

How to Measure Coach Burnout: An Evaluation of Three Burnout Measures

Erik Lundkvist and Andreas Stenling

Department of Psychology, Umeå University, Umeå, Sweden

Henrik Gustafsson

Department of Health and Environmental Sciences, Karlstad University, Karlstad, Sweden

Peter Hassmén

*Department of Psychology, Umeå University, Umeå, Sweden; Faculty of Health,
University of Canberra, Australia*

Although coach burnout has been studied for 30 years, what measure to use in this context has not yet been problematized. This study focuses on evaluating convergent and discriminant validity of three coach burnout measures by using multi-trait/multi-method analysis (CT-C[M-1]) model. We choose Maslach Burnout Inventory (MBI), the two dimensional Oldenburg Burnout Inventory (OLBI), and a coach version of Athlete Burnout Questionnaire (CBQ). Our analysis shows that MBI and OLBI cover similar definitions of exhaustion and depersonalization/disengagement and that CBQ measures somewhat different dimensions. A problem for OLBI is a lack of discriminant validity due to high correlations between exhaustion and disengagement. For lack of personal accomplishment/reduced sense of accomplishment CBQ measures a somewhat different construct than MBI. Although all three measures have advantages and disadvantages, we promote CBQ since it discriminates between dimensions and covers important aspects of burnout in a sports context that the other two do not cover.

Key words: convergent validity, discriminant validity, Maslach burnout inventory, Oldenburg burnout inventory, coach burnout inventory

INTRODUCTION

The work environment for sports coaches is both challenging and stressful. The daily work with athletes, combined with performance-related pressures, budget concerns, and relations with colleagues are a few examples of potential stressors. While some coaches handle it well, others may feel the strain of accumulating stress, which may be problematic. Experiencing these kinds of

stressful work situations over time may create burnout-related problems. These conditions may also lead to some coaches leaving the coaching ranks (Raedeke, 2004).

Since the 1980s, studies on burnout in sports settings have been conducted with both athletes and coaches (Goodger, Gorely, Lavalley, & Harwood, 2007). In contrast to athlete burnout, whereby the concept and how it should be operationalized and measured are regularly debated (Cresswell & Eklund, 2006; Dale & Weinberg, 1990; Raedeke, Arce, Seoane, & De Francisco, 2013; Raedeke, 1997), such issues have seldom been topics of discussion in coach burnout research (Goodger et al., 2007). Therefore, in this article we raise the question of how coach burnout should be defined and operationalized, by comparing three currently used burnout measures.

The most common definition of burnout in occupational research is based on the Maslach Burnout Inventory (MBI; Maslach & Jackson, 1981). In this definition, burnout is described as a syndrome with three dimensions, whereby the interpersonal relationships between provider and recipient in healthcare or educational settings are argued to cause and be symptoms of burnout (Maslach, Schaufeli, & Leiter, 2001). The first dimension is emotional exhaustion, which is seen as the core dimension and defined as exhaustion due to prolonged work life stress, thereby covering the stress part of burnout (Maslach et al., 2001). Depersonalization is the second dimension, focusing on how the individual distances her-/himself from recipients or students and ignores the things that make them human. This is seen as a defense mechanism, whereby the person troubled by burnout makes demands more manageable. Third, lack of personal accomplishment relates to perceived accomplishments, when both doing tasks at work and handling recipients/students (Maslach et al., 2001).

The original MBI builds on this three-dimensional conceptualization of burnout, and is based on research in healthcare settings (labeled MBI-HSS; Maslach & Jackson, 1981). A teacher version was also developed (MBI-ES). These two versions only differ from each other in that the MBI-HSS uses the word “recipients,” whereas the MBI-ES uses “students” (Maslach, Jackson, & Leiter, 1996). A more thoroughly changed context-free general survey was later developed (MBI-GS; Schaufeli, Leiter, Maslach, & Jackson, 1996). The emotional exhaustion dimension in the general survey is the same; however, depersonalization and lack of personal accomplishment were broadened and given new names to fit occupations that are less person-oriented (Maslach et al., 2001). Consequently, the second dimension was labeled cynicism, focusing on cynical feelings toward and disengagement from one’s job instead of recipients/students (Demerouti, Bakker, Vardakou, & Kantas, 2003). The third dimension, professional efficacy, focuses on how a person perceives his/her performance on the job without referring to issues with colleagues or recipients (Schaufeli et al., 1996). Although these changes have broadened the concept, it is argued that the MBI-GS can be used to measure the same dimensions as the original MBI-HSS (Maslach et al., 2001).

The burnout concept has—partly because of the above—received criticism. For example, it has been questioned whether it is a syndrome, since lack of personal accomplishment rarely correlates with the other two dimensions (Lee & Ashforth, 1996). The three dimensions are not intended to be collapsed into one “total burnout” variable, which further underlines the problems with the syndrome argumentation (Shirom, 2005). Further, the development of the MBI has been criticized because of the inductive procedure used, instead of being grounded in theory-driven deductive methods (Shirom, 2005). Another issue relates to the wording; for example, “I feel I treat some

recipients/students as if they were impersonal objects” uses dated language (Demerouti et al., 2003; Lee & Ashforth, 1990).

To remedy some of these shortcomings, the Oldenburg Burnout Inventory (OLBI) was developed. The OLBI includes two dimensions of burnout that build on the emotional exhaustion and cynicism of the MBI-GS. The OLBI covers both physical and emotional aspects of exhaustion, making it suitable for both physically and cognitively challenging jobs. Its second dimension—disengagement—is a development of the cynicism dimension of the MBI-GS that also includes items tapping the individual’s distancing her-/himself from the job itself (Demerouti et al., 2003; Halbesleben & Demerouti, 2005). Although some criticism of the MBI are solved, the problem of the inductive instead of theory-driven approach during the development still stands (Shirom, 2005).

A measure specifically developed for sport environments is the Athlete Burnout Questionnaire (ABQ). This measure was originally developed for athletes, but has recently been reworded to fit a coaching context and renamed the Coach Burnout Questionnaire (CBQ; Harris & Ostrow, 2008; Malinauskas, Malinauskiene, & Dumciene, 2010; Raedeke & Smith, 2001). Although sport-focused, the ABQ/CBQ was developed from the same three dimensions of burnout as the MBI (Raedeke & Smith, 2001). The first dimension is emotional/physical exhaustion, tapping both mental and physical exhaustion. Sports devaluation is the second dimension, measuring the extent to which the coach stops caring about sports and his/her performance as a coach. Finally, the third dimension measures reduced sense of accomplishment, covering coaches’ feelings of reduced skills in their coaching role (Malinauskas et al., 2010; Raedeke & Smith, 2001; Raedeke, 1997).

