, it can be said that the study is a strong study with a high level of reliability.

Results on the Generalized Effect Size of the Reliability of the Satisfaction with Life Scale

After carrying out a search for any signs of publication bias concerning the studies that are part of the study's sample, the generalized effect size should be computed. This computation should be done within the context of the chosen random effects model, taking into consideration the study's sampling frame. Table 2 displays the calculated generalized effect size along with the upper and lower confidence interval values associated with the generalized effect size.

When Table 2 is examined, it is observed that the Cronbach's alpha value or the generalized effect size value of the GES is 0.85 with an error of 0.09. The lower limit for reliability is 0.82 at 95% confidence interval and the upper limit is 0.87. When the point estimation value of 0.85 and the lower and upper values of the confidence interval are interpreted considering the effect size classification of Cohen et al. (2011), it can be said that the Cronbach's Alpha value of the Satisfaction with Life Scale is high.

Table 2. Output on Overall Effect Size

Model	Effect Size	Standart Error of Measurement	Z	p	Lower Confidence Level	Upper Confidence Level
Random Effects	0.85	0.09	21.19	< .001	0.82	0.87

Within the scope of the research, first the heterogeneity statistics in Table 3 were given.

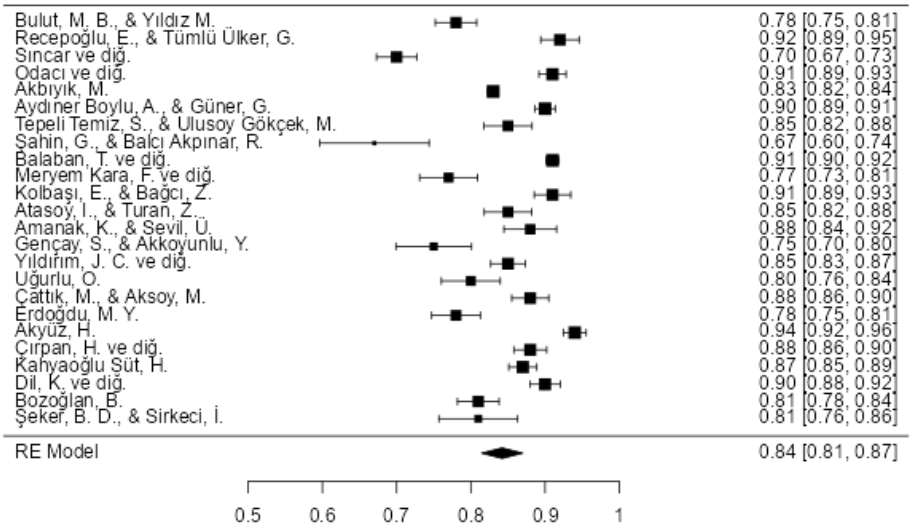
Table 3. Heterogeneity Statistics

I^2	df	Q	p
%97.22	35	668.23	<0.0001

When Table 3 is examined, Cochran's Q Test result is statistically significant (Q ~ (df = 23) ~ = 668.23, p<.0001). In other words, the change in the reliability values of the LSS is larger than expected from the sampling error. According to this result, it can be said that the actual effect size related to reliability varies according to the studies. The I^2 statistic, another statistic that provides information about heterogeneity, shows the rate of change in the observed effect size attributable to sampling error. Table 3 shows that I^2= 97.22%. I^2 provides information about the degree of inconsistency of the findings of the studies within the scope of meta-analysis and reflects the extent to which the confidence intervals obtained from different studies overlap with each other (Borenstein, 2019). The I^2 value obtained within the scope of this study is relatively large and this value means that the reliability value of the LSS varies significantly within the scope of individual studies. In addition to the Q and I^2 statistics, the lower limit value of 0.82 and the upper limit value of 0.87 in the 95% confidence interval for the mean effect size (based on standard deviation) provide information about how widely the reliability values vary between populations. Considering the statistically significant result of the Q statistic, the relatively high I^2 value and the relative width of the prediction interval, it can be said that there is a heterogeneity that needs to be examined. Within the scope of the research, moderator analyses were used to explain the sources of heterogeneity, that is, the sources of variation in the reliability values of individual studies.

