Reciprocal Associations between Burnout and Depression: An Eight-Year Longitudinal Study

István Tóth-Király*
Substantive-Methodological Synergy Research Laboratory, Concordia University, Montreal, Canada

Alexandre J.S. Morin*
Substantive-Methodological Synergy Research Laboratory, Concordia University, Montreal, Canada

Katariina Salmela-Aro
Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

* The first two authors (I.T-K., & A.J.S.M.) contributed equally to this article and their order was determined at random: Both should thus be considered first authors.

Acknowledgements: The first author was supported by a Horizon Postdoctoral Fellowship from Concordia University in the preparation of the manuscript. Preparation of this paper was also supported by a grant from the Social Sciences and Humanities Research Council of Canada (435-2018-0368). The last author was supported by grants from the Academy of Finland 320241 and 308351.

Conflict of interest: None declared.

Corresponding author:
István Tóth-Király, Substantive-Methodological Synergy Research Laboratory
Department of Psychology, Concordia University
7141 Sherbrooke W, Montreal, QC, Canada, H4B 1R6
E-mail.: tothkiralyistvan@gmail.com; istvan.toth-kiraly@concordia.ca

This document is a pre-publication version of the following manuscript:

Tóth-Király, I., Morin, A.J.S., & Salmela-Aro, K. (In Press, Accepted: 16 November 2020). Reciprocal Associations between Burnout and Depression: An Eight-Year Longitudinal Study. *Applied Psychology: An International Review*. doi: 10.1111/apps.12295

Reciprocal Associations between Burnout and Depression: An Eight-Year Longitudinal Study

Abstract

The purpose of the present four-wave longitudinal study was to examine the differentiation and reciprocal associations between burnout and depression, and their associations with a series of correlates related to employees' physical and psychological health (sleep disturbances, somatic symptoms, self-rated subjective health, and life satisfaction). A total of 542 early career Finnish workers filled out questionnaires four times over a period of eight years. First, our results supported the superiority of a bifactor exploratory structural equation modeling (bifactor-ESEM) representation of employees' burnout ratings, and the empirical differentiation between burnout and depression ratings over each measurement occasion. These results further revealed moderate cross-sectional associations between burnout and depression, supporting their inter-related character but also their empirical distinctiveness. Second, autoregressive cross-lagged analyses revealed that both constructs presented a moderate level of stability over time and reciprocal associations that generalized to all time intervals considered. Finally, relations between depression and all correlates measures during the last wave of the study were in the expected direction, whereas burnout was found to be more weakly related to only a subset of these correlates. Taken together, these results thus support the distinctiveness of burnout and depression, and the presence of mutually reinforcing relations between them.

Keywords: burnout; depression; longitudinal; reciprocal effects; bifactor exploratory structural equation modeling (bifactor-ESEM)

Depression has been identified as one of the most prevalent mental disorders (Kessler et al., 2005) with over 300 million people estimated to be suffering from depression in 2015 (World Health Organization, 2017) and economic costs reaching 83 billion USD annually (Greenberg et al., 2003). Depression has also been found to account for substantial proportion of employees' early retirement decisions (Finnish Centre for Pensions, 2019). Sharing conceptual similarities with depression, even though not diagnosable according to formal diagnostic manuals, burnout has been identified as one of the most detrimental work-related psychological state, and is itself known to carry a heavy burden for organizations and employees alike (Maslach, Schaufeli, & Leiter, 2001). Burnout is typically seen as a work-related phenomenon which may sometimes have consequences that spread out of the workplace (Maslach et al., 2001), whereas depression is typically seen as a personal phenomenon with widespread consequences reaching all facets of one's life (American Psychiatric Association, 2013), although both burnout and depression are known to be driven by individual and social factors. However, despite these conceptual differences, the ability to empirically differentiate these two psychological phenomena has long been a subject of debate in the psychological literature, and empirical research has yielded mixed conclusions. The present four-wave longitudinal study seeks to shed light on the empirical differentiation between burnout and depression in two main ways. First, using longitudinal autoregressive cross-lagged analyses, we seek to achieve a better understanding of the directionality of the longitudinal associations between burnout and depression. Second, we examine how burnout and depression are similarly or differentially related to various health-related correlates.

Burnout and Depression

Depression is characterized by a variety of symptoms, encompassing depressive mood, anhedonia, changes in appetite and weight, loss of energy and fatigue, psychomotor agitation or retardation, impaired concentration, sleep difficulties, feelings of guilt and worthlessness, and suicidal ideation (American Psychiatric Association, 2013). Depression has been linked to multiple undesirable outcomes, including decreased social functioning (Kupferberg et al., 2016), coronary heart disease (Rugulies, 2002), and weakened immune system (Reiche et al., 2004).

