



# Initial Validation of the Indonesian Version of the RESTQ-Sport-76: A Psychometric Study among Sport Science Students

Anies Setiowati<sup>1\*</sup>, Sri Sumartiningsi<sup>1</sup>, Setya Rahayu MS<sup>1</sup>, Oktia Woro KH<sup>1</sup>, Zahrah Zakiya A<sup>1</sup>, Ade Bagus Pratama<sup>2</sup>,

<sup>1</sup>Universitas Negeri Semarang, Indonesia

<sup>2</sup>Chulalongkorn University, Thailand

\*Correspondence: E-mail: [setiowatianies@mail.unnes.ac.id](mailto:setiowatianies@mail.unnes.ac.id)

## ABSTRACT

The balance between stress and recovery is essential to maintain sport performance and wellbeing. The Recovery-Stress Questionnaire for Athletes (RESTQ-Sport-76) has been widely used internationally to measure psychosocial dimensions. However, the Indonesian version has not undergone formal psychometric evaluation. This study aimed to conduct an initial validation and reliability testing of the Indonesian version of RESTQ-Sport-76 in sport science students. A cross-sectional quantitative design was used, involving 120 sport science students (74% male, aged 18-23 years) recruited through purposive sampling. The instrument, adapted through forward-backward translation and expert review, was distributed via an online form. Data analysis included item-total correlation for construct validity, Cronbach's alpha for internal consistency, and confirmatory factor analysis (CFA) using SPSS AMOS to assess model fit. Most subscales showed good validity ( $r \geq 0.50$ ) and adequate internal reliability ( $\alpha \geq 0.70$ ). The initial CFA indicated suboptimal fit for general stress and recovery dimensions (CFI = 0.666; RMSEA = 0.102), which improved after model revision (CFI = 0.930; RMSEA = 0.055). Similarly, the sport stress and sport recovery model fit improved from CFI = 0.866 to 0.973 after refinement. In conclusion, the Indonesian version of RESTQ-Sport-76 demonstrated satisfactory initial validity and reliability, supporting its use as a preliminary tool for monitoring stress and recovery among sport students. Nonetheless, several subscales, particularly related to conflict/pressure and sleep quality, require further semantic refinement and longitudinal testing before large-scale practical application.

## ARTICLE INFO

### Article History:

Submitted/Received June 2025

First Revised January 2026

Accepted February 2026

Publication Date April 2026

### Keyword:

RESTQ-Sport, stress and recovery, Validity-reliability

## INTRODUCTION

Maintaining an optimal balance between stress and recovery is crucial for sustaining athletes' performance and overall wellbeing. Inadequate recovery from accumulated stress can increase the risk of overtraining, fatigue, and injury, which ultimately impair athletic development and learning outcomes among sport science students (Kellmann, 2001a; Weiss et al., 2024). Sport science students, who are often required to combine academic responsibilities with intensive physical training, face multidimensional stress arising from academic, physical, and social demands. This condition highlights the importance of early identification and monitoring of stress–recovery balance to prevent performance decline and psychological distress.

Several international studies have emphasized the need for standardized psychometric instruments to monitor the interaction between stress and recovery among athletes (Kadooka et al., 2019; Brauers et al., 2024). One of the most widely used instruments is the Recovery-Stress Questionnaire for Athletes (RESTQ-Sport-76), developed by Kellmann and Kallus (2001). This instrument has been adapted and validated in several countries—such as Japan, Spain, and Germany—with consistent reliability and construct validity across cultural contexts (González-Boto et al., 2008; Kadooka et al., 2019). Despite its broad international use, the RESTQ-Sport-76 has not been formally adapted or validated in Indonesia.

In the Indonesian context, stress among student-athletes has become an emerging concern. Studies have reported that sport science students experience higher psychological pressure, fatigue, and recovery challenges compared to non-athlete peers, particularly during intensive training or academic examinations (Hidayati et al., 2021). However, research on reliable and culturally relevant instruments for measuring stress–recovery balance in this population remains limited. Most available tools are translated informally or used without psychometric testing, raising concerns about conceptual equivalence and accuracy of interpretation.

Cross-cultural adaptation of psychological measurement tools requires not only linguistic translation but also conceptual and semantic equivalence to ensure that the instrument measures the same constructs across cultures (Cruchinho et al., 2024; Jayawickreme et al., 2022). Without formal validation, the use of translated questionnaires may produce biased interpretations and limit their applicability in local contexts. Therefore, validating the Indonesian version of the RESTQ-Sport-76 is essential to provide a standardized, reliable, and culturally appropriate tool for assessing stress and recovery among sport science students and athletes.

Based on this background, the present study was designed to conduct an initial validation and reliability testing of the Indonesian version of the RESTQ-Sport-76. This research is expected to fill the methodological gap in the psychometric evaluation of athlete wellbeing instruments in Indonesia and to contribute to more accurate monitoring of stress and recovery in sports education and training settings.

## **METHODS**

### **Research Design**

This study used a descriptive quantitative design with a cross-sectional approach to evaluate the psychometric validity and reliability of the Indonesian version of the RESTQ-Sport-76 instrument. Data collection was conducted over four weeks, from March to April 2024, coinciding with the mid-term period of the academic semester to obtain representative responses from sport science students actively attending lectures and training. This study is part of a broader project on the cultural and linguistic adaptation of psychological instruments in the context of sports science.

### **Research Sample**

This study involved 129 sport science students from Semarang State University, Indonesia. Participants were recruited purposively, with inclusion criteria being students actively participating in a structured sports training program and enrolled in the faculty. Data collection took place in classrooms and at the Faculty of Sports Science training center to maintain a comfortable and familiar academic environment.

A total of 150 students were invited through faculty announcements and online communication platforms, and 129 completed the questionnaire (response rate: 86%). Participation was voluntary and without financial compensation. To minimize selection bias, participants were informed that the study aimed to validate an assessment tool rather than evaluate stress levels or personal performance. All participants received an explanation of the study's purpose and procedures and provided written informed consent prior to participation.