Another output of the meta-analysis is the forest plot. Figure 3 shows the forest plot.

Figure 3. Forest Graph



Considering the data on the reliability values of the studies included in the study, it is seen that the reliability values of the studies vary between 0.67 (Şahin & Balcı Akpınar, 2016) and 0.94 (Akyüz, 2020). When the statistical results of the reliability values of the studies are evaluated in a holistic manner, it is noticeable that the reliability of all the studies constituting the sample is above the acceptable level. The forest plot also includes study weights. The size of the square representing each study in the forest plot shows the weight of the study. When the study weights in the forest plot are examined, it can be said that the weight of the study conducted by Akbıyık (2020) is the highest.

Results of Moderator Analyses

In the previous stage of this meta-analysis, it was concluded that examining the sources related to the change in the reliability values of individual studies was worth examining statistically. In this framework, the final aim of the study was to determine the differentiation of the reliability values of the LSS in terms of subgroups related to year of publication, sample type, percentage of female students and mean age. For this purpose, first, Analog ANOVA was performed by considering categorical variables and then meta-regression was performed by using continuous variables. The results of the Analog ANOVA are presented in Table 4.

Table 4. Analog ANOVA Results

Moderator Variable	Moderator Variable Levels	Number of Studies	ES_{mean}	ES_{lower}	ES_{upper}	df	Q_B	p
Publication year	2007-2018	16	0.86	0.82	0.88			
	2018 and beyond	8	0.83	0.77	0.88	1	0.79	**<0.0001**
Sample type	University students	10	0.83	0.77	0.87			
	Others	14	0.86	0.83	0.89	1	1.15	**<0.0001**
Percantage of women in the sample	More than half are women	11	0.84	0.80	0.86			
	More than half are men	13	0.86	0.81	0.86	1	1.15	**<0.0001**

The first row of Table 4 is the place to be examined to determine whether the change in the reliability value of the LSS according to publication year subgroups is statistically significant. The first value to be examined to determine the sources of heterogeneity is the Q value. When the moderator analysis result for the first row was examined, it was concluded that the change in the reliability value of the LSS was statistically significant ($Q=0.79$, $p<.0001$) when the categories of the publication year (Between 2007 and 2018/2018 and beyond) variable were taken into consideration. Accordingly, whether the studies were published between 2007 and 2018 or in 2018 and beyond can change the reliability of the LSS. When the average of the reliability values is analyzed, it is seen that the reliability values for the LSS used in the studies published between 2007 and2018 are higher.

The second row of Table 4 is the place to be examined to determine whether the change in the reliability value of the LSS according to the sample type subgroups is statistically significant. When the Q value as a result of the moderator analysis for the second row was analyzed, it was concluded that the change in the reliability value of the LSS was statistically significant when the categories of the sample type (university student/other) variable were taken into consideration ($Q=1.15$, $p<.0001$). According to this, the status of being a university student/not being a university student of the individuals constituting the sample can change the reliability of the LSS. When the average of the reliability values is analyzed, it is seen that the reliability values of the LSS used in the studies whose sample does not consist of university students are higher.

The last row of Table 4 is the place to be examined in order to determine whether the change in the reliability value of the LSS according to the subgroups of the percentage of women in the sample is statistically significant. When the Q value as a result of the moderator analysis for the third row was analyzed, it was concluded that the change in the reliability value of the LSS was statistically significant when the categories

of the percentage of women in the sample (more than half women/more than half men) were taken into consideration (Q=1.15, p<.0001). According to this, in the studies in which the LSS is used, the fact that more than half of the sample is female or male can change the reliability of the LSS. When the average of the reliability values is analyzed, it is seen that the reliability values of the LSS used in the studies in which more than half of the sample was male were higher.

The meta-regression results of the moderator analysis using the moderator variable of mean age are presented in Table 5.