Burnout is generally described as a negative psychological state resulting from work-related strain and characterized by a combination of emotional exhaustion (i.e., experiencing chronic fatigue from overworking), cynicism (i.e., being indifferent toward and detached from work, colleagues, and clients as well as not seeing one's work as meaningful), and a sense of professional inadequacy (i.e., diminished feelings of competence and efficacy, and unsuccessfulness in one's work) (Maslach et al., 1997; Shirom & Melamed, 2006). Despite its work-related nature, the consequences of burnout are not limited to the work context, but spread out into personal life, encompassing absenteeism (Ahola et al., 2008), reductions in work performance (Ruotsalainen et al., 2015), and various health-related issues (Peterson et al., 2008). The present study adopts a dimensional view of burnout and depression as two psychological conditions varying along a continuum of severity (Haslam et al., 2012).

Relations between Burnout and Depression

Despite the theoretical distinction between burnout and depression, the empirical ability to differentiate these two constructs remains, at best, uncertain. Yet, this empirical question caries important implications. If burnout and depression are empirically undifferentiated, then their conceptualization needs to be revised to better capture this similarity and practitioners would need to devise intervention strategies encompassing both. Assuming their distinctiveness, then the question of the directionality of their associations becomes critical. If burnout results in increases in depression over time, then burnout-reducing intervention strategies implemented in the workplace are likely to have benefits in terms of depression prevention. Alternatively, if depression is found to enhance burnout, then this would suggest that burnout prevention efforts could benefit from the early identification and treatment of depressive symptoms among employees.

Many studies have documented substantial cross-sectional associations between burnout and depression, with moderate-to-very high correlations (sometimes reaching .80) between both (e.g., Bianchi & Brisson, 2019; Bianchi, Schonfeld, et al., 2020; Schonfeld & Bianchi, 2016; Schonfeld et al., 2019b; Szigeti et al., 2017). Some (e.g., Bianchi et al., 2015b, 2017) have suggested that burnout might not be a distinct psychological construct but should rather be characterized as a specific form of depression. This proposal was met unfavorably by others who rather argued that discarding the construct of burnout was unwarranted at this stage (e.g., Epstein & Privitera, 2017). A recent meta-analysis (Koutsimani et al., 2019) of 67 published studies having looked at associations between burnout and depression reported a moderate correlation (*r* = .52), suggesting that further research was warranted to better understand interrelations among these two constructs. Importantly, conclusions drawn from cross-sectional research remain limited by the inability to specifically assess the directionality of the associations between these two constructs.

When considering these longitudinal, or time-lagged, associations between burnout and depression, four theoretical models can be considered. First, the *stability* model posits that burnout and depression are related to one another at the same time point, but that neither predicts increases or decreases in the other over time. Even though some studies have provided tentative support for this model (e.g., Bianchi et al., 2015a; Idris et al., 2014), the majority of research suggests otherwise.

Second, the *burnout-as-antecedent* model posits unidirectional paths from burnout to depression and implies that burnout may lead to increases depression, but that depression does not result in increases in burnout levels. This view is aligned with the proposition that burnout represents an early phase in the development of depression (Ahola & Hakanen, 2007) and is consistent with the spillover effect (Hakanen & Schaufeli, 2012) rooted in the conservation of resources theory (Hobfoll, 1998) and the job demands resource model (Demerouti et al., 2001). Because burnout is thought to be a work-related negative state, whereas depression is considered to be a more context-general negative state, it is possible that feeling exhausted, detached, and ineffective at work could spill over and generalize to personal life domains (Kantak et al., 1992). Such spillover effects might happen when employees face a high level of discrepancy between the demands of their job and the resources available to them to support their work, resulting in a net loss of psychological resources and energy (Hobfoll, 1989) that they are no longer able to allocate to other areas of their lives. Some empirical support for the spillover effect has been reported in previous longitudinal studies (Hakanen & Schaufeli, 2012; Hakanen, Schaufeli, & Ahola, 2008; Salmela-Aro & Upadyaya, 2014) showing that work- or school-related burnout predicted subsequent increases in depression. However, only two of these studies systematically contrasted alternative models (Hakanen & Schaufeli, 2012; Hakanen et al., 2008), making it hard to obtain a complete picture of these longitudinal interrelations.

Third, the *burnout-as-consequence* model posits unidirectional paths from depression to burnout (i.e., depression might act as a precursor of the negative work-related feelings associated with burnout) but that burnout itself does not lead to increases in depression. Accordingly, some have argued that suffering from depression might decrease employees' access to the psychological resources needed to adequately meet the demands of their job (Adler et al., 2006; de Lange et al., 2004). Depression is characterized by a negative evaluation of one's life, which is likely to translate into an equally negative assessment of one's work situation and to generate difficulties in concentrating, which are likely to generate work difficulties. Previous longitudinal studies provide tentative support for this model (Armon et al., 2012; Salmela-Aro et al., 2008; Upadyaya et al., 2016), suggesting that showing that depressive symptoms do eventually contaminate work-related psychological states. However, once again, only one of those studies contrasted alternative models (Upadyaya et al., 2016).

