The Convergent Validity of Four Burnout Measures in a Chinese Sample: A Confirmatory Factor-Analytic Approach

Hu Qiao*

Yongkang Nursing School of Zhejiang Province, China

Wilmar B. Schaufeli

Utrecht University, The Netherlands

The purpose of this paper is to study the convergent validity of the four most widely used burnout measures in a sample of Chinese nurses ($N = 717$). First, Structural Equation Modeling was used to investigate the factor structure of scores produced by the Maslach Burnout Inventory-General Survey (MBI-GS), the Burnout Measure (BM), the Shirom-Melamed Burnout Measure (SMBM), and the Oldenburg Burnout Inventory (OLBI). Next, several competing models were tested to investigate the convergent validity of these four burnout instruments. The final results suggest that burnout is best conceived of as a multidimensional construct consisting of exhaustion and withdrawal, which are two related but conceptually distinct aspects. In addition, positively phrased items should be dropped from burnout measures for they constitute a separate factor that is considered to be an artifact.

INTRODUCTION

Although the concept of "burnout" was introduced over 30 years ago, its conceptualisation is still vehemently debated, for instance, as illustrated by a recent special issue of *Work & Stress* (Cox, Tisserand, & Taris, 2005). A conceptual controversy exists about whether burnout should be viewed as work-related exhaustion (Cordes & Dougherty, 1993; Shirom, 1989) or a multidimensional construct that goes beyond mere exhaustion (Maslach & Jackson, 1981; Schaufeli, Leiter, & Maslach, 2009). Another controversy pertains to the measurement of burnout; should the items that are included in the burnout scales all be positively or negatively phrased, or should the items be mixed (Demerouti, Bakker, Vardakou, & Kantas, 2003)? The major objective of the current article is to systematically compare—for the first time—the four most widely used self-report inventories to shed more

---

* Address for correspondence: Hu Qiao, Yongkang Nursing School of Zhejiang Province, 321300 Yongkang City, Zhejiang Province, China. Email: qiaohu2005@yahoo.com.cn

light on these two controversies. In other words, we seek to answer the question: to what extent do these instruments assess a similar burnout construct?

Burnout is usually defined as a psychological response to chronic work stress that is typically characterised by feelings of exhaustion. The concept was introduced in the mid-1970s, and initially a laundry-list of over 100 symptoms was associated with burnout, ranging from anxiety to lack of zeal (Schaufeli & Enzmann, 1998, pp. 20–30). After a pioneer phase of about five years, empirical research on burnout started to flourish following the introduction of short and easy to use self-report questionnaires, most notably the Maslach Burnout Inventory (MBI; Maslach & Jackson, 1981) and Burnout Measures (BM; Pines, Aronson, & Kafry, 1981). These two questionnaires represent two different schools of thought on burnout. The MBI assumes that burnout is a multidimensional construct that involves three distinct but related aspects: emotional exhaustion, depersonalisation (i.e. a callous and cynical attitude towards the people one is working with, such as patients, clients, or students), and reduced personal accomplishment. Originally, burnout as measured with the MBI was restricted to human services professions (e.g. teaching, health care, social work, law enforcement) because interpersonal and emotional demands were considered to be its root cause. In contrast, the BM assumes that burnout is a one-dimensional construct exclusively reflecting exhaustion. Although mental, emotional, and physical exhaustion are distinguished, an overall sum-score on the BM is used to assess "burnout" because it is easier to interpret and to communicate than a test profile (Pines, 1993). Studies on the factorial validity failed to distinguish more than one burnout (exhaustion) dimension in the BM (Corcoran, 1986; Justice, Gold, & Klein, 1981). The study of Schaufeli and Van Dierendonck (1993) revealed that over 50 per cent of the variance of the BM-total score is shared with MBI-emotional exhaustion ($r = .73$, $p < .001$). Thus, effectively, the BM reduces burnout to—different kinds of—exhaustion. In addition, burnout as measured with the BM is believed to occur not only at work, but also in non-occupational contexts such as love and marriage and political activism, because it is caused by long-term involvement in any emotionally demanding situation.

Burnout as a Multidimensional Construct

In time, the MBI became the most popular instrument with which to assess burnout. It was estimated that by the end of the 1990s, the MBI was used in over 90 per cent of the research articles and dissertations on burnout, whereas the BM was the second most widely used instrument with about 5 per cent (Schaufeli & Enzmann, 1998, p. 71). So gradually the concept of burnout became equivalent to the multidimensional way it was assessed by the MBI.

Meanwhile numerous studies using confirmatory factor analyses confirmed the validity of the hypothesised three-factor structure of the MBI (e.g. Golembiewski, Boudreau, Munzenrider, & Luo, 1996; Schaufeli & Van Dierendonck, 1993; Schaufeli, Daamen, & Van Mierlo, 1994; Li & Shi, 2003; Shirom & Melamed, 2006). However, three criticisms of the MBI remained to be voiced.

1. *General Instead of Specific Use.*   It was questioned whether burnout occurs exclusively among those who do "people work" of some kind. By introducing a so-called "General Survey" it was recognised that burnout might also occur outside the human services (Schaufeli, Leiter, Maslach, & Jackson, 1996). The MBI-GS can be applied in any occupational context and includes three generic subscales that parallel those of the original human services version: exhaustion, cynicism, and reduced professional efficacy. Meanwhile, the similarity of the original MBI and the MBI-GS in terms of their three-factor structure has been demonstrated in occupations within and outside the human services (Leiter & Harvie, 1998; Leiter, Harvie, & Frizzell, 1998; Leiter & Schaufeli, 1996).

2. *Two Instead of Three Dimensions.*   Four kinds of criticism have been raised, including lack of professional efficacy as a genuine aspect of burnout. First, from an empirical point of view, most studies show consistently that professional efficacy: (a) correlates relatively poorly with exhaustion and cynicism (Lee & Ashforth, 1996); (b) seems to develop in parallel with exhaustion and cynicism (Leiter, 1992; Taris, Le Blanc, Schaufeli, & Schreurs, 2005); and (c) is related to job resources in particular, whereas both other burnout dimensions are related to job demands as well (see Lee & Ashforth, 1996; Schaufeli & Enzmann, 1998). Second, from a theoretical point of view, it has been argued that exhaustion and cynicism—a form of mental distancing or withdrawal—constitute the core of burnout (Schaufeli & Taris, 2005). In contrast, professional efficacy is considered to be a separate personality factor (Cordes & Dougherty, 1993; Shirom, 2003). In essence, withdrawal or distancing is an adaptive mechanism to cope with job stress and the concomitant feelings of exhaustion. However, in the case of burnout, withdrawal has evolved into a habitual maladaptive pattern, which is characterised by a negative attitude towards one's job. Third, clinical experience with burned-out patients suggests that exhaustion and cynicism appear in tandem, whereas lack of professional efficacy is observed much less frequently (Brenninkmeijer & Van Yperen, 2003; Roelofs, Verbraak, Keijsers, de Bruin, & Schmidt, 2005). Fourth, it has been suggested from a psychometric point of view that the special role of lacking professional efficacy might be due to an artifact (Schaufeli & Salanova, 2007; Bresó, Salanova, & Schaufeli, 2007; Halbesle-

ben & Demerouti, 2005; Demerouti et al., 2003). Thus, all exhaustion and cynicism items are phrased negatively, whereas all professional efficacy items are phrased positively. The positively worded efficacy items are then reversed in order to be indicative of reduced efficacy. Because positive and negative items are likely to cluster in factor analysis (see Anastasi, 1988; Cacioppo, Gardner, & Berntson, 1997), professional efficacy emerges as a distinct factor.

