University of Groningen

Monitoring the recovery-stress states of athletes: Psychometric properties of the Acute Recovery and Stress Scale and Short Recovery and Stress Scale among Dutch and Flemish Athletes

Brauers, Jur J.; den Hartigh, Ruud J.R.; Jakowski, Sarah; Kellmann, Michael; Wylleman, Paul; Lemmink, Koen A.P.M.; Brink, Michel S.

DOI:
10.51224/SRXIV.290

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Early version, also known as pre-print

Publication date:
2023

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Brauers, J. J., den Hartigh, R. J. R., Jakowski, S., Kellmann, M., Wylleman, P., Lemmink, K. A. P. M., & Brink, M. S. (2023). *Monitoring the recovery-stress states of athletes: Psychometric properties of the Acute Recovery and Stress Scale and Short Recovery and Stress Scale among Dutch and Flemish Athletes*. SportRxiv. https://doi.org/10.51224/SRXIV.290

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 03-10-2024

SportRχiv

Part of the Society for Transparency, Openness and Replication in Kinesiology (STORK)

Preprint
not peer reviewed

Monitoring the recovery-stress states of athletes: Psychometric properties of the Acute Recovery and Stress Scale and Short Recovery and Stress Scale among Dutch and Flemish Athletes

Data & supplementary materials: https://doi.org/10.34894/A7GZSJ & https://osf.io/zuyme/?view_only=91b67d07c1ac4f07b92df6f01065f481

For correspondence: j.j.brauers@umcg.nl

Jur J. Brauers[1*], Ruud J.R. Den Hartigh[2], Sarah Jakowski[3], Michael Kellmann[3,4], Paul Wylleman[5,6], Koen A.P.M. Lemmink[1], Michel S. Brink[1]

[1] Center for Human Movement Sciences, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [2] Department of Psychology, University of Groningen, Groningen, the Netherlands. [3] Faculty of Sport Science, Ruhr University Bochum, Bochum, Germany. [4] School of Human Movement and Nutrition Sciences, University of Queensland, Brisbane, Australia. [5] Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels, Belgium. [6] Faculty of Psychology and Educational Sciences, Vrije Universiteit Brussel, Brussels, Belgium

Please cite as: Brauers, J.J., Den Hartigh, R.J.R., Jakowsi, S., Kellmann, M., Wyllemann, P. Lemmink, K.A.P.M., Brink, M.S. (2023). Monitoring the recovery-stress states of athletes: Psychometric properties of the Acute Recovery and Stress Scale and Short Recovery and Stress Scale among Dutch and Flemish Athletes. *SportRχiv*.

All authors have read and approved this version of the manuscript. This article was last modified on March 13, 2023.

J.J. Brauers @BrauersJJ, P. Wylleman @PaulWyllemann, K.A.P.M. Lemmink @lemmink, and M.S. Brink @MichelBrink can be reached on Twitter

ABSTRACT

The Acute Recovery and Stress Scale (ARSS) and the Short Recovery and Stress Scale (SRSS) are recently-introduced instruments to monitor recovery and stress processes in athletes. In this study, our aims were to replicate and extend previous psychometric assessments of the instruments, by incorporating recovery and stress dimensions into one model. Therefore, we conducted five confirmatory factor analyses (CFA) and determined structural validity, internal consistency, cross-cultural validity, and construct validity. Dutch and Flemish athletes (N=385, 213 females, 170 males, 2 others, 21.03±5.44 years) completed the translated ARSS and SRSS, the Recovery Stress Questionnaire for Athletes (RESTQ-Sport-76), and information on their last training. There was a good model fit for the replicated CFA, sub-optimal model fit for the models that incorporated recovery and stress into one model, and satisfactory internal consistency (α=.75 – .87). The correlations within and between the ARSS and SRSS, as well as between the ARSS/SRSS and the RESTQ-Sport-76 (r=.31 – -.77 for the ARSS, r=.28 – -.63 for the SRSS) and information of their last training also supported construct validity. The combined findings support the use of the ARSS and SRSS to assess stress and recovery in sports-related research and practice.

INTRODUCTION

Optimizing the balance between recovery and stress can enhance performance and decrease the risk of injury and illness for athletes, making it a critical aspect of training and coaching.[1] Therefore, researchers and practitioners are constantly searching for methods to capture both recovery and stress. Indeed, it is recommended to closely monitor the physical and psycho-social recovery and stress of athletes during the training process, as both recovery and stress are highly “intertwined and interdependent constructs”.[2]

Monitoring practices can be performed by measuring physiological, psychological, biochemical, and immunological responses.[3] However, it is not feasible to collect these measures on a daily basis, because they are often invasive, costly, and time-consuming (e.g., biochemical markers such as creatine phosphokinase need to be derived from blood samples).[4–8] To tackle these disadvantages, athlete monitoring through self-report was introduced as a valid and time-efficient alternative.[9] Furthermore, it is suggested that self-report is more sensitive to training than physiological, biochemical, and immunological measures.[9,10] Therefore, it is mostly the preferred method for athlete monitoring in practice.[11]

Ideally, self-report measures include the cause, intensity, and frequency of recovery and stress-related activities, or their consequences, such as fatigue, muscle soreness, or mood and concentration disturbances. This allows coaches, staff, and athletes themselves to adjust the cause (e.g., adjust the load), or intervene on the process (e.g., cognitive re-structuring) or the consequences (e.g., adapt the recovery strategies).[3,12] Recently, the *Acute Recovery and Stress Scale* (ARSS) and the *Short Recovery and Stress Scale* (SRSS) were developed to assess the emotional,[13] physical, and mental aspects of recovery and stress on a day-to-day basis. Studies conducted in the UK and Germany are promising in light of the psychometric properties of the scales. For instance, the questionnaires showed satisfactory internal consistency and convergent validity was supported by correlations with the *Recovery Stress Questionnaire for Athletes* (RESTQ-Sport-76).[14,15] In addition, all scales of the ARSS were affected by the changing loads of a field hockey training camp.[16] For the full details of the development and psychometric properties in the German and English cohorts, we refer to the manual of the ARSS and the SRSS.[13]

There are, however, important steps to be made to establish the validity of the ARSS and SRSS. First, previous validation studies proceeded from two separate models, which implies that recovery and stress are two independent and unrelated constructs. However, the so-called 'scissors model', as defined by Kallus and Kellmann (2000), suggests that recovery demands and stress states are interrelated.[17,18] Moreover, from a psychometric perspective, it is important to include the interrelation between the domains in one model to test whether the 'scissors model' fits. Second, given the purpose of the ARSS and SRSS to frequently monitor the load and recovery of athletes, it is also important to determine its validity with respect to the daily load and recovery experienced. Currently, the *Rating of Perceived Exertion* (RPE) and *Total Quality of Recovery* (TQR) are widely used in practice,[3,19] and considered as general measures of exertion and recovery. Hence, in a next validation step, the relations between the ARSS/SRSS and these single-item questions need to be determined. Third, although the English and German questionnaires revealed promising initial results for different samples of athletes, validation in a broader population is warranted.

