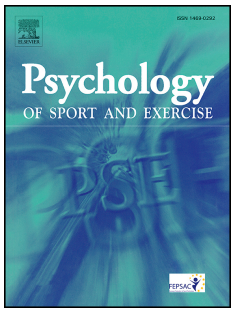
Accepted Manuscript



Usefulness of the Athlete Burnout Questionnaire (ABQ) as a screening tool for the detection of clinically relevant burnout symptoms among young elite athletes

Markus Gerber, Henrik Gustafsson, Harald Seelig, Michael Kellmann, Sebastian Ludyga, Flora Colledge, Serge Brand, Sandrine Isoard-Gautheur, Renzo Bianchi

PII: S1469-0292(18)30122-5

DOI: 10.1016/j.psychsport.2018.08.005

Reference: PSYSPO 1406

To appear in: *Psychology of Sport & Exercise*

Received Date: 5 March 2018

Revised Date: 31 July 2018

Accepted Date: 3 August 2018

Please cite this article as: Gerber, M., Gustafsson, H., Seelig, H., Kellmann, M., Ludyga, S., Colledge, F., Brand, S., Isoard-Gautheur, S., Bianchi, R., Usefulness of the Athlete Burnout Questionnaire (ABQ) as a screening tool for the detection of clinically relevant burnout symptoms among young elite athletes, *Psychology of Sport & Exercise* (2018), doi: 10.1016/j.psychsport.2018.08.005.

Usefulness of the Athlete Burnout Questionnaire (ABQ) as a screening tool for the detection of clinically relevant burnout symptoms among young elite athletes

Markus Gerber*, Henrik Gustafsson,* Harald Seelig, Michael Kellmann, Sebastian Ludyga, Flora Colledge, Serge Brand, Sandrine Isoard-Gautheur, Renzo Bianchi

Markus Gerber works at the Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland. Henrik Gustafsson works at the Department of Health Sciences, Karlstad University, Karlstad, Sweden. Harald Seelig works at the Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland. Michael Kellmann works at the Faculty of Sport Science, Ruhr-University of Bochum, Germany and at the School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, Australia. Sebastian Ludyga works at the Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland. Flora Colledge works at the Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland. Serge Brand works at the Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland. He is also employed at the Psychiatric Clinics (UPK), Center for Affective, Stress and Sleep Disorders, University of Basel, Basel, Switzerland, and at the Kermanshah University of Medical Sciences (KUMS), Psychiatry Department, Substance Abuse Prevention Center and Sleep Disorders Research Center, Kermanshah, Iran. Sandrine Isoard-Gautheur works at Laboratoire Sport et Environnement Social, Université Grenoble Alpes, Grenoble Cedex, France. Renzo Bianchi is employed at the Institut of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland.

Correspondence should be addressed to Prof. Dr. Markus Gerber, University of Basel, Department of Sport, Exercise and Health, St. Jakob-Turm, Birsstrasse 320B, CH-4052 Basel, Switzerland, +41 61 20 747 83 (phone), +41 61 20 747 89 (fax), markus.gerber@unibas.ch

* These authors have contributed equally to this article.

Abstract

Objectives: Having psychometrically sound instruments is essential to the understanding of the determinants and consequences of athlete burnout. Therefore, this study examines the psychometric properties of a German version of the Athlete Burnout Questionnaire (ABQ) and its usefulness as a screening tool for the detection of clinically relevant burnout symptoms.

Design: Prospective study.

Method: 257 young elite athletes were recruited from Swiss Olympic Sport Classes (37% females; *M* = 16.8 years, *SD* = 1.4). 197 students were assessed a second time after six months. All students filled in a standardized questionnaire about domain-specific and domain-unspecific burnout symptoms, depressive symptoms, stress, and life satisfaction.

Results: Confirmatory factor analysis supported the three-factor structure of the ABQ. Moreover, all subscales had acceptable internal consistency. Support was also found for the convergent validity of the ABQ; all subscales were positively correlated with perceived stress, burnout and depression, whereas negative correlations existed with life satisfaction. By contrast, some ABQ subscales shared only limited variance, the three ABQ subscales did not predict each other across time, and none of the ABQ subscales was suitable for the screening of clinically relevant burnout symptoms.

Conclusions: While the factor structure and internal consistency of the ABQ was supported, our study corroborates previous concerns about the psychometric properties and validity of the ABQ. While the ABQ has advanced research on athlete burnout, we hold that further debates about the most suitable way to assess burnout among elite athletes are urgently needed.

Keywords: *adolescents; athlete burnout; confirmatory factor analysis; convergent validity; depression; psychometric properties; sensitivity; specificity; stress*

Usefulness of the Athlete Burnout Questionnaire (ABQ) as a screening tool for the detection of clinically relevant burnout symptoms among young elite athletes

During the past three decades, the issue of burnout among elite athletes has received increasing attention both in the mass media and in the scientific literature (Gustafsson, DeFreese, & Madigan, 2017). While the first reports on the athlete burnout syndrome were published in the early 1980s, more than 100 publications have been identified in a recent review of the literature (Gustafsson, Hancock, & Côté, 2014).

Today, the negative outcomes associated with athlete burnout are well documented. For instance, studies have shown that among elite athletes, elevated burnout symptoms were associated with decreased performance, reduced motivation, and leaving sport (Isoard-Gautheur, Guillet-Descas, & Gustafsson, 2016; Lemyre, Treasure, & Roberts, 2006). Accordingly, gaining a deeper understanding of the issue is essential for coaches, managers, psychologists and sport organizations.

In the scientific literature, athlete burnout has been defined in various ways (Eklund & Cresswell, 2007; Gustafsson, Kenttä, & Hassmén, 2011). According to Goodger et al. (2007), Raedeke's (1997) conceptualization of athlete burnout as a multidimensional syndrome has led to a certain consensus among researchers. Thus, referring to Maslach's and Jackson's (1981) definition of occupational burnout, Raedeke (1997) defined athlete burnout as a gradually developing syndrome consisting of three dimensions including (i) emotional and physical exhaustion (perceived depletion of emotional and physical resources beyond that associated with training and competition), (ii) sport devaluation (development of a cynical attitude towards involvement in elite sport), and (iii) reduced sense of accomplishment (tendency to evaluate oneself negatively in terms of sport abilities and achievement). Gustafsson, DeFreese, et al. (2017) identified three issues that need to be investigated more closely in future research, namely, the lack of information regarding the prevalence of burnout, the over-reliance on cross-sectional and correlational designs, and the lack of research on burnout treatment and prevention. In the present study, we will address two of these shortcomings and expand on previous research in three ways. As a starting point, we

will first examine the psychometric properties of a German version of the Athlete Burnout Questionnaire (ABQ) (Raedeke & Smith, 2001) in a sample of young elite athletes. This examination is relevant because research regarding the psychometric properties of translated versions of the ABQ is still sparse (Raedeke, Arce, De Francisco, Seoane, & Ferraces, 2013). Going beyond the examination of the psychometric properties of the German ABQ, we will in a second step compare the scores on the ABQ with those on the Shirom-Melamed Burnout Measure (SMBM) (Lerman et al., 1999). The SMBM is an internationally accepted instrument, associated with a validated cut-off score to estimate the prevalence of clinically relevant symptoms of burnout in reference to the ICD-10 criteria for 'other reactions to severe stress' (Lundgren-Nilsson, Jonsdottir, Pallant, & Ahlborg, 2012). Finally, we will use longitudinal data to examine how the ABQ dimensions predict each other across time, and to determine whether the ABQ dimensions can help us predict domain-unspecific symptoms of burnout and depression over a six-month period. The latter two points are relevant in order to gauge whether the ABQ dimensions capture processes that contribute to the development of domain-unspecific burnout, and whether the ABQ can be used as a screening tool for clinically relevant levels of burnout symptoms.

Historically, research on athlete burnout has been undermined by a lack of validated, domain-specific measures (Gustafsson, Madigan, & Lundkvist, 2017; Raedeke & Smith, 2001). Nevertheless, having psychometrically sound instruments is essential to the understanding of the determinants and consequences of athlete burnout (Raedeke et al., 2013). Therefore, in the early 1990s, and based on the work of Maslach and Jackson (1981), Eades (1990) developed the Eades Athlete Burnout Inventory (EABI), the first instrument allowing investigators to quantitatively assess athlete burnout symptoms. Based on exploratory factor analysis, Eades identified six subscales in the EABI (emotional/physical exhaustion, psychological withdrawal, devaluation by coach and teammates, negative self-concept of athletic ability, congruent athlete-coach expectations, personal and athletic accomplishment). However, this instrument has been criticized due to its atheoretical nature, and because it presents a mix of burnout symptoms and burnout antecedents (Eklund & Cresswell, 2007).