Tablo 5. Meta-Regression Results

	Estimate	SE	Z	p	R^2	Q_E
Mean Age	0.001	0.001	1.22	0.22	0	557.27

When Table 5 is examined, the mean age variable was not found to be statistically significant according to the random effects model. In other words, the mean age variable is not a statistically significant predictor for the average reliability value of the LSS. At the same time, it is seen in Table 4 that the mean age variable does not contribute to the explained variance (R^2=0). The Q value presented for the residuals, which is one of the meta-regression outputs, was found to be statistically significant ($Q_{(22)}$= 557.27; p<.0001). This means that the assumptions of the fixed effect model are violated, that is, the actual reliability value varies even for studies with participants of the same age in their samples.

Discussion, Conclusions and Recommendations

The aim of this meta-analytic reliability generalization study was to obtain the average reliability of the LSS and to examine the moderating variables that would reveal the variability between studies. For this purpose, a meta-analysis was conducted for 24 studies that used the LSS and met the inclusion criteria. Within the scope of the research, statistical and graphical outputs related to publication bias were examined and it was determined that there was no evidence of publication bias. The average reliability coefficient for 24 studies was .84 [.81-.87]. Based on this value, it can be said that the general estimation of Cronbach's alpha is adequate (Nunnally & Bernstein, 1994). Köker (1991) found that the test-retest coefficient of the scale applied at three-week intervals was 0.85 in a study conducted within the scope of adaptation of the related measurement tool to Turkish culture. It is noticeable that the values obtained in two different studies are very close to each other. As a result of the reliability generalization of the life satisfaction scale, which was conducted by Vassar (2007) and sampled 62 studies, it was concluded that the average reliability was .78. Wallace and Wheeler (2002) conducted a reliability generalization study on the life satisfaction index. 34 studies constituted the sample and as a result of the study, it was concluded that the average reliability coefficient was .79. Considering the average reliability coefficients obtained within the scope of three different studies conducted by this study, Vassar (2007) and Wallace and Wheeler (2002), it was concluded that the three results were above the acceptable reliability value(s) (DeVellis, 1991; Cortina, 1993) and relatively close to each other.

Within the scope of the research, it was concluded that the heterogeneity between the studies was statistically significant. This shows that it would not be appropriate to generalize the reliability coefficients of the LSS as they take different values in different samples. Moderator variables that may be the source of the variation in reliability coefficients, that is, the source of heterogeneity, were analyzed. Within the scope of moderator analyses, three categorical variables (publication year, sample type, percentage of women in the sample) were included as independent variables.

As a result of the moderator analyses, it was concluded that whether the studies were published before 2018 or after 2018 or later could change the reliability of the LSS. When the average of the reliability values is analyzed, it is seen that the reliability values of the LSS used in the studies published before 2018 are higher. When the descriptive statistics of the moderator variables used in the research are examined, it is observed that there are 16 studies published before 2018 and 8 studies published in 2018 and later. The fact that the number of studies before 2018 is twice the number of studies published in 2018 and after is thought to be the main

reason for this situation. At the same time, considering the fact that psychological structures can change over time and that the LSS was developed approximately 37 years ago, it is more understandable that the reliability of the studies conducted before 2018 is high. Although the result is in favor of the studies conducted before 2018 in terms of high reliability, the fact that Cronbach's alpha values for both time periods are very close to each other ($r_{\text{before }2018} = 0.86$ and $r_{2018 \text{ and beyond}} = 0.83$) is another striking result.