This study therefore aims to advance the ongoing process of validating the ARSS and SRSS among Dutch and Flemish athletes. After we translated the ARSS/SRSS, we replicated the analysis of the structural validity according to the analysis done by Kölling et al (2020) for the purpose of comparison with earlier results. Next, we determined the structural validity with five alternative models that included both the recovery and stress dimensions (for the proposed structure of the models, appendix 1). Then, we followed the COnsensus-based Standards for the selection of Health Measurement INstruments (COSMIN) guidelines and analyzed the internal consistency,[20] cross-cultural validity (between females and males and between types of sport), and construct validity with the RESTQ-Sport-76, RPE, and TQR in a large group of athletes.

MATERIALS AND METHODS

Participants

To properly validate the ARSS/SRSS, we aimed to include at least 320 participants aged 16 years or older from various endurance and team sports. This sample size was chosen according to the upper limit of the rule of thumb that Terwee et al. (2007) proposed for factor analysis (#items*10).[21] To ensure a representative sample of the population, we considered all genders, athletes with and without disabilities, different levels of sports, and athletes from all regions in the Netherlands and Flanders, Belgium. Therefore, the research population was recruited through the Dutch Sports Federation, Flemish sport federations, university student athletes, and from the circle of acquaintances of the researchers. All participants were native Dutch speakers.

The study protocol was approved by the Ethics Committee (PSY-1920-S-0513), and informed consent was obtained from all athletes. Of the 850 athletes we contacted, 385 athletes completed the full questionnaires, which were then considered for analysis. The five sports with the highest number of participants were soccer (*n*=74), athletics (*n*=29), field hockey (*n*=28), volleyball (*n*=27), and basketball (*n*=27). Participants competed at the Olympic (*n*=24), continental (*n*=35), national (*n*=266), or regional (*n*=60) levels. The descriptive statistics of the included athletes are shown in Table 1.

Table 1. Description of the included athletes

Characteristics	*N* = 385
Age (years [SD])	21.03 (5.44)
Gender	
Female (*n* [%])	213 (55.3%)
Male (*n* [%])	170 (44.2%)
Other (*n* [%])	2 (0.5%)
Paralympic (*n* [%])	2 (0.5%)
Level	
Regional (*n* [%])	60 (15.6%)
National (*n* [%])	266 (69.1%)
European (*n* [%])	35 (9.1%)
Olympic (*n* [%])	24 (6.2%)
Average training hours per week (hours [SD])	10.71 (8.58)
Average RPE last training (SD)	13.16 (2.10)
Average duration of last training (hours [SD])	1.66 (0.75)

Table 1. Description of the included athletes

Characteristics	*N* = 385
Average hours since last training (hours [SD])	34.57 (31.92)
Average TQR at this moment (SD)	15.75 (2.68)

Note. Age, average training hours per week, average rating of perceived exertion (RPE) last training, average duration of last training, average hours since last training, and average total quality of recovery (TQR) at this moment are given as mean and standard deviation (SD). Other values are given as the number of participants and their percentages (%).

Translation procedure

The English versions of the ARSS/SRSS were translated through a parallel back-translation procedure.[22] Both questionnaires were translated into Dutch by six sports scientists including academic staff of the [Redacted], and experts in endurance and team sports (i.e., rowing and football). All group members individually translated the items of the English ARSS/SRSS. The English version was used because this version is used worldwide, incorporates extra adjectives, and the outcomes can serve as a reference.[13]

First, the agreements and disagreements between the six translations were analyzed. Items that were identical in at least three out of the six translations were considered to have sufficient agreement. In cases of greater variation, different translations were considered. The following ordered procedure was used for the consideration of the items: a) use in sports, b) translation closest to English, c) German equivalent, and d) use of the Dutch version of the RESTQ-Sport-76. After this procedure, a group of nine sports scientists, sports psychologists, and applied sports scientists (including members who translated the questionnaire) had the opportunity to provide feedback on the Dutch translation of the items. In case of disagreement, consensus was reached after one more round of feedback.

After agreement on the Dutch version, the result was translated into English by a near-native English speaker. Then, the original English questionnaire and the new English questionnaire were compared. Any ambiguities were discussed until a consensus was reached. Finally, the Dutch version was pre-tested with a small group of athletes, who were asked to provide feedback on the questionnaire. Their feedback on the items or questions was used to address ambiguities. No items were added or removed compared to the English version.

Design and measures

Participants received a link to an online environment named Qualtrics (2022).[23] This survey included demographics with questions about age, gender, sport type, and sport level. This was followed by questions about the last training, such as duration (in blocks of 15 minutes), the RPE (on a scale of 6 – 20 “no exertion at all – maximal exertion”),[19,24] TQR (on a scale of 6 – 20 “no recovery at all – maximal recovery”),[3] and time since last training (in half hours). Subsequently, participants filled out the ARSS, SRSS, and RESTQ-Sport-76.

The ARSS consists of a list of 32 adjectives related to recovery and stress that are preceded by the sentence: “at this moment I feel/ I am”. Each item describes a different state of recovery or stress (e.g., “*strong*” or “*muscle exhaustion*”). The items are grouped in eight scales, of which four describe the *Recovery* dimension (*Physical Performance Capability, Mental Performance Capability, Emotional Balance, Overall Recovery*). The four other scales describe the *Stress* dimension (*Muscular Stress, Lack of Activation, Negative Emotional State, Overall Stress*). Means and total scores of these scales are calculated. The SRSS is a compact version of the ARSS and consists of eight items that correspond to the eight scales of the ARSS.[14] For the SRSS, the items of the corresponding ARSS scale serve as descriptors for each item (e.g., *Muscular Stress* is described by *muscle soreness* and *muscle stiffness)*. The items of the SRSS are rated in relation to the highest recovery or stress state of the athlete. Both the ARSS and SRSS items are rated on a Likert-type rating scale from 0 (does not apply at all) to 6 (fully applies). For full details of these questionnaires, we refer to the manual by Kellmann and Kölling (2019).[13]

The Dutch RESTQ-Sport-76 is composed of 76 questions that can be answered on a Likert scale from 0 (never) to 6 (always).[2,25] The statements refer to the frequency of perceptions of stress and of recovery activities in the last week (e.g., *last week, I had muscle pain after performance* or *last week, my body felt strong*). The questionnaire consists of 19 scales, which provide insights regarding non-sport and sport-specific aspects of recovery and stress. For further information see the manual.[26]

Statistical analysis

For analysis, we used R with the packages Lavaan, semTools, and semPlot.[27–30] Descriptive statistics were calculated and the means (M) and standard deviations (SD) were determined for all values.

The structural validity of the ARSS was determined with confirmatory factor analysis (CFA) rather than exploratory factor analysis, because the factor structure has been determined previously.[13,31] The analysis of the structural validity was done in two steps. First, we replicated the steps described by Kölling et al. (2020) and performed three CFAs with robust maximum likelihood estimators (first-order model, hierarchical model, and a bifactor model).[31] Second, we conducted five CFA’s (orthogonal first-order, single-factor, bifactor, oblique lower-order, and a higher order model) in which we included all items to assess the proposed multidimensional structure of the ARSS (for the proposed structures, see appendix 1). Initially, in accordance with the models described by Kölling et al (2020),[31] we allowed correlation between the error variances of the items *strong* and *physically capable, muscle exhaustion* and *muscle fatigue*, as well as between *muscle soreness* and *muscle stiffness*. To describe the global fit of the models, we reported the root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis Index (TLI), and standardized root mean square residual (SRMR). Following the recommendations by Credé and Harms (2015),[32] we did not interpret the global fit indices using arbitrary cut-off values (as these were not developed for higher order models) but rather present the change in the $\chi 2$ statistic when comparing different models using the alternative approach as described by Satorra and Bentler (2001) in combination with reporting the RMSEA, CFI, TLI, and SRMR.[33] Ideally, if the $\chi 2$ statistic is *non*-significant at an alpha level of >.01, the CFI is high, and the SRMR and RMSEA are low, the global fit of the model is assumed good.