The study protocol was approved by the Ethics Committee of Semarang State University, ensuring confidentiality, anonymity, and data protection.

### **Research Instruments**

The main instrument used in this study was the Indonesian version of the RESTQ-Sport-76, which underwent forward translation, cultural adaptation, and pilot testing. The RESTQ-Sport-76 consists of 76 items measuring 19 scales, divided into two main domains: stress (general and sport-specific) and recovery (general and sport-specific).

It includes twelve general scales (seven stress scales: General Stress, Emotional Stress, Social Stress, Conflict/Pressure, Fatigue, Lack of Energy, and Physical Complaints; and five recovery scales: Success, Social Recovery, Physical Recovery, General Well-being, and Sleep Quality) and seven sport-specific scales (three stress scales: Rest Disturbance, Emotional Exhaustion, and Injury; and four recovery scales: Getting Fit, Personal Achievement, Self-efficacy, and Self-regulation). All items were measured using a 7-point Likert scale ranging from 0 (never) to 6 (very often) (González-Boto et al., 2008).

### **Language Adaptation Procedure**

The linguistic adaptation of the Indonesian version of the RESTQ-Sport-76 followed international guidelines for cross-cultural adaptation of psychological instruments (Brauers et

al., 2024; Cruchinho et al., 2024). This process involved several steps to ensure linguistic accuracy and cultural equivalence.

An independent bilingual translator with a background in linguistics and sport psychology performed forward translation from English into Indonesian. The translated versions were then reviewed and reconciled by a three-member expert panel (one sport psychology academic, one linguist, and one professional coach) to achieve semantic and conceptual consistency.

Back-translation was performed by another translator who was blinded to the original instrument to verify equivalence, and discrepancies were resolved through consensus. Cognitive debriefing was conducted with 15 sport science students (including student-athletes) to assess clarity, comprehensibility, and cultural relevance. Their feedback was used to refine several items. Subsequently, the expert panel assessed content validity using a structured checklist covering relevance, clarity, and conceptual accuracy.

### **Procedure**

Participants were first informed about the purpose, procedures, and confidentiality of the study and provided written consent prior to participation. The Indonesian version of the RESTQ-Sport-76 questionnaire was distributed online via Google Forms, with the link shared through class WhatsApp groups and official announcements to ensure that only eligible sport science students could access it.

The estimated completion time was 15–20 minutes (maximum 60 minutes). Although the questionnaire was self-administered, researchers monitored responses through timestamps and follow-up communication via WhatsApp when necessary. To maintain data quality, each university account was permitted to submit only one response.

Data were screened for duplicate entries, unusually short completion times (less than five minutes), missing values, and inconsistent response patterns. All data were securely stored in a password-protected file and accessed only by the research team, ensuring confidentiality and data integrity.

### **Data Analysis**

Data analysis was conducted using IBM SPSS Statistics version 22.0 and SPSS AMOS version 26. Prior to analysis, the dataset was examined for completeness and accuracy. Responses with missing values or incomplete data were excluded to maintain data integrity. Screening was also conducted to identify duplicate entries and responses completed in unusually short times (<5 minutes).

Descriptive statistics were used to summarize participants' demographic characteristics. Internal reliability of the Indonesian version of the RESTQ-Sport-76 was assessed using Cronbach's alpha for each of the 19 subscales. A value of  $\alpha \geq 0.70$  was considered acceptable, while values between 0.60–0.69 indicated moderate reliability.

At the item level, corrected item–total correlations were calculated to examine each item's contribution to its corresponding subscale. Items with correlation values below 0.30 were considered weak and flagged for further semantic review. The "Cronbach's Alpha if Item

Deleted" index was also examined to determine whether removing specific items could improve internal consistency.

Construct validity was assessed using Confirmatory Factor Analysis (CFA) with Maximum Likelihood Estimation (MLE) in AMOS. Prior to CFA, assumptions of multivariate normality and absence of multicollinearity were evaluated using skewness, kurtosis, and correlation matrix inspection. Model fit was assessed using standard indices:  $\chi^2/df < 3$ , CFI and TLI  $\geq 0.90$ , GFI  $\geq 0.80$ , and RMSEA  $\leq 0.08$ , as recommended in psychometric literature.

When the initial model fit was suboptimal, model respecification was guided by theoretical considerations and Modification Indices (MI). Items with low factor loadings ( $<0.40$ ) or high residual correlations were reviewed and, if necessary, modified or removed to improve model fit without compromising the theoretical structure.

This study focused on internal reliability and construct validity. Other psychometric properties, such as test–retest reliability, external validity, and discriminant validity, were not included in this phase and are recommended for future research.

All statistical analyses were conducted in accordance with ethical standards approved by the Ethics Committee of Universitas Negeri Semarang (Approval No. 246/KEPK/FK/KLE/2025), ensuring that participants' data were handled securely and confidentially.

## RESULTS

### Descriptive Statistics and Item Analysis

This study involved 129 athletes, consisting of 96 males (74.4%) and 33 females (25.6%). Participants were primarily aged between 18 and 23 years, with a mean age of 19.55 years (SD = 1.17). Most participants were 19 years old, accounting for approximately 50% of the total sample.

The sample included participants from various sports, with the largest proportions coming from football (23.3%), running (18.6%), futsal (13.2%), basketball (7.8%), and volleyball (7%). Other sports, such as badminton, swimming, gymnastics, silat, handball, and Muay Thai, were represented by fewer participants, each comprising less than 5% of the sample.

In terms of athletic background, the majority of participants (72.1%) competed at the university level, while 18.6% had regional-level experience, and only 9.3% had participated in national-level competitions. On average, participants reported  $4.6 \pm 2.0$  years of structured sports training experience, with weekly training frequencies ranging from 3 to 6 sessions depending on their sport discipline and academic schedule.