Of the 30 published coach burnout studies, the authors of 28 stated which instrument they used. In 27 of these 28 studies, MBI was the instrument of choice (20 used a coach-modified original version, six used a coach-modified educator’s version; one did not state the version used). Of the 27 MBI studies, 24 used all subscales; the remaining three only used the emotional exhaustion subscale. The CBQ was used in one study.

The dominant use of MBI in coach burnout research is questionable, because neither the MBI-HSS nor the MBI-ES was developed for such a context. This may create problems, for example in the factor structure. Studies that have performed confirmatory factor analysis (CFA) with the MBI-HSS indeed indicate that it seems to work better in health settings than educational settings, but the results are somewhat ambiguous, with variations from good model fit (Hallberg & Sverke, 2004; Vanheule, Rosseel, & Vlerick, 2007) to more problematic model fit (Schaufeli, Bakker, Hoogduin, & Kladler, 2001). The factor structure of the MBI-ES seems even more problematic (Worley, Vassar, Wheeler, & Barnes, 2008). Based on work characteristics, health and educational occupations differ from each other, and burnout is argued to be a context-specific construct (Eklund & Cresswell, 2007). Since early burnout research stems from health occupations, in which provider and recipient relationships are central (Maslach et al., 2001), the claim that health and educational contexts are similar can therefore be questioned. Although educational jobs also include care-taking provider-recipient relations, they also aim at developing other persons in their learning processes, which is different from work tasks in healthcare occupations. From this point of view, it seems natural that the original MBI fits better in the setting it was developed for than in educational settings. When adapting the MBI-ES to a sports coaching context the occupational characteristics move even further away from health occupations, as most coaches in sports deal with not only stressors that originate from interpersonal relations but also

stressors that can have a more direct relation to sporting results (and ultimately whether they will or will not keep their job). Thus, an evaluation of the burnout instruments currently available for the coaching context is needed. Our aim with this study was to evaluate three burnout measures: the coach versions of the MBI, OLBI, and CBQ. All three are based on the same operationalization of burnout (Halbesleben & Demerouti, 2005; Raedeke et al., 2013), albeit with differences in wording and their respective subscales.

Aims

Our overall aim with this study was to evaluate three related operationalizations of the burnout construct that are available to use in a coach burnout context. First, the MBI, which is the predominant three-dimensional operationalization for coach burnout, grounded in the traditional intra-individual definition of burnout (Maslach et al., 1996). Second, the OLBI, a context-free two-dimensional operationalization with a broader definition of the exhaustion dimension, and disengagement toward the job itself. Third, the sport context-specific CBQ with a broader exhaustion dimension, whereby both devaluation and reduced sense of accomplishment target the job instead of personal relationships in the job (Harris & Ostrow, 2008; Raedeke & Smith, 2001).

The first analysis performed was a test of the factor structure of the three instruments using CFA. The theoretical models for all three instruments were tested individually, and contingent changes were made by deleting items based on theoretical and empirical grounds until the resulting factor structure demonstrated an acceptable model fit.

The second analysis tested whether the three measures represent related or different burnout constructs. A multi-trait/multi-method (MTMM) analysis was used to establish convergent and discriminant validity of the three measures in one structural equation model (Eid, Lischetzke, Nussbeck, & Trierweiler, 2003). Based on the arguments of the researchers who developed the tested instruments, the definitions of the dimensions of burnout in the three tested measures are based on the same construct. We therefore, predict that dimensions related to exhaustion—that is emotional exhaustion (MBI), mental and physical exhaustion (CBQ), and exhaustion (OLBI)—in the three tested methods will show convergent validity. We also predict that dimensions covering depersonalization (MBI), devaluation (CBQ), and disengagement (OLBI) will show convergent validity. Lastly, we predict that sense of accomplishment (MBI) and reduced sense of accomplishment (CBQ) will show convergent validity.

METHOD

Participants and Procedure

Questionnaires were distributed as a web survey to 604 operating email addresses of coaches living in Sweden earning an income as coaches. Of these, 277 completed a usable questionnaire, giving a response rate of 46% (236 coaches were male, 39 female, and two did not want to specify their gender). The population covers coaches from 38 sports, with 28 individual (125 persons) and ten team sports (154 persons). Sporting levels range from the highest Swedish elite to youth

elite, in 13 team and 28 individual sports. The coaches ranged in age from 20 to 82 years (mean 42.9, SD 10.1). Of all the coaches, 149 had coaching as their full time job, 36 were employed 75%–99%, 49 were employed 50%–74%, and 38 were employed 20%–49%, while six did not answer. Of those working solely as coaches or combining coaching with another job, 237 were employed 100% or more, 30 were employed 50%–99%, and two worked less than 25%, while eight did not answer. For total working hours per week, 38 worked less than 39 hours, 132 worked 40–49 hours, and 94 worked more than 50 hours, while 12 did not answer.

Measures

Burnout was measured using three measures, with the definition of burnout based on Maslach (Demerouti et al., 2003; Raedeke et al., 2013). However, the measures differ when it comes to contextual adaption, inclusion of dimensions, and the positive and negative wording of items.

The first tested measure was the MBI-ES, with “students” changed to “athletes” to fit the coaching context (Kelley & Gill, 1993; Vealey, Udry, Zimmerman, & Soliday, 1992). Items were answered on a seven-point scale where 0 = *never*, 1 = *a few times a year or less*, 2 = *once a month or less*, 3 = *a few times a month*, 4 = *once a week*, 5 = *a few times a week*, and 6 = *every day*. The MBI-ES contains three dimensions: Emotional exhaustion, covering feelings of fatigue and tiredness related to working life (Maslach et al., 1996), for example, “I feel used up at the end of the workday;” Depersonalization taps feelings of cynicism toward the athletes, for example, “I feel I treat some athletes as if they were impersonal objects;” Lack of personal accomplishment taps feelings about lost performance in the job. The wording is positively framed but the items are reversed. An example item is, “I feel I’m positively influencing other people’s lives through my work.”

