AÇIK ERİŞİM
OPEN-ACCESS

Gazi Journal of Physical Education and Sports Sciences
Gazi Beden Eğitimi ve Spor Bilimleri Dergisi
2025, 30(3), 74-84

Willpower Level Scale for Athletes (WLSA): A Reliability and Validity Study

Sporcularda İrade Kontrol Düzeyi Ölçeği (SPIRKO): Güvenilirlik ve Geçerlilik Çalışması

Research Article / Araştırma Makalesi

Nuriye Şeyma KARA[1]
Nezaket Bilge UZUN[2]
Gültekin LEKESİZ[3]

[1] Hatay Mustafa Kemal University, Faculty of Sport Sciences Hatay/Türkiye
[2] Mersin University, Faculty of Education Mersin/Türkiye
[3] Mersin University, Faculty of Sport Sciences Mersin/Türkiye

Corresponding Author / Sorumlu Yazar
Assist. Prof. Dr. Nuriye Şeyma KARA
nseymasar@gmail.com

Received / Geliş Tarihi : 07.05.2025
Accepted / Kabul Tarihi : 01.08.2025
Published / Yayın Tarihi : 26.09.2025

Ethical Statement / Etik Bilgilendirme
The study was approved by the Hatay Mustafa Kemal University Social and Human Sciences Ethics Committee on 10.01.2025, with decision number 07.

DOI: 10.53434/gbesbd.1694904

Abstract

This study aimed to develop the Willpower Level Scale for Athletes (WLSA), an instrument designed to assess athletes' capacity to maintain determination, self-discipline, and motivation under challenging conditions. The candidate scale comprised 35 items structured on a 7-point Likert scale ranging from 1 ("Does Not Describe Me at All") to 7 ("Describes Me Completely"). Content validity was established through comprehensive literature review, focus group interviews, and expert consultations during the scale development process. Exploratory factor analysis revealed a three-factor structure (stability, emotional control, and physical control), which collectively explained 60.918% of the total variance. The resultant structure was subsequently validated via confirmatory factor analysis, yielding satisfactory goodness-of-fit indices: $\chi 2/df = 4.46$, RMSEA = .075, CFI = .97, SRMR = .052, GFI = .92, and AGFI = .89. Internal consistency was demonstrated through Cronbach's alpha coefficients of .86 for Stability, .83 for Emotional Control, .75 for Physical Control, and .90 for the overall scale. The findings confirm that the WLSA constitutes a valid and reliable measurement instrument for assessing willpower levels in athletic populations.

Keywords: Sport, Willpower control, Reliability, Validity, Athletes

Öz

Bu çalışma, sporcuların zorlu koşullar altında kararlılık, öz disiplin ve motivasyonlarını sürdürebilme kapasitelerini değerlendirmek amacıyla Sporcularda İrade Kontrol Düzeyi Ölçeği'nin geliştirilmesini hedeflemektedir. Araştırma kapsamında 35 maddeden oluşan denemelik form, 1 = "Beni Hiç Tanımlamıyor" ile 7 = "Tamamen Beni Tanımlıyor" aralığında puanlanan 7 dereceli bir ölçek olarak yapılandırılmıştır. Ölçek geliştirme sürecinde literatür taraması, odak grup görüşmeleri ve uzman görüşlerinden hareketle rasyonel anlamda kapsam geçerliği sağlanmıştır. Keşfedici Faktör Analizi (KFA) sonucunda üç faktörlü bir yapı (İstikrar, Duygusal Kontrol, Fiziksel Kontrol) ortaya çıkmış ve toplam varyansın %60.918'ini açıklamıştır. Elde edilen yapı, Doğrulayıcı Faktör Analizi (DFA) ile test edilmiş; uyum iyiliği indeksleri $\chi^2/sd = 4.46$, RMSEA = .075, CFI = .97, SRMR = .052, GFI = .92 ve AGFI = .89 olarak bulunmuştur. Cronbach alfa katsayıları İstikrar = .86, Duygusal Kontrol = .83, Fiziksel Kontrol = .75 ve toplam ölçek için .90 olarak hesaplanmıştır. Bulgular, ölçeğin geçerli ve güvenilir bir ölçme aracı olduğunu ortaya koymaktadır.

Anahtar Kelimeler: Spor, İrade kontrol, Güvenirlik, Geçerlik, Sporcular

Citation / Atıf : Kara, N.Ş., Uzun, N.B. and Lekesiz, G. (2025). Willpower Level Scale for Athletes (WLSA): A Reliability and Validity Study, *Gazi Journal of Physical Education and Sports Sciences, 30*(3), 74-84.

Introduction

Human beings can be characterized as entities perpetually driven by purpose and the pursuit of motivation to realize their objectives throughout life. The motivation required in this process encompasses the factors that propel individuals to act toward accomplishing tasks (Ryan & Deci, 2000). In this context, achieving a state of motivation is considered crucial for realizing objectives. Indeed, maintaining determination and willpower is comparatively easier in situations that individuals find appealing and intrinsically motivating.

Willpower has been defined as the capacity for choice, decision-making, goal attainment, and self-control (Green & Cohen, 2004). The American Psychological Association (APA, 2012) characterized willpower as the ability to forgo short-term pleasures or rewards to achieve long-term objectives. Based on these definitions, it can be posited that the concept of willpower encompasses both affective and cognitive processes. Conceptually, the constituent processes of willpower can be enumerated as motivation, decision-making, selection, determination, responsibility, self-control, and reasoning (Okan, 2019). Motivation represents the driving force toward satisfying needs (Maslow, 1943). Decision-making behaviour, as part of this process, is the procedure of reducing uncertainty while selecting the most suitable option among alternatives (Tekin et al., 2009). Selection is the process of identifying an alternative from available options to achieve a specific goal (Simon, 1979). Determination is regarded as an indicator of consistency in decision-making processes and behaviours (Kelley, 1973). Responsibility is the capacity of individuals to assume the consequences of their own actions (Lerner & Tetlock, 1999), self-control is the ability of individuals to regulate their behaviours and impulses (Baumeister & Tierney, 2011), and reasoning signifies the capability to evaluate situations, compare potential solutions, and make inferences (Lai, 2011).