Descriptive analysis showed that mean item scores on the RESTQ-Sport-76 ranged from 1.101 to 4.86. No floor or ceiling effects were observed. Each item had a mean score between 1 and nearly 5, indicating sufficient variability in responses. Participants did not consistently select extreme response options (0 or 6), suggesting an adequate distribution of responses across items.

### Initial Construct Validity Test

The construct validity test aimed to determine whether the items within each scale adequately measured the intended constructs (Clark & Watson, 2019a). In this study,

construct validity was initially examined using item–total correlation analysis as a basic indicator of internal construct consistency. This analysis evaluates the extent to which each item correlates with the total score of its respective subscale.

Correlation values were interpreted as follows: values > 0.30 indicate acceptable validity, values > 0.50 indicate strong validity, and values < 0.30 suggest that the item should be reviewed (Raharjanti et al., 2022).

Table 1 presents the construct validity results of the Indonesian version of the RESTQ-Sport-76 based on item–total correlation analysis. Most items showed correlation values above the minimum threshold of 0.30, indicating a positive contribution to their respective constructs.

Subscales such as General Stress, Emotional Stress, Physical Complaints, Fitness/Being in Shape, and General Well-being demonstrated strong construct validity, with item correlations generally exceeding 0.58 and some above 0.80. The General Stress scale showed consistently high correlations (0.697–0.749), indicating strong internal consistency and adequate representation of the general stress construct. Similarly, the Emotional Stress scale demonstrated good validity, with correlations ranging from 0.596 to 0.757, particularly in Items 2 and 3, which showed the highest values (0.755 and 0.757).

However, several subscales, such as Conflicts/Pressure (0.129–0.384) and Sleep Quality (0.006–0.174), included items with very low correlations, in some cases below 0.20. These findings indicate potential misalignment between certain items and their intended constructs, possibly influenced by linguistic or cultural factors affecting respondents' interpretation.

**Table 1. Construct validity test (item–total correlation)**

	Scale	Item 1	Item 2	Item 3	Item 4
1	General Stress	0.697	0.736	0.749	0.707
2	Emotional Stress	0.678	0.755	0.757	0.596
3	Social Stress	0.520	0.710	0.134	0.532
4	Conflicts/Pressure	0.279	0.129	0.384	0.201
5	Fatigue	0.265	0.631	0.644	0.419
6	Lack of Energy	0.715	0.516	0.538	0.435
7	Physical Complaints	0.674	0.715	0.706	0.586
8	Disturbed Breaks	0.669	0.665	0.427	0.597
9	Emotional Exhaustion	0.499	0.127	0.576	0.602
10	Success	0.443	0.480	0.434	0.608
11	Social Recovery	0.442	0.575	0.513	0.637
12	Physical Recovery	0.641	0.668	0.684	0.488
13	General Wellbeing	0.602	0.663	0.725	0.589
14	Sleep Quality	0.006	0.100	0.113	0.174
15	Fitness/Injury	0.539	0.500	0.515	0.571
16	Fitness/Being in Shape	0.817	0.672	0.742	0.791
17	Personal Accomplishment	0.413	0.573	0.559	0.470
18	Self-regulation	0.704	0.470	0.592	0.670
19	Self-efficacy	0.571	0.155	0.597	0.658

Note: Item: question item in RESTQ-Sport-76

Several explanations may account for these low correlations. First, semantic ambiguity may arise, as terms such as “pressure” or “conflict” can have multiple interpretations within the Indonesian cultural and linguistic context, leading to inconsistent understanding. Second, contextual mismatch may occur because some items were originally developed for professional athletes, whereas this study involved university-level sport science students with different academic and social stressors. Third, response variability—particularly in sleep-related items—may reflect individual differences in habits, environments, and daily routines, thereby weakening item–total correlations.

Despite these limitations, the overall findings support the construct validity of the Indonesian version of the RESTQ-Sport-76, indicating that it is a structurally sound instrument that may require only minor semantic and contextual refinements to improve cultural alignment.

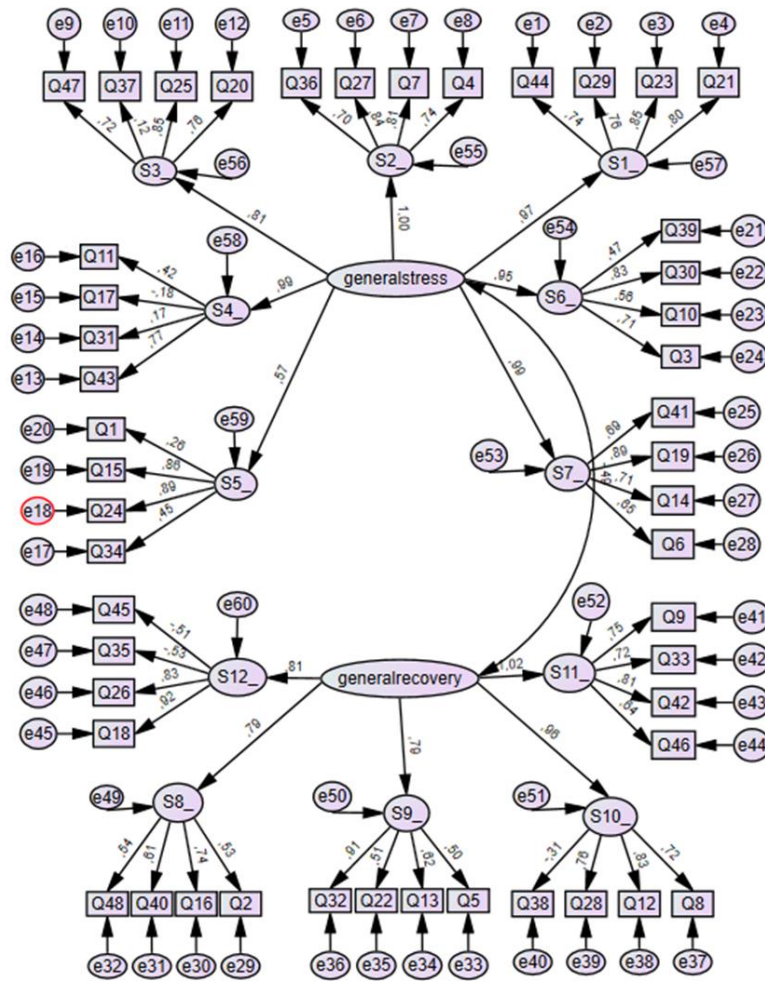
### **Advanced Construct Validity Test (Confirmatory Factor Analysis / CFA)**