The OLBI is a context-free measure that includes two subscales: exhaustion and disengagement. Items are framed both positively and negatively, with the positive items reversed. Items are answered on a four-point scale where 1 = *strongly agree*, 2 = *agree*, 3 = *disagree*, and 4 = *strongly disagree* (Demerouti et al., 2003). Exhaustion taps both physical and mental strain related to working life. An example item is, “During my work, I often feel emotionally drained.” Disengagement covers questions about distancing oneself from one’s job, linked to engagement and identification with the job. Compared to the depersonalization dimension of the MBI-ES, it has a broader focus than cynicism toward athletes (Demerouti et al., 2003). An example item is, “I feel more and more engaged in my work.”

The CBQ is based on the athlete-specific ABQ (Malinauskas et al., 2010; Raedeke & Smith, 2001), with the wording changed to better suit the coaching context. Items are answered on a five-point scale where 1 = *almost never*, 2 = *rarely*, 3 = *sometimes*, 4 = *frequently*, and 5 = *almost always*. Of the 15 items, 13 are framed negatively and two positively. The first dimension, emotional/physical exhaustion, taps both mental and physical exhaustion. An example item is, “I feel tired from coaching.” Sport devaluation covers feelings of lost concern about what is important as a coach, namely the sport itself and one’s performance as a coach. “The effort I spend coaching would be better spent doing other things.” Reduced sense of accomplishment taps feelings of reduced skills in the coaching role. An example item is, “It seems that no matter what I do, I don’t coach as well as I should.”

Statistical Analysis

Statistical analyses were done using Mplus software (version 7.11). In the first step, the three measures were analyzed using CFA to test the hypothesized factor structure for the individual measures. The second step was to combine all three models into one MTMM analysis to test for discriminant validity, convergent validity, and method specificity by combining all three measures in one SEM model. We used the weighted least squares mean and variance-adjusted (WLSMV) estimation method because it performs well when variables are not normally distributed, where sample size is small, and the statistical model is complex. WLSMV estimation is also favorable in SEM modeling with ordinal variables (Nussbeck, Eid, & Lischetzke, 2006).

Chi-square (χ^2), comparative fit index (CFI), root mean square error of approximation (RMSEA), RMSEA with a 90% confidence interval (RMSEA-CI), and Tucker-Lewis Index (TLI) were used to assess model fit. To evaluate model fit values greater than .90 indicate acceptable fit for CFI and TLI. For RMSEA values under .08 indicate acceptable fit (Marsh, 2007).

MTMM Analysis

We chose the correlated trait-correlated method minus one model (CT-C[M-1]), shown in Figure 1. This model allows both trait factors and method factors to correlate. M-1 indicates a chosen comparison standard method that is contrasted against the other tested method factors and is not linked to any of the other method factors. The indicators of the comparison standard method does not have a method factor and is therefore only correlated to the trait factor (Eid et al., 2003). The choice of comparison standard method should be based on previous research results and theoretical considerations (Geiser, Eid, & Nussbeck, 2008). The MBI was chosen as the comparison standard method for this study because it is the most commonly used and serves as a theoretical base for the other tested measures. In the tested model all indicators load on a trait—emotional exhaustion, depersonalization or lack of lack of personal accomplishment—defined by the comparison standard method. The comparison standard is therefore the method that the other method factors are compared with (Eid et al., 2003). The indicators also load on their specific method and cover the variance that is specific to the method and not shared with the comparison standard method. If the factor loading for an indicator is higher on the trait factor than the specific method factor, this means it has high convergent validity with the comparison standard and the measures cover resembling constructs (Grigorenko, Geiser, Slobodskaya, & Francis, 2010). Trait correlations test discriminant validity, with low correlations indicating high discriminant validity, meaning that traits measure different dimensions of burnout; high correlations ($r \geq .70$) indicate that traits cover resembling constructs that are hard to separate (Eid, Nussbeck, Geiser, Cole, Gollwitzer, & Lischetzke, 2008).

Convergent validity was tested with consistency coefficients and method specificity. Consistency coefficients indicate the degree of true variance of an observed indicator of a method factor that is explained by the comparison method. The method specificity coefficients represent the proportion of true variance in an observed indicator that is specific to the particular

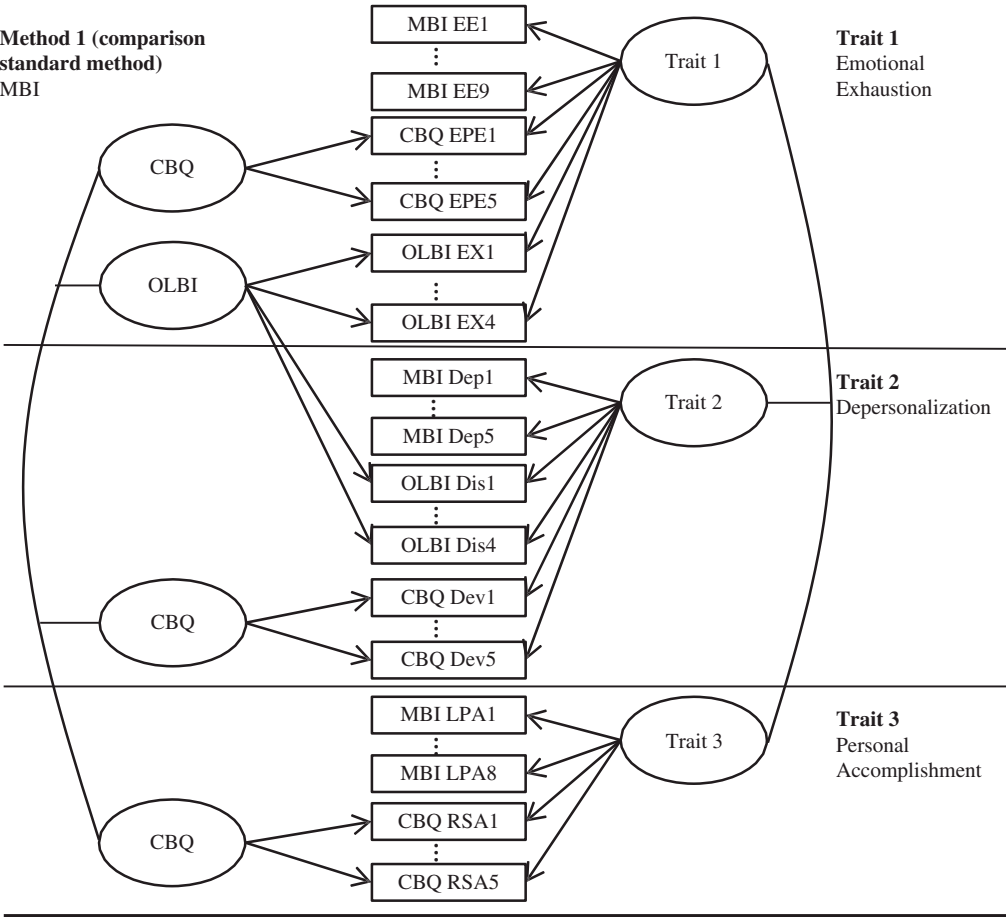


FIGURE 1 CT-C(M-1) model. MBI EE = MBI emotional exhaustion; MBI Dep = MBI depersonalization; MBI LPA = MBI lack of personal accomplishment; OLBI EX = OLBI exhaustion; OLBI Dis = OLBI disengagement; CBQ EPE = CBQ emotional and physical exhaustion; CBQ Dev = CBQ devaluation; CBQ RSA = CBQ reduced sport accomplishment. Residuals of the indicator variables are not shown.