To address some of these issues, Raedeke and Smith (2001) developed the Athlete Burnout Questionnaire (ABQ), a sport-specific adaptation of the Maslach Burnout Inventory (Maslach, Jackson, & Leiter, 1996), including several items of the EABI in their instrument. Today, the ABQ is by far the most widely used questionnaire for assessing athlete burnout symptoms (Gustafsson, Madigan, et al., 2017) and has been translated into several languages such as Chinese (Lu, Chen, & Cho, 2006), French (Isoard-Gautheur, Oger, Guillet, & Martin-Krumm, 2010), German (Ziemainz, Abu-Omar, Raedeke, & Krause, 2004), Norwegian (Lemyre et al., 2006), Portuguese (Guedes & de Souza, 2016), Spanish (Raedeke et al., 2013), and Swedish (Gustafsson & Skoog, 2012). As mentioned previously, the ABQ has been designed as a multidimensional measure (Raedeke & Smith, 2001), based on Raedeke's (1997) conceptualization of athlete burnout. The ABQ consists of 15 items, representing the three (previously described) dimensions of emotional/physical exhaustion, sport devaluation, and reduced sense of accomplishment (with 5 items per subscale). All items are semantically anchored on a 5-point Likert-type scale ranging from 1 (almost never) to 5 (almost always), and means are computed to obtain subscale scores.

Evidence for the reliability and validity of the ABQ has been found in several athlete populations. Specifically, confirmatory factor analyses (CFAs) supported the theoretically assumed three-factor solution (Raedeke & Smith, 2001), with items achieving adequate loadings on the respective factors (Raedeke et al., 2013). Furthermore, single factors had satisfactory internal consistency (Guedes & de Souza, 2016; Isoard-Gautheur et al., 2010; Raedeke et al., 2013; Raedeke & Smith, 2001), and good model fit was found for a first-order and a second-order model (Isoard-Gautheur et al., 2010; Raedeke & Smith, 2001). Prior research also provided evidence for the convergent validity of the ABQ, in the sense that positive associations were found between the ABQ indices and related constructs such as perceived stress, trait anxiety, depression, and amotivation, whereas negative relationships were identified between the ABQ and efficient coping, social support, enjoyment, self-confidence, commitment or intrinsic motivation (Raedeke & Smith, 2001; Sharp, Woodcock, Holland, Duda, & Cummings, 2010). Researchers also observed acceptable test-retest

reliabilities with correlations of $r > .70$ across various periods ranging from one to three weeks (Arce, De Francisco, Andrade, Seoane, & Raedeke, 2012; Raedeke & Smith, 2001). In summary, previous research suggests that the ABQ is a psychometrically acceptable instrument. However, researchers have also highlighted several weaknesses of the ABQ (Gustafsson, Lundkvist, Podlog, & Lundkvist, 2016). According to Gustafsson et al. (2016), the most relevant points of criticism are that (a) the definition of burnout upon which the ABQ is based is founded neither on clinical observation nor on theory, (b) some of the dimensions of the ABQ have considerable overlap with other psychological constructs (e.g., sense of accomplishment with self-efficacy), (c) reduced sense of accomplishment consistently showed lower correlations with the other two dimensions and other psychological constructs than exhaustion and sport devaluation, and (d) the question of how the three ABQ dimensions influence each other over time has rarely been addressed (Lundkvist et al., 2018). With regard to the latter notion, it is currently not possible to reach a consensus, and the existing studies show that the cross-lagged associations between the ABQ dimensions tend to be insignificant, and otherwise only weak (Isoard-Gautheur, Guillet-Descas, Gaudreau, & Chanal, 2015; Lundkvist et al., 2018; Martinent, Louvet, & Decret, 2016). Moreover, one remaining limitation of the ABQ is that no reliable cut-offs have been established to classify participants in terms of burnout symptom severity. Given that most previous studies have focused on athletes with relatively low levels of burnout symptoms (Gustafsson et al., 2011), and that researchers examined athletes that were arbitrarily classified as being ‘at risk’ of elevated burnout (Eklund & Cresswell, 2007), having a more trustworthy cut-off would be important to advance athlete burnout research.

To test the usefulness of this instrument as a screening tool for the detection of clinically relevant burnout symptoms among young elite athletes and to further examine the psychometric properties of the German ABQ, the following aspects will be examined: First, we examine whether the ABQ shows adequate internal consistency. Based on previous studies, we expect inter-item correlations $\geq .20$, Cronbach’s alpha coefficients $\geq .70$, and item-total correlations $\geq .30$ (Guedes & de Souza, 2016; Isoard-Gautheur et al., 2010; Raedeke et al., 2013; Raedeke & Smith, 2001). Second, we examine whether our data provide

support for the three-factor structure of the ABQ (Raedeke et al., 2013). Based on prior research, we expect that adequate model fit will be found for a first- and second-order model (Isoard-Gautheur et al., 2010; Raedeke & Smith, 2001). Based on previous studies (Arce et al., 2012; Raedeke & Smith, 2001), we also assume that significant cross-correlations between ABQ dimensions will be found. In previous studies, these cross-correlations typically varied between r = .30 and .70. Third, we examine whether the ABQ shows adequate convergent validity. Our expectation is that the ABQ subscales will be moderately to strongly (and positively) correlated with the SMBM overall and subscale scores and levels of self-perceived stress, whereas we expect a moderate (negative) relationship between the ABQ indices and participants' satisfaction with life (Isoard-Gautheur et al., 2010; Raedeke et al., 2013; Raedeke & Smith, 2001; Sharp et al., 2010). Fourth, we explored the temporal stability of the ABQ. From the occupational and school burnout literature (Lindwall, Gerber, Jonsdottir, Börjesson, & Ahlborg, 2014; Salmela-Aro, Savolainen, & Holopainen, 2009; Toppinen-Tanner, Kalimo, & Mutanen, 2002), it is known that burnout symptoms are fairly stable across time with auto-correlations often exceeding r = .50. Based on these findings, relatively high correlations between the baseline and follow-up scores can be expected between the three dimensions of athlete burnout, as well. Nevertheless, little research exists so far regarding the stability of the ABQ dimensions. In a previous investigation, Lundkvist et al. (2018) found that the stability of the ABQ subscales was relatively low across an 18-month period. Therefore, we did not have any clear-cut hypotheses regarding the stability of the ABQ subscales. Fifth, we explored how the three ABQ dimensions are associated with each other across a 6-month period, and whether they have the potential to predict domain-unspecific burnout and depressive symptoms over time. Sixth and most importantly, we explored whether the ABQ can be used as a screening tool for the detection of clinically relevant burnout symptoms. This would indicate that the ABQ has a relatively high sensitivity, defined as the ability to correctly identify athletes with clinically relevant burnout symptoms (also known as true positive rate), and a relatively high specificity, defined as the ability of the test to correctly detect participants who are free of clinically relevant burnout symptoms (also called true negative rate) (Altman & Bland, 1994).

Materials and methods

Participants and procedures

Participants were recruited from Swiss Olympic partner schools in the North-Western, German-speaking part of Switzerland. These classes aim at facilitating the combination of school and elite sport (e.g., fewer lessons per week, prolonged school duration). All students obtained information about the general goals of this study, were informed that participation was voluntary, and that all data will be treated confidentially. Contact to schools was sought by the primary investigator. Students were then approached class by class during an official school lesson by a research assistant who provided information about the general goal of the study and the procedure/scope of the data assessment. Students were informed that participation is voluntary, and that not taking part in the data assessment does not have any consequences. Before the beginning of the data assessment, participants were asked to provide informed written consent. Data collection took place during official class time at the students' school in a group setting under the supervision of a trained research assistant. The data assessment lasted from November to December 2016, and was repeated after a follow-up period of six months (May-June 2017). The same instruments were used at baseline and at follow-up. All participants filled in a battery of internationally accepted psychological instruments (see below for more details). The local ethics committee (Ethics committee of Northwestern and Central Switzerland, Nr. 2016-01535b) approved the study, which ensured that all procedures correspond with current Swiss legal requirements. Moreover, the study was carried out in line with the Declaration of Helsinki and its later amendments.

To determine the minimal sample size needed to predict burnout and depressive symptoms at follow-up with the ABQ baseline scores, a power analysis was carried out (using G*Power 3.1). This analysis showed that a minimum of 208 participants is required to detect a weak association (Slope H1 = -0.20, alpha: 0.05, power: 0.80). Assuming a dropout rate of 20%, the target sample size at baseline was 250 students. For this study, the total baseline sample consisted of 257 adolescents (163 males and 94 females; age: M = 16.8 years, SD = 1.4).

197 of these students (125 boys and 73 girls; age: M = 16.83, SD = 1.40) took part in the follow-up data assessment. The dropout analysis revealed that no significant differences existed between dropouts and follow-up completers in any of the study variables (p > .05). At baseline, the age range of the sample was 14 to 24 years. Athletes had participated in competitive sports for M = 7.79 years (SD = 3.07; range: 1-15 years) and invested 21.41 hours per week in training and competition (SD = 8.46; range: 5-56 hours). The athletes engaged in the following sports: Soccer (n = 61), handball (n = 26), volleyball (n = 23), swimming (n = 19), judo (n = 17), track and field (n = 16), tennis (n = 15), uni hockey (n = 12), karate (n = 10), mountain bike (n = 7), golf (n = 5), others (< 5 athletes; n = 46).

