Psychometric Properties of the Short-Form State Self-Esteem Scale (SSES-S) Among Turkish Adolescents: Reliability, Validity, and Multi-Group IRT Analysis

M. Furkan Kurnaz[1] and **Nilüfer Koçtürk**[2]

Abstract

This study examines the psychometric properties of the Short-Form of the State Self-Esteem Scale (SSES-S) among Turkish adolescents. A total of 422 adolescents (60.64% girls) aged 14–18 years (*M* = 15.78, *SD* = 1.01) participated across three studies, which included exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and additional validation analyses. Findings support the reliability and validity of the SSES-S scores, validating its three-factor structure—academic performance, physical appearance, and social success—with acceptable internal consistency and measurement invariance across sex. Multi-group item response theory (MGIRT) results indicated variability in item discrimination and difficulty across sex. For example, Item 11 demonstrated high discrimination (3.69), while Item 8 showed low discrimination (0.56), particularly among boys. Additionally, boys had lower difficulty thresholds for academic performance items, whereas girls showed lower thresholds for social success and physical appearance. Differential item functioning (DIF) was identified in two items–one at a moderate level and one at a large level–suggesting potential bias and the need for item-level refinement. While the SSES-S shows generally acceptable psychometric properties, these item-level findings warrant caution when making comparisons across sex. Continued research is recommended to address identified limitations and explore sociocultural influences on item functioning.

Keywords

state self-esteem, adolescents, multi-group item response theory, differential item functioning, measurement invariance

[1]Department of Guidance and Psychological Counseling, Necmettin Erbakan University, Konya, Türkiye
[2]Department of Guidance and Psychological Counseling, Hacettepe University, Ankara, Türkiye

Corresponding Author:
M. Furkan Kurnaz, Department of Guidance and Psychological Counseling, Necmettin Erbakan University, 10 University Street, Konya, Türkiye.
Email: furkan.kurnaz.mfk@gmail.com

Introduction

Adolescence is a critical developmental period characterized by rapid psychological, emotional, and social changes (Livingstone & Isaacowitz, 2018). During this time, individuals form increasingly complex self-concepts and become highly sensitive to feedback from peers, parents, and individuals within academic environments. Among the many psychological constructs relevant to adolescent development, self-esteem—one's evaluative perception of self-worth—has been shown to play a central role in emotional well-being, social functioning, and academic adjustment (Orth & Robins, 2014; Tsang & Yip, 2006). Low self-esteem in adolescence is linked to a range of adverse outcomes, including depression, anxiety, and poor interpersonal relationships (Harring, 2024; Kurnaz et al., 2020). Given the importance of self-esteem for adolescent mental health and development, it is essential to use psychometrically sound tools to assess it. This study therefore aims to evaluate the psychometric properties of the Short-Form State Self-Esteem Scale (SSES-S) among Turkish adolescents.

Self-Esteem and Adolescent Development

Self-esteem is a crucial psychological construct that begins forming in early childhood and continues to evolve through adolescence into adulthood, shaped by a range of social and environmental influences (Burger & Mortimer, 2024). It is one of the most widely studied self-referential variables in adolescent development (Supervía et al., 2022). Adolescents with higher self-esteem tend to exhibit better psychological functioning (Rey et al., 2011) and experience greater mental well-being later in life (Boden et al., 2008). In contrast, low self-esteem has been associated with a wide range of negative outcomes, including poor academic attitudes (Barragán Martín et al., 2021), loneliness (Lyyra et al., 2021), depressive symptoms (Gu et al., 2024), emotional and behavioral difficulties (Pandey et al., 2025), eating disorders such as loss-of-control eating (Beckers et al., 2025), and anxiety (Han et al., 2025).

Self-esteem is a multifaceted construct with significant implications in personality, social, and clinical psychology. It encompasses beliefs and emotions such as pride, shame, achievement, and hopelessness, reflecting an individual's overall self-worth and personal development (Giri, 2020; Hepper, 2016). According to sociometer theory, self-esteem functions as a monitor of social acceptance, suggesting that it regulates behavior in response to social inclusion or exclusion cues (Bale & Archer, 2013; Wu et al., 2017). Given that adolescents are highly sensitive to peer evaluation, state self-esteem may play a particularly important role in their social and emotional well-being (Sakız et al., 2021; Steinsbekk et al., 2020).

Trait vs. State Self-Esteem

Self-esteem is typically conceptualized as either a trait or a state (McCain et al., 2015). Trait self-esteem refers to an individual's overall and enduring evaluation of their self-worth, which tends to remain relatively stable over time (Hank & Baltes-Götz, 2019; Salice, 2020; Westfall, 2020). In contrast, state self-esteem refers to momentary fluctuations in self-worth that occur in response to immediate experiences or situational factors (Brito et al., 2023; McCain et al., 2015). These two forms are conceptually distinct yet theoretically related: individuals with high trait self-esteem may experience smaller or more regulated fluctuations in state self-esteem, whereas those with low trait self-esteem may exhibit more volatility in their self-worth across contexts (DeHart & Pelham, 2007; Hank & Baltes-Götz, 2019). Adolescents are particularly vulnerable to these fluctuations, as their self-perceptions are shaped by ongoing changes in academic performance, peer feedback, and social media exposure.

Understanding state self-esteem is therefore crucial in adolescence, as it captures the dynamic and situational aspects of self-worth that may not be reflected in trait-level assessments. Given the developmental sensitivity of this period, research focusing on state self-esteem provides valuable insights into adolescents' daily psychological experiences and the contextual factors that influence them (Maldonado et al., 2013; Mertens et al., 2022).

The Short-Form State Self-Esteem Scale (SSES-S)

The Short-Form State Self-Esteem Scale (SSES-S) reflects a sociometer-based conceptualization of self-esteem, which views self-worth as a dynamic, situation-sensitive gauge of perceived social inclusion and evaluation (Leary & Baumeister, 2000). Unlike trait-oriented measures such as the Rosenberg Self-Esteem Scale (RSE), which assess global and relatively stable self-evaluations, the SSES-S is designed to capture moment-to-moment fluctuations in state self-esteem across three domains: academic performance, physical appearance, and social success (Brito et al., 2023). Each domain corresponds to a core developmental task in adolescence: performing in school, managing body image concerns, and navigating peer relationships.

This multidimensional and reactive structure distinguishes the SSES-S from traditional uni-dimensional trait-based scales by offering greater ecological validity—particularly in adolescence, a period marked by identity exploration and social sensitivity. The SSES-S is especially relevant for use in Turkey, where academic achievement is emphasized socially and institutionally, beauty norms are influenced by both traditional values and modern media, and collectivist social expectations shape interpersonal dynamics. These cultural factors may amplify domain-specific self-esteem fluctuations, underscoring the need for a flexible, context responsive tool such as the SSES-S in both research and practice.

Sex and State Self-Esteem

Prior studies have reported sex-based differences in state self-esteem (Aftab & Aslam, 2014; Vitolo, 2022). Girls often report lower self-esteem in appearance and social domains, whereas boys may be more affected by challenges in academic or performance-related settings. These differences may be explained by sex socialization theory, which posits that cultural norms and expectations influence how boys and girls value different aspects of the self (Eccles, 1994). In the Turkish context, girls may experience greater pressure to conform to appearance ideals and relational expectations, while boys may be expected to demonstrate competence and success. These culturally shaped values can result in domain-specific vulnerabilities and fluctuations in state self-esteem, reinforcing the need to assess the scale's measurement invariance across sex.