So taken together, empirical, theoretical, clinical, and psychometric evidence exists for the particular role that professional efficacy plays as the "third dimension" of burnout. For that reason, we will not include professional efficacy as a dimension of burnout in our study.

3. Mixed Instead of Only Negative Items. It has been argued from a psychometric point of view that, in order to avoid answering bias, burnout inventories should include both negatively and positively phrased items (Demerouti et al., 2003). The exhaustion and cynicism scales of the MBI contain only negatively phrased items, which would undermine the validity of these scales.

In order to overcome these three criticisms, an alternative burnout instrument has been proposed: the OLBI (OLdenburg Burnout Inventory; Demerouti & Nachreiner, 1996; Halbesleben & Demerouti, 2005). The OLBI can be used in any occupational context and includes two dimensions: exhaustion and disengagement from work. Moreover, both scales contain positively as well as negatively worded items. Exhaustion is defined as a consequence of intensive physical, affective, and cognitive strain, which corresponds to other definitions of exhaustion (e.g. Pines et al., 1981; Lee & Ashforth, 1993; Shirom, 1989). Unlike exhaustion as operationalised in the MBI, the OLBI covers not only affective aspects of exhaustion, but also physical and cognitive aspects. Disengagement in the OLBI refers to distancing oneself from one's work and experiencing negative attitudes toward the work object, the work content, or one's work in general (i.e. withdrawal or mental distancing).

Exploratory factor analysis with the OLBI in two independent studies showed that—as expected—the exhaustion items of both scales loaded on one factor, whereas the cynicism and disengagement items loaded on another factor (Demerouti et al., 2003; Halbesleben & Demerouti, 2005). The correlations between both method-factors (i.e. the OLBI and MBI) exceeded .70 ($p < .001$) in the two studies, indicating that the two burnout measures share a great deal of similarity. Moreover, the fit to the data of the two-factor model (with exhaustion and disengagement) was superior to that of the wording model (with positively and negatively phrased items). This result led the test author to conclude that the OLBI can not only be used to assess *burnout*, but also *engagement* (Demerouti & Bakker, 2008).

Burnout as Exhaustion

Although the BM was conceived as a context-free, one-dimensional burnout instrument, it appeared to comprise three different but highly correlated factors, dubbed "demoralisation", "exhaustion", and "loss of motive". This three-factor structure was observed in a few studies, using German human services professionals (Enzmann & Kleiber, 1989), Dutch nurses (Schaufeli & Van Dierendonck, 1993), Dutch white-collar workers, human services professionals, and nurses (Enzmann, Schaufeli, Janssen, & Rozeman, 1998), and Chinese teachers (Mei & Li, 2006). Quite remarkably, these factors do not correspond with the three originally hypothesised aspects of exhaustion—emotional, physical, and mental (Pines et al., 1981; Pines & Aronson, 1988). Moreover, "demoralisation" and "exhaustion" are quite highly correlated, whereas "loss of motive"—that is exclusively defined by positively phrased items—correlates least with both other factors. Moreover, a study on the construct validity of the BM and the MBI showed that the three BM factors (i.e. demoralisation, exhaustion, and loss of motive) loaded on the same second-order factor as MBI-exhaustion and two indicators of mental and psychosomatic strain (Schaufeli & Van Dierendonck, 1993). Also, physical and emotional aspects of exhaustion as assumed by Pines et al. (1981) seem to collapse into one factor which reflects fatigue (Enzmann et al., 1998). Hence, the authors of this study concluded that the BM should be considered a general index of psychological strain. In sum, the BM has been criticised for its theoretical indistinctiveness (Schaufeli & Van Dierendonck, 1993) and its imperfect operationalisation (Enzmann et al., 1998). Furthermore, the findings of Shirom and Ezrachi (2003) suggest that the BM may inflate or conceal relations with antecedent stress variables.

Based on Hobfoll's (1989) Conservation of Resources (COR) theory, the SMBM (Shirom-Melamed Burnout Measure) was constructed as an alternative burnout instrument that assesses exhaustion—or the dwindling of energetic resources—regardless of its occupational context (Shirom & Melamed, 2006). More particularly, the SMBM includes three subscales—physical fatigue, emotional exhaustion, and cognitive weariness—that load on a second-order "burnout" factor (Shirom, Nirel, & Vinokur, 2006). According to COR theory, burnout reflects the depletion of energetic resources that results from cumulative exposure to chronic work and life stresses (see Melamed, Shirom, Toker, Berliner, & Shapira, 2006; Hobfoll & Shirom, 1993, 2000). So unlike the BM, the SMBM is based on a theoretical notion about the nature and development of burnout. The SMBM was found to be highly correlated ($r = .74$, $p < .001$) with the Emotional Exhaustion scale of the MBI (Grossi, Perski, Evengard, Blomkvist, & Orth-Gomer, 2003), and with the BM ($r = .74$, $p < .001$) (Soares & Jablonska, 2004). These very substantive correlations suggest that the SMBM, MBI-Emotional Exhaus-

tion, and the BM probably measure a common factor representing emotional and physical exhaustion (see Collins, 1999; Lee & Ashforth, 1996; Melamed et al., 2006). A cross-occupational validity study using the SMBM and the MBI revealed that a two-factor structure of the SMBM (physical fatigue and cognitive weariness) and the three-factor structure of the MBI-GS both were invariant across human service occupations and other professionals (Shirom & Melamed, 2006). In this study, the SMBM and MBI-GS were correlated .74 and .79 in the two groups, respectively, indicating that they share 55 per cent and 62 per cent of their variance. Hence, the test authors concluded that the SMBM has the potential to reveal more information about the core content of burnout—physical, emotional, and cognitive exhaustion (Cordes & Dougherty, 1993) than the MBI.

The Current Study

The current study includes the MBI because it is the most widely used questionnaire to assess burnout. In addition, another often used burnout instrument—the OLBI—is included which claims to solve two problems that are inherent to the MBI (i.e. three instead of two dimensions, and mixed instead of only negative items). Finally, two burnout instruments are included that conceptualise burnout exclusively in terms of exhaustion: the BM is the second most widely used questionnaire to assess burnout, and the SMBM that differentiates between three dimensions of exhaustion (i.e. physical, emotional, and cognitive). By including these four burnout questionnaires we may contribute to the conceptualisation of burnout by addressing validity issues related to the dimensionality of burnout and the role of exhaustion as a core phenomenon (see above). To date, no study on the validity has been conducted that includes all four burnout instruments *simultaneously*. Despite the almost universal acceptance of the MBI and the BM, and to a somewhat lesser degree the more recent SMBM and the OLBI, a careful evaluation of their convergent validity is still lacking. This is particularly important when translated burnout inventories are applied to non-English-speaking national or cultural settings, such as, for instance, China. Because of the rapid economic development of China, job stress and burnout are becoming major problems for the Chinese workforce, especially—but not exclusively—in human service professions like nursing (Cheng, Zhang, & Li, 2008). So there is an urgent need for reliable and valid instruments that may assess burnout in China. Some Chinese studies have confirmed the three-factor structure of the MBI-GS (“*emotional exhaustion*”, “*cynicism*”, and “*reduced professional efficacy*”) (Li & Shi, 2003; Feng, Luo, & Ye, 2004; Jiang, Xu, & Zhang, 2004) and three-factor structure of the BM (“*demoralisation*”, “*exhaustion*”, and “*loss of motive*”) (Mei & Li, 2006). However, validity studies using the OLBI and the SMBM are still lacking in China. The current study was carried out in order to investigate the factorial

validity of the Chinese versions of the four most popular burnout instruments, and to investigate how far these instruments assess the same underlying construct (convergent validity). More particularly we will answer the following two research questions:

1. Can the factorial structure of scores produced by the MBI (exhaustion and cynicism), BM (demoralisation, exhaustion, and loss of motive), SMBM (physical fatigue, emotional exhaustion, and cognitive weariness) and the OLBI (exhaustion and disengagement) be confirmed?
2. Do the MBI, BM, SMBM, and the OLBI assess one underlying burnout construct, or two distinct but related burnout dimensions (i.e. exhaustion and withdrawal)?