Cross-cultural validity of the ARSS/SRSS was determined by multigroup CFA for the subgroups based on gender and type of sport (individual vs. team).[21] Measurement invariance for the multigroup CFA was tested according similar as described by Kölling et al. (2020).[31]

After we determined the structural validity, we assessed the internal consistency with Cronbach's α. Next, the corrected item-total correlation was calculated to assess the strength of the relationship between individual items and the total score of the scale that the item belongs to. Finally, we determined the inter-item correlations between different items within the scale.

Because no gold standard is available, we determined construct validity rather than criterion validity. According to the guidelines proposed by the COSMIN initiative,[20] we formulated hypotheses about the magnitude of the relations within and between the ARSS/SRSS, the RESTQ-Sport-76, RPE, and TQR. Based on previous research,[15,31] we formulated the following hypotheses: 1) there are moderate to large positive correlations within the *Recovery* and *Stress* domains of the ARSS and SRSS, as well as moderate to large negative correlations between the *Recovery* and *Stress* domains of the ARSS and SRSS; 2) there are large to very large positive correlations between the ARSS and SRSS; 3) there are significant positive correlations between the ARSS/SRSS scales with similar dimensions on the RESTQ-Sport-76 and significant negative correlations with opposite scales; and 4) there are significant positive correlations between the exertion and recovery factors (i.e., RPE, TQR) and stress and recovery scales of the ARSS/SRSS. If 75% or more of the proposed hypotheses were confirmed, the concurrent validity of the questionnaire is considered good.[34] The correlation coefficients were determined with Pearson correlations (r) and considered trivial ($r<.1$), small ($.1<r\leq.3$), moderate ($.3<r\leq.5$), large ($.5<r\leq.7$), very large ($.7<r\leq.9$), almost perfect ($r>.9$) or perfect ($r=1$).35 The alpha level was set at .05.

RESULTS

Structural validity

After examining the inter-item correlations, corrected item-total correlations, and the CFA's, item (MPC2; *receptive*) was deleted which had the highest negative contribution to Cronbach's alpha, corrected item-total correlation, and a low factor loading.

First, we replicated the models described by Kölling et al. (2020)(see appendix 2).[31] In addition, we estimated an orthogonal first-order, single-factor, bifactor, oblique lower-order, and a higher order model first-order CFA, a bifactor CFA, and a higher order CFA with all items of the ARSS. Although none of the alternative models reached optimal global fit values, the oblique lower-order model was retained for further analysis as it had the best global fit. Table 2 displays the full details of the factor loadings and appendix 3 the correlation matrix.

Table 2. Results of the confirmatory factor analysis of the Dutch ARSS for the total sample and subsamples.

	Model	χ2	df	p	RMSEA	CFI	TLI	SRMR
Total sample (N = 385)	Orthogonal first-order model	3228.69	430	<.01	.13	.59	.56	.32
	Single-factor model	2549.40	430	<.01	.11	.69	.67	.11
	Bifactor model	1446.57	398	<.01	.08	.85	.82	.09
	Oblique lower-order model	1095.47	402	<.01	.07	.90	.88	.07
	Higher order model	1651.77	420	<.01	.09	.82	.8	.10
Female vs. male sample (n = 383)	Oblique lower-order model	1632.81	804	0	0.07	.88	.86	.07
Team sports vs. individual sports sample (N = 385)	Oblique lower-order model	1669.64	804	0	0.07	.88	.86	.07

Note. ARSS = Acute Recovery and Stress Scale, df = Degrees of Freedom, RMSEA = Root Mean Square Error of Approximation, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, SRMR = Standardized Root Mean Square Residual. Because we did not include the "other" gender in the multigroup CFA, the total number of participants for the multigroup confirmatory factor analysis on gender does not add up to 385.

Internal consistency

The descriptive statistics for the ARSS and SRSS are presented in Tables 3 and 4. Internal consistency of the ARSS scales ranged between α=.59 and α=.87. The corrected item-total correlations ranged between *r*=.47 and *r*=.79, and were all significant. The ARSS also demonstrated good internal consistency for the *Recovery* dimension α=.91 and the Stress dimension α=.90.

As the SRSS is a condensed version of the ARSS, each item was supported by four example adjectives from the ARSS. Internal consistency of the SRSS was good for both *Recovery* (α=.78) and *Stress* (α=.75) dimensions. The corrected item-total correlations ranged from *r*=.55 to *r*=.65 for the *Recovery* dimension and from *r*=.31 to *r*=.72 for the *Stress* dimension.

Table 3. Means, standard deviations (SD), standardized alphas, corrected item-total correlations, and inter-item correlations of the Dutch ARSS for the sample. Scores range between 0 (does not apply) and 6 (fully applies)

Scales; Adjectives		Mean (SD)	α	Corrected item-total correlations	Inter-item correlations
PPC		3.77 (0.99)	.83		
	PPC item 1	3.87 (1.13)		.62	.59
	PPC item 2	4.01 (1.24)		.61	.59
	PPC item 3	3.72 (1.22)		.65	.57
	PPC item 4	3.47 (1.24)		.79	.48
MPC		3.75 (1.03)	.75		
	MPC item 1	3.77 (1.26)		.51	.60
	MPC item 2[a]	3.13 (1.57)		.04	.50
	MPC item 3	3.79 (1.26)		.62	.45
	MPC item 4	3.68 (1.26)		.61	.46
EB		4.14 (0.94)	.77		
	EB item 1	4.54 (1.16)		.55	.48
	EB item 2	4.12 (1.15)		.61	.44
	EB item 3	4.26 (1.19)		.58	.45
	EB item 4	3.64 (1.38)		.56	.47
OR		3.87 (1.08)	.82		
	OR item 1	4.54 (1.26)		.60	.55
	OR item 2	3.48 (1.38)		.55	.58
	OR item 3	3.64 (1.37)		.69	.49
	OR item 4	3.81 (1.35)		.71	.48
MS		1.93 (1.36)	.87		
	MS item 1	2.15 (1.65)		.68	.65
	MS item 2	2.21 (1.58)		.79	.58
	MS item 3	1.66 (1.63)		.78	.58
	MS item 4	1.71 (1.55)		.64	.68
LA		1.66 (1.09)	.77		
	LA item 1	1.28 (1.46)		.47	.52
	LA item 2	1.94 (1.41)		.57	.45
	LA item 3	1.35 (1.33)		.63	.41
	LA item 4	2.06 (1.50)		.60	.43
NES		1.61 (1.18)	.81		
	NES item 1	1.45 (1.45)		.59	.55
	NES item 2	2.10 (1.58)		.58	.56
	NES item 3	1.35 (1.36)		.72	.46
	NES item 4	1.54 (1.49)		.63	.52
OS		1.81 (1.16)	.81		
	OS item 1	2.69 (1.54)		.62	.52
	OS item 2	1.48 (1.37)		.66	.50
	OS item 3	1.45 (1.53)		.51	.60
	OS item 4	1.60 (1.41)		.73	.45

Table 3. Means, standard deviations (SD), standardized alphas, corrected item-total correlations, and inter-item correlations of the Dutch ARSS for the sample. Scores range between 0 (does not apply) and 6 (fully applies)

Scales; Adjectives	Mean (SD)	α	Corrected item-total correlations	Inter-item correlations

Note. PPC = Physical Performance Capability, MPC = Mental Performance Capability, EB = Emotional Balance, OR = Overall Recovery, MS = Muscular Stress, LA = Lack of Activation, NES = Negative Emotional State, OS = Overall Stress, [a] item deleted.