Within the scope of the research, it was concluded that the reliability value changed according to the sample type. When the results are examined, it is seen that the reliability values of the LSS used in studies whose sample does not consist of university students are higher. Özdemir, Yıldırım, and Tan (2020) examined the meta-analytic reliability generalization of the short and long form of the Oxford Happiness Scale (OHS), which measures a construct similar to the construct measured by the LSS, for the Turkish sample and included 95 Cronbach Alpha coefficients obtained from 92 studies in the meta-analysis. As a result of the study, when the sample type variable was examined for the short form of the measurement tool, it was determined that the average α values for studies with and without students were .75 and .77, respectively. The result of this study is similar to the result of Özdemir, Yıldırım, and Tan (2020). In addition to this, Caruso et al. (2001), Vacha-Haase (1998), and Yin and Fan (2000) concluded that sample type affects the overall reliability estimation. It has been determined that the LSS yielded more reliable results in the sample of non-student individuals. However, as the reliability value of .83 for the student sample and the reliability value of .86 for the sample consisting of non-students are acceptable and close to each other, it can be said that the measurement tool can be used in both student and non-student samples for the Turkish sample.

As a result of the research, it was concluded that in the studies in which the LSS was used, the fact that more than half of the sample was male or female could change the reliability of the LSS. When the results are examined, it is seen that the reliability values of the LSS used in studies where more than half of the sample is male are higher. Vassar (2007) concluded that there was a low positive correlation between the percentage of women in the sample and Cronbach's alpha ($r = 0.28$; $p<.05$). Wallace and Wheeler (2002) concluded that there is no statistically significant relationship between the percentage of women in the sample and Cronbach's alpha. The result of this study is not surprising in terms of the premises of need fulfillment and livability theory. Social and institutional structures that advantage men should not be expected to lead to similar subjective welfare levels between men and women. This result in favor of men can be explained by the fact that the samples of the studies examined within the scope of the research are based on a country like Turkey, which is not advanced in terms of gender equality.

In this study, it was observed that the reliability value did not change when the mean age variable was taken into consideration. As a result of the study conducted by Wallace and Wheeler (2016), it was concluded that the reliability value of the life satisfaction index did not vary according to the mean age variable. In the study conducted by Vassar (2007), a negative relationship was found between the reliability and the samples consisting mostly of young people aged between 12 and 16, and in the same study, it was mentioned that the life satisfaction scale may need to be applied to adult populations as a result of this negative relationship. Within the scope of this study, 24 individual studies and 32 individual studies within the scope of the study conducted by Vassar (2007) were included in the meta-analysis. Within the scope of the study conducted by Wallace and Wheeler (2002), 62 studies were included in the meta-analysis. It is thought that the fact that the number of studies included in the meta-analysis within the scope of Vassar (2007) and this study is small and close to each other and that the use of mean age in the two studies and the use of age groups as moderator variables in the study conducted by Wallace and Wheeler (2002) play a role in the different results of the studies.

The disclosure of reliability values related to the study sample is critical for enhancing the validity, generalizability, and quality of the research findings (Wilkinson, 1999; Onwuegbuzie and Daniel, 2002). The notion of treating reliability as a constant attribute of the scale is termed "reliability induction" (Vacha-Haase et. al, 2000). Deditius-Island and Caruso (2002) further expounded on this by identifying two forms of reliability induction. The first form involves researchers referencing a test manual or previously published reliability coefficients and applying these statistics to their own data. In the second, more subtle form, researchers either downplay reliability or quietly assume their own scores to be reliable simply because others have found them to be so. The first method is known as "by report" reliability induction and the second as "by

omission" reliability induction (Shields & Caruso, 2004). Researchers are generally advised against using reliability induction, except in special circumstances.

Within the scope of this study, studies in which Cronbach's alpha value was reported for the reliability of the measurement results related to LSS were included in the sample. Considering some of the assumptions of Cronbach's alpha statistic that are difficult to understand, the use of alpha by academic circles as if there is "only one" reliability coefficient brings to mind the idea that alpha is a marketing concept (Cho & Kim, 2015). Even if researchers are aware of the situations where the alpha coefficient is insufficient, they may be lazy in obtaining information about other reliability coefficients. At the same time, researchers may be afraid of the penalties that may be imposed as a result of the fact that the majority of thesis committees and editors are familiar with the alpha coefficient and are not familiar with the alternatives to the alpha coefficient. It is recommended that academic journal editors and well-known researchers be advised to make frequent use of alternatives to the alpha coefficient. Therefore, it is thought that it would be useful to report reliability coefficients appropriate to the research conditions and assumptions in future studies. Within the scope of this study, the fact that only Cronbach's alpha coefficients were included in the meta-analysis can be considered as a limitation.