Current Study

The SSES was developed by Heatherton and Polivy (1991) to assess domain-specific fluctuations in state self-esteem and originally included three factors: social, performance, and appearance. Brito et al. (2023) later developed the SSES-S, a 12-item version that retained the original factor structure but adapted the dimensions to academic performance, physical appearance, and social success. The SSES-S scores demonstrate good reliability and validity across multiple samples, including adolescents and university students in both Western (Escrivá-Martínez et al., 2025) and non-Western (e.g., Ugwu et al., 2023) contexts.

In Turkey, research on self-esteem has largely relied on trait-based measures, particularly the RSE (Rosenberg, 1965), which was adapted by Çuhadaroğlu (1986). While useful, trait measures may not capture the moment-to-moment fluctuations in self-esteem that adolescents often

experience in response to social, academic, and appearance-related stressors. This is especially relevant in the Turkish context, where strong academic expectations, deeply rooted sex norms, and collectivist cultural values may shape how adolescents perceive themselves in different domains.

Although a Turkish adaptation of the SSES-S has been developed for adults (Kurnaz & Koçtürk, 2025), no version has yet been validated for adolescents— a critical oversight given adolescence is a developmental period marked by high emotional reactivity and identity formation. Understanding how state self-esteem operates within this population and cultural setting can support more accurate assessment and inform culturally responsive interventions.

Given this gap, the primary aim of this study was to examine the psychometric properties of the SSES-S among Turkish adolescents. This included an evaluation of its factor structure, reliability, and validity, along with an investigation of measurement invariance across sex to ensure that the scale functions equivalently for boys and girls. Furthermore, a multi-group item response theory (IRT) analysis was conducted to examine the discriminatory power of the SSES-S items and to explore how item parameters differ by sex. IRT offers a more precise and item-level analysis than classical test theory, providing deeper insights into the scale's functioning across subgroups. By developing and validating a culturally sensitive, psychometrically robust tool for assessing state self-esteem in Turkish adolescents, this study aims to enhance both research and applied practice in adolescent mental health.

Study 1: Exploratory Factor Analysis of SSES-S

Method

Participants. The study initially recruited 212 adolescents; however, two participants were excluded due to 100% missing data on the SSES-S items. The final sample consisted of 210 adolescents (58.57% girls, $n = 123$), aged 14–18 years ($M = 15.67$, $SD = 1.07$). Participants were in grades 9 through 12 ($M = 10.43$, $SD = 1.05$); specifically, 24.29% ($n = 51$) were in 9th grade, 26.19% ($n = 55$) in 10th grade, 31.43% ($n = 66$) in 11th grade, and 18.10% ($n = 38$) in 12th grade. All participants were of Turkish ethnicity and fluent in Turkish, which was their primary language. Regarding socioeconomic status (SES), 92.38% ($n = 194$) of the participants were from a middle SES, 4.76% ($n = 10$) were from a high SES, and 2.86% ($n = 6$) were from a low SES. A sample size of 210 was deemed adequate for conducting exploratory factor analysis (EFA) of the 12-item SSES-S. While traditional guidelines often suggest 5 to 10 participants per item (Fabrigar et al., 1999), more recent and robust recommendations emphasize the importance of absolute sample size and factor-to-item ratios. For instance, Comrey and Lee (1992) classified a sample of 200 as "fair" and 300 as "good" for factor analysis. Additionally, Costello and Osborne (2005) recommended a minimum of 100 participants per factor, which would require at least 300 participants for a three-factor model. Although our EFA sample falls slightly below that stricter threshold, it meets the commonly accepted standards for item-to-participant ratios and exceeds the minimum of 200 often cited as sufficient for stable factor extraction (MacCallum et al., 1999). Therefore, the sample size used in this phase is considered adequate for the intended analysis.

Instruments

Short-Form of State Self-Esteem Scale (SSES-S). The SSES-S, originally developed as a short form of the SSES (Heatherton & Polivy, 1991) by Brito et al. (2023), was adapted into Turkish by Kurnaz and Koçtürk (2025). During the adaptation process, Kurnaz and Koçtürk (2025) conducted a backward translation procedure and implemented a two-phase adaptation process with Turkish adult participants. The Turkish version of the SSES-S demonstrated measurement invariance across sex in adult samples. The scale consists of 12 items across three factors–academic

performance, physical appearance, and social success–with four items per factor. The original version demonstrated high internal consistency in adult samples (ω = .87; α = .80; λ6 = .89). Additionally, confirmatory factor analysis (CFA) indicated that the goodness-of-fit indices were within acceptable ranges in adult populations χ^2 (32) = 81.58 p <. 001; CFI = .96; GFI = .93; TLI = .94; RMSEA = .08, (90% CI: .07 to .11); SRMR = .08. Because the current study examined the scale's psychometric properties in adolescents, the internal consistency of the SSES-S in this sample was assessed separately. The Cronbach's α, McDonald's ω, and Guttman's λ2 values were computed to ensure scale reliability in this population (see Results section).

Procedures. Ethical approval for this study was obtained from the Necmettin Erbakan University Social and Humanities Scientific Research Ethics Committee (Report No: 23066/33). Data were collected from adolescents through high school counselors using a paper-and-pencil adminis-tration of the Short-Form State Self-Esteem Scale (SSES-S). The school counselors, who received standardized training in the administration procedures, facilitated the data collection process by obtaining written parental consent and informed assent from the adolescents. Participants and their parents were informed that counseling services would be available if needed through the counseling center affiliated with the institution where the first author is employed.

Results

Preliminary Analyses

To handle missing data, Little's (1988) Missing Completely at Random (MCAR) test was conducted using SPSS (version 27). The results indicated that data were missing completely at random (χ^2 = 58.08, df = 59, p = .51). As the percentage of missing data was low (<5%), missing values were imputed using the Expectation-Maximization (EM) algorithm. The assumption of normality was assessed by examining skewness and kurtosis values using JASP (version 0.19.3). All skewness values were below 3, and kurtosis values were below 8, confirming that the data followed a normal distribution (Kline, 2016).

Exploratory Factor Analysis (EFA)

Principal axis factoring (PAF) was used to examine the factor structure of the SSES-S for ad-olescents. EFA was conducted using JASP (version 0.19.3). Firstly, PAF was conducted without rotation, and Bartlett's sphericity test and Kaiser-Meyer-Olkin (KMO) test were performed to determine whether the data were suitable for EFA. Bartlett's sphericity test was significant (χ^2 = 53.11; df = 33, p = .015), indicating that the correlation matrix was factorable. Additionally, the overall KMO value was .813, with individual item values ranging from .742 to .943, suggesting that the sample was well-suited for factor analysis (Hair et al., 2009; Tabachnick & Fidell, 2014).

PAF was used to extract latent factors, as it is preferable for potential correlated psychological constructs (Costello & Osborne, 2005). The PAF results suggested a three-factor structure, with eigenvalues greater than 1 (Netemeyer et al., 2003). To confirm this structure, a parallel analysis was conducted. The first three factors had real eigenvalues exceeding those from randomly generated datasets.: Factor 1 — real data = 3.78, simulated = 0.64; Factor 2 — real data = 1.14, simulated = 0.35; and Factor 3 — real data = 0.37, simulated = 0.27. However, the fourth factor's real eigenvalue (0.11) was lower than simulated eigenvalue (0.17), indicating that only the first three factors should be retained. These results were further supported by the scree plot (Figure 1), which showed a distinct inflection point after the third factor. Because the factors were expected to

be correlated, oblimin rotation (an oblique rotation method) was applied (Costello & Osborne, 2005).