Table 4. Means, standard deviations (SD), standardized alphas, corrected item-total correlations, and inter-item correlations of the Dutch SRSS for the sample. Scores range between 0 (does not apply) and 6 (fully applies).

Item	Mean (SD)	Corrected item-total correlations	Inter-item correlations
PPC	3.78 (1.06)	.58	.46
MPC	3.86 (1.17)	.65	.43
EB	3.81 (1.30)	.55	.49
OR	3.92 (1.16)	.55	.48
MS	2.09 (1.47)	.31	.60
LA	1.79 (1.42)	.55	.43
NES	1.71 (1.47)	.64	.37
OS	2.05 (1.38)	.72	.33

Note. PPC = Physical Performance Capability, MPC = Mental Performance Capability, EB = Emotional Balance, OR = Overall Recovery, MS = Muscular Stress, LA = Lack of Activation, NES = Negative Emotional State, OS = Overall Stress.

Cross-cultural validity

Table 5 displays the results of the multigroup CFA for the subgroups gender and type of sport. The configural and metric models showed comparable fit indices; however, in the scalar model, the fit indices were lower than ideal in both multigroup CFAs.

Table 5. Results of multigroup confirmatory factor analysis for the subgroups gender and type of sport

Measurement invariance		χ2	df	p	RMSEA	CFI	SRMR	Δ df	ΔRMSEA	ΔCFI	ΔSRMR
Gender	Configural	1632.81	804	< .001	.073	.880	.07				
	Metric	1658.04	827	< .001	.072	.879	.08	23	-.001	.000	.005
	Scalar	1786.97	850	< .001	.076	.864	.08	2023	.003	-.014	.002
Type of sport	Configural	1669.64	804	< .001	.075	.88	.07				
	Metric	1698.00	827	< .001	.074	.88	.08	23	-.001	-.001	.004
	Scalar	1742.99	850	< .001	.074	.87	.08	23	.000	-.003	.001

Note. CFI = Comparative Fit Index, SRMR = Standardized Root Mean Residual; RMSEA = Root Mean Error of Approximation

Construct validity

Correlations within and between the ARSS and SRSS.

Table 6 presents the correlations within the ARSS scales and table 7 presents the summary of the hypotheses and results for determining construct validity. Within the *Recovery* dimension (table 6: upper left quadrant) of the ARSS, *r* ranged from .50 to .73. Within the *Stress* dimension (table 6: lower right quadrant), *r* ranged from .22 to .69. The correlations between *Recovery* and *Stress* (upper right quadrant) ranged from -.16 to -.73. All correlations were statistically significant ($p < .05$).

The correlations within the items were ranging from .33 to .61 for the *Short Recovery Scale* and from .18 to .67 for the *Short Stress Scale* (Table 6). The correlations between *Recovery* and *Stress* ranged from -.17 to -.68. All correlation coefficients were significant ($p < .05$).

All correlations between the corresponding scales and items of the ARSS and SRSS (Table 6) were significant ($p < .05$), and could be considered as large to very large (*r* ranged from .65 to .77).

Table 6. Pearson correlations within the Dutch ARSS scales, the Dutch SRSS items, and between the scales/items.

		upper data matrix: ARSS							
Scale/ Item	Between ARSS & SRSS	PPC	MPC	EB	OR	MS	LA	NES	OS
PPC	.74*		.73*	.67*	.60*	-.21*	-.59*	-.43*	-.54*
MPC	.65*	.50*		.65*	.54*	-.22*	-.57*	-.55*	-.52*
EB	.67*	.33*	.61*		.50*	-.16*	-.59*	-.67*	-.48*
OR	.72*	.58*	.40*	.37*		-.63*	-.35*	-.37*	-.73*
MS	.72*	-.26*	-.18*	-.17*	-.53*		.22*	.28*	.67*
LA	.66*	-.39*	-.50*	-.41*	-.22*	.18*		.69*	.60*
NES	.77*	-.35*	-.53*	-.68*	-.35*	.23*	.58*		.67*
OS	.65*	-.46*	-.48*	-.52*	-.47*	.39*	.54*	.67*	
		lower data matrix: SRSS							

Note. The upper matrix describers the correlations within the ARSS scales, the lower matrix describes the correlations within the SRSS items. PPC = Physical Performance Capability, MPC = Mental Performance Capability, EB = Emotional Balance, OR = Overall Recovery, MS = Muscular Stress, LA = Lack of Activation, NES = Negative Emotional State, OS = Overall Stress, * = p <.05.

Table 7. Predefined hypotheses and instances in which the hypotheses are confirmed.

Hypothesis	Confirmed in:
There are trivial to large positive correlations of .3 to .7 within the Recovery domain of the ARSS and Stress domains of the ARSS.	6/6
There are trivial to large positive correlations of .3 to .7 within the Recovery domain of the ARSS and Stress domains of the SRSS.	4/6
There are trivial to large positive correlations of .3 to .7 within the Stress domain of the ARSS and Stress domains of the ARSS.	6/6