method factor and not shared with the comparison standard. In our case, between 5% and 100% of the variance in the CBQ and OLBI items is explained by the standard method (MBI). If the consistency coefficient is larger than the method specificity coefficient, this indicates that inter-individual differences in the ratings are primarily determined by trait differences rather than by trait-specific method differences. Higher consistency coefficients indicate higher convergent validity.

RESULTS

CFA

Table 1 shows the results of the CFAs of the proposed models and the revised models for each instrument. Model fit values for the MBI Model A were acceptable. However, Item 20 (emotional exhaustion), “I feel like I’m at the end of my rope,” cross-loaded on the lack of personal accomplishment dimension (modification indices $\Delta\chi^2 = 63.097$) and was removed. Previously, this item has been removed because it covers a very extreme situation of emotional exhaustion (Oh & Lee, 2009). For the CBQ the theoretical model (C) shows acceptable model fit, but there are some problems with Items 1 (I’m accomplishing many worthwhile things in coaching), 14 (I feel successful at coaching), and 15 (I have negative feelings toward coaching). Items 1 and 14 were the two that reversed and shared covariance (modification indices $\Delta\chi^2 = 47.197$) to the degree that Item 1 was removed. Item 15 (DEV) was removed because of cross-loadings (modification indices $\Delta\chi^2 = 38.314$) on the MPE dimension. Model fit for the OLBI was not acceptable for most indices. The eight items that were reverse coded, generally had lower loadings on the proposed dimension and were therefore deleted, giving a measure with eight negatively worded items, four items for each dimension.

MTMM Analyses

The estimated model shown in Figure 1 has the same structure as the tested CT-C(M-1) model (G) presented in Table 1. However, this model does not correspond exactly to the theoretical model. First, in the CFA analysis items were removed from all three methods, as presented in Table 1. The original CT-C(M-1) model displayed an inadmissible solution (i.e., a correlation coefficient larger than 1 between the two OLBI factors). Therefore, they were collapsed into a

TABLE 1
Fit Indices for CFAs on MBI, CBQ, OLBI and CT-C(M-1) Models

Model	χ^2 (WLSMV)	df	CFI	TLI	RMSEA	90% CI	
						Lower	Upper
A	482.079	167	.924	.913	.083	.074	.091
B	393.487	149	.939	.930	.077	.068	.086
C	306.463	87	.942	.930	.095	.084	.107
D	174.692	62	.967	.959	.081	.067	.095
E*	290.181	103	.897	.879	.081	.070	.092
F	61.751	19	.959	.940	.090	.066	.116
G	1,242.182	703	.943	.936	.053	.048	.057

Note: Model A = CFA MBI-ORG (excluding Items 12 and 16); Model B = CFA MBI-REV (excluding Items 12, 16, and 20); Model C = CFA CBQ; Model D = CFA CBQ-REV (excluding Items 1 and 15); Model E = OLBI org; Model F = CFA OLBI (only positively phrased items); Model G = CT-C(M-1) MBI standard (single OLBI factor).

*OLBI loadings reverse coded items: EX: 5 = .539, 10 = .666, 14 = .327, 16 = .441; DIS: 1 = .454, 7 = .479, 13 = .292, 15 = .282.

single OLBI factor. When the proposed changes had been made, the tested model showed good fit with the data.

Discriminant Validity

The correlations between the trait variables (T1-T3) represent the discriminant validity between the different traits at the level of the comparison method MBI (Table 2). These correlations should be not be too high ($r \geq .70$) to justify the instruments measuring different constructs (Eid et al., 2008). Correlations between exhaustion and depersonalization/devaluation traits are high (factor $r = .766$), indicating low discriminant validity. Lack of personal accomplishment shows low correlations to emotional exhaustion (factor $r = .256$) and depersonalization (factor $r = .462$). This indicates that lack of personal accomplishment has high discriminant validity and is justified as a different construct in the scales.

Convergent Validity and Method Specificity

The correlations between trait-specific method factors determine the generalizability of method effects across traits (M11–M13). The correlations between the CBQ method factors (r s from .364 to .504) support the assumption of trait-specific method factors (Table 2). These correlations indicate how strongly an over- or underestimation on one trait-specific method factor is related to the over- or underestimation on another trait-specific method factor of the same method.

The correlations between CBQ method factors and the OLBI method factor (r s from .239 to .658) indicate that devaluation has a high correlation with the combined OLBI factor, exhaustion has a moderate correlation with OLBI, and reduced sense of accomplishment has the lowest correlation with OLBI (Table 2). Since reduced sense of accomplishment is not included in the OLBI, this is to be expected.

The standardized trait and method-specific factor loadings are displayed in Table 3, along with consistency coefficients and method specificity coefficients. In our case, the convergent validity appears to be quite high for exhaustion items, with higher convergent validity for OLBI items than

TABLE 2
Trait and Method Factor Correlations in the CT-C(M-1) Model with MBI as Standard Method

Variable	1	2	3	4	5	6	7
1. T1 (EE)	—						
2. T2 (Dep/Dis/Dev)	.766	—					
3. T3 (PA)	.258	.462	—				
4. M11 (CBQ-E)		<u>-.049</u>	<u>.099</u>	—			
5. M12 (CBQ-Dev)	.198		.209	.504	—		
6. M13 (CBQ-RSA)	.449	.419		.364	.491	—	
7. M2 (OLBI)			.203	.425	.658	.239	—

Note: Underlined correlation coefficients were not significant at $p < .05$.

EE = emotional exhaustion; Dep = depersonalization; Dis = disengagement; Dev = devaluation; PA = lack of personal accomplishment; E = exhaustion; RSA = Reduced sense of accomplishment.