Within this framework, it can be asserted that these concepts are interrelated and collectively influential in the process of exerting willpower. Such concepts are particularly indispensable within the sporting environment. Ultimately, an athlete's success appears contingent not only on athletic performance but also on possessing the motivation to achieve goals, maintaining determination and consistency towards objectives, assuming responsibility, demonstrating self-control, and possessing reasoning abilities. In this context, a high level of willpower, alongside athletic performance, can be considered a significant determinant of success.

This is because willpower is directly related to athletes' ability to adhere to demanding training regimens, exhibit psychological resilience during competitive seasons, and commit to dietary and health-related plans. The concept of motivation is regarded as a crucial element enabling athletes to exercise willpower, a concept closely aligned with self-determination theory (Ryan & Deci, 2000; Hardy et al., 2020). Although there is existing measurement tools related to constructs closely associated with willpower—such as self-control, grit, or delay of gratification—these instruments often fail to capture the multidimensional structure of volitional regulation in sport-specific contexts. Therefore, the development of the current scale addresses a critical gap by offering a more domain-sensitive and theoretically grounded assessment of volitional control.

Developed by Deci and Ryan (1985), Self-Determination Theory (SDT) centres on how individuals' behaviours are guided by intrinsic and extrinsic motivations. The theory posits that fulfilling individuals' basic psychological needs for autonomy, competence, and relatedness enhances intrinsic motivation (Deci & Ryan, 2000). The need for autonomy refers to an individual's sense of control over their choices (Ryan et al., 2008); the need for competence relates to effectively interacting with one's environment (Deci & Ryan, 2000); and the need for relatedness signifies the requirement for meaningful social connections (Baumeister & Leary, 1995). Within the sports context, these three basic needs can be exemplified as follows: athletes acting according to their own volition during competition (autonomy); having opportunities to demonstrate their abilities and receiving resources for development (competence); and establishing effective relationships with teammates and coaches (relatedness). Fulfilling these basic psychological needs, in addition to enhancing athletic performance, may contribute to strengthening willpower and intrinsic motivation (Cheon et al., 2012; Standage & Ryan, 2020). While intrinsic motivation enhances the desire to engage in an activity independent of external rewards, willpower provides the self-discipline and resilience necessary to maintain commitment to these activities (Schunk & Zimmerman, 2008; Vallerand, 1997). In this context, athletes' intrinsic motivation concerning their efforts to improve performance and achieve goals represents an element that can strengthen their willpower (Silva et al., 2011). Stated differently, higher levels of intrinsic motivation among athletes may enable them to exhibit greater willpower (Ryan et al., 2008).

In essence, willpower—a vital component of mental resilience—is crucial for athletes to persevere through challenges, maintain focus, and adhere to rigorous training regimens. Activated by motivation and sustained through various cognitive processes, the effective exercise of willpower significantly influences athletic performance. Recognizing a gap in sports psychology specific to the Turkish context, the present study aimed to develop a valid and reliable measurement instrument designed to determine the extent to which athletes adhere to and sustain the necessary actions to achieve goals for which they are motivated. As a fundamental investigation, this research can contribute to the growing understanding of psychological factors in athletic success and establishes a basis for future work, such as refining the scale or exploring its applications across different sports or populations. The findings derived

from this tool are also intended to empower athletes by helping them identify specific areas where they face challenges in exercising willpower, thus providing opportunities for targeted self-improvement.

Method

Research Type

This research represents a significant effort to develop a scale for assessing willpower levels among athletes affiliated with various sports federations in Turkey. As a fundamental research endeavour, the study employed the summated rating scale construction technique, a well-established method in psychometric research, as outlined by DeVellis (2017). Such an approach involves designing a series of items, or statements, that athletes respond to on a Likert-type scale. The technique is valued for its simplicity, reliability, and ability to capture the intensity of a psychological trait, making it an appropriate choice for this foundational study.

Research Group

The research design incorporated two independent study groups to assess the willpower levels of athletes. This approach was adopted to strengthen the study's reliability and validity by allowing for comparative analysis between groups, reduce potential biases, and enhance the robustness of the findings. Exploratory factor analysis (EFA) was applied to the measurement model to identify underlying factors based on the relationships among variables and to establish an exploratory structure for theory generation. Data for the EFA were collected from the target population of athletes between January 15, 2025, and February 15, 2025. The sample for the EFA comprised 588 actively licensed athletes aged 18 years and older, whose descriptive information is presented in Table 1.

The EFA sample comprised 322 female athletes (54.8%) and 266 male athletes (45.2%). Among them, 305 (51.9%) participated in individual sports, while 283 (48.1%) engaged in team sports. Respecting their years of experience, 222 athletes (37.8%) had been actively licensed for 1-3 years, 133 (22.6%) for 4-6 years, 114 (19.4%) for 7-9 years, and 119 (20.2%) for 10 years or more.

Following EFA, confirmatory factor analysis (CFA) was conducted to evaluate the construct validity, including convergent and divergent validity, of the resulting final form. We aimed to ensure that participants within the independent groups exhibited a heterogeneous distribution concerning the measured trait. The data for the CFA were collected between March 1, 2025, and March 30, 2025, from licensed, actively competing athletes who voluntarily participated in the research. Descriptive statistics for the active athletes included in the CFA sample are listed in Table 1.

Table 1. Descriptive statistics for participants

		EFA samples		CFA samples	
		N	%	N	%
Gender	Female	322	54.8	341	56.1
	Male	266	45.2	267	43.9
Sport Category	Individual	305	51.9	310	51.0
	Team	283	48.1	298	49.0
Years of Athletic Experience	1-3 years	222	37.8	227	37.3
	4-6 years	133	22.6	137	22.5
	7-9 years	114	19.4	118	19.4
	>10 years	119	20.2	126	20.7
Total		588	100.0	608	100.0

To determine the psychometric construct validity of the Willpower Level Scale for Athletes (WLSA), data were collected from 608 active athletes in a separate sample from the EFA group. The CFA sample comprised 341 female (56.1%) and 267 male (43.9%) athletes. Among the participants, 310 (51.0%) competed in individual sports, while 298 (49.0%) engaged in team sports. Concerning athletic experience, 227 athletes (37.3%) had been actively licensed for 1-3 years, 137 (22.5%) for 4-6 years, 118 (19.4%) for 7-9 years, and 126 (20.7%) for 10 years or more. The CFA was conducted to evaluate the construct validity, including convergent and discriminant validity, of the final form derived from the EFA. The analyses were designed to ensure that participants within the independent groups exhibited a heterogeneous distribution concerning the measured trait.