METHOD

Participants and Procedure

Given regional economic and cultural differences, we selected two provinces in China (Zhejiang province—an economically developed province, and Ningxia province—an economically underdeveloped province). A total of 1,000 questionnaires were distributed at random to nurses who worked in different hospital settings, of whom 717 (71.7%) returned the questionnaire. The study sample included 101 nurses from a general hospital, 97 nurses from a maternity hospital, 48 nurses in a traditional Chinese medicine hospital, and 103 nurses in several rural hospitals; all from the Zhejiang province in eastern China (total $n = 349$). In addition, 368 nurses from a general hospital in the Ningxia province in western China were included. The questionnaires were distributed by the head nurse and the survey was accompanied by a letter which explained the nature and the general aim of the study, and emphasised the anonymity of the participants. The biographical details of the sample are displayed in Table 1.

Measures

The Chinese version (Li & Shi, 2003) of the Maslach Burnout Inventory-General Survey (MBI-GS; Schaufeli et al., 1996) was used. Dropping the third component of the MBI—professional efficacy—five emotional exhaustion items as well as five cynicism items remained. Example items are "*I feel emotionally drained from my work*" (emotional exhaustion); "*I have become less enthusiastic about my work*" (cynicism).

The Burnout Measure (BM; Pines & Aronson, 1988) includes 21 items, among them four positively worded items and 17 negatively worded items which are designed to measure physical, emotional, and mental exhaustion. Example items are "*feeling energetic*" (physical exhaustion), "*being emotion-*

TABLE 1
Sample Characteristics ($N = 717$)

Total sample	N	%
Sex:		
Female	717	100
Age groups:		
20–30	380	53.00
31–40	235	32.78
41–50	84	11.72
Above 50	11	1.53
Marital status:		
Unmarried	263	36.68
With spouse	431	60.11
Divorced or Widower	14	1.95
Education:		
Junior Nursing School	178	24.83
Senior Nursing School	389	54.25
University	134	18.69
Tenure:		
1–5 years	296	41.28
5–10 years	92	12.83
10–20 years	223	31.10
20–30 years	84	11.72
Above 30 years	14	1.95
Type of contract:		
Temporary nurse	78	10.88
Contract nurse	152	21.20
Formal nurse	477	66.53

ally exhausted" (emotional exhaustion), "*feeling rejected*" (mental exhaustion). The BM was translated by the first author and the accuracy of the translation was checked by a group of English-speaking Chinese PhD students who are studying in Europe.

The Shirom-Melamed Burnout Measure (SMBM; Shirom, 1989) includes 14 items which cover physical fatigue, emotional exhaustion, and cognitive weariness. Part of the items (i.e. the items of emotional exhaustion) were reworded and have been used before as the Chinese Shirom-Melamed Vigor Measure (SMVM) by a Taiwanese PhD student. For instance, the item "*I feel I am* able *to be sensitive to the needs of co-workers and customers*" (emotional robustness) was modified into "*I feel I am* unable *to be sensitive to the needs of co-workers and patients*" (emotional exhaustion). The remaining items of the SMBM were translated by the first author. Example items are "*I feel physically drained*" (physical fatigue) and "*I have difficulty concentrating*" (cognitive weariness). As with the BM, the accuracy of the translation was checked by the group of Chinese PhD students.

The Oldenburg Burnout Inventory (OLBI; Demerouti et al., 2003) was translated into Chinese and back-translated into English. Differences between the original version and the back-translation were resolved between the test author (Evangelia Demerouti) and the first author. In addition, the accuracy of translation was checked by the Chinese PhD students. The OLBI includes two dimensions (i.e. eight exhaustion items and eight disengagement items) and each dimension consists of four positively worded items and four negatively worded items. Sample items are "*I always find new and interesting aspects in my work*" (disengagement—reversed), "*It happens more and more often that I talk about my work in a negative way*" (disengagement), "*I can tolerate the pressure of my work very well*" (exhaustion—reversed) and "*There are days when I feel tired before I arrive at work*" (exhaustion).

In contrast to the 7-point Likert scale—ranging from 0 ("never") to 6 ("always")—used in the three previous measures of burnout, the English version of the OLBI was originally scored on a 5-point Likert scale. In the current study the answer categories of the OLBI have been adapted to that of the other burnout instruments, because this allowed us to present the items of all four instruments in random order so that answering bias could be avoided. It should be noted that most items are negatively phrased (49 items), whereas only 12 items (four BM and eight OLBI) are positively phrased. In order to avoid inattentive answering or careless responding, the positively and negatively phrased burnout items are grouped together. In the introductory letter we emphasised that positively as well as negatively phrased items have been included in the questionnaire. Within each group of items the order was random.

We used confirmatory factor analyses to test the fit to the data of various models. All analyses were conducted by analysing the covariance matrix, using the maximum likelihood method in AMOS 5.0 (Arbuckle, 2003). The goodness-of-fit of the model was evaluated using three absolute fit indices (see Jöreskog & Sörbom, 1986): χ^2 goodness-of-fit statistic, Goodness-of-Fit Index (GFI), and Root Mean Square Error of Approximation (RMSEA). Browne and Cudeck (1993) proposed that RMSEA values lower than .08 suggest adequate fit and values in the range of .08–.10 suggest mediocre fit. Because χ^2 is sensitive to sample size, meaning that the probability of rejecting a hypothesised model increases with sample size, three relative goodness-of-fit indices were calculated (Bentler, 1990; Hu & Bentler, 1999): the Normed Fit Index (NFI), the Tucker-Lewis Index (TLI), and the Comparative Fit Index (CFI). For these three fit indices, as a rule of thumb, values greater than .90 were originally considered to indicate a good fit (Bentler, 1990). More recently, Hu and Bentler (1999) have suggested CFI values close to .95 or greater. Therefore, in the current study we consider .90 as critical lower bound values for NFI, TLI, and CFI.

TABLE 2
Model Fit of the MBI ($N = 717$)

Model	χ^2	df	p	GFI	RMSEA	RMSEA 90% CI	TLI	NFI	CFI
M_1 (1 factor)	321.49	35	.000	.90	.11	.10–.12	.91	.92	.93
M_2 (2 factors)	146.18	34	.000	.96	.07	.06–.08	.96	.97	.97
Null-model	4132.33	45	.000	.28	.36	.35–.37	–	–	–

Note: χ^2 = chi-square; df = degrees of freedom; p = significance level; GFI = Goodness-of-Fit Index; TLI = Tucker-Lewis Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; RMSEA 90% CI = 90% Confidence Interval of RMSEA.

