THIS ARTICLE HAS BEEN CORRECTED. SEE LAST PAGE

Journal of Occupational Health Psychology

1076-8998/19/$12.00

http://dx.doi.org/10.1037/ocp0000151

Inquiry Into the Correlation Between Burnout and Depression

Irvin Sam Schonfeld
The City College and the Graduate Center of the City University of New York

Jay Verkuilen
The Graduate Center of the City University of New York

Renzo Bianchi
University of Neuchâtel

The extent to which burnout refers to anything other than a depressive condition remains an object of controversy among occupational health specialists. In three studies conducted in two different countries and two different languages, we investigated the discriminant validity of burnout scales by evaluating the magnitude of the correlation between (latent) burnout and (latent) depression. In Study 1 (N = 911), burnout was assessed with the Maslach Burnout Inventory–General Survey's Exhaustion subscale and depression with the depression module of the Patient Health Questionnaire (PHQ). In Study 2 (N = 1,386), the Shirom–Melamed Burnout Measure was used to assess burnout and the PHQ, depression. In Study 3 (N = 734), burnout was assessed with the Maslach Burnout Inventory–Educators Survey and depression, with the PHQ and the Center for Epidemiological Studies Depression scale; additionally, anxiety was measured with the Generalized Anxiety Disorder scale. In each study, we examined the burnout–depression association based on confirmatory factor analysis (CFA), controlling for item-level content overlap. In the three studies, latent exhaustion, the core of burnout, and latent depression were highly correlated (correlations ranging from .83 to .88). In Studies 2 and 3, second-order CFAs indicated that depressive (and anxiety) symptoms and the exhaustion and depersonalization components of burnout are reflective of the same second-order distress/dysphoria factor. Our findings, with their replication across samples, languages, and measures, together with meta-analytic findings, cast serious doubt on the discriminant validity of the burnout construct. The implications of burnout's problematic discriminant validity are discussed.

Keywords: burnout, discriminant validity, confirmatory factor analysis, depression, meta-analysis

Burnout has been regarded as an affliction in which individuals are left exhausted by a long-term confrontation with unmanageable job stressors (Maslach, Schaufeli, & Leiter, 2001; Shirom & Melamed, 2006). Although exhaustion (at physical, cognitive, and emotional levels) is considered the manifestation *par excellence* of burnout, other putative signs of burnout have been documented. One putative sign is withdrawal from one's work, including a distancing from the individuals to whom one's job is connected (e.g., clients, students, patients, and coworkers). This type of withdrawal is known in the burnout literature as depersonalization and cynicism (Maslach et al., 2001; Schaufeli & Taris, 2005). Another putative sign of burnout is a sense of diminished personal accomplishment and professional inefficacy. Because depersonalization and cynicism reflect strategies to cope with exhaustion and diminished personal accomplishment and professional inefficacy refer to long-term by-products of exhaustion (Maslach et al., 2001; Taris, Le Blanc, Schaufeli, & Schreurs, 2005), researchers have advanced the view that the inclusion of these symptoms in the core definition of burnout is not indispensable (Kristensen, Borritz, Villadsen, & Christensen, 2005; Shirom & Melamed, 2006).

Burnout, which has been conceived of both dimensionally and categorically, is thought to constitute a public health problem (Maslach et al., 2001; Shanafelt, Dyrbye, & West, 2017). The interest surrounding burnout, however, has coexisted with marked difficulties in characterizing the syndrome. One of these difficulties concerns the overlap of burnout with depression. Depression, which has also been treated both dimensionally and categorically, has been conceived of as the result of a deficit of positive, rewarding experiences and an excess of negative, punitive experiences (Rolls, 2016); it is considered a basic, universal response to unresolvable stress in human beings (Bianchi, Schonfeld, & Lau-

Irvin Sam Schonfeld, Department of Psychology, The City College of the City University of New York, and Department of Educational Psychology, The Graduate Center of the City University of New York; Jay Verkuilen, Department of Educational Psychology, The Graduate Center of the City University of New York; Renzo Bianchi, Institute of Work and Organizational Psychology, University of Neuchâtel.

Support for this research comes from Professional Staff Congress–City University of New York grant 69065-00 47 to Irvin Sam Schonfeld. As explained in the text, some of the data in this article were analyzed differently in two earlier articles: Bianchi and Schonfeld (2018) and Schonfeld and Bianchi (2016).

Correspondence concerning this article should be addressed to Irvin Sam Schonfeld, Department of Psychology, The City College of the City University of New York, 160 Convent Avenue, New York, NY 10031. E-mail: ischonfeld@ccny.cuny.edu

rent, 2018; Pryce et al., 2011; Willner, Scheel-Krüger, & Belzung, 2013). A good deal of high-quality research has accumulated that causally links adverse working conditions to depressive symptoms and disorders (Schonfeld & Chang, 2017).

There are at least two ways to look at the link between burnout and depression. One is that any burnout–depression redundancy is essentially technical, produced by the fact that depression measures often include symptoms of burnout, thus "artificially" inflating correlations between measures of the two constructs (Maslach & Leiter, 2016). The other is that the overlap may be more profound and reflect the fact that what we mean by burnout is actually a depressive condition (Schonfeld, Bianchi, & Luehring-Jones, 2017).

The Burnout–Depression Connection

The Maslach Burnout Inventory (MBI) is the most widely used burnout measure in the world (Maslach & Leiter, 2016; Schaufeli & Enzmann, 1998). In December 2018, the MBI showed up in 4,375 results in a PsycINFO search. The Copenhagen Burnout Inventory (154 hits), the Oldenburg Burnout Inventory (207 hits), and the Shirom–Melamed Burnout Measure (or Questionnaire; SMBM; 178 hits), for instance, turned up much less often.

Table 1 summarizes research conducted over the past 10 years in which correlations of the scores on the subscales of the MBI and the scores on depression and anxiety scales were estimated. The table is centered on the MBI because it is by far the most commonly used burnout measure. We included anxiety scales in Table 1 because scores on depression and anxiety scales tend to correlate highly (Lang & McTeague, 2009). The correlations were transformed using Fisher's z, meta-analyzed using a random effects model, and then back-transformed (see Table 2; R Core Team, 2013). We tested for heterogeneity, which was clear for all correlations except those with anxiety. Unfortunately, the number of studies involving anxiety was too small for a satisfactory assessment. With regard to the effect of attenuation due to unreliability, we considered the reported αs. Unfortunately, not all studies reported a reliability coefficient, although most αs were between .75 and .90. However, if we consider the effect of disattenuation, the correlations would be about 15%–20% larger. For instance, the correlation between scores on the Emotional Exhaustion subscale of the MBI and scores on measures of depression would be about .70.

The studies enumerated in Table 1 and summarized in Table 2 indicate that scores on the Emotional Exhaustion subscale of the MBI, the core dimension of burnout (Maslach et al., 2001), are, on average, more highly correlated with scores on depression and anxiety measures than with scores on the other two MBI subscales, Depersonalization (or its equivalent, Cynicism) and Personal Accomplishment (or its equivalent, Professional Efficacy). If burnout is (a) distinct from depression and (b) a syndrome consisting of emotional exhaustion, depersonalization, and reduced personal accomplishment, then one would expect that the correlations among scores on the three MBI subscales would be stronger than the Emotional Exhaustion–depression correlation.

Other research reports showed close parallels in the nomological networks of burnout and depression (for a review, see Bianchi et al., 2018). For instance, burnout is related to a history of depressive disorders and current intake of antidepressant medication (Ahola et al., 2007; Bianchi, Schonfeld, & Laurent, 2014; Rössler, Hengartner, Ajdacic-Gross, & Angst, 2015). The personality trait neuroticism—a vulnerability factor for clinical depression—is similarly associated with burnout and depressive symptoms (Hakulinen et al., 2015; Swider & Zimmerman, 2010). Studies have found parallels in the relation of burnout and depressive symptoms to physical activity (Toker & Biron, 2012) and job satisfaction (Faragher, Cass, & Cooper, 2005). Numerous studies have found job stress to be related to both burnout and depression (Melchior et al., 2007; Schaufeli, Bakker, & Van Rhenen, 2009). Oswin (1978), in a pioneering study linking adverse working conditions to depression, identified "professional depression" in nurses who had very trying jobs in understaffed wards serving severely handicapped children. Her research was conducted independently of research on burnout. We note that the findings resulting from nomological network analyses of burnout and depression have shown some heterogeneity (Brenninkmeyer, Van Yperen, & Buunk, 2001), a state of affairs that may be partly due to the "circularity trap" issue (i.e., to the problem of content overlap between independent and dependent variables; see Kasl, 1978). The correlations between measures of burnout and measures of job-related variables (e.g., work stress or work overload) are sometimes inflated because similar items are used in the two types of measures (see Schaufeli & Enzmann, 1998, pp. 73–74).

Disagreement Regarding the Magnitude of the Burnout–Depression Correlation

There has been disagreement regarding the magnitude of the correlation between burnout and depression. The disagreement turns on the question of whether content overlap at the item level inflates the correlation between burnout and depression scales (Maslach & Leiter, 2016). Indeed, burnout measures include items that mimic items on depression measures (see Bianchi, Schonfeld, & Laurent, 2019; see also Rasmussen, Verkuilen, Jayawickreme, Wu, & McCluskey, 2019, for a discussion of between-measure content overlap). The *Diagnostic and Statistical Manual of Mental Disorders*, *Fifth Edition* (*DSM–5*; American Psychiatric Association [APA], 2013) sheds light on the nature of the overlap. In depression "often insomnia or fatigue is the presenting complaint" (APA, 2013, p. 162), although skilled clinicians can recognize the underlying depression (Kahn, 2008). Fatigue is thus a symptom of both burnout and depression.