Measures

Athlete burnout symptoms. Athlete burnout was assessed with a German version (Ziemainz et al., 2004) of the Athlete Burnout Questionnaire (ABQ) (Raedeke & Smith, 2001). The ABQ is a self-report inventory that consists of 15 items. Items are answered on a 5-point Likert scale, with the following stem ("How often do you feel this way?") and the following anchors: (1) almost never, (2) rarely, (3) sometimes, (4) frequently, and (5) almost always. Compared to the version of Ziemainz et al. (2004), we slightly changed the wording of some items to achieve a better fit with the English version, and to facilitate comprehension among young elite athletes (see Table 1 for all German items). Three scores were obtained by calculating separate means over the items of the three subscales related to emotional/physical exhaustion (e.g. "I feel overly tired from my sport participation."), sport devaluation (e.g. "I don't care as much about my sport performance as I used to do."), and reduced sense of accomplishment (e.g. "I am not performing up to my ability in sport."). Two items (item 1 and 14) were inverted before calculating the subscale mean scores.

Domain-unspecific burnout symptoms. A German translation of the 14-item Shirom-Melamed Burnout Measure (SMBM) (Lerman et al., 1999) was used to measure burnout symptoms. The original formulation of the SMBM is composed of three subscales labelled physical fatigue (six items: e.g., "I feel physically drained."), cognitive weariness (five items: e.g., "I feel I am not thinking clearly."), and emotional exhaustion (three items: e.g., "I feel I

am unable to be sensitive to the needs of co-workers and customers."). All items refer to the past four weeks. For the emotional exhaustion subscale, the wording of the items was adapted to increase suitability for adolescents. Thus, a more open formulation was used to refer to people in general instead of co-workers and customers. Answers were given on a 7-point Likert scale ranging from 1 (never or almost never) to 7 (always or almost always). The mean score over all 14 items was calculated to obtain an overall index, with higher scores reflecting higher burnout symptoms. A score of ≥ 4.40 was considered as a clinically relevant burnout level. This cut-off was chosen because Lundgren-Nilsson, Jonsdottir, Pallant, and Ahlborg (2012), in a study with 319 clinical patients and 319 working employees, found that a SMBM score of 4.40 would place 83.4% of the clinical population above the cut, and 86.5% of the working population below the cut. Previous research provided support for adequate psychometric properties of a German translation of the SMBM among German speaking adolescents (Gerber et al., 2015). In the present sample, internal consistency of the overall index was satisfactory, with a Cronbach's alpha of .92.

General perceived stress. Subjectively perceived stress during the past month was assessed with a German version of the widely used 10-item Perceived Stress Scale (PSS) (Cohen, Kamarck, & Mermelstein, 1983). Participants were asked how often they find their lives to be overwhelming, uncontrollable, and unpredictable (e.g., "In the last month, how often have you felt that you were effectively coping with important changes that were occurring in your life?", "In the last month, how often have you been upset because of something that happened unexpectedly?"). Answers were given on a 5-point Likert scale, ranging from 1 (never) to 5 (very often). Four items were inverse-poled and had to be inverted before calculating the (final) sum score. Higher scores reflect higher subjectively perceived stress levels. Evidence for the reliability and validity of the German version of this instrument has been provided previously (Gerber et al., 2013). In the present sample, the Cronbach's alpha was satisfactory (α = .80).

Depressive symptoms: A German version (Gräfe, Zipfel, Herzog, & Löwe, 2004) of the 9-item Patient Health Questionnaire (Kroenke, Spitzer, & Williams, 2001) was used to assess depressive symptoms. The PHQ-9 refers to the DSM-IV diagnosis criteria for major

depressive disorder and can be used to assess severity of depressive symptoms, with scores of >14 reflecting moderately severe depression. Sample items are: “Little interest or pleasure in doing things” or “Feeling down, depressed or hopeless.” Answers are given on a Likert-type scale from 0 (not at all) to 3 (nearly every day), referring to the last two weeks. The sum is built to obtain an overall index. Previous research provided support for the psychometric properties and validity of the PHQ-9 (Kroenke et al., 2001). Adequate psychometric properties have also been found for the German version of the PHQ-9 (Gräfe, Zipfel, Herzog, & Löwe, 2004). The Cronbach’s alpha was .85 in the present sample.

Life satisfaction. A German version (Glaesmer, Grande, Braehler, & Roth, 2011) of the 5-item Satisfaction with Life Scale (SWLS) was employed to obtain an overall judgement of participants’ satisfaction with life (Diener, Emmons, Larsen, & Griffin, 1985). Answers on this instrument were given on a Likert-scale from 1 (strongly disagree) to 7 (strongly agree). A sample item is: “In most ways my life is close to my ideal.” Validity and adequate reliability of the SWLS have been documented previously (Pavot & Diener, 2008), including German speaking adolescents (Glaesmer, Grande, Braehler, & Roth, 2011). The items were summed up to obtain an overall index, with higher scores indicating higher life satisfaction (in the present sample, the Cronbach’s alpha was .82).

Statistical analyses

Correlational analyses were used to examine homogeneity (item-item correlations) and item-total correlations. Cronbach’s alpha coefficients were calculated to test internal consistency. Furthermore, CFAs were applied to examine factorial validity. Based on the assumptions formulated in the ABQ manual and the results of previous confirmatory factor analyses (e.g., Arce et al., 2012; Raedeke & Smith, 2001), we expected that the 15 items would load on three different factors (five items on emotional/physical exhaustion, five items on cognitive weariness, five items on emotional exhaustion). Thus, the 3-factor CFA model was based on 15 observed measures and three latent constructs. Maximum likelihood (ML) was used for parameter estimation, and the fit between the theoretical model and the empirical data was judged via multiple goodness-of-fit indices (McDonald & Ho, 2002). Before carrying out these analyses, multivariate normality was established via Mardia’s (1974) test of multivariate

kurtosis (critical ratio < 5; Byrne, 2010). According to Byrne (2010), adequate model fit is achieved if the normed fit index (NFI) is ≥ .95, the comparative fit index (CFI) is ≥ .95, the Tucker Lewis Index (TLI) is ≥ .95, and root mean square error of approximation (RMSEA) is ≤ .05. According to Comrey and Lee (1992), standardized factor loadings of < .45 should be interpreted as poor, ≥ .45 as fair, ≥ .55 as good, ≥ .63 as very good, and ≥.71 as excellent. Correlation analyses were used to examine convergent validity, as well as test-retest reliability. Correlations of r < were considered small, with r = .30 to .49 as medium, and r ≥ as large (Cohen, 1988).

A cross-lagged panel analysis was performed to examine the three ABQ dimensions across time. Again, ML was used for parameter estimation, and the standards recommended by McDonald and Ho (2002) were employed to assess model fit. Two hierarchical (linear) regression analyses were calculated to find out whether the three ABQ subscales contribute to the prediction of the SMBM overall index and the PHQ-9 score at follow-up, after having controlled for baseline levels of burnout and depressive symptoms. Receiver Operating Characteristic (ROC) analyses were performed to obtain diagnostic efficiency statistics (Youngstrom, 2014). Area under the entire curve (AUC) was used as an indicator of diagnostic accuracy. According to Swets, Dawes and Monahan (2000), accuracy is poor if values are < .70, fair if values are ≥ .70, good if values are ≥ .80, and excellent if values are ≥ .90. In case of values of ≥ .70, we further examined the coordinates of the curve to find out whether there is a cut-off, which provides an acceptable trade-off between sensitivity and specificity. In the optimal case, values for sensitivity and specificity are both close to 1. An alpha level of p < .05 was set throughout all analyses to indicate significant characteristics. Whereas CFA were performed with AMOS® 24 (IBM Corporation, Armonk NY, USA), all other analyses were carried out with SPSS® 24 (IBM Corporation, Armonk NY, USA).

Results

Descriptive statistics

Table 1 shows the descriptive statistics for all ABQ items and the three ABQ subscales. None of the items exceeded the theoretical mean score of M = 3. Skewness and kurtosis were within acceptable limits (limits for skewness < 2, and kurtosis < 7) for all ABQ items according to

the recommendations of West, Finch, and Curran (1995).

Table 2 provides the means and standard deviations for all other constructs used to examine convergent and discriminant validity (both at baseline and follow-up). In the present sample, 12% (n = 30) of the participants had a SMBM overall score of ≥ 4.40 and were thus categorized as having clinically relevant burnout levels.

Internal consistency

The inter-item correlations were greater than r = .20 for all items within each of the three ABQ subscales. Moreover, all item-total correlations within the three ABQ subscales exceeded the critical value of $r \geq .30$. The Cronbach's alpha was .80 for emotional/physical exhaustion, .78 for sport devaluation, .78 for reduced sense of accomplishment.

Factorial validity

The three-factor model fitted well with the empirical data in the CFA. This conclusion applies for both the first-order model, χ^2/df = 1.40, CFI = .98, TLI = .97. RMSEA = .04 (CI = .02, .06), and second-order model, χ^2/df = 1.39, CFI = .98, TLI = .97. RMSEA = .04 (CI = .02, .06). With one exception (item 6), all items showed at least 'good' factor loadings (≥ .55). Figure 1 displays the measurement coefficients of the hypothesized three-factor models. In the first-order model, weak to strong correlations were observed between the three ABQ dimensions (r = .27 to .60, $p < .001$). In the second-order model, the loadings of the first-order factors ranged between .40 and .88 ($p < .001$).