EFA results indicated that items 4 and 8 did not have factor loadings above .40 (Hair et al., 2009; see Table 1). However, these items were retained in subsequent analyses due to their conceptual relevance to the original scale domains and to preserve the structural integrity of the SSES-S. The three-factor, 12-item structure of SSES-S explained approximately 50% of the total variance. The eigenvalues were as follows: social success = 4.43 (33.3% variance), physical appearance = 1.79 (11.7% variance), and academic performance = 1.09 (5.1% variance). Interfactor correlations showed that social success was correlated with physical appearance ($r = .34$) and academic performance ($r = .36$). Additionally, physical appearance was correlated with academic performance ($r = .55$). Moreover, the internal consistency values for each scale were as follows: academic performance ($\alpha = .72$; $\omega = .72$; $\lambda2 = .72$), physical appearance ($\alpha = .72$; $\omega = .74$; $\lambda2 = .73$), and social success ($\alpha = .84$; $\omega = .84$; $\lambda2 = .84$). These reliability coefficients demonstrate strong internal consistency for the scales in an adolescent sample, supporting the suitability of the SSES-S for this age group.

Study 2: Confirmatory Factor Analysis of SSES-S

Method

Participants. The initial sample consisted of 220 adolescents; however, eight participants were excluded due to 100% missing data, resulting in a final sample of 212 adolescents (62.74% girls, $n = 133$), aged 14 to 18 years ($M = 15.87$, $SD = 0.94$). Participants were in grades 9 through 12 ($M = 10.34$, $SD = 0.99$), with 22.64% ($n = 48$) were in 9th grade, 35.38% ($n = 75$) in 10th grade, 27.36% ($n = 58$) in 11th grade, and 14.62% ($n = 31$) in 12th grade. All participants were of Turkish ethnicity and fluent in Turkish, which was their primary language. In terms of socioeconomic status (SES), the majority of participants (83.96%, $n = 178$) were from a middle-income background, while 11.79% ($n = 25$) were from a high-income background, and 2.86% ($n = 6$) were from a low-income background. To ensure an adequate sample size for CFA, we followed recommendations suggesting a minimum of 5–10 participants per item (Kline, 2016). Given that the SSES-S has 12 items, the sample size of 212 participants met these criteria.
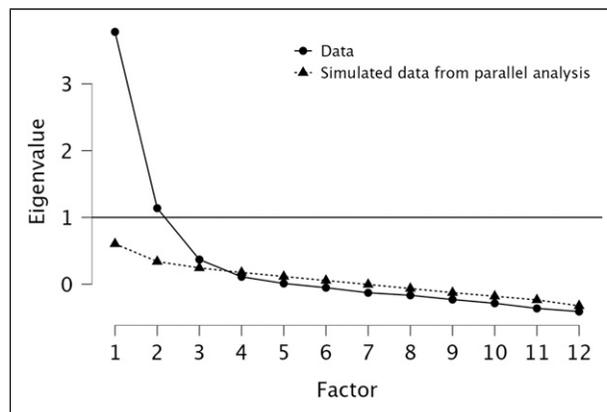

Figure 1. Parallel Analysis

Table 1. EFA and CFA Results for the SSES-S

	EFA			
Items	Factor1	Factor2	Factor3	CFA
Factor 1. Academic performance				
3. I feel confident that I understand things	.744			.666
2. I feel as smart as others	.708			.760
1. I feel confident about my abilities	.473			.673
4. I feel like I am not doing well	.209			.422
Factor 2. Physical appearance				
5. I feel satisfied with the way my body looks right now		.871		.918
7. I am pleased with my appearance right now		.828		.790
6. I am dissatisfied with my weight		.442		.307
8. I feel unattractive		.298		.443
Factor 3. Social success				
11. I feel concerned about the impression I am making			.919	.856
12. I am worried about looking foolish			.766	.692
10. I am worried about what other people think of me			.689	.715
9. I feel self-conscious			.603	.555
% variance explained (eigenvalue)	5.1% (1.09)	11.7% (1.79)	33.3% (4.43)	
Reliability indices				
McDonald's ω	.724	.739	.842	.819
Cronbach's α	.717	.718	.837	.821
Guttman's λ2	.721	.731	.840	.831

Note. EFA loadings are post-rotation values using principal axis factoring with oblimin rotation. CFA loadings are standardized. Percentages represent variance explained by each factor after rotation. Reliability indices were computed separately for each scale.

Instruments

Rosenberg Self-Esteem Scale (RSE). The RSE (Rosenberg, 1965) was adapted into Turkish by Çuhadaroğlu (1986). The scale consists of 10 items (5 positive, 5 negative), with lower scores indicating higher self-esteem. The original Turkish adaptation reported acceptable internal consistency ($\alpha = .74$). In the present study, the RSE demonstrated strong internal consistency ($\alpha = .88$; $\omega = .88$; $\lambda2 = .88$).

Procedures. To validate the factor structure of the SSES-S identified in Study 1, a CFA was conducted with a new sample of high school students. Data collection followed the same procedures as in Study 1, using the paper-and-pencil method through the school counselor. Parental consent and informed assent were obtained prior to participation. Adolescents and their parents were informed that counseling services were available through the counseling center affiliated with the institution where the first author works.

Results

Preliminary Analyses

To assess the impact of missing data, Little's (1988) MCAR test was conducted. The results indicated that missing data were completely at random ($\chi^2 = 84.31$, $df = 79$, $p = .32$) and were subsequently imputed using the EM algorithm. Skewness and kurtosis values were examined to

assess normality. All skewness values were below 3, and kurtosis values were below 8, indicating that the data followed a normal distribution.

Confirmatory Factor Analysis (CFA)

A CFA was conducted using JASP (version 0.19.3) with robust maximum likelihood estimation method to evaluate three-factor structure (academic performance, physical appearance, and social success) identified in Study 1. Model fit was assessed using widely accepted criteria: CFI and TLI values ≥.90 indicate acceptable fit, RMSEA values ≤.08 reflect reasonable error of approximation, and SRMR values ≤.08 are considered acceptable (Hu & Bentler, 1999; Kline, 2016). The results indicated an acceptable model fit for Turkish adolescents χ^2 (51) = 95.93; p < .001; χ^2/df = 1.88; CFI = .943; GFI = .931; TLI = .927; RMSEA = .064, (90% CI: .044 to .084); SRMR = .066. Factor loadings for all items were above .40, except for Item 6, which had a low factor loading (see Figure 2). Despite this, the item was retained due to theoretical significance. Future research may explore whether modifications to Item 6 could improve its alignment with the underlying construct.

Internal Consistency and Evidence Based on Relations to Other Variables

The internal consistency of the SSES-S′ factors was evaluated using Cronbach's α, McDonald's ω, and Guttman λ2. The reliability estimates were acceptable for all three factors: academic performance (α = .712, ω = .728, λ2 = .722), physical appearance (α = .688, ω = .697, λ2 = .699), and social success (α = .793, ω = .798, λ2 = .796). Although some coefficients were in the high .60s, such values are considered acceptable in the context of cross-cultural or group-specific validation studies, particularly within psychological and educational research (Nunnally & Bernstein, 1994).