The *DSM–5* lists nine symptoms (e.g., depressed mood, anhedonia) that are used in diagnosing major depression. Those symptoms are represented in the depression module of the Patient Health Questionnaire (PHQ), a self-report measure of depressive symptoms (Kroenke & Spitzer, 2002). Two PHQ and *DSM–5* symptoms of depression, fatigue and sleep disturbance, map onto the exhaustion component of the MBI (Maslach, Jackson, & Leiter, 2016). Three PHQ and *DSM–5* depressive symptoms, fatigue, sleep disturbance, and a diminished ability to concentrate, are encapsulated in items of another measure of burnout, the SMBM (Shirom & Melamed, 2006). Although burnout is thought to be work related, both Kahn (2008) and Liu and Van Liew (2003) observed that depression is common in the workplace. Research on burnout has largely developed independently of the research conducted in psychiatry, behavioral psychology, and neurobiology on stress-induced conditions such as depression. Consequently, the

Table 1
The Past 10 Years of Research Bearing on Correlations Involving the Maslach Burnout Inventory Subscales and Depression and Related Scales

Study	Sample	Burnout–depression, burnout–anxiety correlations
Bakir, Ozer, Ozcan, Cetin, and Fedai (2010)	377 Turkish military nurses % women = ?; M_{age} = 30	EE, Beck-II, r = .39; EE, DP, r = .60; EE, rPA, r = .43; DP, rPA, r = .48 DP, Beck-II, r = .34; rPA, Beck-II, r = .35
Bakker (2009)	209 Greek teachers; 72% women; M_{age} = 43	Teachers: EE, CES-D, r = .40; DP, CES-D, r = .47; EE, DP, r = .40
	209 spouses; 58% public sector; the rest in private sector or self-employed; M_{age} = 45	Spouses: EE, CES-D, r = .38; Cyn, CES-D, r = .42; EE, Cyn, r = .52
Bianchi and Laurent (2015)	54 French human service workers 31% men; M_{age} = 39	EE, Beck-II, r = .79; EE, HADS-A, r = .67 EE, DP, r = .49; EE, rPA, r = .39 DP, Beck-II, r = .42; DP, rPA, r = .29; DP, HADS-A, r = .38 rPA, Beck-II, r = .38; rPA, HADS-A, r = .32
Bianchi, Boffy, Hingray, Truchot, and Laurent (2013)	1,658 French school teachers 33% men; M_{age} = 41	EE, Beck-II, r = .74; EE, DP, r = .51; EE, rPA, r = .50; DP, rPA, r = .47 DP, Beck-II, r = .44; rPA, Beck-II, r = .45
Bianchi, Schonfeld, and Laurent (2014)	5,575 French school teachers 78% women; M_{age} = 41	EE, PHQ, r = .72; DP, PHQ, r = .40; rPA, PHQ, r = .34 EE, DP, r = .53; EE, rPA, r = .35; DP, rPA, r = .40
De la Fuente-Solana et al. (2017)	101 Spanish oncology nurses 69% women; M_{age} = ?	EE, CECAD-Depr, r = .52 EE, CECAD-Anx, r = .54 DP, CECAD-Depr, r = .35; DP, CECAD-Anx, r = .36 PA, CECAD-Depr, r = −.25; PA, CECAD-Anx, r = −.30 No correlations among EE, DP, and PA subscales presented.
Halpern, Maunder, Schwartz, and Gurevich (2012)	190 Canadian paramedics 62% men; M_{age} = 37.5	EE, CES-D, r = .63; DP, CES-D, r = .28 PA, CES-D, r = −.15 EE, DP, r = .57; EE, PA, r = −.17; DP, PA r = −.09 We thank R.G. Maunder (personal communication, December 2018) for sharing the correlation coefficients from this study.
Jansson-Fröjmark and Lindblom (2010)	General population sample of 1,492 Swedes 54% women; M_{age} = 43	EE, HADS-D, r = .65; Cyn, HADS-D, r = .52; Effic, HADS-D, r = −.35 EE, Cyn, r = .58; EE, Effic, r = −.27; Cyn, Effic, r = −.40 EE, HADS-A, r = .66; Cyn, HADS-A, r = .44; Effic, HADS-A, r = −.25 We thank M. Jansson-Fröjmark (personal communication February 2019) for sharing the full Pearson correlation matrix with us; the correlations in the article involve dichotomized versions of the MBI subscales.
Kroska, Calarge, O'Hara, Deumic, and Dindo (2017)	245 U.S. medical students 51% women; 58% between ages 25–27	EE, IDAS-Depr, r = .68; IDAS-Depr, DP, r = .41; IDAS-Depr, PA, r = −.31 EE, DP, r = .62; EE, PA, r = −.26; DP, PA, r = −.13
Mosing, Butkovic, and Ullén (2018)	N = 6,326 Swedish single twins but not their cotwins; they worked in a variety of occupations; 58% women; M_{age} = 41	EE, Hopkins Symptom Checklist–Depression scale, r = .62 DP and PA not assessed. We thank M. Mosing (personal communication, February 2019) for supplying us with the number of individuals to whom the correlation applies and numbers of men and women in the sample.
Schmidt and Diestel (2014)	195 German nurses 85% women; M_{age} = 37	EE, Beck, r = .71; DP, Beck, r = .42 EE, DP, r = .56; PA not assessed
Tourigny, Baba, and Wang (2010)	239 nurses in Japan; M_{age} = 35 550 nurses in China; M_{age} = 32 99% women in Japan 100% women in China	Japan: EE, CES-D, r = .57; EE, DP, r = .41; EE, rPA, r = −.17 DP, CES-D, r = .40; DP, rPA, r = .02; rPA, CES-D, r = .20 China: EE, CES-D, r = .43; EE, DP, r = .26; EE, rPA, r = .08 DP, CES-D, r = .39; DP, rPA, r = .35; rPA, CES-D, r = .30
Trockel et al. (2018)	250 U.S. physicians 49% women; 51% between 30–39	EE, PROMIS Depression scale, r = .56; EE, PROMIS Anxiety scale, r = .59 DP, PROMIS Depression, r = .31; DP, PROMIS Anxiety, r = .37 PA, PROMIS Depression, r = −.30; PA, PROMIS Anxiety, r = −.33 No correlations among EE, DP, and PA subscales presented.

Note. Maslach Burnout Inventory (MBI) Measures: EE = Emotional Exhaustion subscale; DP = Depersonalization subscale; PA = Personal Accomplishment subscale; rPA = Personal Accomplishment subscale, reverse scored; Cyn = Cynicism subscale (akin to DP); Effic = Professional Efficacy subscale (akin to PA). Depression scales: Beck = Beck Depression Inventory; Beck-II = Beck Depression Inventory-II; CES-D = Center for Epidemiologic Studies Depression scale; HADS-A = Hospital Anxiety and Depression scales–Anxiety scale; PHQ = Patient Health Questionnaire, a nine-item depressive symptom scale keyed to the *DSM–5* symptoms of major depression; *DSM–5* = *Diagnostic and Statistical Manual of Mental Disorders*, *Fifth Edition*; CECAD = Educational-Clinical Questionnaire: Anxiety and Depression scales; Depr = Depression; Anx = Anxiety; HADS-D = Hospital Anxiety and Depression scales–Depression; IDAS-Depr = Inventory of Depression and Anxiety Symptoms–Depression scale; PROMIS = Patient-Reported Outcomes Measurement Information System–Short Form. We considered including Hakanen and Schaufeli's (2012) findings in the table but rejected the idea for the following reason. Although they used a shortened version of the Beck, they shortened the instrument further by dividing the shortened Beck into two briefer (and less reliable) depression subscales.

Table 2
Results of the Meta-Analysis of the Correlational Findings in Table 1

Measures	No. of samples	*r*	95% CI Low	95% CI High
Emotional Exhaustion–Depression	15	.60	.53	.66
Emotional Exhaustion–Depersonalization	12	.49	.42	.55
Depersonalization–Personal Accomplishment	8	.29	.16	.41
Emotional Exhaustion–Personal Accomplishment	8	.30	.19	.40
Depersonalization–Depression	13	.40	.36	.45
Personal Accomplishment–Depression	11	.33	.27	.38
Emotional Exhaustion–Anxiety	4	.59	.55	.62
Depersonalization–Anxiety	4	.41	.37	.46
Personal Accomplishment–Anxiety	4	.29	.25	.33

Note. CI = confidence interval. Depersonalization includes results involving the Depersonalization subscale and its variant, the Cynicism subscale. Personal Accomplishment includes results involving the Personal Accomplishment subscale and its variant, the Professional Efficacy subscale. Some studies used reversed measures of Personal Accomplishment; for the table, we corrected the signs such that all correlations were made to be positive.

question of the overlap of burnout symptoms with depressive symptoms has largely been neglected by the pioneers of burnout research (see Bianchi et al., 2018).

Item overlap obscures the question of the magnitude of the correlation between burnout and depression scales. It is the purpose of the present work to better estimate the burnout–depression correlation. Construct validity requires the establishment of the convergent and discriminant validity of scale scores. The scores on burnout measures tend to have convergent validity (Maslach et al., 2016). To demonstrate convergent validity, there should be a high correlation between the scores on different measures of the same construct. The Pearson correlation between the scores on the MBI–General Survey and the SMBM, for instance, has been found to range from .74 to .79, depending upon the sample examined; the Pearson correlation between the scores on the MBI–General Survey's Exhaustion subscale and SMBM scores has been found to range from .80 to .82 (Shirom & Melamed, 2006). The Pearson correlations between scores on the Exhaustion subscale of the MBI–General Survey and the Exhaustion subscale of the Oldenburg Burnout Inventory have been found to range between .60 and .72 (Demerouti, Bakker, Vardakou, & Kantas, 2003; Halbesleben & Demerouti, 2005).