Scale Development Process

Focus Group Interviews: To create the initial item pool, focus group interviews were conducted with 10 national-level athletes and 4 expert academics in the field. Convenience sampling, a qualitative research sampling method, was employed to select participants for these focus groups.

Essay Writing by Target Audience: To create the item pool, 38 actively licensed athletes (aged 18+) enrolled in the Faculties of Sport Sciences at MEU and MKU (Mersin University and Hatay Mustafa Kemal University) were asked to write essays responding to open-ended questions about their views and experiences as to the willpower levels in athletes.

Literature Review: Scale items from existing literature related to willpower and similar constructs (e.g., discipline, self-control) were systematically examined. This review contributed significantly to the item pool generation. Additionally, item writing was informed by analysing theories relevant to willpower, including Goal Setting Theory and Self-Determination Theory.

Content Validity Ratio (CVR) Analysis: After completing the preceding stages, the item pool and candidate scale were refined. The candidate text was read aloud to 8 active athletes for preliminary feedback. Based on this feedback, a 40-item candidate scale was developed, which was drafted according to

Lawshe's technique and distributed to 8 experts via email for evaluation based on representativeness and clarity. At an α = .05 significance level, the critical CVR value for 8 experts was .612 (Lawshe, 1975), so 3 items failed to meet this criterion, and 2 additional items were removed based on expert recommendations (Lawshe, 1975). Following the expert evaluation and content validity process, 10 items were removed from the initial 40-item form, and 5 items were added based on expert feedback, resulting in a 35-item candidate scale. Experts were also consulted regarding the rating scale format. Based on their input, a 7-point Likert-type rating scale (1: Does not describe me at all to 7: Describes me completely) was deemed most appropriate.

Application of the Candidate Scale: In accordance with expert opinions and the CVR analysis, the 35-item candidate scale using the 7-point rating scale was finalized, and then it was administered both face-to-face and via online platforms to 655 active athletes.

Data Analysis Techniques

Exploratory Factor Analysis and Confirmatory Factor Analysis: We conducted exploratory factor analysis and confirmatory factor analysis as part of the statistical analysis. To evaluate the reliability of the scale, Cronbach's α internal consistency coefficients and composite reliability values were calculated. Besides, assumption testing was conducted prior to the factor analyses. The collected 655 observations were examined for assumptions such as missing data, adequacy of the observation set, outliers, multicollinearity, factorability of the R matrix, normality, and linearity to assess their suitability for factor analysis. These assumption tests were applied separately for both observation groups (EFA and CFA samples).

EFA is often employed to uncover the underlying structure of a set of variables (Tabachnick et al., 2019). Determining an appropriate sample size in this analysis plays a critical role in ensuring the reliability and validity of the obtained results. Several researchers have offered different recommendations regarding the required sample size for EFA. Krichbaum et al. (2011) stated that at least 125 participants are necessary for a 25-item questionnaire, whereas Cheong et al. (2017) suggest a minimum participant-to-item ratio of five-to-one. Vielma-Aguilera et al. (2023) recommend at least 10 participants per scale item for both EFA and CFA, while Comrey and Lee (2013) indicate that more than 500 participants are sufficient for EFA during the scale validation process. In this context, the 655 observations obtained in the present study indicate that the required sample size was satisfied.

Comparison of central tendency measures—median, mode, and arithmetic mean—for the scale items revealed closely aligned values. This finding was interpreted as evidence of normal distribution, indicating that appropriate heterogeneity regarding the measured trait was achieved within the target population. A comprehensive outlier analysis was conducted on the initial 655 observations prior to executing Mahalanobis distance and Z-score analyses. Five observations from both the positive and negative extremes were eliminated, reducing the dataset to 645 observations before further assumption testing. Z-scores were utilized to evaluate the structure in multi-item scales, with scores typically considered acceptable within the range of -3 to +3 (Mertler et al., 2005). Upon examination of these Z-scores, 3 additional observations falling outside this range were excluded from subsequent analyses. The remaining Z-scores were found to range between -2.81 and 2.93, indicating acceptable univariate normality. To detect multivariate outliers, Mahalanobis distances were calculated and compared against the critical threshold derived from the Chi-square distribution ($\chi^2$35, 0.001 = 73.402). Based on this criterion, 54 observations exceeding the critical value were excluded from the analysis. Following these data screening procedures, the final analytical sample consisted of 588 observations, which were utilized for all subsequent analyses.

Testing of Statistical Assumptions: Aligned with Kara et al.'s (2023) assertion that achieving perfect linearity between two variables is virtually impossible; the analyses proceeded under the assumption that the existing relationships were sufficiently linear. Concerning the normality assumption, items were examined individually, with central tendency measures alongside skewness and kurtosis coefficients being evaluated. The proximity of the central tendency measures suggested that univariate normality was achieved (Can, 2018). For the 35 items in the scale, skewness values ranged from -0.957 to -0.172, and kurtosis values ranged from -0.586 to 0.566. Given that skewness coefficients between -3.3 and +3.3 and kurtosis coefficients between -7 and +7 are considered acceptable for normality (Bernstein, 2000), it can be concluded that the data adequately satisfied the normality assumption.

To identify potential multicollinearity issues, Tolerance and Variance Inflation Factor (VIF) values were examined. For the 35 items, Tolerance values ranged from .272 to .693, and VIF values ranged from 1.443 to 3.866. As all Tolerance values exceeded .20 and all VIF values were below 5, it was determined that no multicollinearity problem existed among the items in the dataset, justifying the retention of all items for subsequent analysis (Tabachnick & Fidell, 2015).

In the developed model, correlation between error terms at different points, known as autocorrelation, can increase the risk of Type I error (Jenson et al., 2007). The assessment conducted for this purpose yielded a Durbin-Watson (D-W) value of 2.292. This result, being close to the ideal value of 2 that indicates no autocorrelation, supported the conclusion that the error terms were independent (Kalaycı, 2010).