In this study, the Bonett (2002) transformation was used to adjust the reliability values for the reliability estimator. Future research could explore and compare reliability estimates using other transformation techniques like Fisher's z-score transformation or the Hakstian-Whalen (1976) transformation.

Further reliability generalization research could also be undertaken for other reliability estimates that scrutinize different sources of measurement error. Similar studies can be executed by pinpointing variables such as the mean and standard deviation of the measurements obtained from the assessment tool, the language used in the research, sample size, ethnicity, marital status, type of reliability, research design, etc. as different sources of variability.

Author Contributions: The authors collaborated collectively in conducting this article.

Funding Disclosure: No financial support was received for this work.

Conflicts of Interest: The authors declare no conflicts of interest.

Data Availability: The authors can provide the data upon request.

Ethical Disclosure: Due to the utilization of secondary data for the reliability generalization meta-analysis in this study, ethical approval was not required.

Studies included in the current reliability generalization meta-analysis are marked with an asterisk (*).

References

*Akbıyık, M. (2020). Üniversite öğrencilerinin yaşam doyumlarının çeşitli değişkenler bağlamında incelenmesi [Investigation of university students' life satisfaction in the context of various variables]. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi* , *51*(51) , 280-301.

*Akyüz, H. (2020). rekreasyon bölümü öğrencilerinin rekreasyon fayda ve yaşam doyum düzeyleri arasındaki ilişkinin incelenmesi [Investigation of the relationship between recreation benefits and life satisfaction levels of recreation department students]. *Turkish Studies - Social*, *15*(5) , 2323-2336.

*Amanak, K. & Sevil, Ü. (2020). Üriner inkontinansı olan ve olmayan kadınların yaşam doyumu ve sosyal kaygı düzeylerinin karşılaştırılması [Comparison of life satisfaction and social anxiety levels of women with and without urinary ıncontinence]. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi* , *10*(2) , 217-222.

Andrews, F. M., Crandall, R. (1976). The validity of measures of self-reported well-being. *Soc Indic Res* 3, 1–19.

*Atasoy, İ., & Turan, Z. (2019). Ebe ve hemşirelerde duygusal emek ile yaşam doyumu arasındaki ilişkinin

incelenmesi [Investigation of the relationship between emotional labor and life satisfaction in midwives and nurses]. *Mersin Üniversitesi Tıp Fakültesi Lokman Hekim Tıp Tarihi ve Folklorik Tıp Dergisi* , *9*(3) , 357-366.

*Aydıner Boylu, A. & Günay, G. (2018). Yaşlı bireylerde algılanan sosyal desteğin yaşam doyumu üzerine etkisi [The effect of perceived social support on life satisfaction in elderly individuals]. İnsan ve Toplum Bilimleri Araştırmaları Dergisi , *7*(2) , 1351-1363.

*Balaban, T., Saç, A. & Yıldız, Y. (2021). Rekreasyonel aktivitelere katılan bireylerin serbest zaman özgürlük düzeyleri ile yaşam doyumları ilişkisinin incelenmesi [Investigation of the relationship between leisure time freedom levels and life satisfaction of ındividuals participating in recreational activities]. *Spor ve Performans Araştırmaları Dergisi* , *12*(3) , 245-251.

*Balcı Akpınar, R. & Şahin, G. (2016). Ortopedik engelli bireylerin öfke ifade biçimleri, benlik saygıları ve yaşam doyumları [Anger expression styles, self-esteem and life satisfaction of orthopedically disabled ındividuals]. *Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi* , *19*(3) , 0-0.

Bonett, D. G. (2002). Sample size requirements for estimating intraclass correlations with desired precision. *Statistics in Medicine*, *21*(9), 1331–1335.

Borenstein, M. (2019). *Common Mistakes in Meta-Analysis and How to Avoid Them*. Biostat, Inc, Englewood, NJ.