In order to detect differences in model fit, a change in the GFI, TLI, NFI, and CFI above .01 was used as a general heuristic (Widaman, 1985).

RESULTS

Data were analysed in three steps: (1) testing of the factor structure of scores produced by the four burnout instruments by means of confirmatory factor analysis (CFA); (2) assessment of internal consistency of the subscales of the four burnout instruments (i.e. the MBI-GS, BM, SMBM, OLBI); (3) based on the best fitting model from the previous analyses (step 1), the convergent validity of the four burnout instruments was tested.

The MBI

Two factor-analytic models for the MBI were specified: (1) the one-factor model which assumes that all MBI items load on a general composite burnout factor (M_1); (2) the two-factor model in which exhaustion and cynicism are allowed to be correlated (M_2). As can be seen from Table 2, both models have a large χ^2 in relation to their degrees of freedom. However, this is not unexpected, for Byrne (1994) has pointed out that it is unusual to obtain non-significant χ^2 values for CFA-models when dealing with large sample sizes. Although M_1 fits reasonably well to the data, the fit of M_2 is superior ($\Delta\chi^2(1) = 175.31$, $p < .001$) with all fit indices satisfying their respective criteria. Thus, as expected, the hypothesised two-factor model fits the data well. Both factors are highly correlated: $r = .76$ ($p < .001$).

The BM

In order to assess the factorial validity of the BM, four alternative models were tested: (1) the one-factor model (M_3); (2) the three-factor model as suggested by Pines et al. (1981) (M_4); (3) the two-factor model in which

positively worded items and negatively worded items were grouped in two separate scales (M_5); (4) the three-factor model according to Enzmann and Kleiber (1989) (M_6). Two items were dropped from M_6 (i.e. "*feeling disillusioned and resentful about people*"—mental exhaustion; and "*being weary*"—physical exhaustion) due to semantic ambiguity (Schaufeli & Van Dierendonck, 1993; Enzmann et al., 1998). Table 3 clearly shows that M_3 and M_4 do *not* satisfy the criteria, but M_5 and M_6 have an acceptable fit, satisfying the lower bound criteria for RMSEA, TLI, NFI, and CFI. Since after deleting from M_5 both items that were not included in M_6, both models are nested so that the χ^2 difference test can be used to test their relative fit (Bentler & Bonnet, 1980). It appears that the fit of M_6 is superior to that of $M_{5\text{-respecified}}$ ($\Delta\chi^2$ (3) = 25.69, $p < .001$). The three latent BM factors are highly correlated, ranging from $r = -.56$ ($p < .001$) to $r = .89$ ($p < .001$).

In order to further explore the role of the positively phrased items of the BM, two additional models were tested: (1) the one-factor model without positive items ($M_{3\text{-respecified}}$); (2) the three-factor model of Pines et al. (1981) without positive items ($M_{4\text{-respecified}}$). Both models fit relatively well to the data with all indices except RMSEA meeting their lower bound criteria (RMSEA = 0.10, TLI = 0.90, NFI $\geq$ 0.90, CFI $\geq$ 0.91).

Thus, it seems that the three-factor model as suggested by Enzmann et al. (1989) (M_6) represents the best underlying structure of the BM. However, removing the positive items from the BM leads to a relatively good fit of the one-factor model ($M_{3\text{-respecified}}$), as well as the original three-factor model as proposed by the test authors ($M_{4\text{-respecified}}$).

The SMBM

Two models were tested: (1) the one-factor model (M_7); (2) the three-factor model in which emotional exhaustion, physical fatigue, and cognitive weariness are allowed to be correlated (M_8).

Inspection of the Modification Indices indicated that allowing two unique variances of single items within a particular subscale (physical fatigue) to correlate ("*I feel physically drained*"; "*I feel tired*") would improve the fit of the three-factor model. As can be seen from Table 4, compared with the one-factor model (M_7) and the original three-factor model (M_8), the revised model ($M_{8\text{-respecified}}$) meets the lower bound criteria for model fit ($\Delta\chi^2(4) = 774.19$, $\Delta\chi^2$ (1) = 177.10, $p < .001$). The three latent SMBM factors are highly correlated, ranging from $r = .61$ ($p < .001$) to $r = .73$ ($p < .001$).

The OLBI

Table 5 displays results of the CFA of four plausible models for the OLBI: (1) the one-factor model (M_9); (2) the two-factor model as proposed by the test

TABLE 3
Model Fit of the BM ($N = 717$)

Model	χ^2	df	p	GFI	RMSEA	RMSEA 90% CI	TLI	NFI	CFI
M$_3$ (1 factor)	2018.46	189	.000	.78	.12	.11–.12	.81	.82	.83
M$_{3\text{-respecified}}$ (without positive items)	925.27	90	.000	.85	.10	.09–.10	.90	.90	.91
M$_4$ (3 factors, Pines et al.)	1677.07	186	.000	.79	.11	.10–.11	.85	.85	.86
M$_{4\text{-respecified}}$ (without positive items)	883.91	116	.000	.85	.10	.09–.10	.90	.91	.92
M$_5$ (2 factors, positive–negative)	1167.56	188	.000	.85	.09	.08–.09	.90	.89	.91
M$_{5\text{-respecified}}$ (without two items)	868.58	151	.000	.87	.08	.08–.09	.91	.91	.92
M$_6$ (3 factors, Enzmann et al.)	842.89	149	.000	.87	.08	.08–.09	.92	.91	.93
Null-model	11095.02	210	.000	.15	.28	.27–.28	–	–	–

Note: χ^2 = chi-square; df = degrees of freedom; p = significance level; GFI = Goodness-of-Fit Index; TLI = Tucker-Lewis Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; RMSEA 90% CI = 90% Confidence Interval of RMSEA.

TABLE 4
Model Fit of the SMBM (N = 717)

Model	χ^2	df	p	GFI	RMSEA	RMSEA 90% CI	TLI	NFI	CFI
M7 (1 factor)	1247.25	77	.000	.75	.15	.14–.15	.77	.80	.81
M8 (3 factors)	650.16	74	.000	.88	.10	.10–.11	.88	.89	.90
M8-respecified	473.06	73	.000	.91	.09	.08–.10	.92	.92	.93
Null-model	6084.22	91	.000	.24	.30	.30–.31	–	–	–

Note: χ^2 = chi-square; df = degrees of freedom; p = significance level; GFI = Goodness-of-Fit Index; TLI = Tucker-Lewis Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; RMSEA 90% CI = 90% Confidence Interval of RMSEA.

TABLE 5
Model Fit of the OLBI (N = 717)

Model	χ^2	df	p	GFI	RMSEA	RMSEA 90% CI	TLI	NFI	CFI
M9 (1 factor)	1208.02	104	.000	.75	.11	.12–.13	.73	.75	.76
M10 (2 factors, hypothesised)	1158.26	103	.000	.78	.12	.11–.13	.74	.76	.77
M11 (2 factors, positive–negative)	574.28	103	.000	.90	.08	.07–.09	.88	.88	.90
M12 (4 factors)	433.45	98	.000	.92	.07	.06–.08	.91	.91	.93
Null-model	4755.67	120	.000	.34	.23	.23–.24	–	–	–

Note: χ^2 = chi-square; df = degrees of freedom; p = significance level; GFI = Goodness-of-Fit Index; TLI = Tucker-Lewis Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; RMSEA 90% CI = 90% Confidence Interval of RMSEA.

authors, in which emotional exhaustion and disengagement are allowed to be correlated (M_{10}); (3) the two-factor wording model (positive–negative wording, M_{11}); (4) the four-factor model in which four factors (two negatively worded scales—exhaustion and disengagement, and two positively worded scales—energy and engagement) are allowed to be correlated with each other (M_{12}).