Assessment of Factor Analysis Suitability: In analyses conducted to determine the factorability of the R matrix, a crucial prerequisite for factor analysis, the results of the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test

of Sphericity were evaluated. The calculated KMO value of .91 indicated that the data were highly suitable for factor analysis and that the sampling adequacy was excellent. According to Hutcheson and Sofroniou (1999), KMO values between 0.5 and 0.7 are considered mediocre, 0.7 to 0.8 good, 0.8 to 0.9 great, and 0.9 or above are deemed excellent (Dağlı, 2015). Thus, the obtained KMO value of .91 was concluded to demonstrate excellent sampling adequacy. In addition, Bartlett's Test of Sphericity, which examines whether inter-item correlations differ significantly from zero, rejected the null hypothesis (χ^2 = 4059.149; $p < .05$), confirming that significant correlations existed among the items. This statistical significance ($p < .05$) supports the presence of a structure suitable for factor analysis (Gürbüz & Şahin, 2014). The KMO value of .91 better substantiates that the correlation matrix generated from the study group's data is highly factorable.

Confirmatory Factor Analysis Preparation: Prior to conducting CFA, the scale—now comprising 15 items following EFA—was administered again to the target population of active athletes. Comprehensive assumption analyses were initially performed on the 642 observations obtained. These analyses included checks for missing data, linearity, normality, sample size adequacy, and potential multicollinearity issues.

Following missing data analysis, the normality assumption was evaluated by examining central tendency measures (mode, median, and arithmetic mean) and their relative positions. The proximity of these values confirmed that the condition of univariate normality was met. Outlier analyses identified 5 observations each from the highest and lowest extremes, which were subsequently excluded, leaving 632 observations. Analysis of skewness and kurtosis values for these remaining cases revealed that skewness values generally indicated a negative distribution, ranging from -1.019 to -0.146, while kurtosis values ranged from -0.537 to 0.550. While skewness coefficients are typically recommended to fall within the ±1 range for optimal univariate normality (Göldağ, 2019), values between -3.3 and +3.3 are widely considered acceptable (Bernstein, 2000); thus, the results indicate the normality assumption was adequately satisfied.

To examine outliers to a greater extent, both Mahalanobis distances (for multivariate outliers) and Z-scores (for univariate outliers) were analysed. Based on the critical value for Mahalanobis distances ($\chi^2_{15,0.001}$ = 37.697), an additional 24 observations exceeding this threshold were excluded. The Z-scores for the remaining cases ranged from -3.59 to 2.46, and the analysis proceeded with these 608 observations.

To detect potential multicollinearity problems in the CFA sample, VIF and Tolerance values were calculated. Inter-item VIF values ranged from 1.49 to 2.409, and Tolerance values ranged from .415 to .670. As all Tolerance values exceeded 0.20 and all VIF values remained below 5, it was determined that no multicollinearity problem existed. Based on these comprehensive assumption analyses and considering the criteria recommended by Tabachnick and Fidell (2015), the final dataset of 608 observations was deemed appropriate for CFA implementation. Following these preparatory analyses, CFA was applied using the dataset comprising 608 observations and the 15-item scale form. Within the scope of CFA, analyses were conducted to determine the estimated error variances and standardized loading values for the items, as well as to evaluate goodness-of-fit criteria.

Ethical Statement

The ethical suitability of the research was approved by Nişantaşı University Rectorate Ethics Committee with decision 2023/38 at a meeting dated 29.09.2023. This study was conducted in strict compliance with the ethical standards and principles established by the Declaration of Helsinki, ensuring the protection of participants' rights, autonomy, and well-being throughout the research process.

Findings

EFA Results

Within the scope of the exploratory factor analysis (EFA), 655 observations were initially obtained, however this number was reduced to 588 following assumption testing. The explained communality values, which indicate the extent to which variables are represented by the factors, were determined to range between .437 and .747. In the relevant literature, it is generally accepted that item communality values below 0.10 may indicate an issue (Büyüköztürk, 2022), although it should also be emphasized that relying solely on communality values is not sufficient justification for decision-making. Based on the view that obtaining more comprehensive information about the contribution of items to the measurement is beneficial (Çokluk et al., 2012), additional methods such as the Scree Plot, Percentage of Total Variance Explained, Kaiser's Criterion (Eigenvalue > 1), and the Explained Variance Criterion were included in the analysis process to clarify the factor structure. As stated by Cattell (1966), assuming the plateau reached in the plot represents the start of trivial factors, each point before the plateau indicates a significant factor.

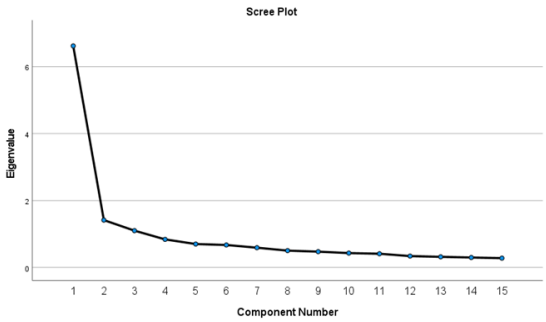

Figure1. Scree Plot

Accordingly, examining the scree plot presented in Figure 1, a rapid decrease was observed from the first component onwards, and the slope markedly decreased after the third point, indicating the beginning of a plateau. According to Cattell's Scree Test, this suggests the prominence of a 3-factor structure. Specifically, the curve beginning to follow a horizontal course from the fourth component onwards indicates that the first three factors carry significant variance and represent the underlying structure of the scale. In this context, it can be stated that the Willpower Level Scale for Athletes (WLSA) in sports consists of a 3-factor structure. To base the evaluations on more objective grounds and pre-empt potential criticisms, the table of explained variance was also included.

The percentage of total variance explained method, one of the techniques used to determine the number of factors, is a frequently employed statistical criterion for assessing the extent to which the overall variability in a data set is explained by different factors. According to this method, when the contribution of an added factor to the total variance drops below 5%, the model is considered to have reached the optimal number of factors (Kalaycı, 2010). The results presented in Table 2 support the presence of a three-factor structure based on this criterion.