To provide evidence based on relations to other variables, SSES-S factor scores were correlated with the RSE scores. These bivariate correlation analyses were conducted using JASP (version 0.19.3). The correlations were statistically significant and negative, consistent with theoretical expectations that momentary fluctuations in self-esteem (state self-esteem) may inversely relate to
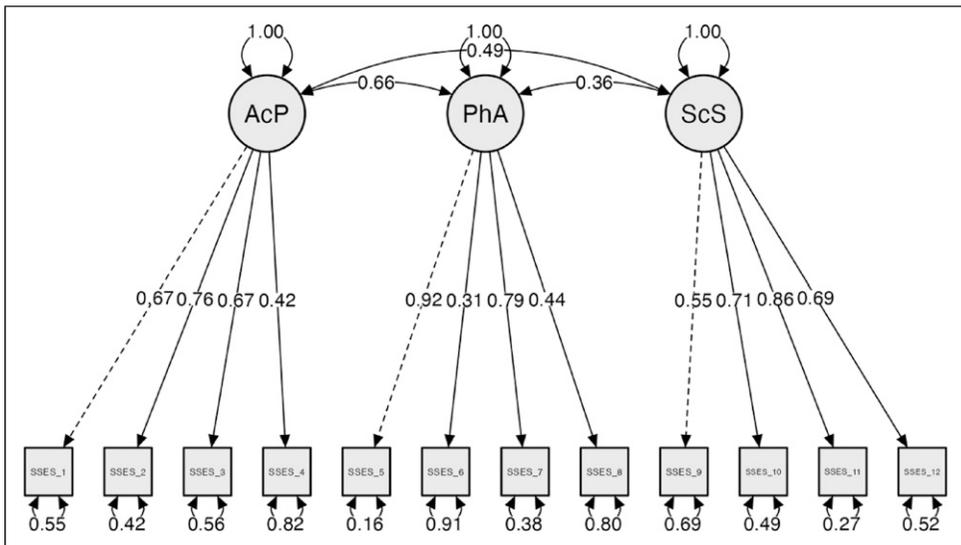
Figure 2. CFA Results of SSES-S

more stable self-appraisals (Heatherton & Polivy, 1991): academic performance ($r = -.55$, $p <$ .001), physical appearance ($r = -.52$, $p < .001$), and social success ($r = -.61$, $p < .001$). According to Cohen's (1988) guidelines, correlations of approximately .10 are considered small, around .30 are moderate, and .50 or higher are large in magnitude. Thus, the observed correlations represent large effect sizes, providing strong validity evidence for the SSES-S scores. The SSES-S is conceptualized as a three-factor structure, consistent with the original model proposed by Brito et al. (2023).

Study 3: Multi-Group CFA, Item Response Theory, and Differential Item Functioning Analysis of the SSES-S

Method

Participants. The final sample consisted of 422 adolescents, of whom 256 (60.64%) were girls, with ages ranging from 14 to 18 years ($M = 15.78$, $SD = 1.01$). In terms of socioeconomic status, the majority of participants (88.15%, $n = 372$) were from a middle-income background, while 8.29% ($n = 35$) were from a high-income background, and 3.32% ($n = 14$) were from a low-income background.

Procedures. To further evaluate the psychometric properties of the SSES-S, data from Study 1 and Study 2 were combined, resulting in a total sample of 422 adolescents. Three types of analyses were conducted to provide a comprehensive assessment of the scale. First, a multi-group CFA was performed to examine whether the factor structure of the SSES-S was invariant across sex. This analysis was conducted using JASP (version 0.19.3). Second, a multi-group IRT analysis was conducted to estimate the discrimination and difficulty parameters of each SSES-S item using the graded response model (GRM). This analysis was performed in R (version 4.4.2; R Core Team, 2024) using the mirt package (Chalmers, 2012) and the psych package (Revelle, 2024). Lastly, a differential item functioning (DIF) analysis was conducted using the lordif package (Choi et al., 2011), applying both the Mantel–Haenszel (MH) and logistic regression (LR) methods to evaluate potential item bias across sex.

Results

Multi-Group Confirmatory Factor Analysis (MG-CFA)

A multi-group CFA was conducted using JASP (version 0.19.3) to examine measurement invariance of the SSES-S across sex. This approach, rooted in CFA (Jöreskog, 1969), allows researchers to determine whether a scale exhibits consistent response patterns across different groups (Milfont & Fischer, 2010; Putnick & Bornstein, 2016). Establishing measurement invariance is critical in psychological research to ensure that self-esteem is interpreted in the same way across sex, allowing for valid comparisons. The findings revealed that Item 6, which states, "I am dissatisfied with my weight," had factor loadings above .40 both girls and boys (see Table 2). Model fit statistics supported measurement invariance across sex, with results indicating configural invariance, as the baseline model fit indices demonstrated acceptable values χ^2 (132) = 310.26; $p < .001$; $\chi^2/df = 2.35$; CFI = .906; RMSEA = .087; SRMR = .075 (see Table 3). While the RMSEA slightly exceeded the commonly accepted cutoff of .08, the overall model fit was considered acceptable given that CFI and SRMR met recommended thresholds (Hu & Bentler, 1999), and RMSEA values can be influenced by model complexity and moderate sample sizes (Kenny et al., 2015; Kline, 2016). Metric invariance was supported, as changes in CFI, RMSEA,

Table 2. Standardized SSES-S Factor Loadings

SSES-S items	Girl	Boy
Academic performance		
1	.716	.584
2	.748	.767
3	.662	.659
4	.447	.387
Physical appearance		
5	.921	.903
6	.306	.293
7	.791	.793
8	.432	.467
Social success		
9	.517	.582
10	.734	.691
11	.805	.923
12	.698	.706

and SRMR were within the established thresholds (ΔCFI <.010, ΔRMSEA <.015, and ΔSRMR <.015; Chen, 2007). Scalar invariance was also achieved, meaning that boys and girls interpreted the SSES-S items equivalently, allowing for valid mean comparisons. In addition, strict invariance was supported, as model fit indices remained within acceptable change thresholds after constraining item residual variances across groups. This suggests that the SSES-S demonstrates strong measurement stability across sex, even at the most restrictive level of invariance.

Multi-Group Item Response Theory (MGIRT)

To further investigate the psychometric properties of the SSES-S items, a multiple-group item response theory (MGIRT) analysis was conducted using the graded response model (GRM). The discrimination and difficulty parameters of each item were analyzed across sex (see Table 4 and Figure 3). The results showed that discrimination parameter estimates varied across items, with the highest value observed for Item 11 ("I feel concerned about the impression I am making"), which had a discrimination estimate of 3.69, indicating strong differentiation between individuals with different levels of self-esteem. In contrast, the lowest discrimination parameter was found for Item 8 ("I feel unattractive"), which had a value of 0.56 for both boys and girls. According to standard guidelines (Baker, 2001), discrimination values below 0.65 indicate very low discrimination, suggesting that Item 8 may not effectively distinguish between adolescents with varying self-esteem levels. Additionally, Item 6 ("I am dissatisfied with my weight") showed low discrimination among girls (0.62), whereas Item 8 had the lowest discrimination parameter among boys (0.56). These results suggest that some items may require modification to improve their ability to measure self-esteem differences accurately.