Convergent validity

With regard to convergent validity, we found moderate to strong correlations between the three ABQ subscales and the SMBM overall index (r = .31 to .52, $p < .001$). The ABQ indices were also positively correlated with our stress measures (r = .29 to .46, $p < .001$). Finally, negative correlations occurred between all ABQ subscales and satisfaction with life (r = -.22 to -.47, $p < .001$).

Test-retest reliability

The three ABQ subscales were strongly correlated with each other across the 6-month study period. The correlation was $r = .57$, $p < .001$, for emotional/physical exhaustion, $r = .64$, $p < .001$, for sport devaluation, and $r = .65$, $p < .001$, for reduced sense of accomplishment.

Prospective associations between the ABQ dimensions

Figure 2 confirms that the three ABQ dimensions have a relatively high temporal stability, with regression weights varying between .65 and .72 ($p < .001$). However, in the cross-lagged panel model, none of the cross-lagged paths from baseline to follow-up was statistically significant ($p > .05$). The examined model showed an acceptable fit with the empirical data, $\chi^2/df = 1.59$, CFI = .92, TLI = .91. RMSEA = .06 (CI = .05, .06).

Prediction of burnout and depressive symptoms

As shown in Table 3, higher levels of emotional/physical exhaustion at baseline predicted higher burnout symptoms at follow-up ($\beta = .14$, $p < .05$, 2.1% of explained variance). By contrast, sport devaluation and reduced sense of accomplishment were without predictive value. In sum, the predictors together explained 31.9% of variance in burnout at follow-up.

With regard to depressive symptoms, the amount of variance explained for the total model was similar (31.7%). However, both increased levels of emotional/physical exhaustion ($\beta = .18$, $p < .01$) and sport devaluation ($\beta = .13$, $p < .05$) predicted depressive symptoms at follow-up (7.3% of explained variance).

Usefulness as a screening tool for clinically relevant burnout levels

To explore whether the ABQ can be used as a tool to identify athletes with clinically relevant burnout symptoms, we first visually inspected the scatterplots showing the correlations between the SMBM overall index and the three ABQ subscales at baseline (Figure 3). As shown in Table 2, the strongest bivariate correlation existed between the SMBM and the ABQ emotional/physical exhaustion subscale ($r = .52$, $p < .001$). In Figure 3, the scatterplot points towards a linear relationship between the SMBM overall index and the ABQ exhaustion subscale. Interestingly, with regard to the ABQ sport devaluation subscale, the scatterplot did not support a linear relationship between the ABQ and the SMBM overall index. Rather, the

majority of athletes with clinically relevant burnout levels did not score high on the ABQ sport devaluation dimension.

The ROC analyses support the suggestion that the accuracy of the ABQ scales to distinguish between athletes with versus without clinically relevant burnout symptoms is limited. The AUC values were .72 for the emotional/physical exhaustion subscale, .55 for sport devaluation, and .59 for reduced sense of accomplishment. The ROC curves for all three ABQ subscales are shown in Figure 4. Table 4 further shows that even for the emotional/physical exhaustion subscale (with the highest AUC score), establishing a meaningful cut-off with an acceptable balance between sensitivity and specificity is difficult. For instance, with a cut-off of ≥ 2.75, the sensitivity score is .71 (that is, 71% of true positive cases are correctly identified). However, with this cut-off, the false positive rate would be .44 (that is, 44% of athletes without clinically relevant burnout symptoms are classified as positives cases).

Discussion

The key finding of the present article is that the three-factor structure of the Athlete Burnout Questionnaire (ABQ) could be confirmed in German speaking junior elite athletes. Our findings also show that the ABQ emotional/physical exhaustion and sport devaluation subscales can be used to describe processes involved in the development of domain-unspecific symptoms of burnout and depression. Nevertheless, our findings also highlight some problematic aspects associated with the ABQ. Most importantly, relatively low correlations were found between some of the ABQ dimensions, and the ABQ dimensions were unrelated across time. Finally, our study also showed that the ABQ cannot be used as a screening tool for clinically relevant burnout symptoms in young elite athletes. Therefore, previous prevalence rates based on the ABQ must be interpreted with utmost caution. Using the SMBM with a clinically validated cut-off score, we found that about one in eight junior elite athletes experienced burnout symptoms of clinical relevance.

In our article, several research questions were addressed, and each of these will be discussed below. First, we tested whether the ABQ would produce adequate internal consistency, and our data did support this. Thus, in line with previous research (e.g., Raedeke

& Smith, 2001; Sharp et al., 2010), inter-item correlations, item-total correlations, and Cronbach's alpha coefficients of the ABQ were in the acceptable range. However, the present results expand upon previous findings, in that this is the first study that systematically examined the psychometric properties of a German version of the ABQ.

Second, we examined whether our data would confirm the factorial validity of the ABQ. Again, our data could confirm this. The hypothesized three-factor model (Raedeke et al., 2013) of the ABQ was corroborated by our empirical data. As shown previously, both a first- and second-order model provided adequate model fit (e.g., Isoard-Gautheur et al., 2010; Smith, Gustafsson, & Hassmén, 2010). Moreover, with reference to Comrey et al.'s (1992) recommendations, all factor loadings were fair (≥ .45). As in previous studies (e.g., Arce et al., 2012; Raedeke et al., 2013), the strongest between-factor correlation was found between sport devaluation and reduced sense of accomplishment. Arce et al. (2012) argued that these factors both have an attitudinal component, which is less salient in the emotional/physical exhaustion factor with a more physiological component. Thus, the other correlations were relatively low, with physical/emotional exhaustion only sharing 4% of variance with reduced sense of accomplishment and 7% with sport devaluation. In fact, the subscales of the ABQ were more strongly correlated with the SMBM global score than with each other. Moreover, although the second-order model fitted with the empirical data, it remains questionable whether the three ABQ dimensions can be grouped under the same label. For instance, such an approach would be at odds with the recommendations in the MBI manual, highlighting that the three burnout dimensions should be measured independently and should not be combined into a single measure. In addition, the factor loadings of the three ABQ dimensions on the higher-order burnout factor varied strongly, which raises further doubts as to why these heterogeneous symptoms should be grouped under a common label. As emphasized by Shirom and Melamed (2006), the theoretical foundations for a combined overall score are weak, and the available meta-analytic studies point out that each component is related to unique precursors and consequences. Given this background, we opted against using an overall ABQ score in the present study (cp. Gustafsson et al., 2016).

Third, we examined the convergent validity of the ABQ. In support of the convergent validity, positive correlations were observed between the ABQ indices, the SMBM, and most of the stress indices, whereas a negative association existed with satisfaction with life (Isoard-Gautheur et al., 2010; Raedeke et al., 2013; Sharp et al., 2010). These findings were expected, as chronic stress is seen as the most important cause of burnout among athletes (Gustafsson & Skoog, 2012). Moreover, the fact that the strongest correlations appeared between the SMBM overall index and the ABQ emotional/physical exhaustion subscale is (a) in line with experts' opinion that exhaustion should be seen as the core component of burnout (Gustafsson, DeFreese, et al., 2017; Shirom & Melamed, 2006) and (b) corresponds well with the fact that the SMBM is based on Shirom and Melamed's (2006) definition of burnout, in which burnout is described as being closely related to individuals' feelings of physical, emotional, and cognitive exhaustion due to the continuous depletion of their energetic coping resources due to chronic stress exposure.

Fourth, we explored the degree to which the ABQ scores are stable across time, and our data suggest that the scores of each subscale are relatively stable. In other words, all ABQ subscales were highly correlated across the 6-month study period. This finding is in line with previous research showing that in the domains of occupation and school, burnout symptoms have a high temporal stability (e.g., Lindwall et al., 2014; Salmela-Aro et al., 2009). Moreover, the auto-correlations of the ABQ subscales were similar compared to the other assessed psychological constructs such as domain-unspecific burnout symptoms, general perceived stress, depressive symptoms or satisfaction with life (all rs > 43, $p < .001$; see Table 2). By contrast, our findings are at odds with a prior study among elite athletes, where the stability was rather low (Lundkvist et al., 2018), which indicates that considerable variability may exist with regard to stability of athlete burnout symptoms in different study populations. To find out whether different study populations really differ with regard to stability of athlete burnout symptoms, more direct comparisons of different athlete populations are needed within the same country, and by employing the same study designs and instruments. Nevertheless, our findings also indicate that the three ABQ dimensions do not predict each other across time. Whereas this supports previous studies (Isoard-Gautheur et al., 2015;

Lundkvist et al., 2018; Martinent et al., 2016), none of the existing hypotheses about how the three burnout dimensions relate to each other temporally could be confirmed (Taris, Le Blanc, Schaufeli, & Schreurs, 2005). Moreover, although our findings corroborate prior research (Arce et al., 2012; Raedeke & Smith, 2001) showing that the emotional/physical and sport devaluation subscales of the ABQ have the potential to contribute to the prediction of domain-unspecific symptoms of burnout and depression, the amount of variance explained by these factors was limited (between 2 and 7%). Given that researchers have criticized the over-reliance on cross-sectional data in athlete burnout research (Gustafsson, DeFreese, et al., 2017), evidence for the predictive validity of the ABQ is important. However, the fact that not all ABQ dimensions were significant predictors of domain-unspecific symptoms of burnout and depression in our study seems to confirm that the symptoms assessed with the ABQ do not reflect a homogeneous syndrome. In line with this notion, Lundkvist et al. (2018) previously showed that the three symptoms develop relatively independent of each other.