Like the one-factor model (M_9), the hypothesised two-factor model (M_{10}) yielded a poor fit. Instead, it appeared that the four-factor model (M_{12}) with exhaustion, energy, disengagement, and engagement fit the data best compared to the other three models ($\Delta\chi^2(6) = 774.57$, $\Delta\chi^2(5) = 724.81$, $\Delta\chi^2(5) = 140.83$, $p < .001$ for M_9, M_{10}, and M_{12}). The four latent OLBI factors are moderately to highly correlated, ranging from $r = -.33$ ($p < .001$) to $r = .64$ ($p < .001$).

Internal Consistency Estimates

The means, standard deviations, internal consistency estimates (Cronbach's α), and subscale correlations of all four burnout measures can be found in Table 6. The internal consistencies of all scales, except the energy scale of the OLBI, were sufficient and exceeded the critical value of .70 (Nunnally & Bernstein, 1994). Correlations between the scales ranged from −.33 to −.59, and from .50 to .91. The mean correlations are −.48 and .75, respectively.

Convergent Validity

Convergent validity was tested progressively by using a series of CFAs. First, the one-factor model (M_{13}) was tested in which all 12 subscales were allowed to load on a single, second-order factor that can be conceived as "burnout". Second, M_{14} was tested, and included four second-order correlated latent variables that correspond with the best fitting models of each of the burnout instruments as identified in previous analyses. Third, in model M_{15} and model M_{16} all subscales were allowed to load on two correlated second-order factors that consist of exhaustion–withdrawal and positively–negatively phrased factors, respectively. The goodness-of-fit estimates of M_{13}–M_{16} are shown in Table 7.

As can be seen from Table 7, M_{16} was the only acceptable fitting model among the four alternative models, suggesting that the measurement model can be divided in terms of positively or negatively phrased items. Considering that positive items play a divergent role in burnout measures (see Introduction), we dropped the positively phrased items from M_{13}, M_{14}, and M_{15}. As a result, the fit of the three respecified models ($M_{13\text{-respecified 1}}$, $M_{14\text{-respecified}}$, and $M_{15\text{-respecified 1}}$) improved significantly compared to their original versions ($\Delta\chi^2(27)$ = 828.08, $\Delta\chi^2(27)$ = 812.46, $\Delta\chi^2(27)$ = 793.07, $p < .001$, respectively). A subsequent comparison among the three respecified models revealed that $M_{15\text{-respecified 1}}$ was the best fitting model as compared to $M_{13\text{-respecified 1}}$ and $M_{14\text{-respecified}}$ ($\Delta\chi^2(1)$ = 102.13, $p < .001$; $\Delta\chi^2(4)$ = 1434.82, $p < .001$, respectively). In addition, based on the Modification Indices, $M_{13\text{-respecified 1}}$ and $M_{15\text{-respecified 1}}$ could be further improved by allowing the errors of the subscales SMBM-CW and SMBM-EX to correlate. Re-estimation of $M_{13\text{-especified 2}}$ and $M_{15\text{-respecified 2}}$ resulted in a significant improvement of the model fit ($\Delta\chi^2(1)$ = 136.54 and 144.95, $p < .001$). All parameter estimates of $M_{15\text{-respecified 2}}$ are significant, so these findings provide evidence for the underlying two-factor structure (exhaustion–withdrawal) of burnout.

DISCUSSION

The present study is the first to examine the factorial validity and the convergent validity of scores produced by the four most widely employed

TABLE 6
Descriptives, Cronbach's α, and Intercorrelations of the MBI, OLBI, BM, SMBM (N = 717)

	Mean	SD	α	1	2	3	4	5	6	7	8	9	10	11
MBI														
1. Exhaustion (5)[a]	11.50	6.74	.88											
2. Cynicism (4)	10.08	6.39	.86	.76***										
OLBI														
3. Exhaustion (4)	10.84	5.11	.78	.84***	.67***									
4. Energy (4)	12.99	4.23	.67	-.44***	-.45***	-.36***								
5. Disengagement (4)	9.43	5.23	.84	.76***	.85***	.71***	-.45***							
6. Engagement (4)	10.68	4.46	.76	-.44***	-.58***	-.33***	.64***	-.56***						
BM														
7. Exhaustion (6)	12.06	7.65	.88	.85***	.80***	.81***	-.45***	.82***	-.47***					
8. Demoralisation (9)	18.89	11.92	.94	.85***	.86***	.78***	-.47***	.86***	-.53***	.89***				
9. Loss of motive (4)	12.05	5.06	.84	-.54***	-.57***	-.45***	.73***	-.57***	.73***	-.56***	-.59***			
SMBM														
10. Physical fatigue (6)	13.17	7.85	.90	.87***	.78***	.83***	-.44***	.81***	-.47***	.91***	.89***	-.56***		
11. Cognitive weariness (5)	8.56	5.28	.85	.71***	.73***	.58***	-.41***	.69***	-.42***	.70***	.77***	-.44***	.69***	
12. Emotional exhaustion (3)	4.56	3.16	.75	.61***	.71***	.50***	-.41***	.66***	-.45***	.63***	.66***	-.45***	.61***	.73***

Note: *** $p < .001$. [a] The item number of subscales.

TABLE 7
Convergent Validity Models of Four Burnout Measures (N = 717)

Model	χ^2	df	p	GFI	RMSEA	RMSEA 90% CI	TLI	NFI	CFI
M$_{13}$ (1 factor)	1505.08	54	.000	.70	.19	.18–20	.82	.85	.85
M$_{13\text{- respecified 1}}$	677.00	27	.000	.79	.18	.17–20	.89	.92	.91
M$_{13\text{- respecified 2}}$	540.46	26	.000	.83	.17	.15–18	.91	.93	.94
M$_{14}$ (four instruments)	2822.15	49	.000,	.68	.28	.27–29	.61	.71	.71
M$_{14\text{-respecified}}$	2009.69	22	.000	.75	.36	.34–37	.60	.75	.75
M$_{15}$ (exhaustion—withdrawal)	1367.94	53	.000	.73	.19	.18–20	.83	.86	.86
M$_{15\text{- respecified 1}}$	574.87	26	.000	.82	.17	.16–18	.91	.93	.93
M$_{15\text{- respecified 2}}$	429.92	25	.000	.87	.15	.14–16	.93	.95	.95
M$_{16}$ (pos–neg items)	822.192	53	.000	.80	.14	.13–15	.90	.92	.92

Note: χ^2 = chi-square; df = degrees of freedom p = significance level; GFI = Goodness-of-Fit Index; TLI = Tucker-Lewis Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; RMSEA 90% CI = 90% Confidence Interval of RMSEA.

measures of burnout (the MBI, BM, SMBM, OLBI) simultaneously. In addition, it introduces three alternative burnout measures to the MBI in China.