Although the convergent validity of the burnout construct has been rather well established, its discriminant validity remains problematic. Leiter and Durup (1994), in a confirmatory factor analytic study, found that burnout and depression factors correlated highly, $r = .72$. These authors obtained such a high correlation despite treating burnout and depression items as continuous indicators when they would be more appropriately treated as ordinal variables—treating the indicators as continuous often induces a degree of attenuation to correlations (Choi, Peters, & Mueller, 2010; Mayer, 1971). Bakker et al. (2000) found that emotional exhaustion and depressed affect were also moderately correlated although the study team broke a 20-item depression scale into smaller, less reliable subscales. Emotional exhaustion scores were nevertheless more strongly related to scores on a truncated depression subscale ($r = .68$) than to Depersonalization ($r = .47$) or Personal Accomplishment ($r = -.42$). Toker and Biron (2012) found that scores on the SMBM and a depression scale were moderately correlated (*r*s ranging from .51 to .54); however, because the SMBM was not broken into its subscales, one could not compare how the scores on the subscales, which represent core (exhaustion) and noncore (depersonalization) facets of burnout, correlated with depression scores and with each other. Other research findings have shown that scores on burnout and depression scales correlate about as strongly as scores on different burnout scales correlate with each other (Bianchi, Verkuilen, Brisson, Schonfeld, & Laurent, 2016; Hätinen, Kinnunen, Pekkonen, & Aro, 2004; Lindblom, Linton, Fedeli, & Bryngelsson, 2006). Maslach and Leiter (2016) advanced the view that the high correlation between burnout and depression scales partly results from item-level content overlap. Because burnout is regarded as an important dependent variable in research in occupational health psychology, burnout scales should have a solid foundation, and discriminant validity of burnout scores vis-à-vis depression scores must be more firmly established. Otherwise, the construct validity of burnout is questionable.

In this article, we capitalize on three data sets that have assessed burnout and depression in educational staff, mostly schoolteachers. Schoolteachers are an apt occupational group to study because their working conditions vary from benign to highly stressful, with greater-than-average exposure to workplace violence and epidemiologic research showing higher-than-average risk for mental disorder (Schonfeld et al., 2017).

Assessing the Burnout–Depression Correlation in Three Data Sets

In Study 1 (Bianchi & Schonfeld, 2018), which was conducted in France, the French-language version of the Exhaustion subscale[1] of the MBI–General Survey was used to assess burnout symptoms. For several reasons, no other MBI–General Survey subscale was used in that study. First, exhaustion has been considered burnout's core (Maslach et al., 2001). Second, exhaustion is the only consensual dimension of burnout (Kristensen et al., 2005; Shirom & Melamed, 2006). Third, exhaustion constitutes the sole symptom of burnout that is not culture specific (Schaufeli, 2017). Fourth, exhaustion has been uniquely linked to a reduction in *objective* job performance (Taris, 2006). Finally, because exhaustion has been viewed as the entry point into the burnout syndrome (Maslach et al., 2001; Taris et al., 2005), with depersonalization/cynicism and diminished personal accomplishment/professional inefficacy constituting repercussions of exhaustion, a focal concern with the relationship between exhaustion and depression is warranted. The eight-item version of the Patient Health Questionnaire (PHQ-8; Kroenke et al., 2009) was used for assessing depressive symptoms in this first study.

In Study 2 (Schonfeld & Bianchi, 2016), which was conducted in the United States, burnout symptoms were assessed with the

[1] For clarification, the MBI–Educators Survey has an "Emotional Exhaustion" subscale that parallels the MBI–General Survey's "Exhaustion" subscale. The core of both versions of the MBI is exhaustion although the MBI–General Survey's Exhaustion subscale is shorter.

14-item version of the SMBM (Toker, Melamed, Berliner, Zeltser, & Shapira, 2012). The SMBM comprises three subscales, namely, Physical Fatigue, Cognitive Weariness, and Emotional Exhaustion. The items in the Physical Fatigue subscale of the SMBM reflect a more general kind of ill-being as much as they reflect actual physical fatigue (e.g., "I feel fed up"; "I feel burned out"; "I feel like my 'batteries' are 'dead'"). This subscale represents the exhaustion-oriented core of burnout that most overlaps with the Exhaustion subscale of the MBI–General Survey and the Emotional Exhaustion subscale of the MBI–Educators Survey. One note on the Emotional Exhaustion subscale of the SMBM: It is measured differently from the (Emotional) Exhaustion subscale of the MBI. Two of its three items are more reflective of depersonalized interpersonal relationships at work (e.g., "I feel that I am not capable of being sensitive to the needs of coworkers and customers" [rewritten for our teacher study as "I feel I am unable to be sensitive to the needs of students and coworkers"]). The nine-item version of the PHQ (PHQ-9; Kroenke & Spitzer, 2002) was used to assess depressive symptoms in this second study (Martin-Subero et al., 2017; Pettersson, Boström, Gustavsson, & Ekselius, 2015).

Study 3 also involved U.S. teachers. Burnout was assessed with the MBI–Educators Survey. We again assessed depression with the PHQ-9; however, we supplemented the PHQ-9 with the 10-item version of the Center for Epidemiological Studies Depression scale (CES-D-10; Cole, Rabin, Smith, & Kaufman, 2004), which would yield more stable estimates of a depression factor. Because there is mounting evidence in the abnormal psychology literature that depressive and anxiety symptoms are highly correlated and probably involve a common distress/dysphoria continuum (Caspi et al., 2014; Dohrenwend, Shrout, Egri, & Mendelsohn, 1980; Kotov et al., 2017; Watson, 2005), we also included the seven-item Generalized Anxiety Disorder scale in Study 3 (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006).

Because we have different measures of burnout symptoms and use samples that speak different languages, there is, to some extent, an opportunity for replication built into the research. We relied on confirmatory factor analysis (CFA) to estimate the magnitude of the correlation between the latent Burnout and Depression factors, and control for measurement error. We proceeded cautiously as per Maslach and Leiter (2016) by calculating PHQ scores without including symptom items that, arguably, are reflected in each of the burnout scales under examination. We omitted one item from the CES-D-10 because the item, possibly, assesses what MBI (Emotional) Exhaustion measures.

In Study 1, in the interest of evaluating the question of the discriminant validity of a burnout measure, we assessed the correlation between the burnout and depression constructs. Based on the exhaustion items of the MBI–General Survey, we created a latent burnout variable. Study 1 has two advantages over our previous work (Bianchi & Schonfeld, 2018). First, we conducted a CFA to better adjust for measurement error. Second, we controlled for item-level overlap.

In Study 2, in the interest of evaluating the question of the discriminant validity of a different burnout measure, our aim was to examine the correlations between latent Physical Fatigue, latent Cognitive Weariness, and latent Emotional Exhaustion, based on the relevant SMBM items, and latent Depression, based on items in the PHQ-9. Study 2 overcomes three limitations of our earlier work (Schonfeld & Bianchi, 2016). First, the current study includes a CFA to better control for measurement error. Second, the current study differentiates the subscales of the SMBM, Physical Fatigue, Cognitive Weariness, and Emotional Exhaustion. Third, the current study controls for item-level overlap.

In Study 3, we bring to bear a new sample in which we assessed burnout with the MBI–Educators Survey, two depression scales, and an anxiety scale. Study 3 provides a further examination of the discriminant validity of burnout with controls for measurement error and content overlap at the item level. In addition, in Studies 2 and 3, we developed models that were sufficiently identified to enable us to conduct higher-order CFAs and determine whether a second-order distress/dysphoria factor emerges on which first-order Exhaustion, Depression, and Anxiety (in Study 3) factors load.

Method

Participants

The Study 1 sample comprised 911 French educational staff members, a majority of whom were schoolteachers. Eighty-three percent were women. The mean age of the sample was 42.4 (SD = 10.2). The mean length of employment in the education field was 15.0 years (SD = 10.2). The study sample is described in the article by Bianchi and Schonfeld (2018). In Study 2, the sample comprised 1,386 U.S. schoolteachers. Seventy-seven percent were women. The mean age was 42.7 (SD = 11.4). The mean number of years teaching was 14.7 (SD = 9.6). The study sample is described in the article by Schonfeld and Bianchi (2016). The sample in Study 3 consisted of 734 U.S. teachers, of whom 78% were women. The mean age was 43.3 (SD = 11.7). The mean number of years teaching was 13.6 (SD = 8.9). The sample has not been previously studied.

Measures

The measures used in each of the three studies are described in detail in Table 3. Because the studies were conducted separately, and had different aims, the measures were treated slightly differently from study to study. Any changes made in the scales are documented in the table.

To summarize, in Study 1, burnout symptoms were assessed with the five-item French-language version (Papineau et al., 2005) of the Exhaustion subscale of the MBI–General Survey and depressive symptoms, with the French version of the PHQ-8 (Arthurs et al., 2012). In Study 2, burnout symptoms were assessed with the three subscales of 14-item version of the SMBM (Toker et al., 2012): Physical Fatigue, Cognitive Weariness, and Emotional Exhaustion. Depressive symptoms were assessed with the PHQ-9 (Kroenke & Spitzer, 2002). In Study 3, burnout symptoms were assessed with the three subscales of the MBI–Educators Survey: Emotional Exhaustion, Depersonalization, and Personal Accomplishment. Depressive symptoms were assessed with the PHQ-9 (Kroenke & Spitzer, 2002) *and* the CES-D-10 (Cole et al., 2004); anxiety symptoms were assessed with the GAD-7 (Spitzer et al., 2006).