The difficulty parameters (b1–b4) of the SSES-S items were also examined. These parameters represent the level of self-esteem required to have a 50% probability of selecting a higher response category. They reflect the relative difficulty of endorsing each successive category and are essential for understanding item functioning in polytomous IRT models (Embretson & Reise, 2013; Samejima, 1969). The results indicated sex differences in item difficulty. For example, items related to academic performance (Items 1, 2, and 3) had lower difficulty parameters for boys,

Table 3. Measurement Invariance Results from Multigroup CFA by Sex

Measurement invariance	CFI	ΔCFI	RMSEA	ΔRMSEA	RMSEA [90% CI]	SRMR	ΔSRMR
Sex: Girl versus boy							
Configural	.906		.087		.074 – .100	.075	
Metric	.906	<.0001	.083	.004	.071 – .096	.078	.003
Scalar	.903	.003	.081	.002	.069 – .093	.074	.004
Strict	.897	.006	.080	.001	.069 – .092	.079	.005

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual. CI = confidence interval. All RMSEA confidence intervals are 90%.

suggesting that boys required a lower level of self-esteem in this domain to endorse higher response categories than girls. In contrast, items related to social success and physical appearance (Items 4, 10, and 11) had lower b1 values among girls compared to boys, whereas items related to physical dissatisfaction (Items 6 and 12) had lower b4 values among boys. These findings indicate that the perception of self-esteem domains varies between sex, potentially influencing the way adolescents respond to self-esteem items.

Differential Item Functioning (DIF)

To assess potential bias in individual items, a differential item functioning (DIF) analysis was conducted using the Mantel-Haenszel (MH) and logistic regression (LR) methods. The results revealed that two items exhibited significant DIF (see Table 5 and Figure 4). Item 1 demonstrated DIF at a C (large) level, indicating substantial response differences across sex that could affect the fairness of the scale. Additionally, Item 2 showed DIF at a B (moderate) level, suggesting that sex-related response patterns may influence its measurement accuracy. According to the Educational Testing Service (ETS) classification (Zwick et al., 1999), DIF magnitude is interpreted as follows: A (negligible), B (moderate), and C (large)—with "C" suggesting potentially meaningful bias that warrants attention. Given that DIF analysis is used to detect whether an item functions differently for different groups despite equal levels of the underlying construct (Li, 2012), these findings suggest that Items 1 and 2 may require further evaluation and potential revision to ensure measurement equivalence across sex.

Discussion

This study examined the psychometric properties of the SSES-S among Turkish adolescents. The three-factor structure—academic performance, physical appearance, and social success—identified through EFA was validated by CFA, replicating the structure reported by Brito et al. (2023) in their original validation of the SSES-S. This structural consistency supports the cross-cultural robustness of the scale. All three scales showed acceptable internal consistency, and multi-group CFA indicated configural, metric, scalar, and strict invariance across sex, suggesting that boys and girls interpret the scale similarly at the latent construct level.

However, item-level analyses provided important qualifications to these findings. MGIRT results indicated variability in item discrimination and difficulty parameters across sex, and DIF was identified in two items (Items 1 and 2). Although the overall model supported measurement invariance, the presence of nonuniform DIF suggests that certain items may be interpreted differently by boys and girls, potentially compromising the fairness of cross-group comparisons. It is possible for a measure to show structural invariance while still containing items that function

Table 4. Multiple-Group IRT Results for the GRM

Item	Girl					Boy				
	a	b1	b2	b3	b4	a	b1	b2	b3	b4
1	1.68 (0.20)	−2.62 (0.31)	−1.60 (0.19)	0.02 (0.11)	1.74 (0.21)	0.85 (0.20)	−5.07 (1.26)	−3.32 (0.76)	−0.76 (0.26)	1.26 (0.34)
2	1.61 (0.21)	−2.33 (0.28)	−1.22 (0.16)	0.08 (0.11)	1.70 (0.21)	0.87 (0.21)	−4.00 (0.92)	−2.80 (0.64)	−0.56 (0.23)	1.47 (0.38)
3	1.15 (0.17)	−3.75 (0.58)	−2.43 (0.35)	−0.69 (0.16)	1.49 (0.23)	0.91 (0.21)	−4.38 (1.04)	−3.41 (0.78)	−1.16 (0.31)	1.15 (0.31)
4	1.22 (0.19)	−2.55 (0.37)	1.25 (0.20)	2.45 (0.36)	1.78 (0.45)	1.56 (0.27)	−1.92 (0.30)	0.77 (0.18)	1.46 (0.25)	1.29 (0.18)
5	2.24 (0.34)	−1.49 (0.17)	−0.92 (0.12)	0.28 (0.10)	1.22 (0.15)	0.98 (0.22)	−2.73 (0.60)	−1.56 (0.36)	0.07 (0.19)	1.44 (0.34)
6	0.62 (0.15)	−3.22 (0.76)	−1.70 (0.43)	0.08 (0.22)	1.85 (0.49)	0.90 (0.20)	−2.58 (0.58)	−1.03 (0.29)	0.20 (0.20)	1.03 (0.29)
7	1.93 (0.27)	−2.16 (0.26)	−1.30 (0.16)	0.12 (0.10)	1.34 (0.16)	0.97 (0.22)	−2.77 (0.59)	−2.01 (0.44)	0.02 (0.19)	1.45 (0.35)
8	1.15 (0.17)	−2.92 (0.43)	−0.51 (0.15)	0.51 (0.15)	1.61 (0.25)	0.56 (0.19)	−5.08 (1.69)	−0.94 (0.42)	0.82 (0.39)	2.55 (0.86)
9	0.90 (0.16)	−2.53 (0.45)	0.05 (0.16)	0.85 (0.21)	1.43 (0.28)	1.48 (0.26)	−1.75 (0.29)	−0.19 (0.14)	0.32 (0.15)	0.80 (0.19)
10	1.44 (0.22)	−1.72 (0.24)	−0.40 (0.13)	0.52 (0.14)	0.93 (0.17)	2.32 (0.39)	−1.60 (0.22)	−0.58 (0.13)	0.12 (0.12)	0.33 (0.13)
11	1.39 (0.22)	−1.90 (0.26)	−0.36 (0.13)	0.16 (0.12)	1.21 (0.19)	3.69 (0.94)	−1.18 (0.16)	−0.50 (0.10)	−0.20 (0.10)	0.38 (0.11)
12	1.56 (0.24)	−1.85 (0.25)	−0.73 (0.14)	−0.46 (0.12)	0.36 (0.12)	2.27 (0.43)	−1.18 (0.19)	−0.54 (0.13)	−0.25 (0.12)	0.18 (0.12)
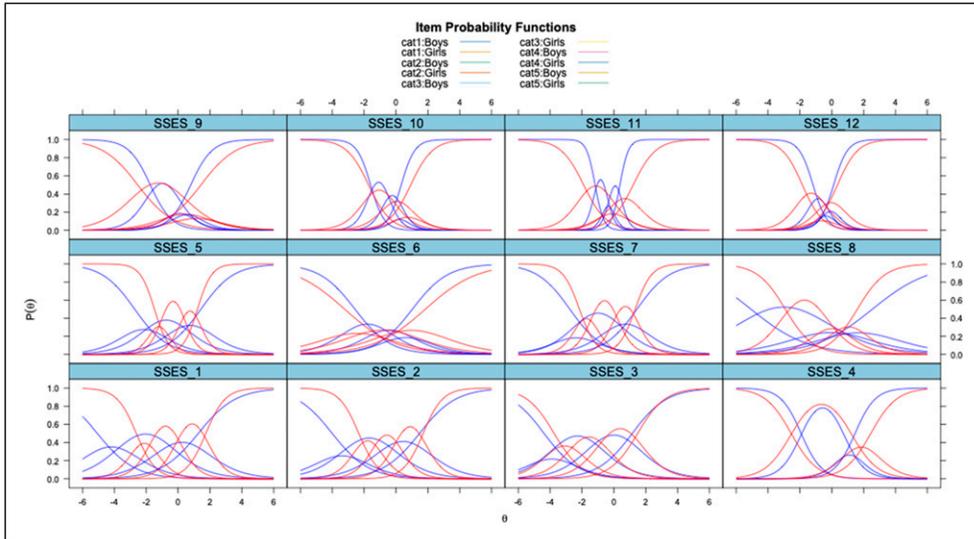
Figure 3. Gender-Based Item Probability Functions.
Note. The Red Curves Represent the boys' Group, while the blue Curves Correspond to the Girls' Group