#### *General Stress and General Recovery*

A CFA analysis of the two main dimensions of the RESTQ-Sport-76, namely General Stress and General Recovery (Figure 1), was conducted to evaluate the construct validity of the measurement model in a sample of sport science students. The model includes seven subscales and more than 30 indicators, which are assumed to reflect the two primary latent constructs. Given the complexity of the model and the multidimensional nature of stress and recovery, an empirical evaluation using CFA was necessary to determine whether the theoretical structure was supported in the Indonesian sport science student population.

The CFA results indicate that the measurement model did not fully meet the criteria for a good fit. Based on AMOS estimation, the chi-square value of 2500.570 ( $p < 0.001$ ) suggests a significant discrepancy between the hypothesized model and the observed data. In addition, other fit indices were below recommended thresholds, namely GFI = 0.573, CFI = 0.666, and TLI = 0.647. The RMSEA value of 0.102 also indicates a substantial approximation error. These findings suggest that the initial model was not adequately aligned with the empirical data, indicating that certain indicators or aspects of the construct structure required refinement before the model could be considered acceptable.

The path diagram shows that most factor loadings for the General Stress and General Recovery constructs were above 0.50. However, several indicators, such as Q17 (–0.18), Q11 (0.42), Q24, and Q15, showed relatively low contributions, which may have affected the overall model fit and therefore required further evaluation, both semantically and statistically. In particular, the negative loading observed for item Q17 suggests a potential conceptual mismatch or ambiguity in respondents’ interpretation, warranting further consideration beyond purely statistical criteria.

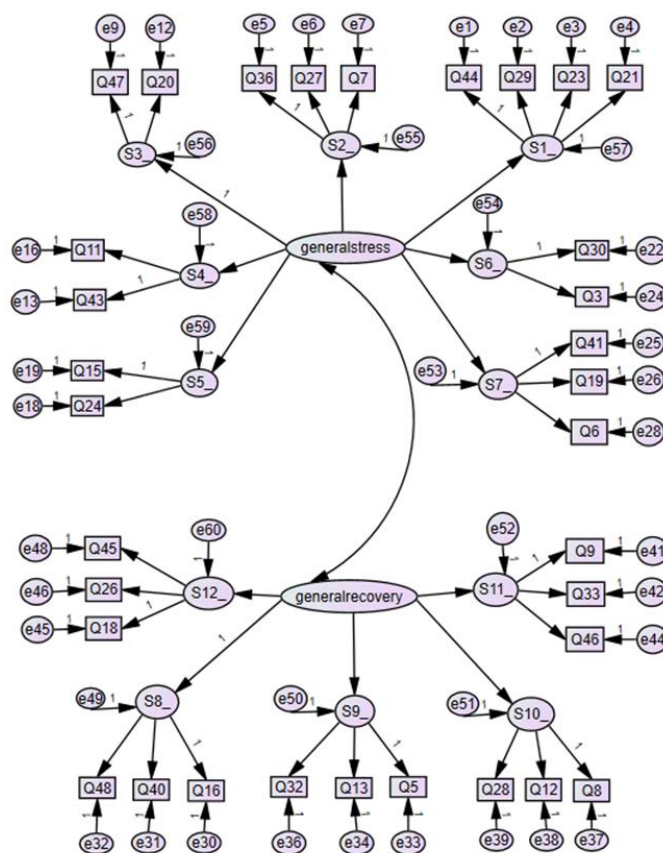


**Figure 1.** Model 1 Path Diagram of Confirmatory Factor Analysis (CFA) for the General Stress and Recovery Dimensions

The CFA model was subsequently revised by removing indicators with negative factor loadings or values below the acceptable threshold (<0.30), as well as by considering Modification Indices (MI) to improve model fit. The decision to remove specific items was guided by both statistical evidence and theoretical considerations, including the conceptual relevance of each item to the latent construct and its semantic clarity within the Indonesian cultural and educational context.

The revised CFA model showed a significant improvement in fit for the General Stress and General Recovery dimensions (Figure 2). Although the chi-square value ( $\chi^2 = 651.4, p < 0.001$ ) remained significant, other fit indices indicated improved model adequacy. It is important to note that the chi-square statistic is highly sensitive to sample size and model complexity; therefore, greater emphasis was placed on alternative fit indices. The CFI (0.930) and TLI (0.921) exceeded the recommended threshold of 0.90, indicating good fit. The GFI value (0.783) remained below the ideal cutoff of 0.90, while the RMSEA (0.055) fell below the acceptable threshold of 0.08, indicating a substantial improvement compared to the initial model.

Although the GFI did not reach the ideal cutoff, this index is known to be sensitive to sample size and the number of estimated parameters. Therefore, the satisfactory values of CFI, TLI, and RMSEA were considered more informative indicators of overall model fit.