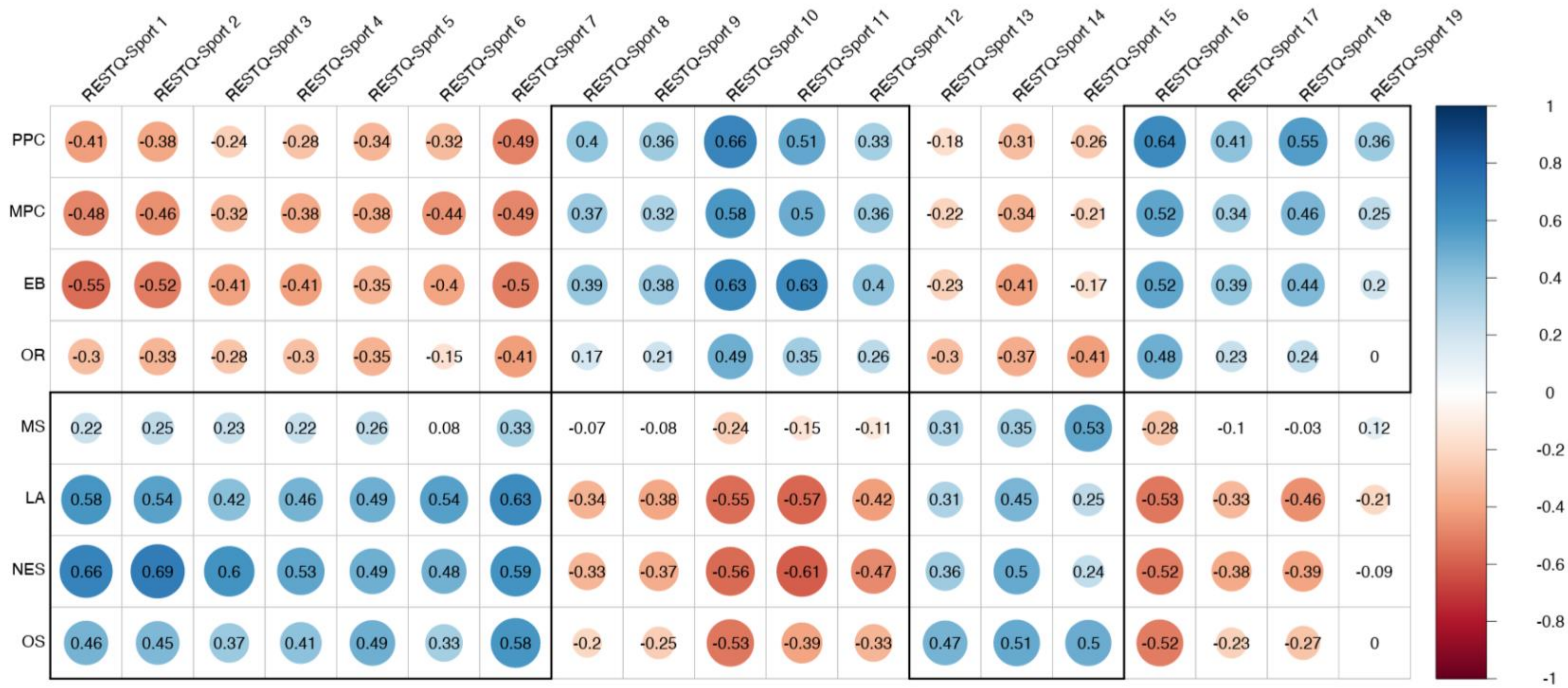
There are trivial to large positive correlations of .3 to .7 within the Stress domain of the ARSS and Stress domains of the SRSS.	4/6
There are trivial to large negative correlations of .3 to .7 between the Recovery and Stress domains of the ARSS.	13/16
There are trivial to large negative correlations of .3 to .7 between the Recovery and Stress domains of the SRSS.	12/16
There are moderate to large positive correlations of .5 to .7 between the ARSS and SRSS;	8/8
There are significant correlations between the ARSS scales and SRSS items with similar dimensions on the RESTQ-Sport-76.	291/304
There are significant positive correlations between the exertion and recovery factors (i.e., RPE, TQR) and stress and recovery scales of the ARSS.	9/16
There are significant positive correlations between the exertion and recovery factors (i.e., RPE, TQR) and stress and recovery items of the SRSS.	6/16
Total confirmed	359/400

Note. Correlations are considered trivial ($r < .1$), small ($.1 < r \leq .3$), moderate ($.3 < r \leq .5$), large ($.5 < r \leq .7$), very large ($.7 < r \leq .9$), almost perfect ($r > .9$) or perfect ($r = 1$).

Correlations between RESTQ-Sport-76, ARSS and SRSS.

Convergent validity of the ARSS and SRSS was assessed by examining their scores in relation to the RESTQ-Sport-76 (Figure 1). The ARSS *Negative Emotional State* showed the largest correlation with the RESTQ-Sport-76 *Emotional Stress* (r=.69). Accordingly, the SRSS *Negative Emotional State* showed the largest correlation with the RESTQ-Sport-76 *Emotional Stress* (r=.66). Overall, 291 of the 304 correlations were significant, and the coefficients were moderate to large, while the coefficients with the SRSS were consistently smaller. Considering the hypothesis-relevant relations between the different questionnaires, a congruent pattern was found for the ARSS and SRSS. This pattern showed positive correlations among the related areas and negative correlations between the opposite areas for both stress and recovery. For example, the ARSS's *Muscular Stress* showed larger coefficients with the RESTQ-Sport-76 *Injury* (r=.53) scale, but not with *Self-efficacy* (r=-.03) or *Self-regulation* (r=-.12).

a


RESTQ-Sport 1
RESTQ-Sport 2
RESTQ-Sport 3
RESTQ-Sport 4
RESTQ-Sport 5
RESTQ-Sport 6
RESTQ-Sport 7
RESTQ-Sport 8
RESTQ-Sport 9
RESTQ-Sport 10
RESTQ-Sport 11
RESTQ-Sport 12
RESTQ-Sport 13
RESTQ-Sport 14
RESTQ-Sport 15
RESTQ-Sport 16
RESTQ-Sport 17
RESTQ-Sport 18
RESTQ-Sport 19
PPC -0.41 -0.38 -0.24 -0.28 -0.34 -0.32 -0.49 0.4 0.36 0.66 0.51 0.33 -0.18 -0.31 -0.26 0.64 0.41 0.55 0.36
MPC -0.48 -0.46 -0.32 -0.38 -0.38 -0.44 -0.49 0.37 0.32 0.58 0.5 0.36 -0.22 -0.34 -0.21 0.52 0.34 0.46 0.25
EB -0.55 -0.52 -0.41 -0.41 -0.35 -0.4 -0.5 0.39 0.38 0.63 0.63 0.4 -0.23 -0.41 -0.17 0.52 0.39 0.44 0.2
OR -0.3 -0.33 -0.28 -0.3 -0.35 -0.15 -0.41 0.17 0.21 0.49 0.35 0.26 -0.3 -0.37 -0.41 0.48 0.23 0.24 0
MS 0.22 0.25 0.23 0.22 0.26 0.08 0.33 -0.07 -0.08 -0.24 -0.15 -0.11 0.31 0.35 0.53 -0.28 -0.1 -0.03 0.12
LA 0.58 0.54 0.42 0.46 0.49 0.54 0.63 -0.34 -0.38 -0.55 -0.57 -0.42 0.31 0.45 0.25 -0.53 -0.33 -0.46 -0.21
NES 0.66 0.69 0.6 0.53 0.49 0.48 0.59 -0.33 -0.37 -0.56 -0.61 -0.47 0.36 0.5 0.24 -0.52 -0.38 -0.39 -0.09
OS 0.46 0.45 0.37 0.41 0.49 0.33 0.58 -0.2 -0.25 -0.53 -0.39 -0.33 0.47 0.51 0.5 -0.52 -0.23 -0.27 0
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