unequally across groups—particularly when DIF-affected items contribute meaningfully to scale scores. As such, while the SSES-S shows generally acceptable psychometric properties at the scale level, researchers should interpret sex-based comparisons with caution and consider item-level functioning in future applications and revisions.

One particularly notable case was Item 6 ("I am dissatisfied with my weight"), which loaded adequately on the physical appearance factor in CFA but exhibited lower discrimination among girls in the MGIRT analysis. This suggests that body dissatisfaction, especially related to weight, may not be uniformly experienced or interpreted across sex. While Turkish girls typically show heightened sensitivity to physical appearance—likely due to sociocultural beauty standards (Avci & Keven Akliman, 2018)—their responses to weight dissatisfaction may be influenced by more complex psychosocial processes. Self-objectification theory (Fredrickson & Roberts, 1997) posits that girls internalize appearance-based expectations, but in the Turkish context, these pressures may manifest in culturally specific ways, such as balancing Western thinness ideals with local norms that value curvier figures. For boys, weight dissatisfaction may reflect concerns with muscularity or physical strength (Grogan, 2017), thus altering the meaning of the item. These results suggest that Item 6 may benefit from sex-sensitive rewording—for example, distinguishing between dissatisfaction related to thinness versus muscularity—or from being split into multiple items to better capture these divergent meanings. However, we acknowledge that item splitting alone may not fully eliminate DIF if underlying interpretations remain group-specific. Therefore, any revised or newly developed items should undergo

Table 5. IRT DIF Probability of Mantel-Haenszel (MH) and Logistic Regression

Item	MHp	LRp	DIF effect
Item 1	.0002	.001	C (Large)
Item 2	.0015	.003	B (Moderate)

Note. MHp: Mantel-Haenszel *p*-value; LRp: Logistic regression *p*-value; DIF: Differential Item Functioning.
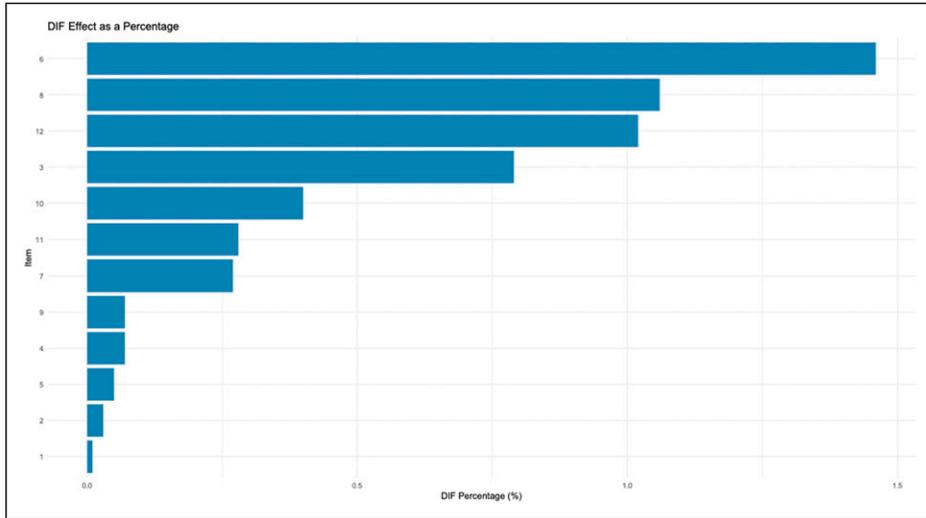
Figure 4. DIF Effect as a Percentage

follow-up validation, including cognitive interviewing and reanalysis of DIF, to assess whether measurement equivalence is improved.

Similarly, the social success factor exhibited greater item discrimination among boys than girls, suggesting that boys may experience or evaluate their social success in more differentiated ways. This aligns with sex role socialization in Turkish society, where boys are encouraged to value independence, assertiveness, and individual accomplishments (Tekke et al., 2020). For girls, social success may be more relational and context-specific, shaped by harmony and peer approval rather than status. These divergent social expectations likely influence how boys and girls interpret the same items. Though the overall factor structure held across sex, these within-factor differences in item functioning underscore the need for culturally and sex-sensitive refinement of the scale.

Beyond individual item interpretation, the presence of DIF in Items 1 and 2—related to academic performance—raises critical concerns for cross-sex comparability. Item 1 showed a large DIF effect (C-level) and Item 2 demonstrated moderate DIF (B-level). Given that these items make up one-sixth of the 12-item scale, this is a nontrivial concern. Although DIF does not necessarily imply item bias (Zumbo, 1999), and further procedures (e.g., expert review, cognitive interviews) are needed to establish bias, the observed nonuniform DIF may challenge the comparability of scale scores across sex. To address these concerns, future versions of the SSES-S might revise or replace these items using sex-neutral language or include parallel items tailored to both boys' and girls' academic self-concepts. Pilot testing these modifications through cognitive interviews could help ensure equivalent interpretation across groups.

While state self-esteem is inherently more variable than trait self-esteem, it also offers a window into how adolescents respond to immediate social and environmental contexts. The Turkish context adds important layers to this construct. In a society where academic performance is highly emphasized, sex roles remain influential, and social comparison is prevalent—particularly through digital media—state self-esteem may fluctuate in patterned ways. Over time, repeated experiences of success, failure, or social reinforcement in these culturally salient domains may contribute to the development of more stable trait-level self-esteem. Thus, longitudinal research is warranted to examine how state self-esteem fluctuations in Turkish adolescents may solidify into enduring self-views across development.

Limitations and Future Directions

While this study provides valuable insights into the psychometric properties of the SSES-S, several limitations should be noted. First, although the paper-and-pencil data collection method allowed for greater ecological validity compared to online surveys, it limited the diversity of the sample, particularly in terms of socioeconomic representation. Future research should examine the measurement invariance of the SSES-S across different socioeconomic groups, ensuring that the scale remains applicable to adolescents from various backgrounds. Second, two of the 12 items (16.66%) displayed nonuniform DIF across sex. While DIF alone does not confirm bias, the presence of nonuniform DIF—especially in a brief measure—may compromise the fairness and comparability of scores across sex. This limitation warrants further investigation through mixed-methods approaches to better understand group-specific interpretations of items. Third, the SSES-S does not establish a critical threshold for classifying different levels of self-esteem. Future studies could employ Rasch modeling or latent class analysis to establish empirically derived cut-off scores that would allow for more precise interpretations of state self-esteem levels. Additionally, since state self-esteem is inherently dynamic, longitudinal research should investigate how self-esteem fluctuates over time in response to academic, social, and cultural experiences. Lastly, measurement invariance across developmental stages should be tested to determine whether the SSES-S remains valid across early adolescence, late adolescence, and young adulthood.