*Bozoğlan, B. (2015). Emeklilik doyum envanterinin Türkçe formunun geçerlik ve güvenirlik çalışması [Validity and reliability study of Turkish form of retirement satisfaction ınventory]. *Türk Geriatri Dergisi* , *18*(3) , 224-230.

*Bulut, M. B. & Yıldız, M. (2020). Üniversite öğrencilerinin kişilik özelliklerinin yaşam doyumuna etkisi [The effect of personality traits of university students on life satisfaction]. *Türkiye Sosyal Araştırmalar Dergisi* , *24*(2) , 397-412.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Belmont, CA: Wadsworth

*Çattık, M & Aksoy, V. (2018). Gelişimsel yetersizliği olan çocukların ebeveynlerinin sosyal destek, öz yeterlik ve yaşam doyum düzeyleri arasındaki ilişkinin incelenmesi [Investigating the relationship between social support, self-efficacy and life satisfaction levels of parents of children with developmental disabilities]. *Eğitim ve Bilim* , *43*(195) 65-77.

Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well known but poorly understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994

*Çırpan, H., Vardarlıer, P. & Koçak, Ö. E. (2019). Kişi - iş uyumu ve yaşam doyumu ilişkisinde işte kendini yetiştirmenin aracı etkisi [The mediating effect of self-cultivation at work in the relationship between person-work fit and life satisfaction]. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* , *22*(41) , 237-252.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal Of Applied Psychology*, *78*(1), 98.

Deditius-Island, H. K., & Caruso, J. C. (2002). An examination of the reliability of scores from Zuckerman's Sensation Seeking Scales, Form V. *Educational and Psychological Measurement, 62*(4), 728 734. https://doi.org/10.1177/0013164402062004012

DeVellis, R. F. (1991). Scale Development: Theory and Applications (Applied Social Research Methods Series, Vol. 26). Newbury Park, CA: Sage Publications.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*(3), 542–575. https://doi.org/10.1037/0033-2909.95.3.542

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75.

*Dil, K., Marcus, K., Dil, S., & Short, S. (2020). Avustralya'daki Türkiyeli göçmenlerin yaşam doyumları ve aktif yurttaşlık durumlarına ilişkin sosyolojik bir değerlendirme [A sociological evaluation of life satisfaction and active citizenship status of Turkish immigrants in Australia]. *Mimar Sinan Güzel Sanatlar Üniversitesi Sosyal Bilimler Dergisi* , *2*(22) , 386-401.

*Erdoğdu, M. Y. (2020). İyimserlik (yaşam yönelimi) ile yaşam doyumu arasındaki ilişkide akademik başarının aracılık rolü [The mediating role of academic achievement in the relationship between optimism (life orientation) and life satisfaction]. *Eğitim ve Bilim* , *46*(205) , 395-407.

Field, A. P. (2003b). The problems in using fixed effects models of metaanalysis on real-world data. *Understanding Statistics*, *2*, 77 – 96.

*Gençay, S., & Akkoyunlu, Y. (2012). A Comparison of Life Satisfaction And Physical Self-Perception of Physical Education And Computer Teacher Candidates. *Nigde University Journal of Physical Education And Sport Sciences* , *6*(3) , 316-323.

Gilman, R., & Huebner, E.S. (2003) A review of life satisfaction research with children and adolescents. *Journal of School Psychology QtlaIte1y*, *2*, 107-116.

Huebner, E. S. (1991). Correlates of life satisfaction in children. *School Psychology Quarterly*, *6*(2), 103–111.

Henson, R. K. ve Thompson, B. (2002). Characterizing Measurement Error in Scores Across Studies: Some Recommendations for Conducting " Reliability Generalization" Studies. Measurement and Evaluation in Counseling and Development, Vol. 35, 113-126

Holland, D. F. (2015). Reliability generalization: a systematic review and evaluation of meta-analytic methodology and reporting practice. Doctoral dissertation, North Texas University, Texas, USA.