b

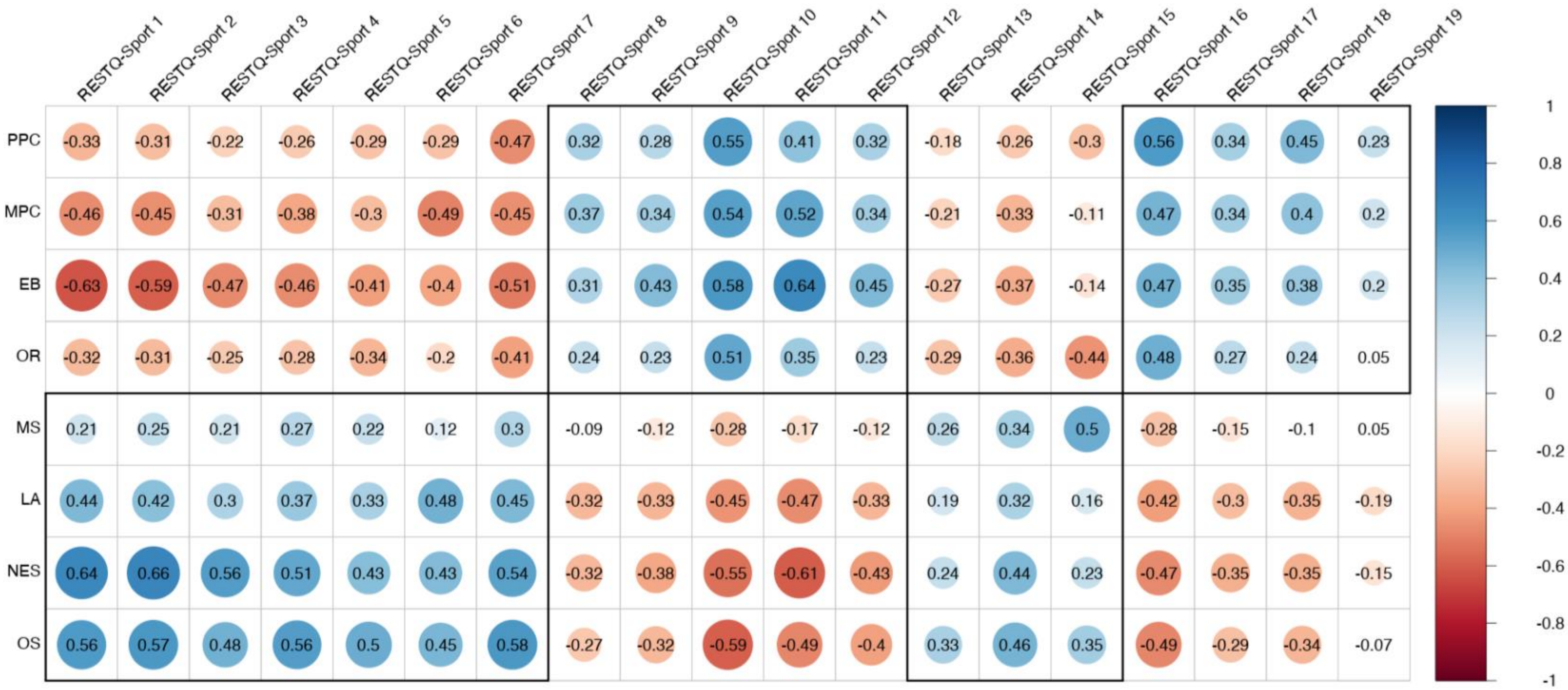


Figure 1. Pearson correlations between the ARSS (a) scales and SRSS (b) items and the RESTQ-Sport-76 scales (significant correlations are colored red (negative) or blue (positive); nonsignificant correlations are colored white). Correlations follow the expected pattern (that is, positive correlations with similar domains, negative correlations with opposite dimensions, and similar magnitudes of correlations as Nässi et al., 2017 and Kölling et al., 2020).

Note. PPC = Physical Performance Capability, MPC = Mental Performance Capability, EB = Emotional Balance, OR = Overall Recovery, MS = Muscular Stress, LA = Lack of Activation, NES = Negative Emotional State, OS = Overall Stress, 1 = General Stress; 2 = Emotional Stress; 3 = Social Stress; 4 = Conflicts/Pressure; 5 = Fatigue; 6 = Lack of Energy; 7 = Physical Complaints; 8 = Success; 9 = Social Recovery; 10 = Physical Recovery; 11 = General Well-being; 12 = Sleep Quality; 13 = Disturbed Breaks; 14 = Emotional Exhaustion; 15 = Injury; 16 = Being in Shape; 17 = Personal Accomplishment; 18 = Self-Efficacy; 19 = Self-Regulation.

Correlation between RPE and TQR and ARSS and SRSS.

Finally, we calculated the correlations between the ARSS/SRSS, RPE, and TQR (N=385). There were significant correlations between the RPE and Overall Recovery (r=-.23) and Muscular Stress (r=.19). All scales of the ARSS except of Lack of Activation were significantly correlated with the TQR, of which Overall Recovery (r=.63), Muscular Stress (r=-.63), and Overall Stress (r=-.51) were the largest. Similar patterns were found for the SRSS, although the magnitude of the correlations was smaller, and fewer correlations were significant (see Appendix 4).

Discussion

This study aimed to determine structural validity, internal consistency, cross-cultural validity, and construct validity of the ARSS and SRSS in a large group of Dutch-speaking athletes. Our findings indicate that these short questionnaires are easy to administer and that their structural validity, internal consistency, cross-cultural validity, and construct validity are sufficient.

The descriptive statistics of the ARSS and SRSS were similar to those previously reported.[14,15,31] When comparing the means and standard deviations of all specific adjectives for the ARSS, the results most closely resembled those of Kölling et al. (2020) and Hitzschke et al. (2016).[14,31] The difference in mean scores between the current study and these studies only exceeded one point for the first item of *Overall Recovery* for the study by Hitschke et al. (2016).[14]

To determine the structural validity, we replicated the analyses as described by Kölling et al. (2020) and retrieved similar global fit indices.[31] Next, we applied multiple CFAs that included both the recovery and stress dimensions. On the whole, at least psychometrically speaking, *Recovery* and *Stress* can be seen as intertwined and interdependent constructs with clear, and correlated, underlying scales but the global fit of this model can still be improved. For instance, during the estimation of the final model, a warning was generated indicating that the variance-covariance matrix of the estimated parameters was not positive definite. This may be caused by multicollinearity or by having too few observations relative to the number of parameters in the model. Despite this warning, the results of the model are presented for the sake of completeness, but should be interpreted with caution. In addition, the scales were mostly stable for gender and type of sport. Results for the subgroups should be interpreted with caution, however, because the sample sizes for these analyses were low.

The deletion of the *Mental Performance Capability* item receptive increased internal consistency, with all scales being above the reliability threshold of α>.70. Next to that, the corrected item-total correlations (r>.47) and inter-item correlations (r>.41) all provided evidence for a good reliability of the scale. These high values are in accordance with validation studies among German and English samples.[16,31] We discussed several possible alternative items and checked with the authors of the original questionnaire whether there were German items that could be considered. However, the items that they proposed were already covered in the final Dutch translation, and there was a consensus among all translators to delete the item in the final translation.

To determine the construct validity, we specified five hypotheses a priori based on previous studies.[15,31] Of all the possible correlations, 90% were as expected. However, 4 out of 19 correlations between the *Muscular Stress* scale of the ARSS and the RESTQ-Sport-76, and 4 out of 8 correlations between the scales of the ARSS and the RESTQ-Sport-76 *Self-regulation* scale were not significant. In addition, there were some small differences with the previous studies. For the ARSS, 51 out of 304 (17%) correlations had a difference of more than .10 (but almost always less than .20) compared with the study by Kölling et al. (2020).[31] Compared to Nässi et al. (2017),[15] 63 out of 304 (21%) correlations differed by more than .10 (but almost always less than .20). The correlations between the SRSS and the RESTQ-Sport-76 showed fairly the same results as the results of the ARSS scales, but were somewhat smaller. Within the *Short Recovery Scale* of the SRSS, the highest correlations were found with the RESTQ-Sport-76 scales of *Overall Recovery* and *Physical Recovery*. The slightly weaker relations for the SRSS items could be explained by the fact that the SRSS exists of one question, whereas the scales of ARSS and RESTQ-Sport-76 are based on four questions, which makes the outcome more robust.[36] A possible explanation for these differences could be that the Dutch RESTQ-Sport-76 refers to the preceding week, whereas the German and English versions refers to the preceding three days/nights.