Finally, based on ROC analyses, our cross-sectional (baseline) data suggest that the ABQ cannot be used to discriminate between athletes who suffer from clinically relevant burnout levels versus peers who do not. In other words, based on the SMBM cut-off of ≥ 4.40, it was not possible to simultaneously achieve an acceptable level of sensitivity and specificity. This is an important insight and underscores Gustafsson, Madigan et al.'s point of view that the choice of instrument should depend on the research questions that a researcher wants to address. Thus, "if the research questions are aimed towards comparing levels with existing data or looking at the changes of the three dimensions over time in a set context, then ABQ would be our measure of choice. But, if the research questions are aimed at burnout as a health problem in athletes, a measure that sets the results you get in comparison with cut-offs of clinical samples, then SMBM would be our recommendation" (Gustafsson, Madigan, et al., 2017, p. 11). Moreover, this finding also highlights that many scales, although developed properly in a scientific manner, may fail in the applied setting, and may not be appropriate as screening tools (Zeek et al., 2017).

A key strength of the present study is the relatively broad and representative sample of athletes from various sports. Taken together, we were able to assess approximately 90% of all

students attending sport classes at Swiss Olympic partner schools in Northwestern Switzerland. Another advantage was that we collected longitudinal data, which permitted us to examine temporal stability and predictive validity. Finally, the use of a domain-unspecific burnout measure associated with a clinically validated cut-off score allowed us to estimate the prevalence of clinically relevant symptoms of burnout more realistically, although more research is needed to reach a consensus on the nosological characterization of burnout. Whereas previous studies among Swiss adolescents showed that regular physical activity is associated with a decreased risk for elevated burnout symptoms (Elliot et al., 2015), our data indicate that young elite athletes are not 'immune' to burnout symptoms, and that psychological disorders are at least as prevalent among elite athletes as in the general population.

Despite these strengths, the findings should be interpreted in light of some limitations. First, most of the participants were male (68%), and only adolescent athletes were included. Moreover, young elite athletes not attending Swiss Olympic partner schools were not sampled. Therefore, further research is needed to examine whether the ABQ is better suited as a screening tool in other elite athlete populations. Finally, we acknowledge that we used a cut-off which was established with adult workers. The best suited cut-off for young people remains to be established in future research. Currently, however, this is the only empirically derived cut-off, and we preferred such a cut-off to an arbitrarily set threshold.

Conclusions

In the present study, we showed for the first time the psychometric properties of a German version of the ABQ. Moreover, the study provides unique insights into whether or not the ABQ can be used as a screening tool for clinically relevant burnout symptoms. In our sample, the factorial validity of the ABQ was supported, with CFA providing a good fit of the presumed three-factor model with the empirical data, satisfactory factor loadings, and adequate internal consistency. Nevertheless, our study corroborates previous concerns about the psychometric properties and validity of the ABQ. Most importantly, some ABQ dimensions shared only limited variance, the three ABQ dimensions did not predict each other across time, and none of the ABQ subscales was suitable for the screening of clinically

relevant burnout symptoms. Although we acknowledge that the development of the ABQ has advanced research on athlete burnout, we hold that this instrument has been adopted too uncritically by the scientific community. We therefore believe that further debates about the most suitable way to assess burnout among elite athletes are urgently warranted.

Acknowledgements

We thank Simone Best, Fabienne Meerstetter and Marco Walter for their support in the assessment of the baseline and follow-up data.

References

Altman, D.G., & Bland, J.M. (1994). Diagnostic tests 1: Sensitivity and specificity. *BMJ, 308*, 1552.

Arce, C., De Francisco, C., Andrade, E., Seoane, G., & Raedeke, T.D. (2012). Adaptation of the Athlete Burnout Questionnaire in a Spanish sample of athletes. *The Spanish Journal of Psychology, 15*, 1529-1536.

Byrne, B.M. (2010). *Structural equation modeling with AMOS. Basic concepts, applications, and programming*. New York: Taylor & Francis.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah: Erlbaum.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*, 385-396.

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis*. Hillsdale: Erlbaum.

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

Eades, A.M. (1990). *An investigation of burnout of intercollegiate athletes: The development of the Eades Athlete Burnout Inventory*. University of California, Berkeley: Unpublished master's thesis.

Eklund, R.C., & Cresswell, S.L. (2007). Athlete burnout. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (pp. 621-641). New York: Wiley & Sons.

513 Elliot, C., Lang, C., Brand, S., Holsboer-Trachsler, E., Pühse, U., & Gerber, M. (2015). The
514 relationship between meeting vigorous physical activity recommendations and
515 burnout symptoms among adolescents: an exploratory study with vocational students.
516 *Journal of Sport and Exercise Psychology*, *37*, 180-192.
517 Gerber, M., Feldmeth, A.K., Lang, C., Brand, S., Elliot, C., Holsboer-Trachsler, E., & Pühse,
518 U. (2015). The relationship between mental toughness, stress, and burnout among
519 adolescents: A longitudinal study with Swiss vocational students. *Psychological*
520 *Reports, 117*, 703-723.
521 Gerber, M., Kalak, N., Lemola, S., Clough, P.J., Perry, J.L., Pühse, U., . . . Brand, S. (2013).
522 Are adolescents with high mental toughness levels more resilient against stress? *Stress*
523 *& Health, 29,* 164-171.
524 Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the
525 Satisfaction with Life Scale (SWLS): Psychometric properties, validity, and
526 population-based norms. *European of Journal of Psychological Assessment, 27*, 127-
527 132.
528 Goodger, K., Gorely, T., Lavallee, D., & Harwood, Ch. (2007). Burnout in sport: a systematic
529 review. *The Sport Psychologist, 21*, 127-151.
530 Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit
531 dem „Gesundheitsfragebogen für Patienten (PHQ-D)“. *Diagnostica, 50,* 171-181.
532 Guedes, D.P., & de Souza, R.O. (2016). Psychometric properties of the Athlete Burnout
533 Questionnaire for young Brazilian athletes. *Journal of Physical Education, 27*, e2708.
534 Gustafsson, H., DeFreese, J.D., & Madigan, D.J. (2017). Athlete burnout: Review and
535 recommendations. *Current Opinion in Psychology, 16*, 109-113.
536 Gustafsson, H., Hancock, D.J., & Côté, J. (2014). Describing citation structures in sport
537 burnout literature: A citation network analysis. *Psychology of Sport and Exercise, 15*,
538 620-626.
539 Gustafsson, H., Kenttä, G., & Hassmén, P. (2011). Athlete burnout: An integrated model and
540 future research directions. *International Review of Sport and Exercise Psychology, 4*,
541 3-24.

Gustafsson, H., Lundkvist, E., Podlog, L., & Lundkvist, C. (2016). Conceptual confusion and potential advances in athlete burnout research. *Perceptual and Motor Skills, 123*, 1-8.

Gustafsson, H., Madigan, D.J., & Lundkvist, E. (2017). Burnout in athletes. In R. Fuchs & M. Gerber (Eds.), *Stressregulation und Sport*. Heidelberg: Springer.

Gustafsson, H., & Skoog, T. (2012). The mediational role of perceived stress in the relation between optimism and burnout in competitive athletes. *Anxiety, Stress, & Coping, 25*, 183-199.

Isoard-Gautheur, S., Guillet-Descas, E., Gaudreau, P., & Chanal, J. . (2015). Development of burnout perceptions during adolescence among high-level athletes: A developmental and gendered perspective. *Journal of Sport and Exercise Psychology, 37*, 436-448.

Isoard-Gautheur, S., Guillet-Descas, E., & Gustafsson, H. (2016). Athlete burnout and the risk of dropout among young elite handball players. *The Sport Psychologist, 30*, 123-130.

Isoard-Gautheur, S., Oger, M., Guillet, E., & Martin-Krumm, C. (2010). Validation of a French version of the Athlete Burnout Questionnaire (ABQ): In competitive sport and physical education context. *European Journal of Psychological Assessment, 26*, 203-211.

Kroenke, K., Spitzer, R.L., & Williams, J.B.W. (2001). The PHQ-9. Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606-613.

Lemyre, P.-N., Treasure, D. C., & Roberts, G.C. (2006). Influence of variability in motivation and affect on elite athlete burnout susceptibility. *Journal of Sport and Exercise Psychology, 28*, 32-48.