Table 3
Instruments Used in Studies 1, 2, and 3

Measures	Explanation of the measures
	Study 1
MBI-GS's Exhaustion subscale	French version of the Maslach Burnout Inventory–General Survey's (MBI-GS) Exhaustion subscale Five items; e.g., "Je me sens épuisé(e) à la fin de ma journée de travail." [I feel exhausted at the end of my workday.] The MBI's response format, which ordinarily applies to the last year, was modified to match the PHQ's 0–3 format (*not at all–nearly every day over the last two weeks*). The mean score was calculated. $M = 1.25$; $SD = .84$; $\alpha = .89$ No other subscales of the MBI-GS were used in Study 1.
PHQ-8	French version of the eight-item Patient Health Questionnaire (PHQ), Depression module. The ninth and last item, a suicidal ideation item, was not used in the study. Exemplary item: "Être triste, déprimé(e) ou désespéré(e)." [To be sad, depressed or hopeless.] For the scale's format, see above. $M = 1.01$; $SD = .72$; $\alpha = .885$
PHQ-6	French version of six-item adaptation of the abovementioned PHQ-8 We omitted two fatigue-related items (sleep problems and tiredness) from the eight-item version because the two items arguably cover some of the same content as the MBI Emotional Exhaustion subscale. $M = .85$; $SD = .72$; $\alpha = .85$ *Note*: 1. The original time frame for MBI-GS items is 1 year and the time frame for PHQ items 2 weeks. To mitigate the influence of format differences on the Exhaustion–PHQ correlation, we adjusted the time frame of the MBI-GS's Exhaustion subscale to make it more compatible with that of the PHQ. 2. Instead of presenting total scores as is usually done with the full nine-item version of the PHQ, we calculated the item mean score to better understand the comparability of the scores on the eight- and six-item versions.
	Study 2
SMBM	Shirom–Melamed Burnout Measure (SMBM) comprises 14 items and three subscales. The response format for all SMBM items is as follows: *never or almost never* [0] . . . *sometimes* [3] . . . *always or almost always* [6]. The scale applies to the previous month. We calculated the mean score for each scale.
1. Physical Fatigue	Six items; e.g., "I have no energy for going to work in the morning." $M = 4.17$, $SD = 1.62$, $\alpha = .94$
2. Cognitive Weariness	Five items; e.g., "My thinking process is slow." $M = 3.31$, $SD = 1.51$, $\alpha = .95$
3. Emotional Exhaustion	Three items; e.g., "I feel I am unable to be sensitive to the needs of coworkers and students." $M = 2.92$, $SD = 1.42$, $\alpha = .83$ *Note*: 1. The content of the SMBM's Emotional Exhaustion subscale resembles that of the MBI's Depersonalization subscale more than the content of MBI's Emotional Exhaustion subscale. 2. The content of the SMBM's Physical Fatigue subscale more closely resembles that of the MBI's Emotional Exhaustion subscale.
PHQ-9	Full nine-item Patient Health Questionnaire depression module; e.g., "Feeling down, depressed, or hopeless." We calculated mean scores for the PHQ-9 and −6 (below). $M = .98$, $SD = .67$, $\alpha = .88$
PHQ-6	Six-item version of the abovementioned PHQ-9. Three items from the PHQ-9 were omitted; two were the fatigue-related items (sleep problems and tiredness) because of potential content overlap with the SMBM's Physical Exhaustion subscale; one additional deleted item was the "trouble concentrating" item because of potential content overlap with the SMBM's Cognitive Weariness subscale. $M = .80$, $SD = .66$, $\alpha = .82$ *Note*: 1. Because the PHQ applies to the last 2 weeks and the SMBM applies to the last month, time spans that are close, we did not manipulate the response formats of either scale. 2. Instead of presenting total scores as is usually done with the PHQ, we calculated the item mean score to better understand the comparability of the scoring of the nine- and six-item versions.
	Study 3
MBI-ES	Maslach Burnout Inventory–Educators Survey (MBI-ES). It comprises 22 items and three subscales. The response format for all items is a 7-point scale that covers the previous year (from 0 [*for never*] to 6 [*for every day*]). We calculated the mean score on each subscale
1. Emotional Exhaustion	Nine items; e.g., "I feel emotionally drained by work." $M = 3.56$, $SD = 1.42$, $\alpha = .93$

(*table continues*)

Table 3 (*continued*)

Measures	Explanation of the measures
2. Depersonalization	Five items; e.g., "I feel I treat some students as if they were impersonal objects." $M = 1.88$, $SD = 1.30$, $\alpha = .75$
3. Personal Accomplishment	Eight items; e.g., "I deal effectively with the problems of my students." $M = 4.29$, $SD = .92$, $\alpha = .80$
CES-D-10	10-item version of the Center for Epidemiological Depression scale; e.g., "I felt my life had been a failure." $M = 2.35$, $SD = 1.16$, $\alpha = .86$
CES-D-9	Nine-item version of the abovementioned scale. One item, "I feel that everything I do is an effort," was omitted because of potential content overlap with the Emotional Exhaustion subscale of the MBI. $M = 2.58$, $SD = .80$, $\alpha = .84$ *Note*: To mitigate the influence of format differences on the MBI–CES-D correlations, we modified the CES-D's response format, which ordinarily applies to the last week, to match the MBI's 1-year response format. We did not change the format of the PHQ, the other depression scale.
PHQ-9	Full nine-item Patient Health Questionnaire depression module We calculated mean scores for the PHQ-9 and −7 (below). $M = .98$, $SD = .67$, $\alpha = .88$
PHQ-7	Seven-item version of the abovementioned PHQ-9; two fatigue-related items (sleep problems and tiredness) were omitted from the PHQ-9 because of potential content overlap with the MBI's Emotional Exhaustion subscale. $M = .80$, $SD = .66$, $\alpha = .82$ *Note*: Instead of presenting total scores as is usually done with the PHQ, we calculated the item mean score to better understand the comparability of the scoring of the nine- and seven-item versions.
GAD-7	Seven-item Generalized Anxiety Disorder scale; e.g., "Feeling nervous, anxious or on edge." $M = 8.36$, $SD = 6.00$, $\alpha = .93$ As per the usual scoring, we used the total score. Like the PHQ-9, the time frame for the GAD-7 is the last two weeks.

Procedure

In Study 1, educational staff members in France were reached by e-mail contacts with school administrators. School administrators were invited to themselves complete an Internet survey and to forward our cover e-mail to staff members. The e-mail briefly described the nature of the study and contained a weblink to the online survey. In Study 2, school administrators in 18 U.S. states were contacted by e-mail, and asked to forward the e-mail to their teachers, with one exception as explained below. As in France, the e-mail briefly described the nature of Study 2, and asked the teachers to participate, except in a locality in which we were required, with the principal's permission, to contact teachers by placing in their in-school mailboxes flyers describing the nature of the study and asking them to participate. In Study 3, we used procedures in 22 states like those used in Study 2.

Our recruitment procedures did not permit us to calculate the teachers' response rate. Barriers to calculating the response rate included having administrators at each school act as gatekeepers to our access to teachers. In most cases, we could not find out if the administrator gave us access to teachers or not. Because we conducted an analytic, rather than a descriptive, study (see Kristensen, 1995, p. 21), we were not concerned with estimating the prevalence of burnout or depression. Rather our goal was to estimate the burnout–depression correlation in a sample in which we had low, medium, and high scorers on the relevant scales.

Each study was conducted in accordance with the ethical standards of the investigator's home university's institutional research committee (and that of the institutional review board of one U.S. locality's education department). In both countries, potential respondents were told that participation was voluntary and that they could withdraw from the study at any time and for any reason. In the three studies, confidentiality was assured. We note that Internet surveys are as reliable and valid as paper-and-pencil questionnaires (Gosling, Vazire, Srivastava, & John, 2004; Jones, Ferny-hough, de-Wit, & Meins, 2008; Ritter, Lorig, Laurent, & Matthews, 2004).

Data Analyses

We conducted a CFA in each sample, using the weighted least squares method in Mplus 8.1 (Muthén & Muthén, 1998–2017). We treated the items as ordinal in view of their skew. As might be expected in a nonclinical sample, depression items were skewed toward the bottom. This is less true for the other items. Regardless, it is current recommended practice to use an ordinal estimation method when analyzing ordinal items. The recent literature recommends weighted least squares or maximum likelihood by numerical integration. Not adjusting for ordinality, especially for skewed items, can distort estimates (Li, 2016). It tends to inflate the number of latent variables needed because so-called "difficulty factors" emerge that reflect similarities among items with similar skew. In Study 1, we created latent Depression and latent Burnout factors. In Study 2, we created latent Depression and latent Physical Fatigue, Cognitive Weariness, and Emotional Exhaustion factors based on the items that belong to their respective scales. In Study 3, we created latent Depression, Anxiety, Emotional Exhaustion, Depersonalization, and Personal Accomplishment factors based on the items that belong respectively to the original scales.

In view of Maslach and Leiter's (2016) earlier mentioned concern about content overlap at the item level, in Study 1 we deliberately excluded two depression items that pertain to fatigue, PHQ Items 3 ("Trouble falling or staying asleep, or sleeping too

much") and 4 ("Feeling tired or having little energy"). Because there was the potential for even greater content-level item overlap involving the PHQ-9 and the SMBM, we deliberately excluded three depression items of the PHQ in Study 2. The items were PHQ Items 3, 4, and 7 ("Trouble concentrating on things, such as reading the newspaper or watching television"). With regard to Item 7 on the PHQ, we were concerned about overlap with the SMBM's Cognitive Weariness items. In Study 3, we deliberately omitted three depression items that pertain to fatigue, PHQ Items 3 and 4 and CES-D-10 Item 5 ("I felt that everything I did was an effort").

Results

In the CFA involving the Study 1 French data, we let PHQ-8 Items 1, 2, 5, 6, 7, and 8 load on latent Depression. All five of the MBI–General Survey's Exhaustion items were allowed to load on latent Burnout. Without cross-loadings or correlated residuals, latent Burnout and latent Depression were highly correlated (r = .86). The chi-square statistic was 308.12, df = 43. Although three (comparative fit index [CFI] = 0.983, Tucker–Lewis index [TLI] = 0.978, standardized root mean square residual [SRMR] = 0.033) of the other four fit statistics Mplus yields indicated a good fit, one statistic (root mean square error of approximation [RMSEA] = 0.082) could have been better. The fact that RMSEA tends to be systematically higher relative to its benchmarks compared to the other fit indices, sometimes artificially so, has been noted in the literature (Browne, MacCallum, Kim, Andersen, & Glaser, 2002). The effect identified by Browne et al. (2002) is most notable when indicators have relatively high reliability, as is the case here. To improve the fit in this CFA, we used rational decision-making supported by the modification indices in allowing three pairs of residuals to correlate, leading to a chi-square of 229.29, df = 40: PHQ Items 1 and 2 (the two cardinal symptoms of depression, anhedonia and depressed mood, which often co-occur), PHQ Items 7 and 8 (attention problems and psychomotor retardation or restlessness), and MBI-GS's Exhaustion subscale Items 1 and 2 (two items with similar meanings, "emotionally drained" and "used up"). The RMSEA fell to 0.072. As expected, the other fit indices remained satisfactory (CFI = 0.988, TLI = 0.983, SRMR = 0.028); the correlation between latent Burnout and latent Depression remained very large (r = .88). We saw no point in continuing efforts to improve the RMSEA. These correlated residuals primarily served to manage misfit generated by local dependence (i.e., doublets) and did not alter the substantive conclusions. If there were additional indicators (e.g., the full MBI–General Survey), we could have developed a second-order CFA. However, given the data in Study 1, it was not possible because a model with a second-order factor would be underidentified.