To assess whether the ARSS and SRSS scores were related to previous training, we studied the association between the preceding training RPE and subsequent recovery. There were some significant but trivial correlations between the RPE and ARSS and SRSS. An explanation for the trivial correlations is most likely that enough time elapsed for full recovery between the end of the training and the moment that the athlete completed the questionnaire.[37] Another explanation could be the reduced precision of RPE as a result of recall bias.[38] Kölling et al. (2015),[16] found a relation between the intensity rated by the coach and the ARSS. However, they collected the ARSS twice a day, before and after training, during a 5-day training camp, thus excluding recall bias. Furthermore, the present study neglects the timeframe between training and measurement of stress and recovery states as the questionnaires were filled out independent of training. Thus, it is unknown what kind of activities an athlete has engaged in since their last training. This means that athletes can be exposed to stressors in daily life, or perform recovery-enhancing activities after the training.39 This could theoretically influence the relation between the preceding RPE and the score on the ARSS or SRSS at a later moment.[39]

However, there were more significant relations between the TQR and the ARSS and SRSS. One could argue that TQR encompasses recovery and integrates different dimensions into one question. This suggests that TQR could serve as an early marker and ARSS could be used to distinguish between recovery dimensions.[3] Because the TQR was completed at the same time as the ARSS and SRSS, and there was also no influence of other stressors since the last training.[3] Consequently, the relation between TQR and the ARSS and SRSS was stronger than the relation between RPE and the ARSS and SRSS as discussed above.

Limitations, and future research

This study has limitations that must be considered. First, we received 385 valid responses while 850 athletes received a link to the questionnaire. Therefore, a selection bias could not be ruled out. However, the absolute number of respondents was above the target (N=320). This means that with the current number of participants, the study design is adequate to assess construct validity.[20] For the multigroup CFA, however, the required sample size (n=200 per subgroup) was not met and therefore, these results should be interpreted with caution.[21]

For future research, it would be fruitful to better understand the reasons behind the lower fit of the models including both the recovery and stress dimensions. These reasons can be diverse and hence relate to, amongst others, instruction, translation, grouping of items, anchors used, and scoring. For instance, the ARSS instruction asks athletes to rate their current state, whereas the SRSS the instruction asks athletes to rate their current state compared to their best recovery or stress state ever. These different instructions could have led to confusion among respondents when filling out both questionnaires. Additionally, the CFA revealed high factor loadings between some Dutch items and their underlying factor (appendix 3; seven correlations >.8), which makes it difficult to determine the unique contributions of each item. This could mean that some correlations should be omitted from the model, or that items with high factor loadings within a scale should be collapsed.[40] If confirmed in other studies, the number of Dutch items may be reduced. Future studies could determine whether these choices affect the measured construct and could add new items or remove items to reduce multicollinearity, and to determine if the global model fit of a model that includes both recovery and stress could be improved.

Conclusion

This study provides a new step in the validation process of promising new scales to measure recovery and stress: the ARSS and SRSS. Our psychometric assessment revealed novel evidence that recovery and stress could be considered as intertwined constructs which is in line with the theorized model. However, the model fit of the models that include both constructs is less than ideal. In addition, we found valuable evidence that the Dutch translations of the ARSS and SRSS show sufficient construct and convergent validity, and are both correlated with total quality of recovery. Therefore, these questionnaires can be used by coaches and athletes to assess the perception of recovery and stress in sports-related research and practice. For instance, to measure recovery and stress before and after training. Future research could focus on unravelling the underlying relations between recovery and stress to further improve the structure of the questionnaires.

Contributions

Contributed to conception and design: JB, RDH, MB
Contributed to acquisition of data: JB, MB, PW
Contributed to analysis and interpretation of data: JB, RDH, MB
Drafted and/or revised the article: RDH, SJ, MK, PW, KL, MB
Approved the submitted version for publication: JB, RDH, SJ, MK, PW, KL, MB

Acknowledgements

We thank Niels de Vries, Koen Levels, and Steffie van der Steen for their translations of the questionnaire. We thank NOC*NSF for helping with data collection, and NOC*NSF athletes and all other athletes for completing the questionnaire. In addition, we thank Berdien Prins, Liset Schuitert, Mare Mathijssen, Suzan Blijlevens (NOC*NSF/Vrije Universiteit Brussel), and Simon Defruyt (Vrije Universiteit Brussel) for their help with the data collection. Finally, we thank Susan Niessen for providing valuable psychometric insights into the analysis.

Funding information

This work was supported by the Netherlands Organization for Health Research and Development (ZonMw) under Grant [546003004]

Data and Supplementary Material Accessibility

The full dataset and scripts are available in DataverseNL at https://doi.org/10.34894/A7GZSJ. The supplementary material is available at OSF (https://osf.io/zuyme/?view_only=91b67d07c1ac4f07b92df6f01065f481):
Appendix 1. Proposed structure of the ARSS.
Appendix 2. Replication of the analysis by Kölling et al. (2020).
Appendix 3. Standardized factor loadings and correlation matrix of the ARSS.
Appendix 4. Pearson correlations between the Dutch ARSS scales, SRSS items, and RPE and TQR.

REFERENCES

1. Smith DJ. A Framework for Understanding the Training Process Leading to Elite Performance: *Sports Med*. 2003;33(15):1103-1126. doi:10.2165/00007256-200333150-00003

2. Kellmann M, Kallus K. *Recovery-Stress Questionnaire for Athletes: User Manual*. Human Kinetics; 2001.

3. Herrera G, Hassmén P. Overtraining and Recovery: A Conceptual Model. *Sports Med*. 1998;26(1):1-16. doi:10.2165/00007256-199826010-00001

4. Hug M, Mullis PE, Vogt M, Ventura N, Hoppeler H. Training Modalities: Over-reaching and Over-training in Athletes, Including a Study of the Role of Hormones. *Best Pract Res Clin Endocrinol Metab*. 2003;17(2):191-209. doi:10.1016/S1521-690X(02)00104-5

5. Lehmann M, Foster C, Dickhuth HH, Gastmann U. Autonomic Imbalance Hypothesis and Overtraining Syndrome. *Med Sci Sports Exerc*. 1998;30(7):1140-1145. doi:10.1097/00005768-199807000-00019

6. Petibois C, Cazorla G, Poortmans JR, Déléris G. Biochemical Aspects of Overtraining in Endurance Sports: A Review. *Sports Med Auckl NZ*. 2002;32(13):867-878. doi:10.2165/00007256-200232130-00005

7. Robson PJ. Elucidating the Unexplained Underperformance Syndrome in Endurance Athletes: The Interleukin-6 Hypothesis. *Sports Med*. 2003;33(10):771-781. doi:10.2165/00007256-200333100-00004

8. Smith LL. Cytokine Hypothesis of Overtraining: A Physiological Adaptation to Excessive Stress? *Med Sci Sports Exerc*. 2000;32(2):317. doi:10.1097/00005768-200002000-00011

9. Saw AE, Main LC, Gastin PB. Monitoring Athletes Through Self-Report: Factors Influencing Implementation. *J Sports Sci Med*. 2015;14(1):137-146.