The MBI

The hypothesised two-factor structure of the MBI (exhaustion–cynicism) was clearly confirmed in our sample of Chinese nurses. Both MBI scales are highly correlated ($r = .76$; see Table 6), which agrees with the notion of burnout as a *syndrome* (Maslach, 1993)—that is, the association or co-occurrence of particular symptoms. By definition, the correlation between both latent factors is even higher: $r = .67$; which means that, when measurement unreliability is taken into account, both subscales share about 88 per cent of their variance.

The SMBM

Various researchers (Pines et al., 1981; Shinn, 1982; Shirom, 1989; Demerouti et al., 2003) have suggested that the exhaustion component of burnout should not only include emotional exhaustion—as in the MBI—but other aspects of exhaustion as well, such as cognitive and physical exhaustion. This is done in the SMBM that assesses exhaustion or "burnout" more broadly by including emotional, physical, and cognitive exhaustion. As hypothesised by the test author (Shirom, 1989), our results show that these three aspects of exhaustion can be distinguished, albeit that the fit of the three-factor model improves when the errors of two items are allowed to correlate ("*I feel physically drained*"; "*I feel tired*"). Obviously, this common variance is caused by overlapping item content—both refer to drained physical energy. Moreover, it appeared that the fit of the two-factor overall burnout model with exhaustion and withdrawal (M_{15}) improved when the errors of emotional and cognitive exhaustion, as measured by the SMBM, were allowed to be correlated ($M_{15\text{-respecified 2}}$). This illustrates that both aspects of exhaustion share some common variance above and beyond the variance that is explained by the common latent exhaustion factor. So obviously cognitive and emotional exhaustion are difficult to distinguish.

The test authors assume that exhaustion, as assessed by the SMBM, does not overlap with cynicism or detachment because this is not deemed to be typical for burnout (Shirom & Melamed, 2006). This is true in the sense that, indeed, a two-factor solution with exhaustion and withdrawal (see Figure 1) fit better to the data than a one-factor model. But on the other hand, both factors are so highly correlated ($r = .94$; see Figure 1) that it is difficult to deny that *together* they constitute the burnout syndrome.

Moreover, it is interesting to note that the emotional exhaustion scale of the SMBM includes items such as "*I feel I am not capable of being sympathetic*
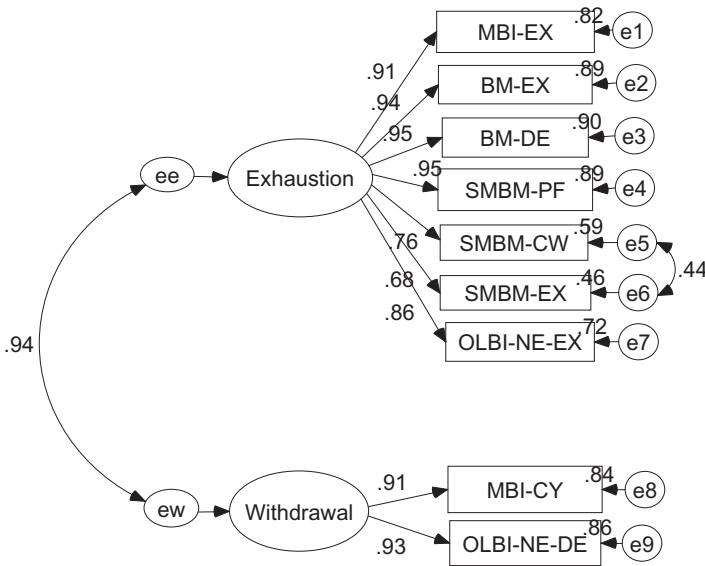
FIGURE 1.   The final two-factor model M$_{15\text{-respecified 2}}$.

Note: MBI-EX = Exhaustion of MBI, MBI-CY = Cynicism of MBI, OLBI-NE-EX = Exhaustion of the negatively worded scales of OLBI, OLBI-NE-DE = Disengagement of the negatively worded scales of OLBI, SMBM-PF = Physical Fatigue of SMBM, SMBM-CW = Cognitive Weariness of SMBM, SMBM-EX = Emotional Exhaustion of SMBM, BM-EX = Emotional Exhaustion of BM, BM-DE = Demoralisation of BM, all parameter estimates *** $p < .001$.

The BM

The factorial validity of the BM as suggested by Pines et al. (1981) is somewhat equivocal. Although the three-factor model (M$_4$, with the hypothesised factors "emotional exhaustion", "physical exhaustion", and "mental exhaustion") fit significantly better to the data than the one-factor model (M$_3$), which concurs with the results of Gold, Bachelor, and Michael (1989), Byrne (1991), and Schaufeli and Van Dierendonck (1993), the fit of both models (M$_3$ and M$_4$) is rather poor. It is noteworthy that the fit of M$_3$ and M$_4$ improved markedly

when the positively framed items were deleted. So, obviously, they measure something different. The three-factor structure ("demoralisation", "exhaustion", and "loss of motive"; M_6) that has previously been found in Germany (Enzmann & Kleiber, 1989), the Netherlands (Schaufeli & Van Dierendonck, 1993; Enzmann et al., 1998), and China (Mei & Li, 2006) was clearly replicated in our sample of Chinese nurses. Unlike M_3 and M_4, in M_6 the four positive items constitute one factor (dubbed—after reversion—"loss of motive") and consequently the fit of M_6 is superior to both other models. So it seems that the positively worded items play a divergent role in the BM, which seems the root cause of the imperfect operationalisation of the one-factor model and original three-factor model as suggested by Pines et al. (1981).

The OLBI

The OLBI was deliberately based on positively and negatively worded items to measure exhaustion and disengagement. The positively worded items are reversed in order to achieve an indicator of exhaustion or disengagement. However, including positively and negatively phrased items may lead to interpretational problems, because positive and negative affective states have been shown to have different antecedents (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Moreover, research on the structure of affect (Lloret & González-Romá, 2003) has demonstrated that a high score on positive affect is not equivalent to a low score on negative affect, and vice versa. In a similar vein, our study showed that the four positively worded BM items do not load on the same factor as the negatively worded items.

Demerouti et al. (2003) suggested that individuals not only respond to the content of the items, they are also sensitive to how the content is presented (the positive or the negative framing of the items). Usually in studies with the OLBI (Demerouti et al., 2003; Halbesleben & Demerouti, 2005; Demerouti & Bakker, 2008), positive items are alternated with negative items in order to force respondents to reflect carefully on the item content. Typically, these studies show that a two-factor model with exhaustion and disengagement (and mixed positive and negative items) fits better to the data than an alternative two-factor model with positively and negatively phrased (exhaustion or disengagement) items. In contrast, in our study the positive and negative items presented separately and we found that a four-factor model (with Exhaustion, Energy, Disengagement, Engagement; M_{12}) fit better to the data than the two-factor models with positively and negatively worded items (M_{11}), and with Exhaustion and Disengagement (M_{10}), respectively. This suggests that the underlying factor structure of the OLBI may depend on the framing of the items (i.e. alternating positive and negative items versus clustered positive and negative items).