In the CFA we conducted using the Study 2 U.S. data, we let PHQ-9 Items 1, 2, 5, 6, 8, and 9 load on the Depression factor; we let the appropriate SMBM items load on latent Physical Fatigue, Cognitive Weariness, and Emotional Exhaustion factors. Without cross-loadings or correlated residuals (chi-square = 1457.69, df = 164), the correlations among the latent variables were high (see Table 4 for the final correlations), and three of the other four fit statistics indicated a good fit (CFI = 0.984, TLI = 0.982, SRMR = 0.029). One fit statistic, however, could be improved upon (RMSEA = 0.075). As before, to test whether misfit induced by local dependence mattered, we allowed three pairs of residuals (e.g., Physical Fatigue Items 4 and 6—"fed up" and "burned out") to correlate on both rational (similar meaning) and empirical (modification indices) grounds. The chi-square statistic was 1090.55, df = 161. The fit was improved (CFI = 0.989, TLI = 0.987, SRMR = 0.027, RMSEA = 0.065). The correlations in Table 4 did not change, or changed by .01, suggesting that the misfit indicated by RMSEA did not materially influence the results.

Table 4
Correlations Among the First-Order Factors in the Confirmatory Factor Analysis in Study 2

Factor	Depression	Physical Fatigue	Cognitive Weariness
Physical Fatigue	.83		
Cognitive Weariness	.75	.84	
Emotional Exhaustion	.64	.72	.73

Next, we took advantage of our having four factors by conducting a second-order CFA, which we could not do with the two-factor solution we obtained with the Study 1 data. We let the four factors load on a single second-order distress/dysphoria factor and allowed the same pairs of residuals to correlate as we did above (chi-square = 1125.70, df = 163). The other fit statistics were satisfactory (CFI = 0.988, TLI = 0.986, SRMR = 0.028, RMSEA = 0.065). Each first-order factor's standardized loading on the higher-order factor was 0.78 or greater (M = 0.87).

In Study 3, without cross-loadings or correlated residuals (chi-square = 4806.43, df = 935), we examined the five factors of interest, latent Depression, Anxiety, Emotional Exhaustion, Depersonalization, and Personal Accomplishment (CFI = 0.934, TLI = 0.930, SRMR = 0.061, RMSEA = 0.075). We improved model fit by allowing correlated residuals of items with similar meanings. Reasoning that exhaustion is the core of burnout, and supported by the modification indices, we also allowed cross-loadings. For example, the MBI Personal Accomplishment item (Item 12), "I feel energetic," cross-loaded negatively on the Emotional Exhaustion factor. The model fit improved (CFI = 0.960, TLI = 0.957, SRMR = 0.051, RMSEA = 0.058). The chi-square statistic was 3255.55, df = 927. The final correlations among the first-order factors are found in Table 5. The correlations among the Depression, Anxiety, and Emotional Exhaustion factors found in the table differ slightly from the correlations obtained in the original model (on average between .01 and .02). The correlations involving the Personal Accomplishment factor changed more noticeably (from .07 to .11). This was because some individual items (e.g., "energetic") better reflected a different factor (Emotional Exhaustion) than the factor originally specified as per the make-up of the MBI (Personal Accomplishment subscale). The two positively worded CES-D items (e.g., hopeful about the future) loaded on the Personal Accomplishment factor, and *all* the Personal Accomplishment items are positively worded, underlining local wording-related dependencies.

We conducted a second-order CFA with the Study 3 data (chi-square = 3562.06, df = 932). In view of the high correlations among the Emotional Exhaustion, Depersonalization,

Table 5
Correlations Among the Study 3 First-Order Factors

Factor	Depression	Anxiety	Emotional Exhaustion	Depersonalization
Anxiety	.84			
Emotional Exhaustion	.85	.77		
Depersonalization	.73	.65	.79	
Personal Accomplishment	−.49	−.35	−.47	−.67

Depression, and Anxiety factors, we next let those four factors (but not Personal Accomplishment) load on the second-order distress/dysphoria factor. The fit of this higher-order model was satisfactory (CFI = 0.955, TLI = 0.952, SRMR = 0.055, RMSEA = 0.062). The first-order factors' standardized loadings on the higher order factor were .84 or greater (M = .88).

Discussion

We evaluated, in three samples, the magnitude of the correlation between burnout and depression, adjusting for measurement error and item-level content overlap. To be conservative, we omitted, as per Maslach and Leiter (2016), fatigue-related items from the depression scales under scrutiny. We examined the correlation between latent Burnout, as manifest in the MBI–General Survey's Exhaustion items (Study 1), SMBM items (Study 2), and MBI–Educators Survey items (Study 3), and latent Depression and, in Study 3 only, latent Anxiety. We observed that the correlations between latent Exhaustion and latent Depression were high (rs ≥ .83), and the latent Exhaustion–Anxiety correlation was almost as high (r = .77). These results cast serious doubt on the discriminant validity of burnout scales. The results, combined with the earlier mentioned research on the parallel nomological networks of burnout and depression (Bianchi et al., 2018), suggest that burnout scales essentially measure what depression (and anxiety) scales measure.

Regarding Study 1, latent Burnout and latent Depression correlated .86–.88 depending on whether we released some constraints on the tested model, even with fatigue-related items omitted from the pool of PHQ-8 items. In Study 2, latent Physical Fatigue was found to correlate .83 with latent Depression. Latent Depression had a slightly lower correlation with latent Cognitive Weariness and a still lower correlation with latent Emotional Exhaustion. Our second-order CFA indicated that all four factors loaded highly on the second-order factor, suggesting that Depression, Physical Fatigue, Cognitive Weariness, and Emotional Exhaustion were essentially manifestations of the same construct. Given that Physical Fatigue and Cognitive Weariness (e.g., in terms of difficulties paying attention) are part of the symptomatology of depression, and that we deliberately omitted those symptoms in developing the latent Depression construct, the higher-order construct is more fully reflective of distress/dysphoria. That SMBM-type Emotional Exhaustion, which is more reflective of depersonalized interpersonal relationships at work than of fatigue, loaded on the higher order construct is not surprising because ancillary symptoms of depression include irritability, distancing oneself from others, and "not caring anymore" (APA, 2013).

In Study 3, latent Depression and latent Anxiety were highly related to latent Emotional Exhaustion (correlations were .85 and .77, respectively); latent Depression and latent Anxiety correlated highly with one another (.84). Secondarily, latent Depression, latent Anxiety, and latent Emotional Exhaustion were almost as highly correlated with latent Depersonalization. All four first-order factors loaded on the second-order distress/dysphoria factor. The second-order factor that emerged in Study 3 reflects a dimensional conceptualization of psychopathology, a conceptualization that includes distressful/dysphoric symptoms such as those of depression and anxiety (Kotov et al., 2017).

One potential objection to our interpretation of the results could be that if depression and burnout are "the same thing," then a one-factor model should fit the data well. Put differently, without a one-factor model, the claim that burnout is encompassed by depression would not be strong. Such an objection, though possibly tempting at first sight, is not justified. Factor analysis is a method that accounts for shared variance among a set of items by positing latent variables, which are missing, nonobservable independent variables in a multivariate regression. Upon conditioning on these latent variables, the items should be independent. Ideally, the latent variables have substantively meaningful interpretations. However, it is nearly inevitable that there are one or more small factors that make the one-factor model less plausible in practice and thus undermine the desired model. For example, items that are part of the same scale tend to share some variance due to wording effects common to how the items are written, meaning that the items are likely to make up a factor of their own. Negatively and positively worded items, for instance, tend to share a bit more common variance with each other than would be consistent with a single-factor model. This point is notably elaborated in Podsakoff, MacKenzie, Lee, and Podsakoff (2003).

Maslach and Leiter (2016) argued that part of the reason why burnout and depression scales correlate as highly as they do is because of content overlap at the level of fatigue items in the two types of scales. The findings from our three studies consistently suggest that such content overlap does not explain the high correlations among burnout and depression measures. The results suggest that a better explanation for the high correlation between burnout and depression is that burnout and depression scales tap the same phenomenon, namely, distress/dysphoria.

Another potential concurrent explanation of our findings might be that burnout and depression are comorbid conditions. The comorbidity explanation is problematic because discriminant validity findings bearing on burnout would have to first show that burnout is, indeed, separate from depression; the findings from the three studies fail to establish burnout's discriminant validity. Moreover, it should be recalled that burnout is *not* characterized nosologically and diagnostically. By definition, comorbidity involves (at least) two nosologically separate, diagnosable entities.

Our findings can be put into perspective with those of previous CFA studies. In an influential study of 307 health care workers, Leiter and Durup (1994) conducted a CFA involving the MBI (22-item version), the Beck Depression Inventory, and the Profile of Mood States. The authors concluded that burnout and depression were best-modeled as two second-order factors, while recognizing that the correlation between the two factors (.72) was strong. However, the study had important limitations that undermined the validity of the two- (higher-order) factor solution. First, the final model had a poor fit (adjusted goodness-of-fit index or AGFI = .810). Second, the authors excluded almost half the depression items for reasons of skewness. Skewness was not a problem in our work because we treated the items as ordinal variables. Third, in Leiter and Durup's (1994) study, burnout and depressive symptoms were assessed within highly different time windows (see Bianchi et al., 2016), which was not the case in Study 1 and largely not the case in Studies 2 and 3.

Bakker et al. (2000) published another frequently cited CFA study of burnout and depression measures. This study involved 154 Dutch teachers. These authors focused on (a modified 20-item version of) the MBI and the 20-item version of the CES-D. Bakker et al. (2000) concluded that their CFAs provided "strong evidence for the discriminant validity of burnout and depression" (p. 261). This conclusion, however, is also questionable. In Bakker et al.'s (2000) study, emotional exhaustion correlated more strongly with depression than with depersonalization and personal accomplishment, the two other putative components of burnout. A second problem, which mirrors a problem found in the study by Leiter and Durup (1994), is that the measures of burnout and depressive symptoms covered highly different periods of time (1 year vs. 1 week). In contrast to Leiter and Durup (1994) and Bakker et al. (2000), our research (a) mitigated the problem of incommensurate time periods for symptom assessment, (b) relied on large samples, and (c) used more advanced, technically stronger analytic techniques.