Table 2. Total variance explained

Component	*Initial Eigenvalues*			*Extraction Sums of Squared Loadings*		
	Total	*% of Variance*	*Cumulative %*	*Total*	*% of Variance*	*Cumulative %*
1	6.620	44.134	44.134	3.637	24.245	24.245
2	1.416	9.443	53.577	2.965	19.766	44.011
3	1.101	7.340	60.918	2.536	16.907	60.918
15	.279	1.862	100.000			

According to Kaiser's criterion, factors with an eigenvalue greater than 1 are recommended for inclusion in the model. Although four factors in the current dataset appear to have eigenvalues greater than 1, the eigenvalue of the fourth factor is very close to this threshold and is marginal. Therefore, the significance of the fourth factor was considered debatable. Çokluk et al. (2012) state that in cases where an eigenvalue is borderline, one should not rely solely on numerical thresholds; it is sounder for the researcher to decide by considering the theoretical framework, graphical breaks—such as those in the scree plot, and the overall measurement model.

The result of the factor analysis, where the three-factor structure explained 60.918% of the total variance, indicates that it falls within acceptable limits for the social sciences (Tabachnick & Fidell, 2013). The first factor explains 44.134% of the variance, the second factor 9.443%, and the third factor 7.340%; each has an eigenvalue greater than 1. These findings are consistent with the variance criteria suggested in the literature (Büyüköztürk, 2018; Demir, 2023) and demonstrate that the scale provides a sufficient level of evidence for construct validity. The items not included in the analysis and the reasons for their exclusion are presented in Table 3.

Table 3. Items Removed from Exploratory Factor Analysis and Reasons for Removal

Items with Communalities < 0.30	*Items with Factor Loadings < 0.40*	*Cross-loading Items (Difference between highest two loadings < 0.10)*	*Rational Reasons (Factor Naming, Phrasing/Clarity)*
17,18,26	16,19,30	12,13,14,21,22,24,25 27,28,31,32	33,34,35

Within the scope of the EFA conducted, various items had to be excluded from the analysis to enhance the structural integrity of the measurement tool and achieve a statistically more robust structure. Items numbered '17, 18, and 26', which had communality coefficients below 0.30, and items '16, 19, and 30', which did not show sufficient loading (below 0.45) on any factor, were removed. In addition, items '12, 13, 14, 21, 22, 24, 25, 27, 28, 31, and 32', which loaded at similar levels on different factors (i.e., the difference in loadings was less than 0.10), making discrimination between structures difficult, were considered cross-loading items and were excluded from the analysis. Besides, items '33, 34, and 35' were also deemed appropriate for removal from the scale due to reasons such as content inconsistency or weak conceptual fit (related to factor naming, language, and expression). Therefore, the semantic integrity and measurement validity of the statements in the scale were supported, and a more robust factor structure was established. The final list of items obtained, along with the communality values for each retained item, are presented in detail in Table 4.

Table 4. Item communalities, factor loadings, and factor assignments

Item No	*Item*	*Factor 1*	*Factor 2*	*Factor 3*	*Common Factor Variance (h^2)*	*Item Factor Total Test Correlation*	*Cronbach's Alfa if Item Deleted*
1	I continue with my training even if the program is challenging			.831	.712	.598	.901
3	I adhere to the training program even if my performance declines.			.750	.670	.710	.897
4	I persistently work on techniques that are difficult to execute.			.714	.672	.716	.898
29	I participate in training sessions even when tired.			.681	.564	.610	.901
15	I adhere to my training program even when feeling demoralized.			.628	.560	.678	.898
2	Negative criticism I receive does not deter me from training.			.536	.450	.538	.904
9	Negative spectator pressure during a match does not affect my performance.		.783		.684	.607	.901
11	I am not daunted by my opponent's strengths.		.714		.643	.657	.899
10	Problems experienced with other athletes do not daunt me.		.696		.683	.869	.898
23	Provocations from my opponents do not affect my performance.		.640		604	.681	.899
20	I am not affected by unfavorable referee decisions.		.607		.524	.515	.904
7	I adhere to the dietary rules I am required to follow.	.812			.686	.499	.905
6	I resolutely continue training to maintain my physical fitness.	.708			.747	.716	.897
5	I strive to maintain my weight.	.689			.501	.375	.910
8	I abandon detrimental habits for the sake of my athletic success.	.592			.437	.491	.905
	Explained Variance Values (%)	%16.907	%19.766	%24.245	%60.918		
	Cronbach's Alpha Values	%75	% 83	%86	%90		

The communalities (h^2), factor loading values for the items, and their respective factor structure are presented in detail in Table 5. According to the analysis findings, the total variance explained was 60.918%; it can be stated that this value indicates the factor structure adequately explains the total variance. Regarding the explained variance proportions, factor 1 contributed 24.245%, factor 2 19.766%, and factor 3 16.907%. This indicates that the items in the scale loaded significantly onto the factor structure and that the sub-dimensions are distinctly structured (Kline, 2015).

The Corrected Item-Total Correlation values indicate the degree to which each item correlates with the total score excluding that item. Values above .40 are generally considered acceptable, reflecting adequate item discrimination (DeVellis, 2017). In this analysis, all items—except Item 5 (.375)—met or exceeded this threshold. The Cronbach's Alpha if Item Deleted values demonstrate the internal consistency of the scale if a given item were removed. Since the alpha values remain stable across items and do not substantially increase when any item is deleted, all items appear to contribute meaningfully to the overall reliability (Tavşancıl, 2014).

In the reliability assessments, Cronbach's α coefficients were calculated as .86 for the first factor, .83 for the second factor, .75 for the third factor, and .90 for the overall scale. From a psychometric perspective, these values indicate that the items in the scale possess a high level of internal consistency (Tabachnick & Fidell, 2013). The fact that both the levels of explained variance and the internal consistency coefficients fall within acceptable limits supports the structural validity of the developed scale. The naming of each sub-dimension, the number of items per factor, and their reliability levels, determined by considering the items' relationships with their respective factors, their wording, and their conceptual coherence with the literature, are presented in Table 5.

Table 5. Factor names and reliability coefficients

Factor	*Factor Name*	*Number of Items*	*Cronbach's Alpha*
Factor 1	Stability	6	.86
Factor 2	Emotional control	5	.83
Factor 3	Physical control	4	.75
Overall Scale / Total			.90

The data presented in Table 5 display the reliability coefficients along with the number of items designated for the three sub-dimensions of the scale. The Cronbach's α value of .86 obtained for the Stability sub-dimension indicates that this dimension is highly reliable. The coefficient for the Emotional Control dimension was .83, a value considered sufficient for measurement consistency. The reliability coefficient for the Physical Control sub-dimension was .75, which is generally regarded as an acceptable threshold in psychological measurement instruments. The calculated α value for the entire scale was .90, demonstrating that the scale possesses a reliable structure overall. In psychological measurement instruments, α coefficients above .70 are generally considered adequate; values of .80 and above are regarded as high, while values of .90 and above indicate a very strong level of internal consistency (Tavşancıl, 2014; Büyüköztürk, 2022).