Lerman, Y., Melamed, S., Shragin, Y., Kushnir, T., Rotgoltz, Y., Shirom, A., & Aronson, M. (1999). Association between burnout at work and leukocyte adhesiveness/aggregation. *Psychosomatic Medicine, 61*, 828-833.

Lindwall, M., Gerber, M., Jonsdottir, I., Börjesson, M., & Ahlborg, G. Jr. (2014). The relationships of change in physical activity with change in depression, anxiety, and burnout: A longitudinal study of Swedish healthcare workers. *Health Psychology, 33*, 1309-1318.

Lu, J.H., Chen, L.H., & Cho, K.H. (2006). Revision of Raedeke and Smith's Athlete Burnout Questionnaire (ABQ): Analyses of validity and reliability of Chinese version. *Physical Education Journal, 39*, 83-94.

Lundgren-Nilsson, A., Jonsdottir, I.H., Pallant, J., & Ahlborg, G. (2012). Internal construct validity of the Shirom-Melamed Burnout Questionnaire (SMBQ). *BMC Public Health*, doi:10.1186/1471-2458-1112-1181.

Lundkvist, E., Gustafsson, H., Davis, P.A., Holmström, S., Lemyre, N., & Ivarsson, A. (2018). The temporal relations across burnout dimensions in athletes. *Scandinavian Journal of Medicine and Science in Sports*, doi:10.111/sms.13000.

Mardia, M.V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya, 36*, 115-128.

Martinent, G., Louvet, B., & Decret, J.C. (2016). Longitudinal trajectories of athlete burnout among young table tennis players: A 3-wave study. *Journal of Sport and Health Science*.

Maslach, C., & Jackson, S. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior, 2*, 99-113.

Maslach, C., Jackson, S.E., & Leiter, M.P. (1996). *Maslach Burnout Inventory Manual*. Palo Alto: Consulting Psychology Press.

McDonald, R.P., & Ho, R.M. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 65-69.

Pavot, W., & Diener, E. (2008). The Satisfaction With Life Scale and the emerging construct of life satisfaction. *Journal of Positive Psychology, 3*, 137-152.

Raedeke, T.D. (1997). Is athlete burnout more than stress? A commitment perspective. *Journal of Sport and Exercise Psychology, 19*, 396-417.

Raedeke, T.D., Arce, C., De Francisco, C., Seoane, G., & Ferraces, M.J. (2013). The construct validity of the Spanish version of the ABQ using a multi-trait/multi-method approach. *Anales de Psicologia, 29*, 693-700.

Raedeke, T.D., & Smith, A.L. (2001). Development and preliminary validation of an athlete burnout measure. *Journal of Sport and Exercise Psychology, 23*, 281-306.

Salmela-Aro, K., Savolainen, H., & Holopainen, L. (2009). Depressive symptoms and school burnout during adolescence: Evidence from two cross-lagged longitudinal studies. *Journal of Youth and Adolescence, 38*, 1316-1327. doi:1310.1007/s10964-10008-19334-10963.

Sharp, L.-A., Woodcock, C., Holland, M.J.G., Duda, J.L., & Cummings, J. (2010). Validation of the Athlete Burnout Questionnaire with youth athletes. *Journal of Sport and Exercise Psychology, 28*, S218-219.

Shirom, A., & Melamed, S. (2006). A comparison of the construct validity of two burnout measures in two groups of professionals. *International Journal of Stress Management, 13*, 176-200.

Smith, A.L., Gustafsson, H., & Hassmén, P. (2010). Peer motivational climate and burnout perceptions of adolescent athletes. *Psychology of Sport and Exercise, 11*, 453-460.

Swets, J.A., Dawes, R.M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*, 1-26.

Taris, T.W., Le Blanc, P.M., Schaufeli, W.B., & Schreurs, P.J.G. (2005). Are there causal relationships between the dimensions of the Maslach Burnout Inventory? A review and two longitudinal tests. *Work & Stress, 19*, 238-255.

Toppinen-Tanner, S., Kalimo, R., & Mutanen, P. (2002). The process of burnout in white-collar and blue-collar jobs: Eight-year prospective study of exhaustion. *Journal of Organizational Behavior, 23*, 555-570.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues, and applications* (pp. 56-75). Thousand Oakes: Sage.

Youngstrom, E.A. (2014). A primer on Receiver Operating Characteristics analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology, 39*, 204-221.

Zeek, A., Schlegel, S., Giel, K.E., Junne, F., Kopp, C., Joos, A., . . . Hartmann, A. (2017). Validation of the German version of the Commitment to Exercise Scale. *Psychopathology, 50*, 146-156.

628 Ziemainz, H., Abu-Omar, K., Raedeke, T.D., & Krause, K. (2004). Burnout im Sport. Zur
629 Prävalenz von Burnout aus bedingungsbezogener Perspektive. *Leistungssport, 34*, 12-
630 17.

Table 1. Descriptive statistics and corrected item-total correlations of the 15-item Athlete Burnout Questionnaire (ABQ)

	M	*SD*	Range	Skew	Kurt	r_{it}
Emotional/physical exhaustion	**2.66**	**0.73**	**1-5**	**0.25**	**0.00**	
Item 2: Mein Training ermüdet mich so stark, dass ich kaum noch Energie für andere Dinge habe. (*English: I feel so tired from my training that I have trouble finding energy to do other things.*)	2.88	1.00	1-5	0.24	-0.35	.59
Item 4: Durch meine sportliche Betätigung fühle ich mich erschöpft. *(English: I feel overly tired from my sport participation.)*	2.53	0.97	1-5	0.22	-0.46	.48
Item 8: Ich fühle mich ausgelaugt von meiner Sportart. *(English: I feel „wiped out" from sport.)*	2.22	1.00	1-5	0.53	-0.25	.58
Item 10: Durch meine Sportart fühle ich mich körperlich erschöpft. *(English: I feel physically worn out from sport.)*	2.93	0.96	1-5	-0.11	-0.34	.59
Item 12: Die geistigen und körperlichen Anforderungen in meiner Sportart erschöpfen mich. *(English: I am exhausted by the mental and physical demands of sport.)*	2.74	0.95	1-5	-0.03	-0.47	.66
Sport devaluation	**1.67**	**0.68**	**1-5**	**1.83**	**4.98**	
Item 3: Ich habe das Gefühl, dass ich die Energie, die ich in meine Sportart stecke, besser für andere Dinge verwenden sollte. *(English: The effort I spend in sport would be better spent doing other things.)*	1.62	0.83	1-5	1.51	2.65	.54
Item 6: Ich sorge mich in meiner Sportart nicht mehr gleich viel um meine Leistungen wie früher. *(English: I don't care as much about my sport performance as I used to.)*	1.88	1.12	1-5	1.16	0.35	.49
Item 9: Ich habe das Gefühl, mich nicht mehr gleich viel für meine Sportart zu interessieren wie früher. *(English: I'm not into sport like I used to be.)*	1.54	0.91	1-5	1.89	3.45	.67
Item 11: In meiner Sportart ist mir der Erfolg nicht mehr so wichtig, wie er es einmal war. *(English: I feel less concerned about being successful in sport than I used to.)*	1.51	0.90	1-5	1.94	3.23	.63
Item 15: Wenn ich an meine Sportart denke, habe ich zum Teil negative Gefühle. *(English: I have negative feelings toward sport.)*	1.80	0.90	1-5	0.90	0.22	.43
Reduced sense of accomplishment	**2.28**	**0.64**	**1-5**	**0.67**	**0.87**	
Item 1: Ich habe das Gefühl, in meiner Sportart erstrebenswerte Dinge zu leisten.[a] *(English: I'm accomplishing many worthwhile things in sport.)*	2.19	0.73	1-5 [a]	0.46	0.56	.49
Item 5: Ich habe das Gefühl, in meiner Sportart meine Ziele nicht zu erreichen. *(English: I am not achieving much in sport.)*	2.26	0.78	1-5 [a]	0.72	1.36	.59
Item 7: In meiner Sportart schaffe ich es nicht, meine Leistungsfähigkeit voll auszuschöpfen. *(English: I am not performing up to my ability in sport.)*	2.43	1.03	1-5	0.44	-0.37	.53
Item 13: Ich habe das Gefühl, egal was ich tue, nicht das zu leisten, was ich leisten sollte. *(English: It seems that no matter what I do, I don't perform as well as I should.)*	2.18	1.04	1-5	0.70	-0.09	.57
Item 14: In meiner Sportart fühle ich mich erfolgreich.[a] *(English: I feel succesful at sport.)*	2.37	0.81	1-5 [a]	0.29	-0.09	.64

Note. [a]Inverted scale (1→5, 2→4, 3=3, 4→2, 5→1).