**Figure 2.** Model 2 Path Diagram of Confirmatory Factor Analysis (CFA) for the General Stress and General Recovery Dimensions

### Sport Stress and Sport Recovery

Figure 3 presents the Confirmatory Factor Analysis (CFA) results for the Sport Stress and Sport Recovery dimensions of the RESTQ-Sport-76, showing that the measurement model achieved an acceptable fit, although it has not yet reached the optimal level. The chi-square value ( $\chi^2 = 590.875$ ,  $p < 0.001$ ) indicates that the model differs significantly from the observed data; however, this statistic is known to be sensitive to sample size. Therefore, the evaluation of model adequacy was not based solely on the chi-square value but also on complementary fit indices that are less affected by sample size.

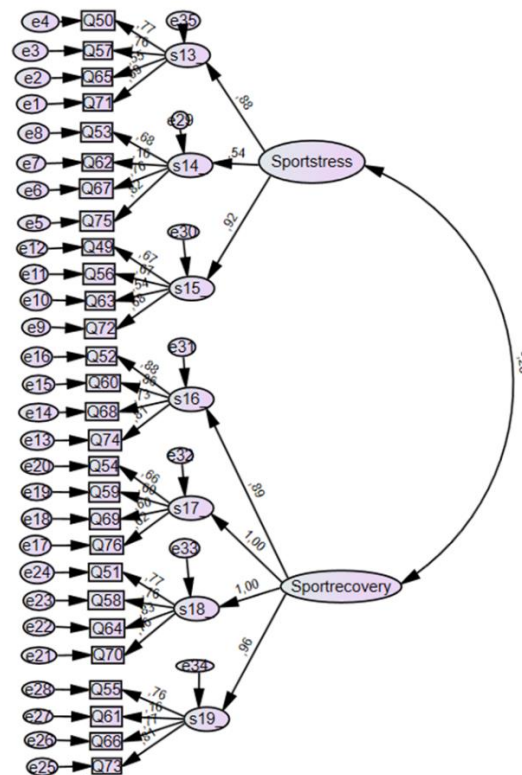
The fit indices showed that the CFI (0.866) and TLI (0.851) were close to the recommended threshold of  $\geq 0.90$ , indicating a near-acceptable model fit, although not yet optimal (Hu & Bentler, 1999). The RMSEA value (0.075) was within the acceptable range ( $< 0.08$ ), suggesting that the model's approximation error was relatively low. Meanwhile, the GFI value (0.763) indicates that the overall model fit still requires improvement, particularly in terms of global fit. These results suggest that, although the model structure was generally acceptable, further refinement was needed to achieve a more optimal representation of the latent constructs.

Regarding the factor structure, the path diagram shows that most indicators within the Sport Stress and Sport Recovery dimensions had factor loadings ranging from 0.60 to 0.88, indicating strong contributions to their respective latent constructs. All indicators loaded positively, with no negative loadings observed, and no parameter identification issues were detected. This suggests that the model was more stable than the General Stress and General Recovery model. This stability indicates that sport-specific stress and recovery constructs may be more clearly defined and contextually relevant for sport science students compared to general stress constructs.

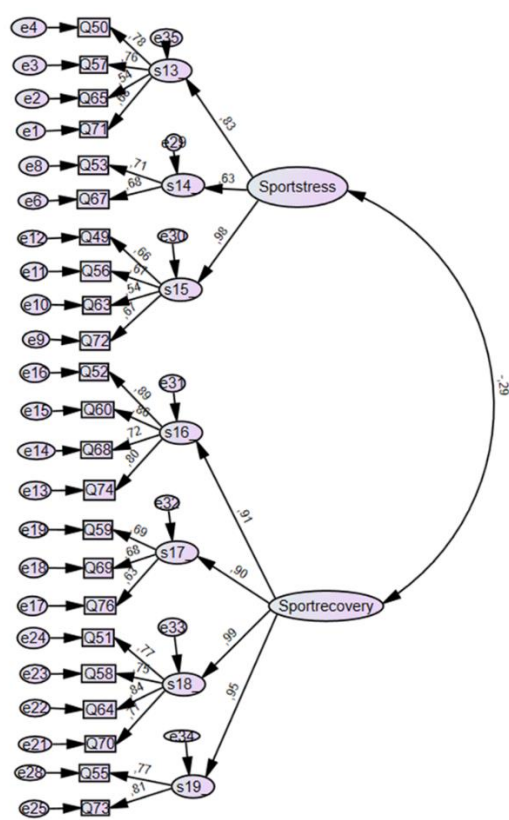
After examining the Modification Indices (MI), error covariances between selected items were added sequentially until an acceptable model fit was achieved. The revised CFA results for the Sport Stress and Sport Recovery dimensions (Figure 4) showed a substantial improvement in model fit. The CFI (0.973), TLI (0.969), and RMSEA (0.037) indicate a good fit, while the GFI (0.855) reached an acceptable level.

All indicators demonstrated strong factor loadings (0.54–0.88), with no negative or weak contributions to the constructs. The factor structure remained stable, with no parameter identification issues. These findings confirm that the revised model demonstrates good construct validity and reliability for measuring sport-specific stress and recovery in the Indonesian sport science student population.

Overall, these improvements indicate that the revised model achieves both statistical robustness and theoretical coherence, supporting its suitability for assessing sport-specific stress and recovery.



**Figure 3.** Model 1 Path Diagram of Confirmatory Factor Analysis (CFA) for the Sport Stress and Sport Recovery Dimensions



**Figure 4.** Model 2 Path Diagram of Confirmatory Factor Analysis (CFA) for the Sports Stress and Sports Recovery Dimensions

### Internal Reliability Test

The internal reliability test was conducted to evaluate the consistency among items within each scale of the Indonesian version of the RESTQ-Sport-76 instrument. Internal reliability was assessed using Cronbach’s alpha coefficient, which is generally considered adequate if the value is  $\geq 0.70$  (Greco et al., 2018). However, in the context of early exploratory research or newly adapted instruments, alpha values  $\geq 0.60$  are considered acceptable (Taber, 2018). A higher Cronbach’s alpha coefficient indicates that the items within a scale consistently measure the same construct.