Limitations

The present work has at least three limitations. First, the studies were cross-sectional. There is a need for longitudinal research that would show the stability of burnout and depressive symptoms over time. Second, the sample was limited to one occupational group, educational staff. Research on burnout–depression involving other occupational groups is needed for purposes of external validity. This being mentioned, the evidence enumerated in Table 1 shows that in a variety of occupational groups, and in one general population sample, scores on the MBI's Emotional Exhaustion subscale were more highly related to scores on measures of depression and anxiety than to scores on measures of the other two putative MBI-related components of burnout. These findings were sustained by our meta-analysis. In addition, Ahola, Hakanen, Perhoniemi, and Mutanen's (2014) research with dentists and Wurm et al.'s (2016) research with physicians also tie burnout and depression closely together. A third limitation is that the study samples were predominantly female; it might be feared that the findings do not apply to males. Such a concern is unjustified, however. For instance, the study by Jansson-Fröjmark and Lindblom (2010), included in our meta-analytical review, had substantial numbers of men. In a similar vein, research conducted by Brenninkmeyer et al. (2001) and Halpern, Maunder, Schwartz, and Gurevich (2012) involving predominantly male samples also found that Emotional Exhaustion was more highly related to depressive symptoms than to the other components of the MBI.

Strengths

One strength of our work is that we conducted the research with three different burnout measures and two different depression measures in two different countries and in two different languages, building a degree of replicability in the findings. Second, we used two different depressive symptom scales as well as an anxiety measure. Building an anxiety measure into Study 3 strengthened the case that burnout scales largely measure distress/dysphoria. Third, we used the most commonly employed burnout measure, the MBI, and an alternative measure, the SMBM. Fourth, we relied on advanced analytic techniques that allowed us to overcome limitations of past factor analytic studies. Finally, we controlled item-level content overlap. This analytical strategy allowed us to effectively address recently raised issues surrounding the origin of burnout–depression redundancy (Maslach & Leiter, 2016).

Legal Ramifications of Burnout–Depression Overlap

At clinical levels of severity, depression is a disabling mental disorder that can be diagnosed by occupational health specialists (Kahn, 2008; Liu & Van Liew, 2003). The Americans with Disabilities Act of 1990 (ADA; Public Law No: 101–336; 2008 Amendment to the law, i.e., ADA Amendments Act of 2008, Public Law 110–325) bars discrimination against workers based on their disabilities. A disability is defined as "a physical or *mental* [emphasis ours] impairment that substantially limits one or more major life activities of such individual." The law also requires that workplaces make "reasonable accommodations and modifications" for workers with disabilities. Regarding the ADA, Liu and Van Liew (2003) wrote,

> reasonable accommodations [for major depression] may include providing time off for counseling and stress management groups, providing a mentor, providing additional training to learn new skills and responsibilities, scheduling regular meetings with the supervisor to discuss workplace issues, and parceling a large task into smaller ones so tasks do not seem so overwhelming. (p. 447)

Because burnout is not a diagnostic category in *DSM–5* or the *International Classification of Disease and Related Health Problems*, 10th rev., it is not covered by the law. Because fatigue is very often the presenting problem when an individual seeks help from a health specialist (APA, 2013), the specialist is likely to be alert to the possibility that the individual is suffering from depression. A diagnosis of depression is more likely to provide the individual with a measure of ADA-related protection.

Concluding Thoughts

It may be claimed that burnout causes depression or depression causes burnout. Such claims, however, are hardly justifiable. Because it is extremely difficult to locate a clear difference between

burnout and depression, either at a symptom or an etiological level (Bianchi et al., 2019), staking out a hypothesis that one entity causes the other is of questionable value. As demonstrated by taxometric research (Haslam, Holland, & Kuppens, 2012), depression is best conceived of as a *dimensional* variable, with clinical depression only representing a section—the high end—of the depressive continuum. When adopting a dimensional approach to depression, no theoretical space is left to the notion that burnout is a phase in the development of depression (Bianchi et al., 2018). When examining burnout and depression consistently by adopting a dimensional approach to *both* entities, the continuum of burnout appears to parallel the continuum of depression. Claiming that burnout is a phase in the development of depression implies that one mistakenly reduces depression to its clinical stage, at the upper end of its continuum. Because depression is best conceived of as a dimensional variable, such a reduction constitutes an inacceptable contraction of the phenomenon of depression.

The balance of evidence from the research presented here undermines the view that the discriminant validity of burnout scales is satisfactory (Bakker et al., 2000; Leiter & Durup, 1994), notably because our results were obtained controlling for item-level content overlap. The replication of our findings in three different samples, combined with research that shows that the nomological networks (e.g., relationship to job adversity, stressful life events, social support, work–nonwork interference, attentional, interpretative, and memory biases in the processing of emotional information) for burnout and depression scales are highly parallel (Bianchi et al., 2019), reinforces that view. Our work suggests that research on occupational health could rely on the depression construct to yield critical results.

It should be emphasized that the use of the burnout construct in occupational health research is problematic for several reasons (Bianchi et al., 2018). Nosologically speaking, burnout is undefined; despite more than 40 years of sustained research, the syndrome cannot be diagnosed (Bianchi et al., 2019; Rotenstein et al., 2018). This prevents burnout researchers from getting a clear view of workers' health status. Conclusions from burnout assessments are typically vague and clinically foundationless (Bianchi, 2017; Schears, 2017; Schwenk & Gold, 2018). Growing numbers of studies suggest that burnout and depression are etiologically linked to the same occupational *and* nonoccupational factors (Gauche, de Beer, & Brink, 2017; Verweij et al., 2017). A recent study found that burnout and depressive symptoms were attributed to work to a similar extent by affected individuals (Bianchi & Brisson, 2017). Interestingly, there is evidence that burnout is more strongly related to personality traits such as neuroticism than to occupational-level factors (Bianchi, 2018). In a large meta-analysis, Swider and Zimmerman (2010) found correlations around .50 between the personality traits of the five-factor model and the components of burnout. By comparison, in a study involving 6,815 participants, Leiter and Maslach (2004) found an average correlation of only .26 between the MBI–General Survey and the Areas of Worklife Scale, an instrument designed to assess "the major organizational antecedents of burnout" (Maslach et al., 2001, p. 414). All in all, the burnout construct does not offer a reliable window on *job-elicited* symptoms.

Liu and Van Liew (2003) observed that,

> the term burnout is used so frequently that it has lost much of its original meaning. As originally used, burnout meant a mild degree of stress-induced unhappiness . . . Ultimately, it was used to describe everything from fatigue to a major depression and now seems to have become an alternative word for depression, but with a less serious significance. (p. 434)

Beck and Alford (2009) noted that depression has been recognized in human society for more than 2,000 years. Depression, in its various forms, has been extensively studied and is both diagnosable and treatable. It is well-established that depressive symptoms constitute *basic responses* to unresolvable stress in *Homo sapiens*, including individuals with no noticeable susceptibility to mood disorders (Dura, Stukenberg, & Kiecolt-Glaser, 1990; Pryce et al., 2011; Willner et al., 2013). Depression can be approached from both an individual and a social standpoint, and methods for assessing the weight of occupational factors in the development of depression are available (Bianchi, Schonfeld, & Laurent, 2017). On these bases, we recommend that occupational health specialists focus on depression rather than burnout to more effectively identify and help suffering workers.

References

Ahola, K., Hakanen, J., Perhoniemi, R., & Mutanen, P. (2014). Relationship between burnout and depressive symptoms: A study using the person-centred approach. *Burnout Research, 1,* 29–37. http://dx.doi.org/10.1016/j.burn.2014.03.003

Ahola, K., Honkonen, T., Virtanen, M., Kivimäki, M., Isometsä, E., Aromaa, A., & Lönnqvist, J. (2007). Interventions in relation to occupational burnout: The population-based health 2000 study. *Journal of Occupational and Environmental Medicine, 49,* 943–952. http://dx.doi.org/10.1097/JOM.0b013e31813736e3

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Americans with Disabilities Act (ADA) of 1990 (Public Law 101–336). *Clerk of the Library of the House of Representatives.* Retrieved from http://library.clerk.house.gov/reference-files/PPL_101_336_AmericansWithDisabilities.pdf

Americans with Disabilities Act (ADA) of 1990, as Amended. ADA Amendments Act of 2008 (Public Law 110–325). *The United States Code*. Retrieved from https://www.ada.gov/pubs/adastatute08.htm#12201h

Arthurs, E., Steele, R. J., Hudson, M., Baron, M., Thombs, B. D., & The (CSRG) Canadian Scleroderma Research Group. (2012). Are scores on English and French versions of the PHQ-9 comparable? An assessment of differential item functioning. *PLoS ONE, 7,* e52028. http://dx.doi.org/10.1371/journal.pone.0052028

Bakir, B., Ozer, M., Ozcan, C. T., Cetin, M., & Fedai, T. (2010). The association between burnout, and depressive symptoms in a Turkish military nurse sample. *Klinik Psikofarmakoloji Bülteni-Bulletin of Clinical Psychopharmacology, 20,* 160–163. http://dx.doi.org/10.1080/10177833.2010.11790651

Bakker, A. B. (2009). The crossover of burnout and its relation to partner health. *Stress and Health, 25,* 343–353. http://dx.doi.org/10.1002/smi.1278

Bakker, A. B., Schaufeli, W. B., Demerouti, E., Janssen, P. P. M., Van Der Hulst, R., & Brouwer, J. (2000). Using equity theory to examine the difference between burnout and depression. *Anxiety, Stress, and Coping, 13,* 247–268. http://dx.doi.org/10.1080/10615800008549265

Beck, A. T., & Alford, B. A. (2009). *Depression: Causes and treatment* (2nd ed.). Philadelphia: University of Pennsylvania Press.