Table 2. Correlations between dimensions of athlete burnout with the overall burnout, general perceived stress and school-based stress at baseline

	Baseline (N=257)		*Follow-up (N=197)*											
	M	*SD*	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Emotional/Physical Exhaustion (ABQ)	2.66	0.73	2.59	0.76	.65***									
2. Sport Devaluation (ABQ)	1.67	0.68	1.83	0.84	.26***	.57***								
3. Reduced Sense of Accomplishment (ABQ)	2.28	0.64	2.33	0.67	.21**	.52***	.65***							
4. Overall burnout (SMBM)	3.32	0.96	3.25	1.04	.52***	.36***	.31***	.56***						
5. Physical fatigue	3.55	1.13	3.46	1.17	.57***	.35***	.31***	.89***	.58***					
6. Cognitive weariness	3.34	1.08	3.24	1.16	.36***	.33***	.31***	.86***	.65***	.52***				
7. Emotional exhaustion	2.80	1.16	2.84	1.17	.29***	.22***	.13*	.71***	.49***	.47***	.43***			
8. General perceived stress (PSS)	2.64	0.49	2.59	0.50	.33***	.29***	.46***	.55***	.48***	.48***	.37***	.64***		
9. Depressive symptoms (PHQ-9)	7.38	4.70	5.38	3.90	.19**	.17**	.28***.	.39***	.38***	.37***	.23***	.29***	.49***	
10. Satisfaction with life (SWLS)	27.52	4.42	27.50	4.30	-.22***	-.30***	-.47***	-.43***	-.36***	-.34***	-.35***	-.50***	-.29***	.57***

Note. ABQ=Athlete Burnout Questionnaire, SMBM=Shirom-Melamed Burnout Measure, PSS=Perceived Stress Scale, PHQ-9=Patient Health Questionnaire 9, SWLS=Satisfaction with Life Scale. Auto-correlations between variables are displayed in the diagonal.
*p<.05. **p<.01. ***p<.001.

Table 3. Hierarchical multiple regression analyses predicting burnout symptoms and depressive symptoms at the 6-month follow-up with baseline scores on the ABQ subscales

	Burnout Symptoms (SMBM)		Depressive Symptoms (PHQ9)	
	ΔR^2	β	ΔR^2	β
Step 1	.298***		.244***	
Baseline Levels		.45***		.40***
Step 2	.021*		.073***	
ABQ Emotional/Physical Exhaustion		.14*		.18***
ABQ Sport Devaluation		.02		.13***
ABQ Reduced Sense of Accomplishment		.05		.07
Total R^2	.319***		.317***	

Note. Regression weights are presented as they are after step 2.
***p<.001. **p<.01. *p<.05.

Table 4. Coordinates of the Receiver Operating Characteristics (ROC) curve for the ABQ physical/emotional exhaustion subscale

Positive if ≥ to:	Sensitivity (true positive rate)	1-Specificity (false positive rate)
1.10	1.00	.99
1.30	1.00	.97
1.50	1.00	.94
1.70	1.00	.91
1.90	.97	.84
2.10	.90	.76
2.25	.87	.66
2.35	.87	.66
2.50	.87	.54
2.65	.74	.44
2.75	.71	.44
2.85	.58	.32
2.95	.58	.32
3.05	.48	.20
3.15	.48	.20
3.25	.48	.13
3.35	.48	.12
3.50	.45	.09
3.70	.32	.06
3.85	.19	.04
3.95	.16	.04
4.10	.10	.04
4.30	.03	.00
4.70	.03	.00

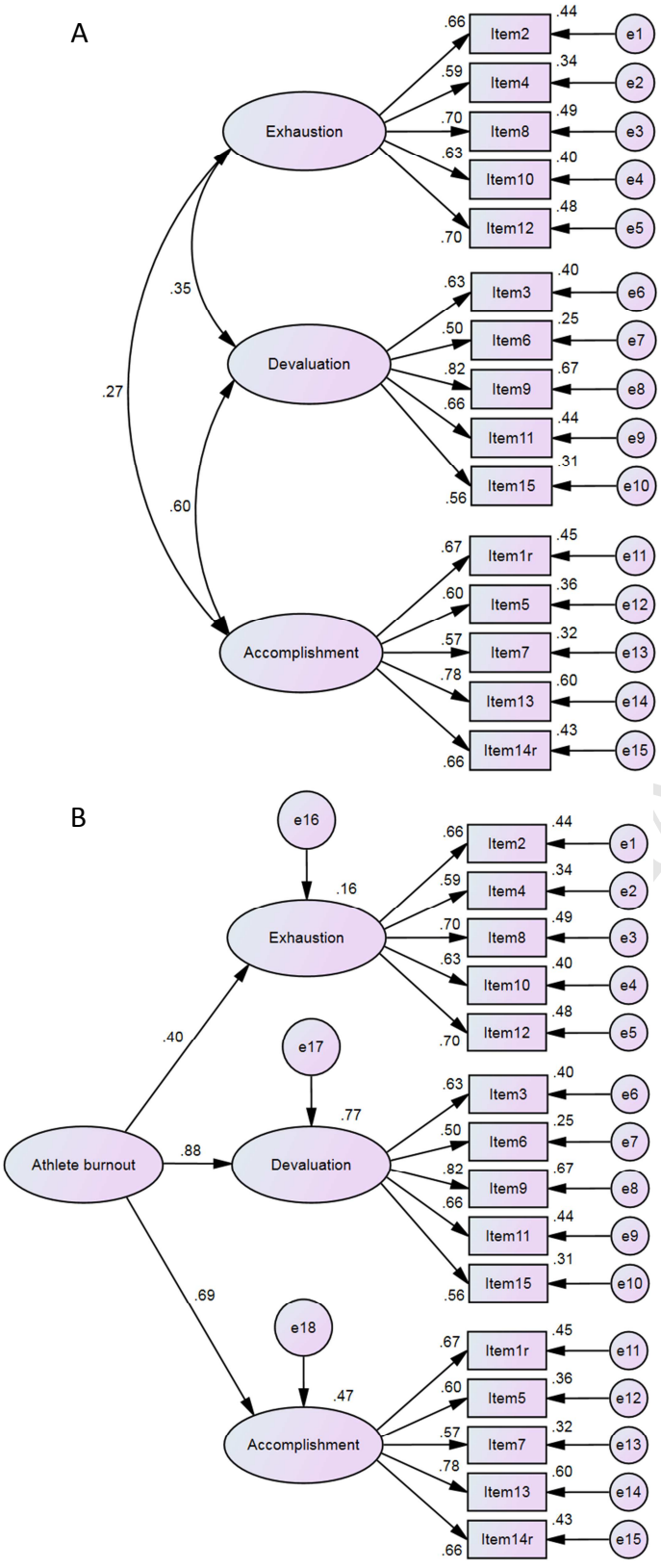


Figure 1. Factor structure of the Athlete Burnout Questionnaire and factor loadings for the first-order (A) and second-order analysis (B) at baseline
Note. r=reverse-scored, e=residual error