CFA Results

Based on the confirmatory factor analysis (CFA), the standardized loading values ranged from .536 to .831 for factor 1, from .607 to .783 for factor 2, and from .592 to .812 for factor 3. The t-values for all items in the three-factor CFA model exceeded the ±1.96 threshold, indicating statistical significance. This demonstrates that the items within the *Stability*, *Emotional Control*, and *Physical Control* dimensions loaded significantly onto their respective factors. As noted in the literature, t-values outside this threshold are considered statistically significant (Kara et al., 2023). Therefore, each item exhibits a strong relationship with its corresponding factor, reflecting high item discriminability, which further supports the construct validity of the measurement instrument and confirms that it accurately assesses the intended theoretical constructs. The t-values for the 15 items included in the analysis demonstrate a consistent and valid structure within the model, as illustrated in the path diagrams below—the standardized loading values and t-values for the items are presented in Figure 2 and Figure 3, respectively, to support these analysis results.

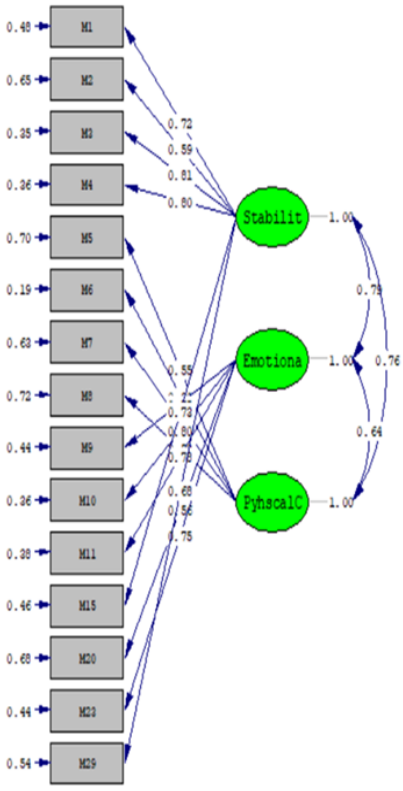

Figure 2. Standardized estimates for the tested model

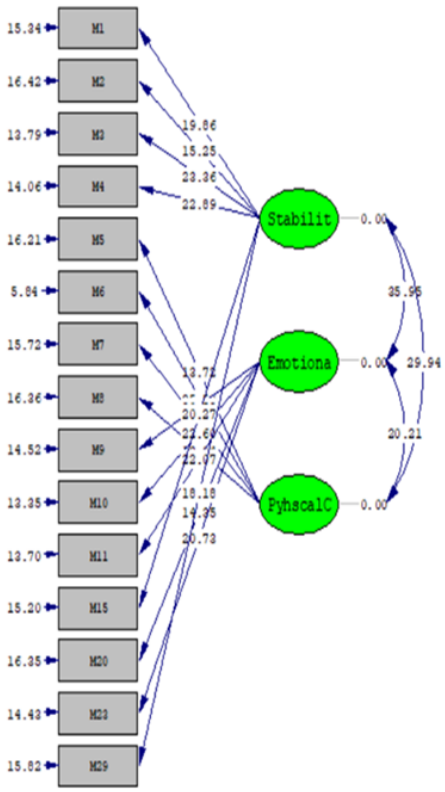

Figure 3. T-values for parameter estimates (p<=.05)

In the model presented in Figure 2, each item exhibits a significant and consistent relationship with its corresponding factor. Similarly, Figure 3 indicates that all t-values for these items exceed the ±1.96 threshold, confirming their statistical significance. Collectively, these findings demonstrate the scale's discriminative power at the item level and provide evidence of its construct validity (Hair et al., 2014). Likewise, a comprehensive evaluation of the model's goodness-of-fit criteria and the overall analysis results strongly confirms the proposed structure within the study sample.

Table 6. Goodness-of-fit criteria and obtained values

Goodness-of-fit index	*Perfect Fit*	*Good or Acceptable Fit*	*Value Achieved*
χ^2/df	< 2	< 5	4.46
RMSEA	0 ≤ RMSEA ≤ .05	.05 < RMSEA ≤ .08	.07
SRMR	0 ≤ SRMR ≤ .05	.05 < SRMR ≤ .10	.05
NFI	.95 ≤ NFI ≤ 1.00	.90 ≤ NFI < .95	.97
NNFI	.97 ≤ NNFI ≤ 1.00	.95 ≤ NNFI < .97	.97
CFI	.97 ≤ CFI ≤ 1.00	.95 ≤ CFI < .97	.97
GFI	.95 ≤ GFI ≤ 1.00	.90 ≤ GFI < .95	.92
AGFI	.90 ≤ AGFI ≤ 1.00	.85 ≤ AGFI < .90	.89

The evaluation of model fit for the proposed confirmatory factor analysis yielded a χ^2/df ratio of 4.46 (χ^2 = 387.86, df = 87). Although the Chi-square statistic is known to be sensitive to sample size, this ratio falls within acceptable limits, as supported by established guidelines (Kline, 2014; Sümer, 2000). Complementing this finding, additional goodness-of-fit indices confirm the adequacy of the model. The values obtained were as follows: Root Mean Square Error of Approximation (RMSEA) = .075, Standardized Root Mean Square Residual (SRMR) = .052, Normed Fit Index (NFI) = .97, Non-Normed Fit Index (NNFI) = .97, Comparative Fit Index (CFI) = .97, Goodness of Fit Index (GFI) = .92, and Adjusted Goodness of Fit Index (AGFI) = .89. These indices collectively fall within ranges indicative of a good fit, thereby confirming the validity of the model for the given sample (Jöreskog & Sörbom, 1993). Consequently, the measurement instrument, designed with a 3-factor structure comprising 15 items, demonstrates a significant and adequate level of model fit.