TABLE 3
Standardized Loadings in CT-C(M-1) Model with MBI as Standard Method

	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>M11</i>	<i>M12</i>	<i>M13</i>	<i>M2</i>	<i>Consistency Coefficient</i>	<i>Method Specificity</i>
Emotional exhaustion									
MBI									
1	.834								
2	.803								
3	.802								
6	.575								
8	.811								
13	.627								
14	.556								
CBQ									
2	.736			.182				.78	.22
4	.785			.340				.53	.47
8	.764			.500				.33	.67
10	.737			.412				.41	.59
12	.743			.490				.33	.67
OLBI									
2	.615						.212	.89	.11
4	.637						.075	.99	.01
8	.618						.489	.61	.39
12	.786						.047	1.00	.00
Depersonalization/devaluation/disengagement									
MBI									
5		.608							
10		.664							
11		.656							
15		.454							
22		.542							
CBQ									
3		.485			.327			.69	.31
6		.226			.741			.08	.92
9		.312			.690			.17	.83
11		.144			.616			.05	.95
OLBI									
3		.583					.404	.68	.32
6		.351					.524	.31	.69
9		.613					.615	.50	.50
11		.567					.254	.83	.17
Lack of personal accomplishment/reduced sense of accomplishment									
MBI									
7			.392						
9			.694						

(Continued)

TABLE 3
(Continued)

	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>M11</i>	<i>M12</i>	<i>M13</i>	<i>M2</i>	<i>Consistency Coefficient</i>	<i>Method Specificity</i>
17			.816						
18			.642						
19			.756						
CBQ									
5			.389			.536		.35	.65
7			.239			.618		.13	.87
13			.405			.808		.20	.80
14			.559			.360		.71	.29

Note: Underlined loadings were not significant at $p < .05$.

M11 = CBQ-E; M12 = CBQ-Dev; M13 = CBQ-RSA; M2 = OLBI.

CBQ items. The convergent validity is low for CBQ items for devaluation and reduced sense of accomplishment, while the convergent validity for OLBI items for disengagement is fairly high.

For CBQ exhaustion items, two have a consistency coefficient over 50% (r s from .53 to .78) and three less than 50% (r s from .33 to .41). This indicates that there is some consistency with the comparison method's definition of exhaustion. However, the CBQ also contains items with more method specificity. These items are, "I feel wiped out from coaching" (Item 8), "I feel physically worn out from coaching" (Item 10), and "I am exhausted by the mental and physical demands of coaching" (Item 12). This is somewhat expected, since these three cover both mental and physical exhaustion, whereas the two other items only cover tiredness, which is closer to the MBI definition of exhaustion. The OLBI items that cover exhaustion show higher consistency coefficients than method specificity (r s from .61 to 1), indicating that the OLBI items are closer to the MBI items covering exhaustion than the CBQ items are.

For the depersonalization trait, only one CBQ item has a consistency coefficient of more than 50% (.68), and the other three items—"I don't care as much about my coaching performance as I used to," "I'm not into coaching like I used to be," and "I feel less concerned about being successful in coaching than I used to"—have consistency coefficients less than 50% (r s from .08 to .17). The OLBI items have higher values for their consistency coefficients than the CBQ items on the depersonalization trait, whereby one item has a consistency coefficient less than 50% (.31) and three have more than 50% (r s from .50 to .83). This indicates that the OLBI items cover a trait that is more consistent with the MBI than the CBQ. However, the large correlations between exhaustion and depersonalization indicate that these traits do not discriminate well between each other.

One of four CBQ items covering the lack of personal accomplishment trait, "I feel successful at coaching," has a consistency coefficient higher than 50% (.71). The other three, "I am not achieving much in coaching," "I am not performing up to my ability in coaching," and "It seems that no matter what I do, I don't coach as well as I should" have consistency coefficients lower than 50% (r s from .13 to .35). This indicates that the CBQ items are not consistent with the MBI comparison standard method and instead have high method specificity.

To summarize, discriminant validity was low for exhaustion and devaluation ($r = .766$) but high between exhaustion and reduced sense of accomplishment ($r = .258$), and devaluation and

reduced sense of accomplishment ($r = .462$). Convergent validity was quite high for exhaustion items, albeit higher for the OLBI than the CBQ. The CBQ displayed low convergent validity for devaluation and reduced sense of accomplishment, while the OLBI displayed fairly high convergent validity for disengagement.

DISCUSSION

In this study we compared three coach burnout measures. Our aim was to evaluate the MBI-ES, OLBI, and CBQ by examining convergent and discriminant validity. We used the MBI-ES as a comparison method because it is the measure that the others theoretically build on. Although the OLBI has its theoretical base in the MBI-GS rather than the MBI-ES, the two MBI instruments are claimed to measure the same construct with the same factor structure (Maslach et al., 2001). The CBQ also stems from the MBI-HSS/MBI-ES, and is argued to have a similar theoretical ground as the MBI-GS even though it was developed independently (Raedeke et al., 2013). The key findings in our study are that burnout measured with either the MBI or OLBI seems to entail problems discriminating between emotional exhaustion and depersonalization dimensions. Furthermore, although the measures claim to have the same theoretical base, the CBQ seems to measure different aspects of these dimensions, especially when it comes to devaluation and reduced sense of accomplishment.

The CFA analyses identified different problems for the three measures, and we had to modify all of them to varying extents to attain acceptable model fit. For the MBI three items were removed, coinciding with two (Items 12 and 16) that have previously been recommended to be removed in European contexts (Maslach et al., 1996). In addition, Item 20, "I feel like I'm at the end of my rope," was removed because it cross-loaded on emotional exhaustion and lack of personal accomplishment. This item has been removed in CFA analysis in the past because of low correlation with emotional exhaustion; the authors claimed that this was because it tapped extreme feelings of emotional exhaustion (Oh & Lee, 2009). However, we argue that it is an ambiguous statement that can be interpreted differently, since its wording does not specifically cover anything necessarily having to do with exhaustion. We also removed two items from the CBQ: Item 1, "I'm accomplishing many worthwhile things in coaching," since it shared covariance with Item 14, "I feel successful at coaching;" Further, Item 15, "I have negative feelings toward coaching," from the devaluation dimension was removed because it cross-loaded on the mental and physical exhaustion dimensions. Eight items were removed from the OLBI, as they all were reversely coded and did not correlate with the proposed dimensions as much as the non-reversed items did. Similar results when using a four-factor model have been presented, showing that model fit increased when negatively and positively worded items were separated (Qiao & Schaufeli, 2011).