10. Montull L, Slapšinskaitė-Dackevičienė A, Kiely J, Hristovski R, Balagué N. Integrative Proposals of Sports Monitoring: Subjective Outperforms Objective Monitoring. *Sports Med - Open*. 2022;8(1):41. doi:10.1186/s40798-022-00432-z

11. Taylor KL, Chapman D, Cronin J, Newton M, Gill N. Fatigue Monitoring in High Performance Sport: A Survey of Current Trends. *J Aust Strength Cond*. 2012;20:12-23.

12. Fry RW, Grove JR, Morton AR, Zeroni PM, Gaudieri S, Keast D. Psychological and Immunological Correlates of Acute Overtraining. *Br J Sports Med*. 1994;28(4):241-246. doi:10.1136/bjsm.28.4.241

13. Kellmann M, Kölling S. *Recovery and Stress in Sport: A Manual for Testing and Assessment*.; 2019.

14. Hitzschke B, Holst T, Ferrauti A, Meyer T, Pfeiffer M, Kellmann M. Entwicklung des Akutmaßes zur Erfassung von Erholung und Beanspruchung im Sport [Development of the Acute Recovery and Stress Scale]. *Diagnostica*. 2016;62(4):212-226. doi:10.1026/0012-1924/a000155

15. Nässi A, Ferrauti A, Meyer T, Pfeiffer M, Kellmann M. Development of Two Short Measures for Recovery and Stress in Sport. *Eur J Sport Sci*. 2017;17(7):894-903. doi:10/gg6q9r

16. Kölling S, Hitzschke B, Holst T, et al. Validity of the Acute Recovery and Stress Scale: Training Monitoring of the German Junior National Field Hockey Team. *Int J Sports Sci Coach*. 2015;10(2-3):529-542. doi:10.1260/1747-9541.10.2-3.529

17. Kellmann M. Preventing Overtraining in Athletes in High-Intensity Sports and Stress/Recovery Monitoring: Preventing Overtraining. *Scand J Med Sci Sports*. 2010;20:95-102. doi:10.1111/j.1600-0838.2010.01192.x

18. Kallus W, Kellmann M. Burnout in Athletes and Coaches. In: *Emotions in Sport*. Human Kinetics; 2000:209-230.

19. Borg GAV. Psychophysical Bases of Perceived Exertion. *Med Sci Sports Exerc*. 1982;14(5):377-381.

20. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Published online July 2019:32.

21. Terwee CB, Bot SDM, de Boer MR, et al. Quality Criteria Were Proposed for Measurement Properties of Health Status Questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012

22. Vallerand RJ. Vers Une Méthodologie De Validation Trans-Culturelle De Questionnaires Psychologiques: Implications Pour La Recherche En Langue Française. [toward a Methodology for the Transcultural Validation of Psychological Questionnaires: Implications for Research in the French Language.]. *Can Psychol Psychol Can*. 1989;30(4):662-680. doi:10.1037/h0079856

23. Qualtrics. Qualtrics. Published online March 8, 2022. https://www.qualtrics.com

24. Borg E, Borg GAV. A New Generation of Scaling Methods: Level-Anchored Ratio Scaling. *Psychologica*. 2001;28(1):15-45.

25. Nederhof E, Brink M, Lemmink KAPM. Reliability and Validity of the Dutch Recovery Stress Questionnaire for Athletes. *Int J Sport Psychol*. 2008;39:301-311.

26. Kallus W, Kellmann M. *The Recovery-Stress Questionnaires: User Manual*.; 2016.

27. Epskamp S, Stuber S, Nak J, Veenman M, Jorgensen TD. semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output. Published online August 10, 2022. Accessed October 21, 2022. https://CRAN.R-project.org/package=semPlot

28. Jorgensen TD, Pornprasertmanit S, Schoemann AM, et al. semTools: Useful Tools for Structural Equation Modeling. Published online May 10, 2022. Accessed October 21, 2022. https://CRAN.R-project.org/package=semTools

29. Rosseel Y. lavaan: An *R* Package for Structural Equation Modeling. *J Stat Softw*. 2012;48(2). doi:10.18637/jss.v048.i02

30. R Core Team. R: A Language and Environment for Statistical Computing. Published online 2020. https://www.r-project.org

31. Kölling S, Schaffran P, Bibbey A, et al. Validation of the Acute Recovery and Stress Scale (arss) and the Short Recovery and Stress Scale (srss) in Three English-Speaking Regions. *J Sports Sci*. 2020;38(2):130-139. doi:10.1080/02640414.2019.1684790

32. Credé M, Harms PD. 25 Years of Higher-Order Confirmatory Factor Analysis in the Organizational Sciences: A Critical Review and Development of Reporting Recommendations. *J Organ Behav*. 2015;36(6):845-872. doi:10.1002/job.2008

33. Satorra A, Bentler PM. A Scaled Difference Chi-Square Test Statistic for Moment Structure Analysis. *Psychometrika*. 2001;66(4):507-514. doi:10.1007/BF02296192

34. Terwee CB, Prinsen CAC, Chiarotto A, et al. Cosmin Methodology for Evaluating the Content Validity of Patient-Reported Outcome Measures: A Delphi Study. *Qual Life Res*. 2018;27(5):1159-1170. doi:10.1007/s11136-018-1829-0

35. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive Statistics for Studies in Sports Medicine and Exercise Science. *Med Sci Sports Exerc*. 2009;41(1):3-12. doi:10.1249/MSS.0b013e31818cb278

36. Diamantopoulos A, Sarstedt M, Fuchs C, Wilczynski P, Kaiser S. Guidelines for Choosing Between Multi-Item and Single-Item Scales for Construct Measurement: A Predictive Validity Perspective. *J Acad Mark Sci*. 2012;40(3):434-449. doi:10.1007/s11747-011-0300-3

37. Doeven SH, Brink MS, Kosse SJ, Lemmink KAPM. Postmatch Recovery of Physical Performance and Biochemical Markers in Team Ball Sports: A Systematic Review. *BMJ Open Sport Exerc Med*. 2018;4(1):e000264. doi:10.1136/bmjsem-2017-000264

38. Scantlebury S, Till K, Sawczuk T, Phibbs P, Jones B. Validity of Retrospective Session Rating of Perceived Exertion to Quantify Training Load in youth Athletes. *J Strength Cond Res*. 2018;32(7):1975-1980. doi:10.1519/JSC.0000000000002099

39. Otter R, Brink M, Diercks R, Lemmink K. A Negative Life Event Impairs Psychosocial Stress, Recovery and Running Economy of Runners. *Int J Sports Med*. 2015;37(03):224-229. doi:10/gkzx2t

40. Chen F, Bollen KA, Paxton P, Curran PJ, Kirby JB. Improper Solutions in Structural Equation Models: Causes, Consequences, and Strategies. *Sociol Methods Res*. 2001;29(4):468-508. doi:10.1177/0049124101029004003