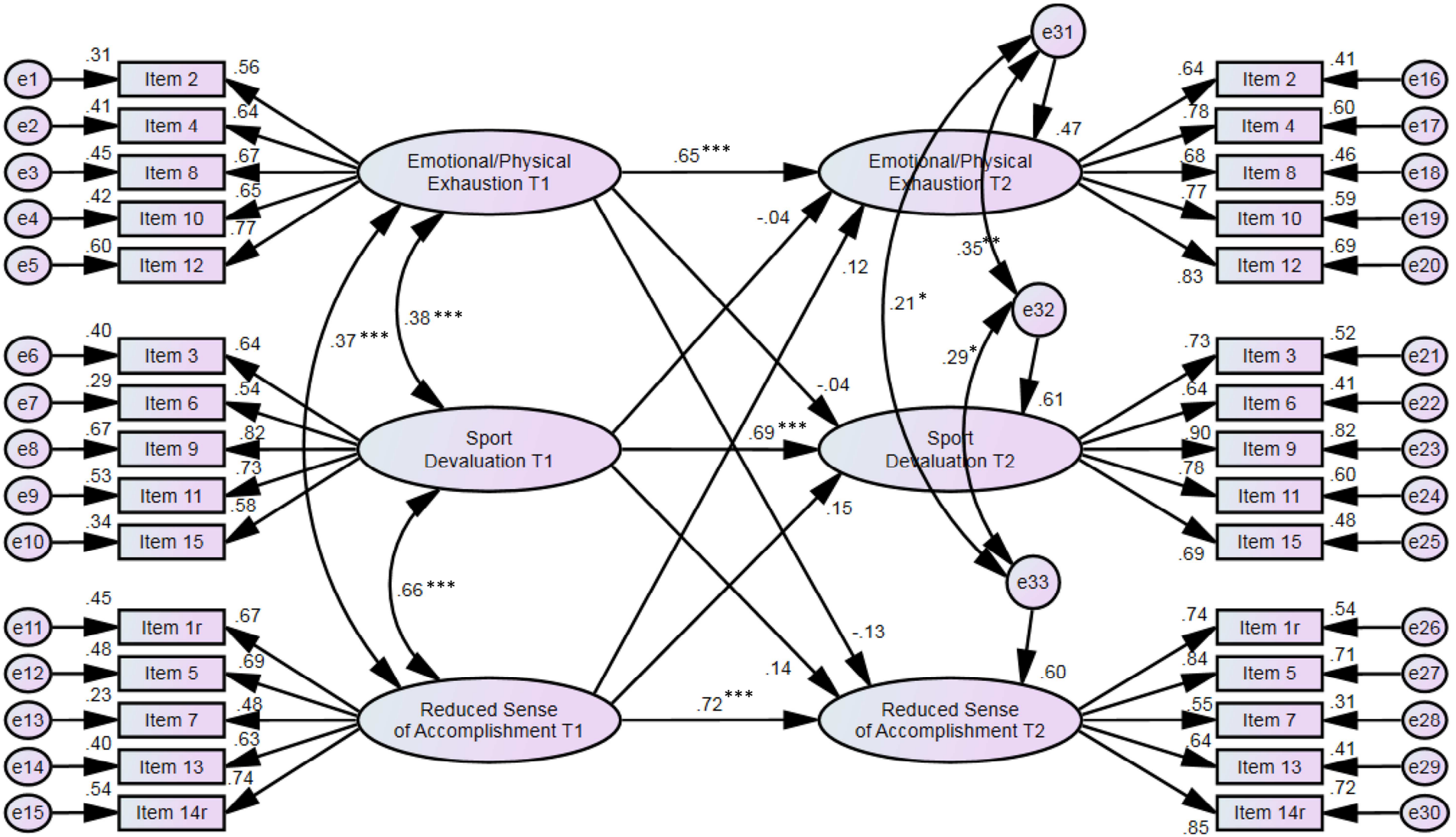
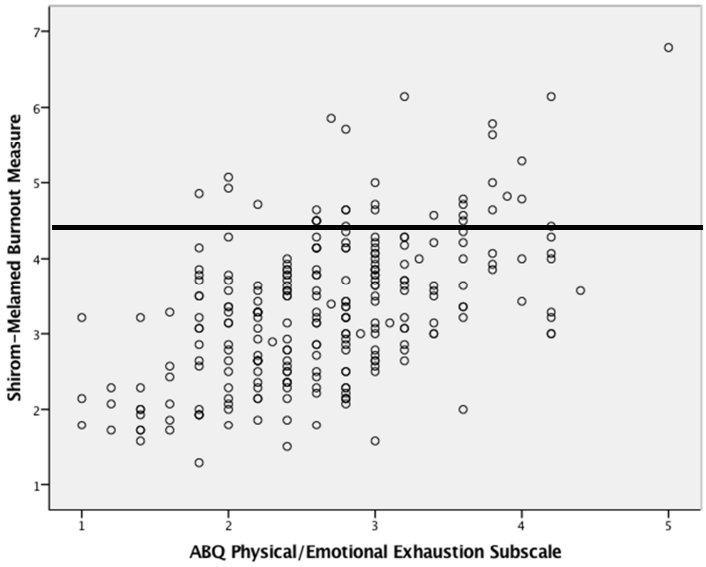
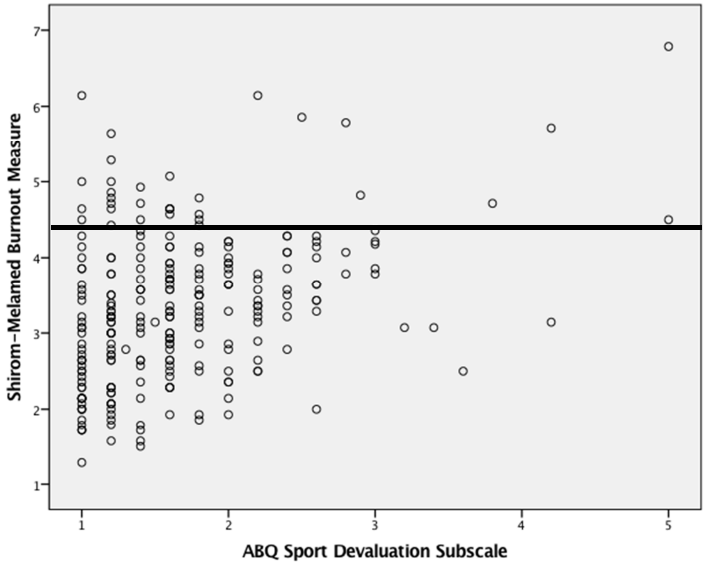


Figure 2. Cross-lagged panel analyses examining the temporal interplay between the three ABQ dimensions. *p<.05. **p<.01. ***p<.001.





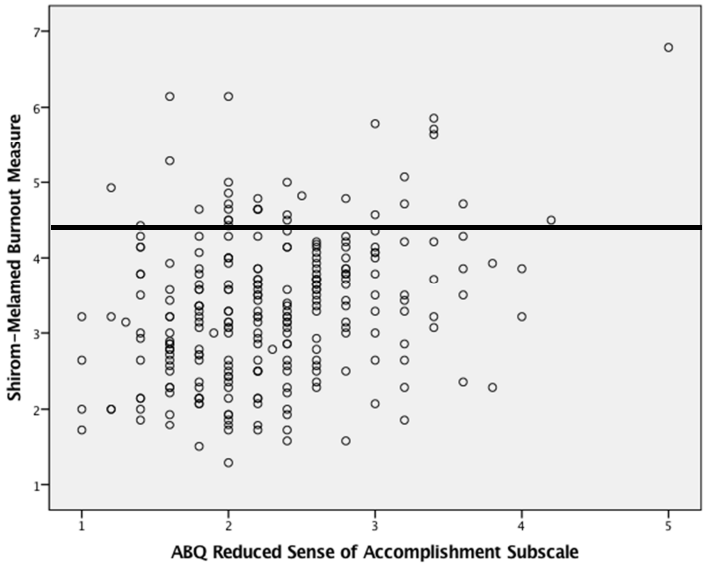


Figure 3. Scatterplot representing the correlations between the Athlete Burnout Questionnaire (ABQ) indices and the Shirom-Melamed Burnout Measure (SMBM) overall index (at baseline).
Note. Black horizontal line indicates SMBM cut-off for clinically relevant burnout symptoms.

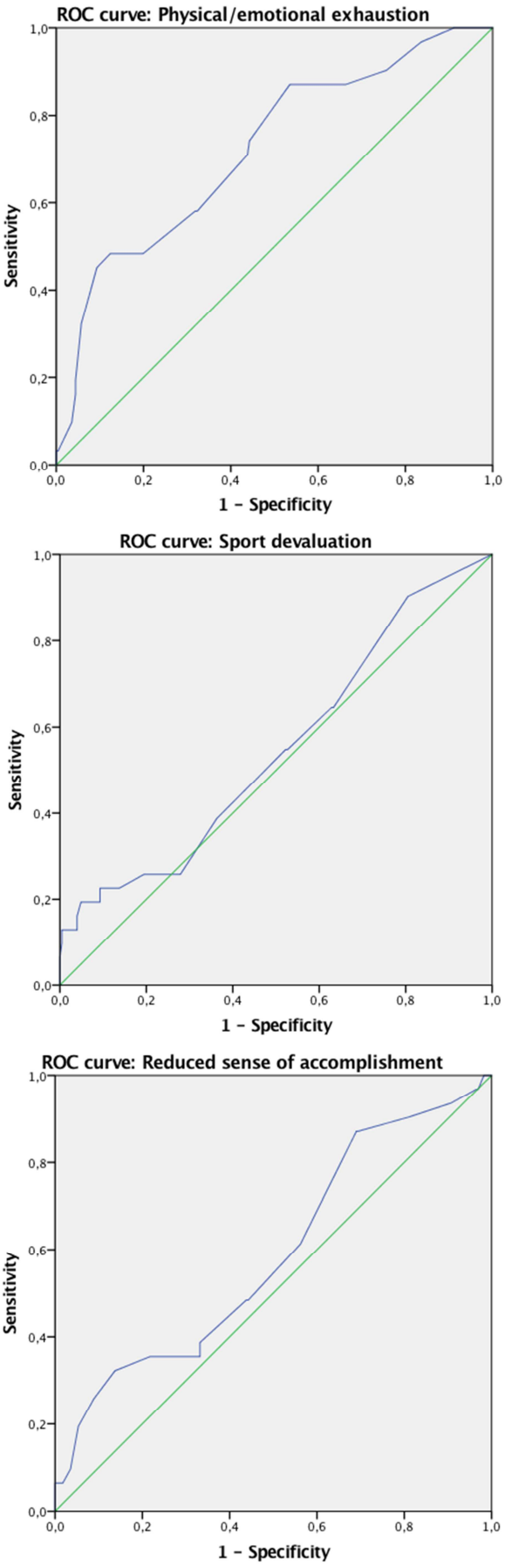


Figure 4. Receiver Operating Characteristics (ROC) curves, separately for the three Athlete Burnout Questionnaire (ABQ) subscales

Highlights

- Prospective study with 257 young elite athletes, assessed twice across a 6-month period
- Confirmatory factor analysis supported the three-factor structure of the ABQ.
- All ABQ subscales had acceptable internal consistency and show acceptable convergent validity
- However, some ABQ subscales shared only limited variance and did not predict each other across time
- None of the ABQ subscales was suitable for the screening of clinically relevant burnout symptoms