The MTMM analysis revealed some interesting points about the burnout construct. The correlation between traits shows that the exhaustion and depersonalization (devaluation) traits were highly correlated. This indicates that these two dimensions may not represent different dimensions of burnout, because they hardly discriminate from each other. This was further indicated when the two OLBI subscales had to be merged into one dimension on account of high correlations between exhaustion and devaluation. In a study using the MBI-HSS in a military setting, emotional exhaustion and depersonalization were collapsed into one dimension (Leiter, Clark,

& Durup, 1994). The authors argued that this had to do with the military context, a setting not appropriate for the MBI-HSS. Conclusively, then several researchers urge the user to, when using burnout measures based on the MBI, be aware of and observe correlations between exhaustion and depersonalization (devaluation/disengagement). Further research is warranted to explain why these dimensions are sometimes hard to discriminate between.

The MTMM analysis demonstrated that the exhaustion dimension in the MBI and CBQ showed convergent validity, thereby seemingly measuring similar types of the exhaustion construct. Since the OLBI had to be merged into one dimension, convergent validity was not specifically examined for the exhaustion dimension. In calculations of consistency coefficients and method specificity items from the CBQ, method specificity was higher than the consistency coefficients. This indicates that the CBQ covers a somewhat different definition of exhaustion than the MBI. For the OLBI items trait consistency was high, indicating that these items measure a concept very close to what is measured with the MBI.

For the dimension labeled depersonalization/disengagement/devaluation, the convergent validity with the MBI was lower compared to exhaustion for the CBQ, indicating that the devaluation items do not cover the same aspects as the depersonalization items do. Only one item had a higher consistency coefficient with the depersonalization trait than method specificity. This means that the CBQ devaluation factor covers a construct that is different from MBI depersonalization. Although it is argued that devaluation is equivalent to depersonalization translated into a sport context (Raedeke, 1997), the items covering these two dimensions share less than 50% of the variance. It can therefore be argued that devaluation fits better in a coaching context, because the burnout concept has to be adjusted to the context in which it is used (Maslach & Schaufeli, 1993). The OLBI, on the other hand, had three items with a higher consistency coefficient than method specificity, covering a construct similar to MBI depersonalization. This shows that the claim that aiming at depersonalization from recipients or devaluation of one's job with these items seems to cover similar dimensions of the burnout construct. However, the high correlations between the traits and the OLBI factor indicate a lack of discriminant validity for the exhaustion and depersonalization/disengagement/devaluation dimensions. The CBQ is slightly better for discriminating between exhaustion and devaluation.

The lack of personal accomplishment trait of the MBI and reduced sense of accomplishment from the CBQ cover different aspects related to perceived performance at work: the former focuses on interpersonal relations, whereas the CBQ focuses on coach competence. In a coaching context, sport-related performance can seldom be ignored since the coach may be out of a job if the athletes or team does not perform over certain standards (Lundkvist, Gustafsson, Hjälm, & Hassmén, 2012). When studying burnout in a coaching context, the lack of personal accomplishment dimension of the construct might miss a very important part of the burnout experience by only using items that tap coaches' experiences of working with athletes, and never ask about their performance as coaches. Therefore, it could be valuable for the burnout construct in a coaching context to use a broader definition that also includes the coaching performance rather than more directly focusing on relations with athletes. These are issues that a large number of coaches with contract-based employment cannot ignore. In interviews with coaches describing their burnout experiences, team performance issues were brought up as an important experienced cause of their burnout problems (Lundkvist et al., 2012). A broader definition of lack of personal accomplishment is warranted in a broader context as well, not least because Western societies increasingly promote organizational productivity in both the private and public sectors (Cooper,

1999); interpersonal relations may also become less important from a burnout perspective in recipient-oriented jobs like education, healthcare, and coaching.

Whereas, the MBI-HSS aims at the coach's relationship with the athletes, both the CBQ and OLBI are more focused on the relation to the job. For the CBQ, this is shown in the MTMM analysis by high loadings on the method factors. The differences between the CBQ and MBI-HSS lie in whether the coaching profession should be defined as a human service profession in which relationships with athletes are main dimensions of burnout, or a profession in which burnout dimensions tap one's relation to the work itself (Maslach et al., 1996). The MBI manual defines human service professions as jobs with direct personal contact with recipients (Maslach et al., 1996); the coaching profession fits this definition. However, it is plausible that a burnout definition with a recipient focus may miss certain important aspects of the burnout construct, especially in a coaching context, which most likely differs from nursing or other helping professions.

One solution to this might be to use a sport-specific part, whereby reduced sense of accomplishment covers performance linked to the results of athletes, an aspect of burnout that is otherwise forgotten—specifically because burnout is known to be context-specific (Eklund & Cresswell, 2007; Maslach, 2003). On the other hand, a coaching job does not entail the direct stress in one's working situation as does, for example, being a nurse or a doctor. Therefore, devaluation of the job may cover another central part of burnout for sport coaches, as smaller taxing pressures could be the perceived stressors leading to burnout (Hjälml, Kenttä, Hassmén, & Gustafsson, 2007). Opposite situations, with a coach working at a college or in junior settings and performing tasks more similar to a teaching job, could warrant the use of the MBI and a recipient-targeted definition of burnout.

Because context is important when measuring burnout, it is important to consider the transformation of the MBI from its original focus on healthcare professions to one on teaching environments. The similarity between these two types of occupations can be debated. Whereas, healthcare occupations have a helping focus, teaching (and coaching) jobs focus on developing individuals and groups (teams). Previous research on the MBI factor structure supports this distinction, with studies in healthcare settings showing better factor structure than teaching environments (Hallberg & Sverke, 2004; Schaufeli et al., 2001; Vanheule et al., 2007; Worley et al., 2008). Therefore, researchers studying coach burnout should be sensitive to the context in which the coaches are active.

Although it is claimed that these three measures are related and measure the same construct (Raedeke et al., 2013), our analysis shows that this claim is questionable. We acknowledge that it can be problematic to change the definition of burnout to fit different occupational settings, as this undermines the idea of burnout as an important health threat to individuals in different occupational settings. One way to cover a broad burnout concept is to include items from all three tested measures. For example, using items from the reduced sense of accomplishment dimension of the CBQ when using any of the other two measures could pinpoint parts the MBI-HSS misses that are important in coach settings.