Conclusions

The findings of this study provide preliminary but encouraging evidence for the validity and reliability of the SSES-S scores among Turkish adolescents. The scale's three-factor structure (academic performance, physical appearance, and social success) was supported, and its scores demonstrated adequate internal consistency and measurement invariance across sex. However, the presence of nonuniform DIF in two items–constituting one-sixth of the scale–raises important concerns about cross-sex comparability. These findings suggest that some items may not function equivalently for boys and girls and warrant refinement in future versions.

From a practical standpoint, the SSES-S shows promise as a brief and flexible tool for capturing momentary fluctuations in adolescents' self-esteem in educational, psychological, and clinical contexts. Its state-based design allows researchers and practitioners to assess self-esteem in real time, making it potentially useful for monitoring emotional well-being, identifying students in need of support, and evaluating intervention outcomes. Moreover, the observed sex-related differences, particularly in the physical appearance and social success domains, may inform the development of sex-sensitive practices in school-based mental health programs. Continued research is needed to further enhance the cultural and psychometric robustness of the instrument.

ORCID iD

M. Furkan Kurnaz ⓘ https://orcid.org/0000-0003-3773-9418

CRediT Author Contributions

M. Furkan Kurnaz: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing **Nilüfer Koçtürk:** Visualization, Writing – review & editing.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

Data will be made available on request.

References

Aftab, S., & Aslam, S. (2014). Gender differences on state self-esteem in young adults. *Pakistan Journal of Clinical Psychology*, *13*(2), 47–57.

Avci, M., & Keven Akliman, C. (2018). Gender differences in perception of body, expressions of body image and body image coping strategies among Turkish adolescents. *International Journal of Curriculum and Instruction*, *10*(2), 30–44.

Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). ERIC Publications.

Bale, C., & Archer, J. (2013). Self-perceived attractiveness, romantic desirability and self-esteem: A mating sociometer perspective. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, *11*(1), 68–84. https://doi.org/10.1177/147470491301100107

Barragán Martín, A. B., Pérez-Fuentes, M. D. C., Molero Jurado, M. D. M., Martos Martínez, Á., Simón Márquez, M. D. M., Sisto, M., & Gázquez Linares, J. J. (2021). Emotional intelligence and academic engagement in adolescents: The mediating role of self-esteem. *Psychology Research and Behavior Management*, *14*, 307–316. https://doi.org/10.2147/prbm.s302697

Beckers, D., Burk, W. J., Larsen, J. K., & Cillessen, A. H. N. (2025). Friends' influence on adolescents' loss of control eating: The moderating role of self-esteem, fear of negative evaluation, and body dissatisfaction. *Journal of Research on Adolescence: The Official Journal of the Society for Research on Adolescence*, *35*(1), Article e70005. https://doi.org/10.1111/jora.70005

Boden, J. M., Fergusson, D. M., & Horwood, L. J. (2008). Does adolescent self-esteem predict later life outcomes? A test of the causal role of self-esteem. *Development and Psychopathology*, *20*(1), 319–339. https://doi.org/10.1017/S0954579408000151

Brito, T. R. S., Pereira, C. R., Santos, F. Á., & Nery, N. N. F. (2023). Measuring the general and specific domains of self-esteem: The short-form of the state self-esteem scale. *Psychological Reports*, *126*(6), 3123–3149. https://doi.org/10.1177/00332941221077909

Burger, K., & Mortimer, J. (2024). Self-esteem. In *Reference module in encyclopedia of adolescence* (pp. 431–443). Academic Press. https://doi.org/10.1016/b978-0-323-96023-6.00058-0

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

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

Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software*, *39*(8), 1–30. https://doi.org/10.18637/jss.v039.i08

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Comrey, A., & Lee, H. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Earlbaum Associates.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1–9. https://doi.org/10.7275/jyj1-4868

Çuhadaroğlu, F. (1986). *Self-esteem in adolescents*. [Unpublished doctoral dissertation]. Hacettepe University.

DeHart, T., & Pelham, B. W. (2007). Fluctuations in state implicit self-esteem in response to daily negative events. *Journal of Experimental Social Psychology*, *43*(1), 157–165. https://doi.org/10.1016/j.jesp.2006.01.002

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related Choices. *Psychology of Women Quarterly*, *18*(4), 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Psychology Press. https://doi.org/10.4324/9781410605269

Escrivá-Martínez, T., Molinari, G., Ciudad-Fernández, V., Corno, G., & Baños, R. (2025). A validation study of the state self-esteem scale-20 (SSES-20) and the state self-esteem scale-6 (SSES-6) in a Spanish-speaking sample. *Anales de Psicología / Annals of Psychology*, *41*(1), 29–40. https://doi.org/10.6018/analesps.603481

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Fredrickson, B. L., & Roberts, T. A. (1997). Objectification theory. Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly*, *21*(2), 173–206. https://doi.org/10.1111/j.1471-6402.1997.tb00108.x

Giri, A. (2020). Self-harm behavior among adolescent students of higher secondary schools of Kathmandu Valley, Nepal. *Journal of Health and Allied Sciences*, *10*(1), 52–58. https://doi.org/10.37107/jhas.149

Grogan, S. (2017). *Body image: Understanding body dissatisfaction in men, women, and children* (3rd ed.). Routledge.

Gu, H., Zhang, P., & Li, J. (2024). The effect of self-esteem on depressive symptoms among adolescents: The mediating roles of hope and anxiety. *Humanities and Social Sciences Communications*, *11*(1), 932. https://doi.org/10.1057/s41599-024-03249-1

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Prentice-Hall.

Han, Z., Chen, S., Zhou, Y., Liu, Y., & Guo, C. (2025). The longitudinal relationship between self-esteem, life satisfaction, and depressive and anxiety symptoms among Chinese adolescents: Within- and between-person effects. *Behavioral Sciences*, *15*(2), 182. https://doi.org/10.3390/bs15020182

Hank, P., & Baltes-Götz, B. (2019). The stability of self-esteem variability: A real-time assessment. *Journal of Research in Personality*, *79*, 143–150. https://doi.org/10.1016/j.jrp.2019.03.004

Harring, M. (2024). Coping with crises in youth. In *Contributions to political science* (pp. 1–7). Springer. https://doi.org/10.1007/978-3-031-58626-2_1

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology*, *60*(6), 895–910. https://doi.org/10.1037/0022-3514.60.6.895

Hepper, E. (2016). Self-esteem. In *Encyclopedia of mental health* (pp. 80–91). Elsevier. https://doi.org/10.1016/b978-0-12-397045-9.00076-8

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*(2), 183–202. https://doi.org/10.1007/BF02289343

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kurnaz, M. F., & Koçtürk, N. (2025). Adaptation of the short-form of state self-esteem scale into Turkish: A study of validity and reliability. *Ankara University Journal of Faculty of Educational Sciences*, *58*(1), 51–85. https://doi.org/10.30964/auebfd.1500410