Table 7. Convergent and discriminant validity and CR Values for the scale

Factors	*AVE*	*MSV*	*ASV*	*CR*
1	0.56	0.43	0.12	0.86
2	0.58	0.46	0.12	0.83
3	0.50	0.29	0.12	0.75
Index	AVE>.50 CR>AVE	MSV<AVE	ASV<MSV	CR>.70

(Cronbach's alpha for the entire scale: .90)

To confirm convergent validity, it is generally recommended that the Average Variance Extracted (AVE) exceeds 0.50 and that the Composite Reliability (CR) is higher than the AVE (Yaşlıoğlu, 2017). In the current study, this criterion (CR > AVE) was fulfilled across all three dimensions. Regarding discriminant validity, especially in multi-dimensional scales, several criteria are applicable. One such criterion, which requires that the Average Shared Variance (ASV) remains lower than the Maximum Shared Variance (MSV), was also satisfied (ASV < MSV). These findings collectively indicate that discriminant validity was achieved. Moreover, all CR values surpassed the commonly accepted threshold of 0.70.

Conclusion

This study was designed to develop a measure to assess athletes' abilities to sustain determination, self-discipline, and motivation when faced with challenging conditions. The development of the scale was thorough, involving focus group interviews, extensive literature reviews, and expert evaluations to ensure a robust foundation. Through exploratory factor analysis, a three-factor structure emerged—comprising Stability, Emotional Control, and Physical Control—which was later validated using confirmatory factor analysis.

Stability sub-dimension captures an athlete's capacity to exhibit consistency and persistence in pursuing their goals. It reflects the ability to engage in planned and sustained efforts toward long-term objectives, as noted by Kelley (1973). Items within this factor demonstrate that athletes remain dedicated to their training programs despite difficulties, adhere to their goals even when motivation dips, and maintain steady training efforts over time.

Emotional Control addresses an athlete's ability to manage emotional responses under stress and pressure, thereby preserving performance levels (Baumeister, Vohs, & Tice, 2007). The items here emphasize maintaining emotional equilibrium and focus in the presence of external challenges, such as spectator pressure, superior opponents, or unfavourable referee decisions.

Physical Control sub-dimension relates to an athlete's discipline in maintaining lifestyle habits—such as sleep, nutrition, and avoidance of harmful behaviours—essential for physical preparedness (Bandura, 1991). Items in this category highlight adherence to training and nutrition plans, avoidance of unhealthy habits, and self-regulation of physical preparation processes.

The resulting scale utilizes a 7-point Likert format, ranging from 1 ("Does Not Describe Me at All") to 7 ("Describes Me Completely"), enabling precise self-assessment of these attributes (DeVellis, 2017; Tavşancıl, 2014). The scale's construct validity is supported by acceptable model fit indices: $\chi^2/df = 4.46$, RMSEA = .075, CFI = .97, SRMR = .052, NFI = .97, NNFI = .97, GFI = .92, and AGFI = .89. The scale exhibits strong internal consistency, with Cronbach's α coefficients of .86 for Stability, .83 for Emotional Control, .75 for Physical Control, and .90 for the overall scale. These results establish the scale as a reliable and valid instrument for measuring willpower levels among athletes.

Recommendations

The Willpower Level Scale for Athletes (WLSA), developed in this study, is a valid and reliable instrument for assessing an athlete's determination, self-discipline, and self-regulation skills across three dimensions. Future research could conduct comparative examinations of how this scale functions across different age groups and performance levels. An investigation into the relationship between motivation levels, within the framework of self-determination theory, and willpower levels may also prove insightful.

Financial Sources

The preparation and writing of this study were completed without financial support from any institution or organization.

Conflict of Interest

There is no conflict of interest among the authors related to publication of this article.

Author Contributions

Research Idea: NSK; Research Design: NSK, NBU; Data Collection: NSK, GL; Data Analysis: NSK, NBU; Writing: NSK, NBU, GL; Critical Review: NBU, NSK, GL

References

1. **APA. (2012).** *What Americans think of willpower*? Copyright American Psychological Association.
2. **Bandura, A. (1991).** Social cognitive theory of self-regulation. *Organizational Behaviour and Human Decision Processes*, 50(2), 248–287.
3. **Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007).** The strength model of self-control. Current Directions in Psychological Science, 16(6), 351–355.
4. **Baumeister. R. F.. & Leary. M. R. (1995).** The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin.* 117(3). 497-529.
5. **Baumeister. R. F.. & Tierney. J. (2011).** *Willpower: Rediscovering the greatest human strength*. Penguin Press.
6. **Bernstein, I. (2000, April).** *Some consequences of violating SEM's assumptions*. Paper presented at the annual meeting of the Southwestern Psychological Association, Dallas, TX.
7. **Büyüköztürk. Ş. (2018).** *Sosyal bilimler için veri analizi el kitabı: İstatistik. araştırma deseni. SPSS uygulamaları ve yorum* (24. Baskı). Ankara: Pegem Akademi.
8. **Büyüköztürk. Ş. (2022).** *Sosyal Bilimler İçin Veri Analizi El Kitabı* (29. Baskı). Ankara: Pegem Akademi.
9. **Can, A. (2018).** *Sosyal bilimlerde veri analizi: SPSS ile araştırma süreci* (6. baskı). Ankara: Pegem Akademi Yayıncılık.
10. **Cattell, R. B. (1966).** The scree test for the number of factors. *Multivariate Behavioural Research*, 1(2), 245–276.
11. **Cheon. S. H.. Reeve. J.. & Moon. I. (2012).** A classroom-based intervention to increase students' motivation: An experimental study of the effects of autonomy-supportive teaching on students' intrinsic motivation. *Learning and Instruction.* 22(4). 193-204.
12. **Cheong. A. T.. Chinna. K.. Khoo. E. M.. & Liew. S. M. (2017).** Determinants for cardiovascular disease health check questionnaire: *A validation study. Plos one*. 12(11). e0188259.
13. **Çokluk. Ö.. Şekercioğlu. G.. & Büyüköztürk. Ş. (2012).** *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (Vol. 2). Ankara: Pegem Akademi.
14. **Comrey. A. L.. & Lee. H. B. (2013).** *A first course in factor analysis*. Psychology press.
15. **Dağlı, A. (2015).** Örgütsel muhalefet ölçeğinin türkçe'ye uyarlanması: geçerlik ve güvenirlik çalışması. *Elektronik Sosyal Bilimler Dergisi*, 14(53), 198-218.
16. **Deci. E. L.. & Ryan. R. M. (1985).** *Intrinsic motivation and self-determination in human behaviour*. Plenum Press.
17. **Deci. E. L.. & Ryan. R. M. (2000).** The "what" and "why" of goal pursuits: Human needs and the self-determination of behaviour. *Psychological Inquiry.* 11(4). 227-268.
18. **Demir. A. (2023).** Ölçek geliştirme sürecinde açıklanan varyans yüzdesi kriterinin değerlendirilmesi. *Eğitimde Araştırmalar ve Uygulamalar Dergisi*. 7(2). 89–102.
19. **DeVellis, R. F. (2017).** *Scale development: Theory and applications* (4th ed.). SAGE Publications.
20. **Greene, J., & Cohen, J. (2004).** For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*(1451), 1775
21. **Gürbüz, S., & Şahin, F. (2014).** *Sosyal bilimlerde araştırma yöntemleri*. Ankara: Seçkin Yayıncılık, 271.
22. **Hair, J. F., Babin, B. J., Anderson, R. E., & Black, W. C. (2014).** *Multivariate data analysis* (7th ed.). Pearson
23. **Hardy, J., Barlow, M., Evans, L., Rees, T., Woodman, T., & Taylor, J. (2020).** The role of self-determination theory in understanding motivation in sport and exercise. *International Review of Sport and Exercise Psychology*, 13(1), 194–220.
24. **Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007).** Statistical reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in the Schools*, 44(5), 483-493.
25. **Jöreskog, K. G., & Sörbom, D. (1993).** *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International.
26. **Kalaycı, Ş. (2010).** *SPSS uygulamalı çok değişkenli istatistik teknikleri* (Vol. 5, p. 359). Ankara, Turkey: Asil Yayın Dağıtım.
27. **Kara, M., Kara, N. Ş., & Uzun, N. B. (2023).** Level of competition forgiveness scale in sport (FCLSS): validity and reliability study. *International Innovative Education Researcher* (IEDRES), 3(3).
28. **Kelley, H. H. (1973).** Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution theory: Perception of the causes of behaviour* (pp. 1–26). General Learning Press.