Further, it has been argued that the three-dimensional definition grounded in the MBI is too simplistic in terms of causal relations of exhaustion to the other two dimensions. Both depersonalization/cynicism and lack of personal accomplishment can be seen as antecedents to burnout. And depending on personality and perceived causes of burnout, the two dimensions have different effects on exhaustion (Moore, 2000); this in combination with the fact that other similar antecedents, like perceived resources, self-esteem, perceived control over causes of exhaustion

and different possible causal effects of the three burnout dimensions (Lee & Ashforth, 1993, 1996; Moore, 2000), make the thought of burnout as a problematic syndrome. Longitudinal studies need to further investigate these relations.

LIMITATIONS

Although this study has several strengths, it also has limitations. First, our population is limited to Swedish coaches and their environment. Although the sample contains coaches from a wide range of individual and team sports in a wide range of ages, specific environmental differences can nevertheless limit generalizability. Second, most persons in the sample were male (86%); unfortunately, this reflects the distribution in the current coaching profession. Clearly, a sample containing more women would have been fruitful. Third, it cannot be ignored that the modifications we have performed in the CFAs have to do with *capitalization of chance*, meaning that this data-driven approach to model fit improvement is due to the sample we used rather than the measures. That is, the same analyses done in other samples could be modified differently. Therefore, these results should be viewed with caution (MacCallum, Roznowski, & Necowitz, 1992) and should be cross validated in future research. However, the modifications we needed to perform indicate that all tested measures indeed have problems. This is not unique to this study, however; when CFAs have been conducted on all three tested measures (especially MBI), cross-validation of items, removal of items and error term covariances have almost always been included since acceptable model fit would otherwise have been impossible to attain (Worley et al., 2008). This further indicates the problems these measures have. Therefore, replicating this study with other samples will provide further knowledge about the burnout construct. We also encourage further discussion on how burnout in sport contexts should be defined, operationalized, and measured.

CONCLUSIONS

Our study shows that although burnout has been studied since the 1980s, there are still problems involved with how to measure and define the construct. All three tested measures have psychometric shortcomings, and all were modified in the MTMM analysis. This is possibly a consequence of the lack of consensus on the definition and measure of burnout in occupational settings, which need further development (Shirom, 2005). Choosing between the three measures included in this study we argue that, based on this psychometric evaluation and theoretical concerns, the CBQ is the measure of choice. Although the CBQ measures dimensions that are different from those the MBI-HSS measures, especially when it comes to devaluation and reduced sense of accomplishment, the MBI-HSS overlooks central aspects of how burnout was defined by Maslach (1981). This is especially true for lack of personal accomplishment. Only tapping interpersonal relationships is problematic, because coaches may have other stressors more directly related to their performance at work. Another argument for using the CBQ is the suggestion that burnout is a syndrome; thus, combining the dimensions into one combined latent factor may be the way to go. If a researcher prefers a combined latent factor, this is not recommended when using the MBI (Maslach et al., 1996); combining the CBQ dimensions into a “global index,” however, is considered theoretically sound (Raedeke & Smith, 2004).

The burnout construct for coaches needs to be further researched, and we encourage researchers to look at the dimensionality of the three dimensions longitudinally to further study how they relate to each other. Including concepts like depression and anxiety (see Cresswell & Eklund, 2006; Raedeke et al., 2013) would also be interesting because these concepts are related and examining if these concepts discriminate from/are related to one another in a coach context would be fruitful. We also believe that testing other measures of burnout, like the Shirom-Melamed Burnout Measure and the Copenhagen Burnout Inventory, which both have other ways of defining burnout (Kristensen, Borritz, Villadsen, & Christensen, 2005; Shirom & Melamed, 2006), could be fruitful for taking coach burnout research forward toward a better understanding of the burnout construct.

REFERENCES

- Cooper, C. (1999). Can we live with the changing nature of work? *Journal of Managerial Psychology*, 14(7), 569–572. doi:10.1108/02683949910292169
- Cresswell, S. L., & Eklund, R. C. (2006). The convergent and discriminant validity of burnout measures in sport: A multi-trait/multi-method analysis. *Journal of Sports Sciences*, 24(2), 209–20. doi:10.1080/02640410500131431
- Dale, J., & Weinberg, R. (1990). Burnout in sport: A review and critique. *Journal of Applied Sport Psychology*, 2(1), 67–83.
- Demerouti, E., Bakker, A. B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments. *European Journal of Psychological Assessment*, 19(1), 12–23. doi:10.1027//1015-5759.19.1.12
- Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods*, 8(1), 38–60. doi:10.1037/1082-989X.8.1.38
- Eid, M., Nussbeck, F. W., Geiser, C., Cole, D. A., Gollwitzer, M., & Lischetzke, T. (2008). Structural equation modeling of multitrait-multimethod data: Different models for different types of methods. *Psychological Methods*, 13(3), 230–253. doi:10.1037/a0013219
- Eklund, R. C., & Cresswell, S. L. (2007). Athlete burnout. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 621–641). Hoboken, NJ: John Wiley & Sons, Ltd.
- Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the meaning of the latent variables in the CT-C(M-1) model: A comment on Maydeu-Olivares and Coffman (2006). *Psychological Methods*, 13(1), 49–57. doi:10.1037/1082-989X.13.1.49
- Goodger, K., Gorely, T., Lavallee, D., & Harwood, C. (2007). Burnout in sport: A Systematic Review. *The Sport Psychologist*, 22(2), 127–151.
- Grigorenko, E. L., Geiser, C., Slobodskaya, H. R., & Francis, D. J. (2010). Cross-informant symptoms from CBCL, TRF, and YSR: Trait and method variance in a normative sample of Russian youths. *Psychological Assessment*, 22(4), 893–911. doi:10.1037/a0020703
- Halbesleben, J. R. B., & Demerouti, E. (2005). The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. *Work & Stress*, 19(3), 208–220. doi:10.1080/02678370500340728
- Hallberg, U. E., & Sverke, M. (2004). Construct validity of the Maslach Burnout Inventory: Two Swedish health care samples. *European Journal of Psychological Assessment*, 20(4), 320–338. doi:10.1027/1015-5759.20.4.320
- Harris, B., & Ostrow, A. (2008). Coach and athlete burnout: The role of coaches' decision-making style. In S. R. Bakere (Ed.), *Hot topics in sports and athletics* (pp. 143–157). New York: Nova Science Publishers.
- Hjälms, S., Kenttä, G., Hassmén, P., & Gustafsson, H. (2007). Burnout among elite soccer coaches. *Journal of Sport Behavior*, 30(4), 415–428.
- Kelley, B. C., & Gill, D. L. (1993). An examination of personal-situational variables, stress appraisal, and burnout in collegiate teachers-coaches. *Research Quarterly for Exercise & Sport*, 64(1), 94–102.
- Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, 19(3), 192–207. doi:10.1080/02678370500297720