*Kahyaoğlu Süt, H., Küçükkaya B., Cumur E., & Özdemir E. (2019). Hemşirelik öğrencilerinin sosyal ve entelektüel aktiviteleri ile akademik başarı durumları ve yaşam doyumları arasındaki ilişki [The relationship between nursing students' social and ıntellectual activities, academic achievement status and life satisfaction]. *Yükseköğretim ve Bilim Dergisi/Journal of Higher Education and Science* , *9*(2) , 326-334.

*Kara, F. M., Gürbüz, B., Küçük Kılıç, S., & Öncü, E. (2018). Beden eğitimi öğretmenleri adaylarının serbest zaman sıkılma algısı, yaşam doyumu ve sosyal bağlılık düzeylerinin incelenmesi [Investigation of physical education teacher candidates' perception of leisure time boredom, life satisfaction and social commitment levels]. *Journal of Computer and Education Research* , *6*(12) , 342-357.

*Kolbaşı, E. & Bağcı, Z. (2019). İş doyumu ve yaşam doyumu ilişkisi: bir kamu kurumu üzerinde analitik bir araştırma [Relationship between job satisfaction and life satisfaction: An analytical research on a public ınstitution]. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi* , *7*(4) , 119-123.

Köker, S. (1991). Normal ve sorunlu ergenlerin yaşam doyumu düzeylerinin karşılaştırılması [Comparison of life satisfaction levels of normal and troubled adolescents]. Yayınlanmamış yüksek lisans tezi. Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Köse, M. F, Çobanoğlu, G., & Mercan Sarı, R. (2022). Yaşam memnuniyeti ölçeğinin geliştirilmesi ve psikometrik özellikleri [Life satisfaction scale and psychometric properties]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, *55,* 324-346 .

Lavallee, L. F., Hatch, P. M., Michalos, A. C., & McKinley, T. (2007). Development of the contentment with life assessment scale (CLAS): Using daily life experiences to verify levels of self-reported life satisfaction. *Social Indicators Research*, *83*(2), 201–244

Lightsey, O. R., Jr., McGhee, R., Ervin, A., Gharghani, G. G., Rarey, E. B., Daigle, R. P., Wright, K. F., Constantin, D., & Powell, K. (2013). Self-efficacy for affect regulation as a predictor of future life satisfaction and moderator of the negative affect—Life satisfaction relationship. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being*, *14*(1), 1–18

Neugarten, B. L., Havighurst, R. J., & Tobin, S. S. (1961). The measurement of life satisfaction. *Journal of Gerontology, 16,* 134–143. https://doi.org/10.1093/geronj/16.2.134

Onwuegbuzie, A. J., & Daniel, L. G. (2002). A framework for reporting and interpreting internal consistency reliability estimates. *Measurement and Evaluation in Counseling and Development, 35*(2), 89–103.

Onwuegbuzie, A. J., & Daniel, L. G. (2000, November 17). Reliability Generalization: The Importance of Considering Sample Specificity, Confident Intervals, and Subgroup Differences. Paper presented at the Annual meeting of the Mid-South Educational Research Association. Bowling Green, KY.

*Odacı, H., Kaya, F., & Kınık, Ö. (2021). Akademik yaşam doyumu ölçeği'nin Türkçe versiyonu [Turkish version of the satisfaction with academic life scale]. *Ege Eğitim Dergisi* , 22 (1), 1-15.

Özdemir, V., Yıldırım, Y., & Tan, Ş. (2020). A meta-analytic reliability generalization study of the Oxford Happiness Scale in Turkish sample. *Journal of Measurement and Evaluation in Education and Psychology*, *11*(4), 374-404. doi: 10.21031/epod.766266.

Pavot, W. G., Diener, E., Colvin, C. R., & Sandvik, E. (1991). Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment*, *57*(1), 149–161.

Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, *5*(2), 164–172

*Recepoğlu, E., & Ülker Tümlü, G. (2015). Üniversite Akademik Personelinin Mesleki ve Yaşam Doyumları Arasındaki İlişkinin İncelenmesi [Investigation of the Relationship Between Professional and Life Satisfaction of University Academic Staff]. *Kastomonu Eğitim Dergisi* , *23*(4) , 1851-1868.