In psychometric validation studies, particularly during the initial stages of cross-cultural adaptation, variations in reliability across subscales are expected and may reflect differences in construct clarity, cultural interpretation, or item relevance within the target population.

Based on the internal reliability results presented in Table 2, most subscales of the Indonesian version of the RESTQ-Sport-76 demonstrated relatively high Cronbach’s alpha values, indicating good internal consistency across the majority of scales. The highest reliability coefficients were observed in the General Stress ( $\alpha = 0.866$ ) and Fitness/Being in Shape ( $\alpha = 0.887$ ) subscales, suggesting that the items consistently measure the constructs of general stress and physical fitness. The General Well-being subscale also showed strong reliability ( $\alpha = 0.820$ ), indicating that aspects of overall well-being are measured consistently within this instrument. These findings support the consistency of core stress and recovery

constructs among Indonesian sport science students and reinforce the instrument's applicability for general monitoring purposes.

**Table 2.** Subscales of the RESTQ-Sport-76, Mean Scores (M), Standard Deviations (SD), and Cronbach's Alpha

	Scale	S/R	N	Example Statements	Mean	Cronbach $\alpha$
1	General Stress	s	4	<i>I feel anxious or worried</i>	1.829	0.866
2	Emotional Stress	s	4	<i>I feel sad or depressed</i>	1.969	0.853
3	Social Stress	s	4	<i>I feel lonely</i>	2.210	0.667
4	Conflicts/Pressure	s	4	<i>I feel pressured by other people's expectations</i>	2.959	0.440
5	Fatigue	s	4	<i>I feel very tired</i>	2.093	0.647
6	Lack of Energy	s	4	<i>I feel low on energy</i>	1.983	0.759
7	Physical Complaints	s	4	<i>I have headaches</i>	1.857	0.836
8	Disturbed Breaks	s	4	<i>I don't get enough rest</i>	2.180	0.781
9	Emotional Exhaustion	s	4	<i>I feel emotionally exhausted</i>	1.527	0.696
10	Success	r	4	<i>I feel successful in my training</i>	3.459	0.752
11	Social Recovery	r	4	<i>I feel supported by my friends</i>	3.793	0.746
12	Physical Recovery	r	4	<i>I feel physically recovered</i>	3.333	0.803
13	General Wellbeing	r	4	<i>I feel happy</i>	3.517	0.820
14	Sleep Quality	r	4	<i>I sleep well</i>	2.516	0.122
15	Fitness/Injury	s	4	<i>I feel my body is injured</i>	2.004	0.736
16	Fitness/Being in Shape	r	4	<i>I feel fit and in good shape</i>	3.048	0.887
17	Personal Accomplishment	r	4	<i>I feel satisfied with my personal achievements</i>	3.378	0.716
18	Self-regulation	r	4	<i>I am able to manage myself</i>	3.484	0.859
19	Self-efficacy	r	4	<i>I believe in my own abilities</i>	3.436	0.691

**Note:** s=stress; r=recovery; N=number of items; values represent mean scores, standard deviations, and Cronbach's alpha coefficients.

Several other subscales also demonstrated good reliability, including Emotional Stress ( $\alpha = 0.853$ ), Physical Recovery ( $\alpha = 0.803$ ), and Physical Complaints ( $\alpha = 0.836$ ), indicating that these dimensions are reliably measured within this population. The Self-regulation ( $\alpha = 0.859$ ) and Success ( $\alpha = 0.752$ ) subscales also showed acceptable internal consistency, supporting their use in assessing athletes' self-regulatory capacity and perceived achievement. From a practical perspective, these subscales may be used by coaches and sport psychologists to assess emotional and physical conditions related to training demands and performance readiness.

Several subscales showed moderate reliability, including Fatigue ( $\alpha = 0.647$ ), Emotional Exhaustion ( $\alpha = 0.696$ ), and Social Stress ( $\alpha = 0.667$ ). These values indicate acceptable but limited internal consistency and should therefore be interpreted with caution. Variability in participants' perceptions of fatigue and social stress may have contributed to these results. In applied contexts, these subscales are better used alongside additional indicators or repeated measurements to support more reliable decision-making.

The Conflicts/Pressure ( $\alpha = 0.440$ ) and Sleep Quality ( $\alpha = 0.122$ ) subscales demonstrated low reliability. The low coefficient for Conflicts/Pressure suggests that the items within this subscale may lack homogeneity or that participants interpreted the concept of conflict inconsistently. Similarly, the very low reliability of the Sleep Quality subscale indicates that

the construct may not be consistently captured by the current items. This may be due to variability in individual sleep patterns, environmental factors, or differences in interpretation of sleep-related items.

These findings indicate that the Conflicts/Pressure and Sleep Quality subscales require particular attention in future research, including semantic refinement, item revision, or reconsideration of their conceptual relevance for university-level sport science students. In practical applications, the use of these subscales as standalone indicators is not recommended until further validation is conducted, as low internal consistency may compromise the accuracy of stress and recovery assessments.

## DISCUSSION

The results of this study indicate that the Indonesian version of the RESTQ-Sport-76 demonstrates adequate psychometric properties for monitoring stress and recovery among sport science students in Indonesia. Most subscales showed strong construct validity, as indicated by item–total correlations  $> 0.50$ , and good internal reliability ( $\alpha > 0.70$ ). These findings are consistent with those of Kadooka et al. (2019), who reported similar results in validating the Japanese version of the RESTQ-Sport-76. This consistency supports the reliability of the instrument across different cultural contexts, as effective cross-cultural adaptation requires not only linguistic translation but also conceptual and semantic equivalence (Cruchinho et al., 2024b; Zaragoza-Salcedo et al., 2023). Thus, the present findings suggest that the Indonesian version generally preserves the core psychometric structure of the original instrument while remaining sensitive to local population characteristics.