Kurnaz, M. F., Teke, E., & Günaydın, H. A. (2020). Relationship between self-esteem and life satisfaction: A meta-analysis study. *Research on Education and Psychology*, *4*(2), 236–253.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In *Advances in experimental social psychology* (pp. 1–62). Elsevier. https://doi.org/10.1016/s0065-2601(00)80003-9

Li, Y. (2012). Examining the impact of drifted polytomous anchor items on test characteristic curve (TCC) linking and IRT true score equating. *ETS Research Report Series*, *2012*(1), 22. https://doi.org/10.1002/j.2333-8504.2012.tb02291.x

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Livingstone, K. M., & Isaacowitz, D. M. (2018). The roles of age and attention in general emotion regulation, reappraisal, and expressive suppression. *Psychology and Aging*, *33*(3), 373–383. https://doi.org/10.1037/pag0000240

Lyyra, N., Thorsteinsson, E. B., Eriksson, C., Madsen, K. R., Tolvanen, A., Löfstedt, P., & Välimaa, R. (2021). The association between loneliness, mental well-being, and self-esteem among adolescents in four Nordic countries. *International Journal of Environmental Research and Public Health*, *18*(14), 7405. https://doi.org/10.3390/ijerph18147405

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Maldonado, L., Huang, Y., Chen, R., Kasen, S., Cohen, P., & Chen, H. (2013). Impact of early adolescent anxiety disorders on self-esteem development from adolescence to young adulthood. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *53*(2), 287–292. https://doi.org/10.1016/j.jadohealth.2013.02.025

McCain, J. L., Jonason, P. K., Foster, J. D., & Campbell, W. K. (2015). The bifactor structure and the "dark nomological network" of the State Self-Esteem Scale. *Personality and Individual Differences*, *72*, 1–6. https://doi.org/10.1016/j.paid.2014.08.006

Mertens, E. C. A., Deković, M., Van Londen, M., & Reitz, E. (2022). Parallel changes in positive youth development and self-awareness: The role of emotional self-regulation, self-esteem, and self-reflection. *Prevention Science: The Official Journal of the Society for Prevention Research*, *23*(4), 502–512. https://doi.org/10.1007/s11121-022-01345-9

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111–121. https://doi.org/10.21500/20112084.857

Netemeyer, R., Bearden, W., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Sage. https://doi.org/10.4135/9781412985772

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science*, *23*(5), 381–387. https://doi.org/10.1177/0963721414547414

Pandey, M., Dwivedi, K., & Behera, N. (2025). Self-esteem mediates the effect of yoga-induced mindfulness on adolescents' emotional and behavioral problems and pro-social behavior. *Youth & Society*, *57*(6), 0044118X241313065. https://doi.org/10.1177/0044118X241313065

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2024). *R: A language and environment for statistical computing (Version 4.4.2) [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://doi.org/10.32614/CRAN.package.psych

Rey, L., Extremera, N., & Pena, M. (2011). Perceived emotional intelligence, self-esteem and life satisfaction in adolescents. *Psychosocial Intervention*, *20*(2), 227–234. https://doi.org/10.5093/in2011v20n2a10

Rosenberg, M. (1965). *Rosenberg self-esteem scale (RSES) [Database record]*. APA PsycTests. https://doi.org/10.1037/t01038-000

Sakız, H., Mert, A., & Sarıçam, H. (2021). Self-esteem and perceived social competence protect adolescent students against ostracism and loneliness. *Journal of Psychologists and Counsellors in Schools*, *31*(1), 94–109. https://doi.org/10.1017/jgc.2020.25

Salice, A. (2020). Self-esteem, social esteem, and pride. *Emotion Review*, *12*(3), 193–205. https://doi.org/10.1177/1754073920930788

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, *34*(S1), 1–97. https://doi.org/10.1007/bf03372160

Steinsbekk, S., Wichstrøm, L., Stenseng, F., Nesi, J., Hygen, B. W., & Skalická, V. (2020). The impact of social media use on appearance self-esteem from childhood to adolescence – a 3-wave community study. *Computers in Human Behavior*, *114*, Article 106528. https://doi.org/10.1016/j.chb.2020.106528

Supervía, P. U., Bordás, C. S., Robres, A. Q., Blasco, R. L., & Cosculluela, C. L. (2022). Empathy, self-esteem and satisfaction with life in adolescent. *Children and Youth Services Review*, *144*, Article 106755. https://doi.org/10.1016/j.childyouth.2022.106755

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th Ed.). Pearson Education.

Tekke, M., Ghani, F. A., & Kassim, R. M. (2020). Evaluation of gender roles in Turkish society: Cultural impact on education perspective. In *3rd international conference on research of educational administration and management (ICREAM 2019)* (pp. 245–248). Atlantis Press.

Tsang, S., & Yip, F. (2006). Positive identity as a positive youth development construct: Conceptual bases and implications for curriculum development. *International Journal of Adolescent Medicine and Health*, *18*(3), 459–466. https://doi.org/10.1515/IJAMH.2006.18.3.459

Ugwu, L. E., Idemudia, E. S., Chukwu, O. O., & Onyedibe, M. C. C. (2023). Measuring the impact of social media on young people's mental health: Development and validation of the Social Media-Induced Tendency Scale. *Depression Research and Treatment*, *2023*, Article 8677521. https://doi.org/10.1155/2023/8677521

Vitolo, S. (2022). *The impact of social media use on state self-esteem and differences between gender identities*. [Doctoral dissertation]. Indiana University of Pennsylvania.

Westfall, R. S. (2020). Self-esteem, theory of. In B. J. Carducci, C. S. Nave, J. S. Mio, & R. E. Riggio (Eds.), *The Wiley Encylopedia of Personality and Individual Differences* (pp. 381–385). Wiley. https://doi.org/10.1002/9781119547143.ch63

Wu, X., Chen, Y., Chen, B., Guan, L., & Zhao, Y. (2017). The relationship between regional gray matter volume of social exclusion regions and personal self-esteem is moderated by collective self-esteem. *Frontiers in Psychology*, *8*, 1989. https://doi.org/10.3389/fpsyg.2017.01989

Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF)*. National Defense Headquarters.

Zwick, R., Thayer, D. T., & Lewis, C. (1999). An empirical bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement*, *36*(1), 1–28. https://doi.org/10.1111/j.1745-3984.1999.tb00543.x

Appendix

Durumluluk Benlik Saygısı Ölçeği Kısa Formu (Ergen)

Akademik Performans
1. Yeteneklerim konusunda kendime güveniyorum.
2. Kendimi başkaları kadar zeki hissediyorum.
3. Bir şeyleri iyi anlamak konusunda kendime güveniyorum.
4. İyi yapamadığımı hissediyorum.*

Fiziksel Görünüm
5. Şu anda vücudumun görünüşünden memnunum.
6. Kilomdan memnun değilim.*
7. Şu anda görünüşümden memnunum.
8. Kendimi çekici hissetmiyorum.*

Sosyal Başarı
9. Kendimi utangaç hissediyorum.*
10. Diğer insanların benimle ilgili ne düşündükleri hakkında endişeliyim.*
11. Yarattığım izlenim konusunda endişe duyuyorum.*
12. Aptal gibi görünmekten endişe duyuyorum.*

(*) Maddeler ters kodlanmaktadır.
Her madde 5'li Likert tipi puanlamaya sahiptir. "1-beni hiç tanımlamıyor" ve "5-beni tamamen tanımlıyor"