Convergent Validity

Of the initially tested four convergent validity models, the two-factor model with positively and negatively worded items (M_{16}) fit best. The fit of the one-factor model (M_{13}), the four instruments model (M_{14}), and the exhaustion–withdrawal model (M_{15}) was relatively poor. This is consistent with the notion that positively and negatively worded items might have different functions (see above). It should be noted that using the Bivariate Evaluations and Ambivalence Measures (BEAMs) for measuring positive and negative feelings activated by different attitude stimuli, the study of Cacioppo et al. (1997) revealed the two phrased items—antonym pairs (e.g. good–bad, happy–unhappy) which arranged in two forms separately could load highly the same factor—that is, positively worded items demonstrated positive feelings and vice versa. Note especially that instead of loading on burnout, positive items (i.e. professional efficacy) also load on the opposite, positive concept of work engagement (Schaufeli, Martínez, Marqués Pinto, Salanova, & Bakker, 2002). Reversed positive efficacy scores yield different results from genuine negatively phrased items that are used to measure the same construct (Schaufeli & Salanova, 2007; Bresó et al., 2007). The negatively worded items represented burnout, whereas the positively worded items represented engagement, thus representing two independent yet negatively correlated states of mind (Schaufeli & Bakker, 2010).

Thus, we put our focus on only the negatively phrased items while eliminating the positively phrased items. The results of CFAs showed that all respecified models have an acceptable fit, except the four-instruments model ($M_{14\text{-respecified}}$). Moreover, the exhaustion–withdrawal model ($M_{15\text{-respecified 1}}$) fits significantly better to the data than the one-factor model ($M_{13\text{-respecified 1}}$), thus confirming that burnout consists of two highly correlated but different dimensions (exhaustion and withdrawal). This agrees with the theoretical notion that exhaustion (or the inability to work) and withdrawal (or the unwillingness to work) constitute two inseparable parts of the burnout experience (Schaufeli & Taris, 2005). Exhaustion refers to the energy aspect of burnout, whereas withdrawal refers to its motivational nature. Empirical research, particularly with the MBI, has consistently found substantial correlations between emotional exhaustion and cynicism (depersonalisation), usually with rs > 0.50 (Maslach, Jackson, & Leiter, 1996; Lee & Ashforth, 1996; Maslach, Leiter, & Schaufeli, 2008).

In sum, our study is important in that it provides further insight into the construct of burnout and the role that those positive items play in burnout measures. The convergent validity results confirmed the assumption that both exhaustion and withdrawal are the core element of burnout—i.e. two related but conceptually distinct aspects. In addition, the separate factorial analysis of the four burnout instruments as well as the simultaneous convergent factorial

analysis suggested that positively phrased items should be dropped since they constitute a separate factor that is considered to be an artifact.

Study Limitations

The current study has some limitations that should be mentioned. The sample under study is homogeneous (i.e. female) and occupation- and nation-specific (i.e. Chinese nurses). Therefore, the findings should be interpreted with caution and further research is needed in order to generalise our findings to other professions and countries.

Practical Implications

Our convergent validity analyses showed that burnout is best represented by two underlying, strongly related factors: exhaustion and withdrawal. In addition, confirmatory factor analytic analyses of the separate individual burnout questionnaires, as well as simultaneous convergent validity analyses of all questionnaires suggests that positively phrased items should be dropped since they constitute a separate factor that is considered to be an artifact. This means that for assessing burnout, either the MBI (with exhaustion and cynicism) or the OLBI (with negative exhaustion and negative disengagement) can be used. For assessing exhaustion, the SMBM (with physical fatigue, cognitive weariness, and emotional exhaustion) or the BM (with exhaustion and demoralisation) can be used.

REFERENCES

Anastasi, A. (1988). *Psychological testing*. New York: Macmillan.

Arbuckle J.L. (2003). *AMOS 5.0* (Computer Software). Chicago, IL: SPSS.

Armon, G., Shirom, A., Shapira, I., & Melamed, S. (2008). On the nature of burnout–insomnia relationships: A prospective study of employed adults. *Journal of Psychosomatic Research*, *65*, 5–12.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C., & Vohs, K.D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bentler, P.M., & Bonnet, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structure. *Psychological Bulletin*, *88*, 588–606

Brenninkmeijer, V., & Van Yperen, N. (2003). How to conduct research on burnout: Advantages and disadvantages of a unidimensional approach to burnout. *Occupational and Environmental Medicine*, *60*(Suppl. 1), 16–21.

Bresó, E., Salanova, M., & Schaufeli, W.B. (2007). In search of the third dimension of burnout: Efficacy or inefficacy? *Applied Psychology: An International Review*, *56*, 460–478.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J. Scott Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Byrne, B.M. (1991). Burnout: Investigating the impact of background variables for elementary, intermediate, secondary, and university educators. *Teaching and Teacher Education*, 7, 197–209.

Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS/Windows*. Thousand Oaks, CA: Sage Publications.

Cacioppo, J.T., Gardner, W.L., & Berntson, G.G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1, 3–25.

Cheng, S.Z., Zhang, H., & Li, S.B. (2008) Work stress of nurses increasing rapidly by alternative work schedule and job fatigue. *Yangcheng Evening News* (in Chinese). 12 May.

Collins, V.A. (1999). A meta-analysis of burnout and occupational stress. Unpublished doctoral dissertation, University of North Texas, Texas, USA. University Microfilm Accession Number: AAT 9945794.

Corcoran, K.J. (1986). Measuring burnout: A reliability and convergent validity study. *Journal of Social Behavior and Personality*, 1, 107–112.

Cordes, C.L., & Dougherty, T.W. (1993). A review and an integration of research on job burnout. *Academy of Management Review*, 18, 621–656.

Cox, T., Tisserand, M.C., & Taris, T.W. (2005). Editorial: The conceptualization and measurement of burnout. *Work & Stress*, 19, 187–191.

Demerouti, E., & Bakker, A.B. (2008). The Oldenburg Burnout Inventory: A good alternative to measure burnout and engagement. In J. Halbesleben (Ed.), *Stress and burnout in health care* (pp. 65–78). Hauppauge, NY: Nova Science Publishers.

Demerouti, E., Bakker, A.B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments: A multitrait–multimethod analysis. *European Journal of Psychological Assessment*, 19, 12–23.

Demerouti, E., & Nachreiner, F. (1996). Reliabilität und Validität des Maslach Burnout Inventory (MBI): Eine kritische Betrachtung [Reliability and validity of the Maslach Burnout Inventory: A critical view]. *Zeitschrift für Arbeitswissenschaft*, 50, 32–38.

Enzmann, D., & Kleiber, D. (1989). *Helfer-Leiden: Stress und Burnout in psychosozialen Berufen* [Stress and burnout in the human services]. Heidelberg: Asanger.

Enzmann, D., Schaufeli, W.B., Janssen, P., & Rozeman, A. (1998). Dimensionality and validity of the Burnout Measure. *Journal of Occupational and Organizational Psychology*, 71, 331–351.

Feng, Y., Luo, H., & Ye, Z.H. (2004). Job burnout of nurses. *Chinese Mental Health Journal* (in Chinese), 7, 477–479.

Gold, Y., Bachelor, P., & Michael, W.B. (1989). The dimensionality of a modified form of the Maslach Burnout Inventory for university students in a teacher training program. *Educational and Psychological Measurement*, 49, 549–561.

Golembiewski, R.T., Boudreau, R.A., Munzenrider, R.F., & Luo, H. (1996). *Global burnout: A worldwide pandemic explored by the phase model*. London: JAI Press.

Grossi, G., Perski, A., Evengard, B., Blomkvist, V., & Orth-Gomer, K. (2003). Physiological correlates of burnout among women. *Journal of Psychosomatic Research*, *55*, 309–316.