The low validity and reliability observed in subscales such as Conflicts/Pressure and Sleep Quality may be attributed to cultural differences in interpreting these constructs. Cultural factors significantly influence how individuals perceive and respond to psychological items, which is critical for accurate assessment. This is supported by studies emphasizing the importance of culturally adapting psychological measures (Clark & Watson, 2019b; Jayawickreme et al., 2022; Shan & Ji, 2024). In the Indonesian collectivistic context, psychological pressure is often perceived implicitly and closely associated with social expectations, academic demands, and interpersonal harmony rather than explicit interpersonal conflict or competitive stress. These cultural characteristics may influence how sport science students interpret items related to conflict and pressure, resulting in lower item correlations compared to professional athlete samples in Western contexts.

Further construct validity analysis of the two main dimensions of the RESTQ-Sport-76, namely General Stress and General Recovery, was conducted using Confirmatory Factor Analysis (CFA) in AMOS. In Model 1, the analysis was performed on the initial structure consisting of 12 subscales and 48 items. The results indicated that the model did not meet acceptable fit criteria. The CFI (0.666), TLI (0.647), and RMSEA (0.102) values suggest that the theoretical factor structure was not well aligned with the empirical data. In addition, several indicators showed low ( $<0.40$ ) or even negative factor loadings, contributing to model instability and poor fit indices. CFA models with a large number of parameters and problematic indicators tend to yield poor model fit and may lead to identification issues (Brown, 2015). Previous research also shows that overly complex models with method effects may result in low or insignificant factor loadings, indicating a mismatch between the

theoretical structure and observed data (Geiser et al., 2015). These findings highlight that the theoretical structure of the RESTQ-Sport-76 cannot be assumed to function equivalently across different populations without empirical validation.

Based on these findings, Model 2 was revised by removing 16 items with low factor loadings or weak semantic contributions to the latent constructs. This revision resulted in a substantial improvement in model fit, with CFI increasing to 0.930, TLI to 0.921, and RMSEA decreasing to 0.055, all indicating good model fit. These improvements suggest that the revised model better represents the underlying latent structure and is statistically more stable (Jeon et al., 2018; Kline, 2016). Importantly, item removal decisions were guided not only by statistical criteria but also by theoretical relevance and contextual appropriateness for sport science students, ensuring conceptual coherence.

The removal of problematic items is a common procedure in CFA to improve construct validity when supported by both statistical and theoretical considerations (Byrne, 2010). Modifications based on Modification Indices (MI) are acceptable as long as they remain theoretically justified (Xiong et al., 2025). In some cases, constraining parameters relative to a general factor may reduce model complexity without compromising theoretical integrity, thereby enhancing interpretability and statistical stability (Jeon et al., 2018). Nevertheless, item removal may reduce construct coverage, as certain nuanced aspects of stress or recovery may become underrepresented. Therefore, the modifications in this study should be viewed as provisional refinements aimed at improving measurement precision rather than permanent exclusions of conceptual domains.

Previous psychometric studies of the RESTQ-Sport have also reported limitations in the original English version (Davis et al., 2007; Martinet et al., 2014) as well as in translated versions, including the French adaptation. However, most studies suggest that these limitations can be addressed through restructuring of subscales. Although alternative models have been proposed (Davis et al., 2006; Davis et al., 2007), they have rarely been empirically tested in subsequent research.

The Sport Stress and Sport Recovery dimensions assess psychophysiological aspects directly related to stress experiences and recovery processes in sport settings. Initial CFA results for these dimensions indicated moderate model fit (CFI = 0.866; TLI = 0.851; RMSEA = 0.075), suggesting an acceptable but improvable structure. Several indicators showed low factor loadings and high residual correlations, which may reflect content redundancy or differences in semantic interpretation. Practically, these findings indicate that sport science students are generally able to distinguish between sport-related stress and recovery processes, although item interpretation remains influenced by academic and social contexts.

Following model revision—through the removal of items Q1, Q5, Q54, Q62, Q66, and Q75 due to low statistical and semantic contributions, as well as adjustments based on Modification Indices—the model fit improved substantially (CFI = 0.973; TLI = 0.969; RMSEA = 0.037), indicating excellent fit.

Theoretically, Sport Stress reflects pressures arising from training demands, competition, performance expectations, and physical and emotional fatigue, whereas Sport Recovery represents psychological, physical, and social resources that support regeneration (Kellmann, 2001). The strong validity and internal consistency of the revised subscales suggest that sport

science students can meaningfully differentiate between sport-related stress and recovery strategies within the Indonesian cultural context.

These findings reinforce the importance of empirically testing factor structures to ensure that theoretical constructs are adequately represented by their indicators (Prudon, 2015). The more parsimonious and statistically valid revised structure also improves practical efficiency in assessment settings, particularly for sport science students. However, item reduction should be followed by further evaluation of reliability and external validity to ensure comprehensive construct representation (Nye, 2023). Items in this study were removed not only due to low statistical loadings but also based on expert evaluation in linguistics and sport psychology to preserve theoretical integrity.

Several limitations of this study should be acknowledged. First, the relatively small sample size ( $n = 120$ ) and the inclusion of participants from a single institution may limit generalizability and affect model stability. Kline (2016) recommends a minimum sample size of 200 for complex CFA models to ensure stable parameter estimation. Second, the cross-sectional design does not allow for the assessment of temporal stability (test–retest reliability) of the RESTQ-Sport-76, limiting conclusions about changes over time. Third, although the language adaptation followed standard procedures, cultural and semantic influences on stress and recovery perceptions were not explored qualitatively.

Future research is therefore recommended to employ longitudinal designs and qualitative approaches, such as cognitive interviews, to further examine item interpretation and cultural relevance.