- Lee, R. T., & Ashforth, B. E. (1990). On the meaning of Maslach's three dimensions of burnout. *Journal of Applied Psychology*, 75(6), 743–747. doi:10.1037/0021-9010.75.6.743
- Lee, R. T., & Ashforth, B. E. (1993). A longitudinal study of burnout among supervisors and managers: Comparisons between the Leiter and Maslach (1988) and Golembiewski et al. (1986) models. *Organizational Behavior and Human Decision Processes*, 54, 369–398.
- Lee, R. T., & Ashforth, B. E. (1996). A meta-analytic examination of the correlates of the three dimensions of job burnout. *The Journal of Applied Psychology*, 81(2), 123–133. doi:10.1037/0021-9010.81.2.123
- Leiter, M. P., Clark, D., & Durup, J. (1994). Distinct models of burnout and commitment among men and women in the military. *The Journal of Applied Behavioral Science*, 30(1), 63–82. doi:10.1177/0021886394301004
- Lundkvist, E., Gustafsson, H., Hjälm, S., & Hassmén, P. (2012). An interpretative phenomenological analysis of burnout and recovery in elite soccer coaches. *Qualitative Research in Sport, Exercise, and Health*, 4(3), 400–419. doi:10.1080/2159676X.2012.693526
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111(3), 490–504. doi:10.1037/0033-2909.111.3.490
- Malinauskas, R., Malinauskiene, V., & Dumciene, A. (2010). Burnout and perceived stress among university coaches in Lithuania. *Journal of Occupational Health*, 52(5), 302–307. doi:10.1539/joh.O10006
- Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport and exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of Sport Psychology* (3rd ed., pp. 774–798). Hoboken, NJ: John Wiley & Sons, Ltd.
- Maslach, C. (2003). Job burnout: New directions in research and intervention. *Current Directions in Psychological Science*, 12(5), 189–192. doi:10.1111/1467-8721.01258
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, 2(2), 99–113. doi:10.1002/job.4030020205
- Maslach, C., Jackson, S., & Leiter, M. P. (1996). *Maslach Burnout Inventory: Manual* (3rd ed.). Paulo Alto: Consulting Psychologist Press.
- Maslach, C., & Schaufeli, W. B. (1993). Historical and conceptual development of burnout. In *Professional burnout: Recent developments in theory and research* (pp. 1–16). Washington, DC: Taylor & Francis.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Reviews Psychology*, 52, 397–422.
- Moore, J. (2000). Why is this happening? A causal attribution approach to work exhaustion consequences. *Academy of Management Review*, 25(2), 335–349. doi:10.5465/AMR.2000.3312920
- Nussbeck, F. W., Eid, M., & Lischetzke, T. (2006). Analysing multitrait-multimethod data with structural equation models for ordinal variables applying the WLSMV estimator: What sample size is needed for valid results? *The British Journal of Mathematical and Statistical Psychology*, 59(1), 195–213. doi:10.1348/000711005X67490
- Oh, S.-H., & Lee, M. (2009). Examining the psychometric properties of the Maslach Burnout Inventory with a sample of child protective service workers in Korea. *Children and Youth Services Review*, 31(2), 206–210. doi:10.1016/j.childyouth.2008.07.012
- Qiao, H., & Schaufeli, W. B. (2011). The convergent validity of four burnout measures in a chinese sample: A confirmatory factor-analytic approach. *Applied Psychology: An International Review*, 60(1), 87–111. doi:10.1111/j.1464-0597.2010.00428.x
- Raedeke, T. D. (1997). Is athlete burnout more than just stress? A sport commitment perspective. *Journal of Sport & Exercise Psychology*, 19(4), 396–417.
- Raedeke, T. D. (2004). Coach commitment and burnout: A one-year follow-up. *Journal of Applied Sport Psychology*, 16(4), 333–349. doi:10.1080/10413200490517995
- Raedeke, T. D., Arce, C., Seoane, G., & De Francisco, C. (2013). The construct validity of the Spanish version of the ABQ using a multi-trait/multi-method approach. *Anales de Psicología*, 29(3), 693–700. Doi:http://dx.doi.org/10.6018/analesps.29.3.175831
- Raedeke, T. D., & Smith, A. L. (2001). Development and preliminary validation of an athlete burnout measure. *Journal of Sport & Exercise Psychology*, 23(4), 281–306.
- Raedeke, T. D., & Smith, A. L. (2004). Coping resources and athlete burnout: An examination of stress mediated and moderation hypotheses. *Journal of Sport & Exercise Psychology*, 26(4), 525–541.
- Schaufeli, W. B., Bakker, A. B., Hoogduin, K., & Kladler, A. (2001). On the clinical validity of the Maslach Burnout Inventory and the Burnout Measure. *Psychology and Health*, 16, 565–582. doi:10.1080/08870440108405527

- Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). Maslach Burnout Inventory-General Survey. In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *Maslach Burnout Inventory, Manual* (3rd ed., pp. 19–26). Paulo Alto, CA: Consulting Psychologists Press.
- Shirom, A. (2005). Reflections on the study of burnout. *Work & Stress*, 19(3), 263–270. doi:10.1080/02678370500376649
- 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(2), 176–200. doi:10.1037/1072-5245.13.2.176
- Vanheule, S., Rosseel, Y., & Vlerick, P. (2007). The factorial validity and measurement invariance of the Maslach Burnout Inventory for human services. *Stress and Health*, 23(2), 87–91. doi:10.1002/smi.1124
- Vealey, R. S., Udry, E. M., Zimmerman, V., & Soliday, J. (1992). Intrapersonal and situational predictors of coaching burnout. *Journal of Sport and Exercise Psychology*, 14(1), 40–58.
- Worley, J. A., Vassar, M., Wheeler, D. L., & Barnes, L. L. B. (2008). Factor structure of scores from the Maslach Burnout Inventory: A review and meta-analysis of 45 exploratory and confirmatory factor-analytic studies. *Educational and Psychological Measurement*, 68(5), 797–823. doi:10.1177/0013164408315268

Copyright of Measurement in Physical Education & Exercise Science is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.