Bianchi, R. (2017). Is the "burnout epidemic" an academic fiction? *British Medical Journal, 358,* j4389. http://dx.doi.org/10.1136/bmj.j4389

Bianchi, R. (2018). Burnout is more strongly linked to neuroticism than to work-contextualized factors. *Psychiatry Research, 270,* 901–905. http://dx.doi.org/10.1016/j.psychres.2018.11.015

Bianchi, R., Boffy, C., Hingray, C., Truchot, D., & Laurent, E. (2013). Comparative symptomatology of burnout and depression. *Journal of Health Psychology, 18,* 782–787. http://dx.doi.org/10.1177/1359105313481079

Bianchi, R., & Brisson, R. (2017). Burnout and depression: Causal attributions and construct overlap. *Journal of Health Psychology*. Advance online publication. http://dx.doi.org/10.1177/1359105317740415

Bianchi, R., & Laurent, E. (2015). Emotional information processing in depression and burnout: An eye-tracking study. *European Archives of Psychiatry and Clinical Neuroscience, 265,* 27–34. http://dx.doi.org/10.1007/s00406-014-0549-x

Bianchi, R., & Schonfeld, I. S. (2018). Burnout-depression overlap: Nomological network examination and factor-analytic approach. *Scandinavian Journal of Psychology, 59,* 532–539. http://dx.doi.org/10.1111/sjop.12460

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2014). Is burnout a depressive disorder? A reexamination with special focus on atypical depression. *International Journal of Stress Management, 21,* 307–324. http://dx.doi.org/10.1037/a0037906

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2017). Burnout or depression: Both individual and social issue. *The Lancet, 390,* 230. http://dx.doi.org/10.1016/S0140-6736(17)31606-9

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2018). Burnout syndrome and depression. In Y.-K. Kim (Ed.), *Understanding depression: Vol. 2. Clinical manifestations, diagnosis and treatment* (pp. 187–202). Singapore: Springer Nature. http://dx.doi.org/10.1007/978-981-10-6577-4_14

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2019). Burnout: Moving beyond the status quo. *International Journal of Stress Management, 26,* 36–45. http://dx.doi.org/10.1037/str0000088

Bianchi, R., Verkuilen, J., Brisson, R., Schonfeld, I. S., & Laurent, E. (2016). Burnout and depression: Label-related stigma, help-seeking, and syndrome overlap. *Psychiatry Research, 245,* 91–98. http://dx.doi.org/10.1016/j.psychres.2016.08.025

Brenninkmeyer, V., Van Yperen, N. W., & Buunk, B. P. (2001). Burnout and depression are not identical twins: Is decline of superiority a distinguishing feature? *Personality and Individual Differences, 30,* 873–880. http://dx.doi.org/10.1016/S0191-8869(00)00079-9

Browne, M. W., MacCallum, R. C., Kim, C.-T., Andersen, B. L., & Glaser, R. (2002). When fit indices and residuals are incompatible. *Psychological Methods, 7,* 403–421. http://dx.doi.org/10.1037/1082-989X.7.4.403

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., . . . Moffitt, T. E. (2014). The P factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2,* 119–137. http://dx.doi.org/10.1177/2167702613497473

Choi, J., Peters, M., & Mueller, R. O. (2010). Correlational analysis of ordinal data: From Pearson's r to Bayesian polychoric correlation. *Asia Pacific Education Review, 11,* 459–466. http://dx.doi.org/10.1007/s12564-010-9096-y

Cole, J. C., Rabin, A. S., Smith, T. L., & Kaufman, A. S. (2004). Development and validation of a Rasch-derived CES-D short form. *Psychological Assessment, 16,* 360–372. http://dx.doi.org/10.1037/1040-3590.16.4.360

De la Fuente-Solana, E. I., Gómez-Urquiza, J. L., Cañadas, G. R., Albendín-García, L., Ortega-Campos, E., & Cañadas-De la Fuente, G. A. (2017). Burnout and its relationship with personality factors in oncology nurses. *European Journal of Oncology Nursing, 30,* 91–96. http://dx.doi.org/10.1016/j.ejon.2017.08.004

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. http://dx.doi.org/10.1027//1015-5759.19.1.12

Dohrenwend, B. P., Shrout, P. E., Egri, G., & Mendelsohn, F. S. (1980). Nonspecific psychological distress and other dimensions of psychopathology. Measures for use in the general population. *Archives of General Psychiatry, 37,* 1229–1236. http://dx.doi.org/10.1001/archpsyc.1980.01780240027003

Dura, J. R., Stukenberg, K. W., & Kiecolt-Glaser, J. K. (1990). Chronic stress and depressive disorders in older adults. *Journal of Abnormal Psychology, 99,* 284–290. http://dx.doi.org/10.1037/0021-843X.99.3.284

Faragher, E. B., Cass, M., & Cooper, C. L. (2005). The relationship between job satisfaction and health: A meta-analysis. *Occupational and Environmental Medicine, 62,* 105–112. http://dx.doi.org/10.1136/oem.2002.006734

Gauche, C., de Beer, L. T., & Brink, L. (2017). Exploring demands from the perspective of employees identified as being at risk of burnout. *International Journal of Qualitative Studies on Health and Well-being, 12,* 1361783. http://dx.doi.org/10.1080/17482631.2017.1361783

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59,* 93–104. http://dx.doi.org/10.1037/0003-066X.59.2.93

Hakanen, J. J., & Schaufeli, W. B. (2012). Do burnout and work engagement predict depressive symptoms and life satisfaction? A three-wave seven-year prospective study. *Journal of Affective Disorders, 141,* 415–424. http://dx.doi.org/10.1016/j.jad.2012.02.043

Hakulinen, C., Elovainio, M., Pulkki-Råback, L., Virtanen, M., Kivimäki, M., & Jokela, M. (2015). Personality and depressive symptoms: Individual participant meta-analysis of 10 cohort studies. *Depression and Anxiety, 32,* 461–470. http://dx.doi.org/10.1002/da.22376

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 and Stress, 19,* 208–220. http://dx.doi.org/10.1080/02678370500340728

Halpern, J., Maunder, R. G., Schwartz, B., & Gurevich, M. (2012). Identifying, describing, and expressing emotions after critical incidents in paramedics. *Journal of Traumatic Stress, 25,* 111–114. http://dx.doi.org/10.1002/jts.21662

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine, 42,* 903–920. http://dx.doi.org/10.1017/S0033291711001966

Hätinen, M., Kinnunen, U., Pekkonen, M., & Aro, A. (2004). Burnout patterns in rehabilitation: Short-term changes in job conditions, personal resources, and health. *Journal of Occupational Health Psychology, 9,* 220–237. http://dx.doi.org/10.1037/1076-8998.9.3.220

Jansson-Fröjmark, M., & Lindblom, K. (2010). Is there a bidirectional link between insomnia and burnout? A prospective study in the Swedish workforce. *International Journal of Behavioral Medicine, 17,* 306–313. http://dx.doi.org/10.1007/s12529-010-9107-8

Jones, S. R., Fernyhough, C., de-Wit, L., & Meins, E. (2008). A message in the medium? Assessing the reliability of psychopathology e-questionnaires. *Personality and Individual Differences, 44,* 349–359. http://dx.doi.org/10.1016/j.paid.2007.08.014

Kahn, J. P. (2008). Diagnosis and referral of workplace depression. *Journal of Occupational and Environmental Medicine, 50,* 396–400. http://dx.doi.org/10.1097/JOM.0b013e31816c0043

Kasl, S. V. (1978). Epidemiological contributions to the study of work stress. In C. L. Cooper & R. L. Payne (Eds.), *Stress at work* (pp. 3–38). Chichester, United Kingdom: Wiley.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional

nosologies. *Journal of Abnormal Psychology, 126,* 454–477. http://dx.doi.org/10.1037/abn0000258

Kristensen, T. S. (1995). The demand-control-support model: Methodological challenges for future research. *Stress Medicine, 11,* 17–26. http://dx.doi.org/10.1002/smi.2460110104

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work and Stress, 19,* 192–207. http://dx.doi.org/10.1080/02678370500297720

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32,* 509–515. http://dx.doi.org/10.3928/0048-5713-20020901-06

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders, 114,* 163–173. http://dx.doi.org/10.1016/j.jad.2008.06.026

Kroska, E. B., Calarge, C., O'Hara, M. W., Deumic, E., & Dindo, L. (2017). Burnout and depression in medical students: Relations with avoidance and disengagement. *Journal of Contextual Behavioral Science, 6,* 404–408. http://dx.doi.org/10.1016/j.jcbs.2017.08.003

Lang, P. J., & McTeague, L. M. (2009). The anxiety disorder spectrum: Fear imagery, physiological reactivity, and differential diagnosis. *Anxiety, Stress, and Coping, 22,* 5–25. http://dx.doi.org/10.1080/10615800802478247

Leiter, M. P., & Durup, J. (1994). The discriminant validity of burnout and depression: A confirmatory factor analytic study. *Anxiety, Stress, and Coping, 7,* 357–373. http://dx.doi.org/10.1080/10615809408249357

Leiter, M. P., & Maslach, C. (2004). Areas of worklife: A structured approach to organizational predictors of job burnout. In P. L. Perrewé & D. C. Ganster (Eds.), *Emotional and physiological processes and positive intervention strategies* (Vol. 3, pp. 91–134). Amsterdam, the Netherlands: Elsevier Science/JAI Press.

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48,* 936–949. http://dx.doi.org/10.3758/s13428-015-0619-7

Lindblom, K. M., Linton, S. J., Fedeli, C., & Bryngelsson, I.-L. (2006). Burnout in the working population: Relations to psychosocial work factors. *International Journal of Behavioral Medicine, 13,* 51–59. http://dx.doi.org/10.1207/s15327558ijbm1301_7

Liu, P. M., & Van Liew, D. A. (2003). Depression and burnout. In J. P. Kahn & A. M. Langlieb (Eds.), *Mental health and productivity in the workplace: A handbook for organizations and clinicians* (pp. 433–457). San Francisco, CA: Jossey-Bass.

Martin-Subero, M., Kroenke, K., Diez-Quevedo, C., Rangil, T., de Antonio, M., Morillas, R. M., . . . Navarro, R. (2017). Depression as measured by PHQ-9 versus clinical diagnosis as an independent predictor of long-term mortality in a prospective cohort of medical inpatients. *Psychosomatic Medicine, 79,* 273–282. http://dx.doi.org/10.1097/PSY.0000000000000390

Maslach, C., Jackson, S. E., & Leiter, M. P. (2016). *Maslach Burnout Inventory manual* (4th ed.). Menlo Park, CA: Mind Garden.