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*, 208–220.

Hobfoll, S.E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*, 513–524.

Hobfoll, S.E., & Shirom, A. (1993). Stress and burnout in work organizations: Conservation of resources. In R.T. Golembiewski (Ed.), *Handbook of organization behavio*r (pp. 41–60). New York: Marcel Dekker.

Hobfoll, S.E., & Shirom, A. (2000). Conservation of resources theory: Applications to stress and management in the workplace. In R.T. Golembiewski (Ed.), *Handbook of organization behavior* (2nd edn., pp. 57–81). New York: Marcel Dekker.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Jiang, J., Xu, Y., & Zhang, X.C. (2004). Job burnout and the degree of related stress among bank employees. *Chinese Journal of Clinical Psychology* (in Chinese), *2*, 178–180.

Jöreskog, K.G., & Sörbom, D. (1986). *LISREL user guide version VI* (4th edn.). Mooresville, IL: Scientific Software International.

Justice, B., Gold, R.S., & Klein, J.T. (1981). Life events and burnout. *Journal of Psychology*, *108*, 219–226.

Lee, R.T., & Ashforth, B.E. (1993). A further examination of management burnout: Toward an integrated model. *Journal of Organizational Behavior*, *14*, 3–20.

Lee, R.T., & Ashforth, B.E. (1996). A meta-analytic examination of the correlates of the three dimensions of job burnout. *Journal of Applied Psychology*, *81*, 123–133.

Leiter, M.P. (1992). Burnout as a crisis in self-efficacy: Conceptual and practical implications. *Work & Stress*, *6*, 107–115.

Leiter, M., & Harvie, P. (1998). Conditions for staff acceptance of organizational change: Burnout as a mediating construct. *Anxiety, Stress, & Coping*, *11*, 1–25.

Leiter, M., Harvie, P., & Frizzell, C. (1998). The correspondence of patient satisfaction and nurse burnout. *Social Science and Medicine*, *47*, 1611–1617.

Leiter, M.P., & Schaufeli, W.B. (1996). Consistency of the burnout construct across occupations. *Anxiety, Stress and Coping*, *9*, 229–243.

Li, C.P., & Shi, K. (2003). The influence of distributive justice and procedural justice on job burnout. *Acta psychologica Sinica* (in Chinese), *35*, 677–684.

Lloret, S., & González-Romá, V. (2003). How do respondents construe ambiguous response formats of affect items? *Journal of Personality and Social Psychology*, *85*, 956–968.

Maslach, C. (1993). Burnout: A multidimensional perspective. In W.B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 19–23). Washington, DC: Taylor & Francis.

Maslach, C., & Jackson, S.E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, *2*, 99–115.

Maslach, C., Jackson, S.E., & Leiter, M.P. (1996). *Maslach Burnout Inventory: Manual* (3rd edn.). Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Leiter, M.P., & Schaufeli, W.B. (2008). Measuring burnout. In C.L. Cooper & S. Cartwright (Eds.), *The Oxford handbook of organizational well-being* (pp. 86–108). Oxford: Oxford University Press.

Mei, M.J., & Li, Y.X. (2006). Construction research on the burnout measure. *Psychological Science* (in Chinese), *29*, 409–411.

Melamed, S., Shirom, A., Toker, S., Berliner, S., & Shapira, I. (2006). Burnout and risk of cardiovascular disease: Evidence, possible causal paths, and promising research directions. *Psychological Bulletin*, *132*, 327–353.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd edn.). New York: McGraw-Hill.

Pines, A. (1993). Burnout. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress* (2nd edn., pp. 386–403). New York: The Free Press.

Pines, A., & Aronson, E. (1988). *Career burnout: Causes and cures* (2nd edn., pp. 80–96). New York: Free Press.

Pines, A., Aronson, E., & Kafry, D. (1981). *Burnout: From tedium to personal growth*. New York: Free Press.

Roelofs, J., Verbraak, M., Keijsers, G.P.J., de Bruin, M.B.N., & Schmidt, A.J.M. (2005). Psychometric properties of a Dutch version of the Maslach Burnout Inventory-General Survey (MBI-DV) in individuals with and without clinical burnout. *Stress and Health*, *21*, 17–25.

Schaufeli, W.B., & Bakker, A.B. (2010). The conceptualization and measurement of work engagement: A review. In A.B. Bakker & M.P. Leiter (Eds.), *Work engagement: A handbook of essential theory and research* (pp. 10–24). New York: Psychology Press.

Schaufeli, W.B., Daamen, J., & Van Mierlo, H. (1994). Burnout among Dutch teachers: An MBI-validity study. *Educational and Psychological Measurement*, *54*, 803–812.

Schaufeli, W.B., & Enzmann, D. (1998). *The burnout companion to study and practice: A critical analysis*. London: Taylor & Francis.

Schaufeli, W.B., Leiter, M.P., & Maslach, C. (2009). Burnout: 35 years of reseach and practice. *Career Development International*, *14*, 204–220.

Schaufeli, W.B., Leiter, M.P., Maslach, C., & Jackson, S.E. (1996). *MBI-General Survey*. Palo Alto, CA: Consulting Psychologists Press.

Schaufeli, W.B., Martínez, I.M., Marqués Pinto, A., Salanova, M., & Bakker, A.B. (2002). Burnout and engagement in university students: A cross-national study. *Journal of Cross-Cultural Psychology*, *33*, 464–481.

Schaufeli, W.B., & Salanova, M. (2007). Efficacy or inefficacy, that is the question: Burnout and work engagement, and their relationship with efficacy beliefs. *Anxiety, Stress & Coping*, *20*, 177–196.

Schaufeli, W.B., & Taris, T.W. (2005). The conceptualization and measurement of burnout: Common ground and worlds apart. *Work & Stress*, *19*, 356–262.

Schaufeli, W.B., & Van Dierendonck, D. (1993). The construct validity of two burnout measures. *Journal of Organizational Behavior*, *14*, 631–647.

Shinn, M. (1982). Methodological issues: Evaluating and using information. In W.S. Paine (Ed.), *Job stress and burnout: Research, theory, and intervention perspectives* (pp. 61–79). Beverly Hills, CA: Sage.

Shirom, A. (1989). Burnout in work organizations. In C.L. Cooper & I. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 26–48). New York: Wiley.

Shirom, A. (2003). Job-related burnout. In J.C. Quick & L.E. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 245–265). Washington, DC: American Psychological Association.

Shirom, A., & Ezrachi, Y. (2003). On the discriminant validity of burnout, depression, and anxiety: A re-examination of the burnout measure. *Anxiety, Stress and Coping*, *16*, 83–99.

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

Shirom, A., Nirel, N., & Vinokur, A. (2006). Overload, autonomy, and burnout as predictors of physicians' quality of care. *Journal of Occupational Health Psychology*, *11*, 328–342.

Soares, J.J.F., & Jablonska, B. (2004). Psychological experiences among primary care patients with and without musculoskeletal pain. *European Journal of Pain*, *8*, 79–89.

Taris, T.W., Le Blanc, P., Schaufeli, W.B., & Schreurs, P.J.G. (2005). Are there relationships between the dimensions of the Maslach Burnout Inventory? A review and two longitudinal tests. *Work & Stress*, *19*, 256–262.

Widaman, K.F. (1985). Hierarchically nested covariance structure models for multitrait—multimethod data. *Applied Psychological Measurement*, *9*, 1–26.