## **CONCLUSION**

This study demonstrates that the Indonesian version of the RESTQ-Sport-76 has satisfactory construct validity and internal reliability for use among university-level sport science students. Revisions to the Confirmatory Factor Analysis (CFA) model substantially improved model fit for both general and sport-specific stress–recovery dimensions. Nevertheless, several subscales—particularly Conflicts/Pressure and Sleep Quality—exhibited psychometric weaknesses, indicating the need for further semantic and contextual refinement within the Indonesian cultural context.

From a practical perspective, this instrument can be used by coaches, sport psychologists, and educators as a monitoring tool to assess stress–recovery balance and support athlete well-being and performance. The main limitations of this study include its cross-sectional design and the restricted sample scope. Therefore, future research employing longitudinal designs, more diverse athlete populations, and complementary qualitative approaches is recommended.

Overall, this study provides an important initial contribution to the development of a culturally sensitive and psychometrically sound instrument for assessing stress–recovery balance in sport science and applied settings in Indonesia.

## **ACKNOWLEDGMENT**

The authors would like to thank all individuals who contributed to the writing of this article. The authors are especially grateful to the Faculty of Sports Science and the Faculty of Medicine, Universitas Negeri Semarang.

### CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this article. The authors also confirmed that this paper is free of plagiarism.

### REFERENCES

- Brauers, J. J., Den Hartigh, R. J., Jakowski, S., Kellmann, M., Wylleman, P., Lemmink, K. A., & Brink, M. S. (2024). Monitoring the recovery-stress states of athletes: Psychometric properties of the acute recovery and stress scale and short recovery stress scale among Dutch and Flemish athletes. *Journal of sports sciences*, 42(2), 189-199.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Byrne, B. M. (2001). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*. Britania Raya: Taylor & Francis.
- Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological assessment*, 31(12), 1412.
- Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., Miranda da Silva, M., ... & Gaspar, F. (2024). Translation, cross-cultural adaptation, and validation of measurement instruments: a practical guideline for novice researchers. *Journal of Multidisciplinary Healthcare*, 2701-2728.
- Davis IV, H., Orzeck, T., & Keelan, P. (2007). Psychometric item evaluations of the Recovery-Stress Questionnaire for athletes. *Psychology of sport and exercise*, 8(6), 917-938.
- Geiser, C., Bishop, J., & Lockhart, G. (2015). Collapsing factors in multitrait-multimethod models: Examining consequences of a mismatch between measurement design and model. *Frontiers in Psychology*, 6, 946.
- González-Boto, R., Salguero, A., Tuero, C., & Márquez, S. (2009). Validez concurrente de la versión española del Cuestionario de Recuperación-Estrés para Deportistas (RESTQ-Sport). *Revista de psicología del deporte*, 18(1), 53-72.
- Gonzalez-Boto, R., Salguero, A., Tuero, C., Marquez, S., & Kellmann, M. (2008). Spanish adaptation and analysis by structural equation modeling of an instrument for monitoring overtraining: the recovery-stress questionnaire (RESTQ-SPORT). *Social Behavior and Personality: an international journal*, 36(5), 635-650.
- Greco, L. M., O'Boyle, E. H., Cockburn, B. S., & Yuan, Z. (2018). Meta-analysis of coefficient alpha: A reliability generalization study. *Journal of Management Studies*, 55(4), 583-618.
- Hidayati, H. N., Hayat, B., & Rahayu, W. (2021). Assessment of the validity and reliability of mental health instruments of high school student in Indonesia. *European Journal of Educational Research*, 10(2), 729-742.
- Jayawickreme, N., Knight, S., Rasmussen, A., & Ryder, A. G. (2022). Cultural Issues in Psychological Assessment. In *Comprehensive Clinical Psychology* (pp. 285-300). Elsevier.
- Jeon, M., Rijmen, F., & Rabe-Hesketh, S. (2018). CFA models with a general factor and multiple sets of secondary factors. *psychometrika*, 83(4), 785-808.

- Kadooka, S., Horisawa, S., Araki, M., & Sugo, T. (2019). Reliability and validity of the Japanese version of the Recovery-Stress Questionnaire for Athletes 76. *Japanese Journal of Physical Fitness and Sports Medicine*, 68(2), 133-143.
- Kellmann, M., & K. K. W. (2001). *Recovery-Stress Questionnaire for athletes: User manual*. Champaign, IL: Human Kinetics.
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (D. A. Kenny & T. D. Little, Eds.; Fourth). The Guilford Press.
- Nye, C. D. (2023). Reviewer resources: Confirmatory factor analysis. *Organizational Research Methods*, 26(4), 608-628.
- Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: a critique. *Comprehensive Psychology*, 4, 03-CP.
- Raharjanti, N. W., Wiguna, T., Purwadianto, A., Soemantri, D., Indriatmi, W., Poerwandari, E. K., ... & Levania, M. K. (2022). Translation, validity and reliability of decision style scale in forensic psychiatric setting in Indonesia. *Heliyon*, 8(7).
- Shan, Y., & Ji, M. (2024). Cultural Relevance of Mental Health Scales. In *Cultural Adaptation in Chinese Mental Health Translation* (pp. 27-31). Singapore: Springer Nature Singapore.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education*, 48(6), 1273-1296.
- Xiong, Z., Xia, H., Ni, J., & Hu, H. (2025, February). Basic assumptions, core connotations, and path methods of model modification—using confirmatory factor analysis as an example. In *Frontiers in Education* (Vol. 10, p. 1506415). Frontiers Media SA.
- Zaragoza-Salcedo, A., Oroviogoicochea, C., Saracíbar-Razquin, M. I., & Osácar, E. (2023). The significance of exploring conceptual equivalence within the process of the cross-cultural adaptation of tools: The case of the patient's perception of feeling known by their nurses scale. *Journal of Nursing Scholarship*, 55(6), 1268-1279.