Maslach, C., & Leiter, M. P. (2016). Understanding the burnout experience: Recent research and its implications for psychiatry. *World Psychiatry, 15,* 103–111. http://dx.doi.org/10.1002/wps.20311

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52,* 397–422. http://dx.doi.org/10.1146/annurev.psych.52.1.397

Mayer, L. (1971). A note on treating ordinal data as interval data. *American Sociological Review, 36,* 519–520. http://dx.doi.org/10.2307/2093098

Melchior, M., Caspi, A., Milne, B. J., Danese, A., Poulton, R., & Moffitt, T. E. (2007). Work stress precipitates depression and anxiety in young, working women and men. *Psychological Medicine, 37,* 1119–1129. http://dx.doi.org/10.1017/S0033291707000414

Mosing, M. A., Butkovic, A., & Ullén, F. (2018). Can flow experiences be protective of work-related depressive symptoms and burnout? A genetically informative approach. *Journal of Affective Disorders, 226,* 6–11. http://dx.doi.org/10.1016/j.jad.2017.09.017

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Oswin, M. (1978). *Children living in long stay hospitals*. London, UK: Heinemann.

Papineau, M., Morin, A., Legault, L., Demers, C., Chevrier, N., & Côté, A. (2005). *MBI-GS, version française* [MBI-GS, French version]. Groupe de recherche sur l'épuisement professionnel (GREP). Université de Sherbrooke, Québec.

Pettersson, A., Boström, K. B., Gustavsson, P., & Ekselius, L. (2015). Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review. *Nordic Journal of Psychiatry, 69,* 497–508. http://dx.doi.org/10.3109/08039488.2015.1008568

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88,* 879–903. http://dx.doi.org/10.1037/0021-9010.88.5.879

Pryce, C. R., Azzinnari, D., Spinelli, S., Seifritz, E., Tegethoff, M., & Meinlschmidt, G. (2011). Helplessness: A systematic translational review of theory and evidence for its relevance to understanding and treating depression. *Pharmacology and Therapeutics, 132,* 242–267. http://dx.doi.org/10.1016/j.pharmthera.2011.06.006

Rasmussen, A., Verkuilen, J., Jayawickreme, N., Wu, Z., & McCluskey, S. T. (2019). When did post-traumatic stress disorder get so many factors? Confirmatory factor models since DSM–5. *Clinical Psychological Science, 7,* 234–248. http://dx.doi.org/10.1177/2167702618809370

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Ritter, P., Lorig, K., Laurent, D., & Matthews, K. (2004). Internet versus mailed questionnaires: A randomized comparison. *Journal of Medical Internet Research, 6,* e29. http://dx.doi.org/10.2196/jmir.6.3.e29

Rolls, E. T. (2016). A non-reward attractor theory of depression. *Neuroscience and Biobehavioral Reviews, 68,* 47–58. http://dx.doi.org/10.1016/j.neubiorev.2016.05.007

Rössler, W., Hengartner, M. P., Ajdacic-Gross, V., & Angst, J. (2015). Predictors of burnout: Results from a prospective community study. *European Archives of Psychiatry and Clinical Neuroscience, 265,* 19–25. http://dx.doi.org/10.1007/s00406-014-0512-x

Rotenstein, L. S., Torre, M., Ramos, M. A., Rosales, R. C., Guille, C., Sen, S., & Mata, D. A. (2018). Prevalence of burnout among physicians: A systematic review. *Journal of the American Medical Association, 320,* 1131–1150. http://dx.doi.org/10.1001/jama.2018.12777

Schaufeli, W. B. (2017). Burnout: A short socio-cultural history. In S. Neckel, A. K. Schaffner, & G. Wagner (Eds.), *Burnout, fatigue, exhaustion: An interdisciplinary perspective on a modern affliction* (pp. 105–127). Cham, Switzerland: Palgrave Macmillan. http://dx.doi.org/10.1007/978-3-319-52887-8_5

Schaufeli, W. B., Bakker, A. B., & Van Rhenen, W. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior, 30,* 893–917. http://dx.doi.org/10.1002/job.595

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

Schaufeli, W. B., & Taris, T. W. (2005). The conceptualization and measurement of burnout: Common ground and worlds apart. *Work and Stress, 19,* 256–262. http://dx.doi.org/10.1080/02678370500385913

Schears, R. M. (2017). Defining physician burnout, and differentiating between burnout and depression—II. *Mayo Clinic Proceedings, 92,* 1455–1456. http://dx.doi.org/10.1016/j.mayocp.2017.07.006

Schmidt, K.-H., & Diestel, S. (2014). Are emotional labour strategies by nurses associated with psychological costs? A cross-sectional survey. *International Journal of Nursing Studies, 51,* 1450–1461. http://dx.doi.org/10.1016/j.ijnurstu.2014.03.003

Schonfeld, I. S., & Bianchi, R. (2016). Burnout and depression: Two entities or one. *Journal of Clinical Psychology, 72,* 22–37. http://dx.doi.org/10.1002/jclp.22229

Schonfeld, I. S., Bianchi, R., & Luehring-Jones, P. (2017). Consequences of job stress for the psychological well-being of teachers. In T. M. McIntyre, S. E. McIntyre, & D. J. Francis (Eds.), *Educator stress - An occupational health perspective* (pp. 55–75). Cham, Switzerland: Springer International Publishing.

Schonfeld, I. S., & Chang, C.-H. (2017). *Occupational health psychology: Work, stress, and health*. New York, NY: Springer.

Schwenk, T. L., & Gold, K. J. (2018). Physician burnout—A serious symptom, but of what? *Journal of the American Medical Association, 320,* 1109–1110. http://dx.doi.org/10.1001/jama.2018.11703

Shanafelt, T. D., Dyrbye, L. N., & West, C. P. (2017). Addressing physician burnout: The way forward. *Journal of the American Medical Association, 317,* 901–902. http://dx.doi.org/10.1001/jama.2017.0076

Shirom, A., & Melamed, S. (2006). A comparison of the construct validity of two burnout measures in two groups of professionals. *International Journal of Stress Management, 13,* 176–200. http://dx.doi.org/10.1037/1072-5245.13.2.176

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166,* 1092–1097. http://dx.doi.org/10.1001/archinte.166.10.1092

Swider, B. W., & Zimmerman, R. D. (2010). Born to burnout: A meta-analytic path model of personality, job burnout, and work outcomes. *Journal of Vocational Behavior, 76,* 487–506. http://dx.doi.org/10.1016/j.jvb.2010.01.003

Taris, T. W. (2006). Is there a relationship between burnout and objective performance? A critical review of 16 studies. *Work and Stress, 20,* 316–334. http://dx.doi.org/10.1080/02678370601065893

Taris, T. W., Le Blanc, P. M., Schaufeli, W. B., & Schreurs, P. J. G. (2005). Are there causal relationships between the dimensions of the Maslach Burnout Inventory? A review and two longitudinal tests. *Work and Stress, 19,* 238–255. http://dx.doi.org/10.1080/02678370500270453

Toker, S., & Biron, M. (2012). Job burnout and depression: Unraveling their temporal relationship and considering the role of physical activity. *Journal of Applied Psychology, 97,* 699–710. http://dx.doi.org/10.1037/a0026914

Toker, S., Melamed, S., Berliner, S., Zeltser, D., & Shapira, I. (2012). Burnout and risk of coronary heart disease: A prospective study of 8838 employees. *Psychosomatic Medicine, 74,* 840–847. http://dx.doi.org/10.1097/PSY.0b013e31826c3174

Tourigny, L., Baba, V. V., & Wang, X. (2010). Burnout and depression among nurses in Japan and China: The moderating effects of job satisfaction and absence. *The International Journal of Human Resource Management, 21,* 2741–2761. http://dx.doi.org/10.1080/09585192.2010.528656

Trockel, M., Bohman, B., Lesure, E., Hamidi, M. S., Welle, D., Roberts, L., & Shanafelt, T. (2018). A brief instrument to assess both burnout and professional fulfillment in physicians: Reliability and validity, including correlation with self-reported medical errors, in a sample of resident and practicing physicians. *Academic Psychiatry, 42,* 11–24. http://dx.doi.org/10.1007/s40596-017-0849-3

Verweij, H., van der Heijden, F. M. M. A., van Hooff, M. L. M., Prins, J. T., Lagro-Janssen, A. L. M., van Ravesteijn, H., & Speckens, A. E. M. (2017). The contribution of work characteristics, home characteristics and gender to burnout in medical residents. *Advances in Health Sciences Education, 22,* 803–818. http://dx.doi.org/10.1007/s10459-016-9710-9

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for *DSM-V*. *Journal of Abnormal Psychology, 114,* 522–536. http://dx.doi.org/10.1037/0021-843X.114.4.522

Willner, P., Scheel-Krüger, J., & Belzung, C. (2013). The neurobiology of depression and antidepressant action. *Neuroscience and Biobehavioral Reviews, 37,* 2331–2371. http://dx.doi.org/10.1016/j.neubiorev.2012.12.007

Wurm, W., Vogel, K., Holl, A., Ebner, C., Bayer, D., Mörkl, S., . . . Hofmann, P. (2016). Depression-burnout overlap in physicians. *PLoS ONE, 11,* e0149913. http://dx.doi.org/10.1371/journal.pone.0149913

Received September 18, 2018
Revision received February 9, 2019
Accepted February 13, 2019 ■

Correction to Schonfeld et al. (2019)

In the article "Inquiry Into the Correlation Between Burnout and Depression" by Irvin Sam Schonfeld, Jay Verkuilen, and Renzo Bianchi (*Journal of Occupational Health Psychology,* advance online publication, April 4, 2019, http://dx.doi.org/10.1037/ocp0000151), there were wording errors in the Results section. Specifically, we referred to imposing "constraints" when we meant the opposite, namely, relaxing constraints, such as allowing residuals to correlate. We corrected the wording errors, added chi-square statistics, and corrected four small typographic errors bearing on fit statistics (three changes of .001 and one change of .003). The results remain fundamentally the same. All versions of this article have been corrected.

http://dx.doi.org/10.1037/ocp0000165