29. **Kelley, H. H. (1973).** The processes of causal attribution. American Psychologist, 28(2), 107–128.
30. **Kline, R. B. (2015).** *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
31. **Krichbaum. K. E.. Peden-McAlpine. C.. Diemert. C.. Koenig. P.. Mueller. C.. & Savik. K. (2011).** Designing a measure of complexity compression in registered nurses. *Western Journal of Nursing Research*. 33(1). 7-25.
32. **Lai. E. R. (2011).** *Critical thinking*: A literature review. Pearson.
33. **Lawshe. C. H. (1975).** A quantitative approach to content validity. *Personnel psychology*. 28(4).
34. **Lerner, J. S., & Tetlock, P. E. (1999).** Accounting for the effects of accountability. *Psychological bulletin*, *125*(2), 255.
35. **Maslow. A. H. (1943).** A theory of human motivation. *Psychological Review*. 50(4). 370-396.
36. **Mertler, C. A., & Vannatta, R. A. (2005).** *Advanced and multivariate statistical methods: Practical application and interpretation* (3rd ed.). Los Angeles, CA: Pyrczak Publishing.
37. **Okan. N. (2019).** *İrade ve ahlaki bütünlüğün olumlu psikolojik çıktıları üzerine bir karma yöntem araştırması*. (Yayımlanmış Doktora Tezi). Marmara Üniversitesi. Eğitim Bilimleri Enstitüsü. İstanbul.
38. **Ryan. R. M.. & Deci. E. L. (2000).** Self-determination theory and the facilitation of intrinsic motivation. social development. and well-being. *American Psychologist.* 55(1). 68-78.
39. **Ryan. R. M.. Patrick. H.. Deci. E. L.. & Williams. G. C. (2008).** Facilitating health behaviour change and its maintenance: Interventions based on self-determination theory. *The European Health Psychologist.* 11(2). 2-5.
40. **Schunk, D. H., & Zimmerman, B. J. (2008).** *Motivation and self-regulated learning: Theory, research, and applications*. Routledge.
41. **Silva, M. N., Marques, M. M., & Coutinho, P. (2011).** Self-determination theory in physical activity and sport: An overview of the current literature. *Psychology of Sport and Exercise*, 12(2), 157–165.
42. **Simon. H. A. (1979).** Rational decision making in business organizations. *The American Economic Review.* 69(4). 493-513.
43. **Sofroniou, N., & Hutcheson, G. (1999).** *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage Publications.
44. **Standage. M.. & Ryan. R. M. (2020).** Self-determination theory in sport and exercise. *Handbook of sport psychology*. 37-56.
45. **Tabachnick, B. G., & Fidell, L. S. (2013).** *Using multivariate statistics* (6th ed.). Pearson Education.
46. **Tabachnick, B. G., & Fidell, L. S. (2015).** *Using multivariate statistics* (6th ed.). Boston: Pearson.
47. **Tabachnick. B. G.. Fidell. L. S.. & Ullman. J. B. (2019).** *Using multivariate statistics* (Vol. 6. pp. 497-516).
48. **Tavşancıl. E. (2014).** *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. Ankara: Nobel Yayıncılık.
49. **Tekin. M.. Özmutlu. İ. & Erhan. S. E. (2009).** Özel yetenek sınavlarına katılan öğrencilerin karar verme ve düşünme stillerinin incelenmesi. *Atatürk Üniversitesi Beden Eğitimi ve Spor Bilimleri Dergisi.* 11 (3). 42-56.
50. **Vallerand. R. J. (1997).** Toward a hierarchical model of intrinsic and extrinsic motivation. *Advances in Experimental Social Psychology*. 29. 271-360.
51. **Vielma-Aguilera. A.. Bustos. C.. Saldivia. S.. & Grandón. P. (2023).** Psychometric properties of the attitudes scale of health care professionals' toward people with a diagnosis of mental illness (EAPS-TM). *Current Psychology*. 42(7). 5851-5863.
52. **Yaşlıoğlu, M. M. (2017).** Sosyal bilimlerde faktör analizi ve geçerlilik: Keşfedici ve doğrulayıcı faktör analizlerinin kullanılması. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, *46*, 74-85.