Sánchez-Meca, J., López-López, J. A., & López-Pina, J. A. (2013). Some recommended statistical analytic practices when reliability generalization studies are conducted. *British Journal of Mathematical and Statistical Psychology*, *66*(3), 402-425

Shields, A. L., & Caruso, J. C. (2004). A reliability ınduction and reliability generalization study of the cage questionnaire. *Educational and Psychological Measurement*, *64*(2), 254–270. https://doi.org/10.1177/0013164403261814

*Sincar, S., Sohbet, R., Birimoglu Okuyan, C. & Karasu, F. (2020). Üniversite Öğrencilerinde Yaşam Doyumu ve Depresyon: Kesitsel Bir Çalışma [Satisfaction with life and depression in university students: A cross-sectional study]. *Balıkesir Sağlık Bilimleri Dergisi* , *9*(2) , 93-100.

*Şeker, B. D., & Sirkeci, İ. (2014). Birleşik Krallık'daki Türkiye kökenli kadınlarda yaşam doyumu: Kimlik, kültürleşme ve ayrımcılık [Life satisfaction among women of Turkish origin in the United Kingdom: Identity, acculturation and discrimination]. *Türk Psikoloji Yazıları* , *17*(34) , 69-81.

Taylor, R. T. (2012). Review of the Motivated Strategies for Learning Questionnaire (MSLQ) using reliability generalization techniques to assess scale reliability [Unpublished doctoral dissertation]. Auburn University.

*Tepeli Temiz, S., & Ulusoy Gökçek, V. (2020). Yurtta kalan üniversite öğrencilerinin anksiyete, sosyal destek ve yaşam doyum düzeyleri ile baş etme stillerinin incelenmesi [Investigation of anxiety, social support and life satisfaction levels and coping styles of university students living in dormitory]. *Fatih Sultan Mehmet İlmi Araştırmalar İnsan ve Toplum Bilimleri Dergisi* , *15*, 431-458.

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, *60*(2), 174–195.

Traub, R. R., & Rowley, G. L. (1991). Understanding reliability. *Educational Measurement:Issues and Practice*, *10*(1), 37-45.

*Uğurlu, O. (2013). Pozitif ve Negatif duyguların iyimserlik-karamsarlık ve yaşam doyumu arasındaki ilişkide ara bulucu etkisi [Mediating effect of positive and negative emotions on the relationship between optimism-pessimism and life satisfaction]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* , *28*(2) ,

497-504.

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement*, *58*(1), 6–20.

Vacha-Haase, T., Henson, R. K., & Caruso, J. C. (2002). Reliability generalization: Moving toward improved understanding and use of score reliability. *Educational and Psychological Measurement*, *62*(4), 562–569.

Vacha-Haase, T., Kogan, L. R., & Thompson, B. (2000). Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. *Educational and Psychological Measurement*, *60*(4), 509–522. https://doi.org/10.1177/00131640021970682

Vassar, M. (2007). A note on the score reliability for the satisfaction with life scale: An RG study. *Social Indicators Research, 86*(1), 47–57. https://doi.org/10.1007/s11205-007-9113-7

*Yıldırım, J. C., Kozak, E. D., & Türkarslan, K. K. (2021). Utandıran diğeri ölçeği-2'nin Türkçe Formunun psikometrik özellikleri [Psychometric properties of the Turkish version of the embarrassing other scale-2]. *Klinik Psikoloji Dergisi* , *5*(1) , 13-26.

Wallace, K. A., & Wheeler, A. J. (2002). Reliability generalization of the Life Satisfaction Index. *Educational and Psychological Measurement, 62*(4), 674–684. https://doi.org/10.1177/0013164402062004009

Wilkinson, L. (1999). The task force on statistical ınference. *Statistical Methods in Psychology Journals: Guidelines and explanations. American Psychologist*, *54*(8), 594-604. https://doi.org/10.1037/0003-066X.54.8.594