Fourth, the *reciprocal effects* model assumes that both of the previous models are correct, thus suggesting that burnout and depression both act as a precursor and an outcome of each other. In line with this proposition, the conservation of resources theory (Hobfoll, 1998) further suggests that any loss of psychological resource suffered either in the work domain or in one's personal life may start a downward spiral. In other words, burnout may spillover into one's personal life, leading to increases in depression, which may in turn generalize to the workspace, leading to further increases in burnout. Likewise, depression may lead to reduction in the psychological resources available for employees at work, leading to difficulties in coping with job demands and to increases in burnout symptoms. These symptoms, in turn, may lead to further reductions in personal resources, and thus to further increases in depression. As for the previous models, some previous longitudinal studies have provided support for the reciprocal effects model (e.g., Ahola & Hakanen, 2007; Salmela-Aro, Savolainen, et al., 2009; Toker & Biron, 2012), suggesting mutually reinforcing relations between burnout and depression.

To formally contrast these four alternative models, autoregressive cross-lagged (ARCL) models are needed. ARCL provide a way to consider all possible longitudinal associations between constructs (i.e., the cross-lagged component) while accounting for the longitudinal stability of each construct (i.e., the autoregressive component). Accounting for longitudinal stability as part of this autoregressive component allows the cross-lagged component to directly and explicitly reflect how each construct relate to increases or decreases over time in the other constructs (over and above their own stability).

When working with ARCL using more than two measurement points, additional verifications are required to assess whether the predictive system has reached a state of equilibrium (Cole & Maxwell, 2003). Testing for predictive equilibrium involves the verification of whether the predictive paths generalize across time points. Apart from the statistical advantages of achieving predictive equilibrium (i.e., the resulting model is more parsimonious, leading to more stable and trustworthy estimates), this verification is necessary to ascertain the generalizability of the results across time periods (Cole & Maxwell, 2003). Despite its theoretical importance, to the best of our knowledge, none of the previous studies have examined predictive equilibrium with respect to the relations between burnout and depression. This limitation could explain some of their diverging results, especially when diverging conclusions are limited to a subset of time intervals (Salmela-Aro, Savolainen, et al., 2009).

Hypothesis 1. Burnout and depression will form distinct, yet correlated, factors.

Hypothesis 2. The time-specific associations between burnout and depression will be moderate-to-high (r = .400 to .700), but not high enough to suggest conceptual redundancies (r < .800).

Hypothesis 3. The reciprocal effects model will be supported when compared to the alternative theoretical models (i.e., stability, burnout-as-antecedent, and burnout-as-consequence).

Hypothesis 4. The predictive effects will be similar over time (predictive equilibrium).

Correlates of Burnout and Depression

Another way to examine the distinctness of burnout and depression is to assess their differential associations with correlates. Even assuming that the theoretical differentiation between these constructs could be empirically supported, this differentiation becomes less relevant if both constructs yield identical consequences. In fact, observing such matching effects would suggest that the empirical differentiation between burnout and depression may be more suggestive of burnout being a distinct subtype of depression (Bianchi et al., 2015b, 2017) rather than an entirely distinct phenomenon.

Given that burnout and depression have several undesirable health-related implications (e.g., Shirom et al., 2005; Penninx et al., 2013), we investigate their relations with indicators of physical and psychological health. For physical health, we consider sleep disturbances, somatic symptoms, and self-reported subjective health. Despite occurring within one’s personal life, these manifestations are accompanied by a variety of work-related consequences, including decreased work effectiveness, and sick leaves (e.g., Eriksen et al., 1998). We also considered life satisfaction as an additional correlate, given its longstanding recognition as a focal indicator of psychological wellbeing and a global indicator of the quality of life (Diener et al., 1985).

Thus far, research has generally supported the role of burnout (e.g., Honkonen et al., 2006; Pikó, 2006) and depression (e.g., Bianchi & Mirkovic, 2020; Chang-Quan et al., 2010; Han, 2002) in increasing the risk of physical health difficulties. Research has also supported the presence of negative associations between life satisfaction, burnout (e.g., Ozkan & Ozdevecioglu, 2013; Raiziene et al., 2014; Salmela-Aro & Upadyaya, 2014) and depression (e.g., Busseri & Peck, 2015; Cordeiro et al., 2016; Upadyaya et al., 2016). Although direct evidence regarding the differential role of both psychological constructs in the prediction of these correlates (i.e., as assessed within the same study) remains limited (Hakanen et al., 2008; Upadyaya et al., 2016), examination of the bulk of previous research tentatively suggests that context-general correlates (i.e., not work-related) tend to present stronger associations with depression than with burnout. This differential effect could be explained by the theoretically more widespread nature of depression relative to burnout. Likewise, the fact that these correlates are not work-related could also explain this observation of slightly larger associations.

Hypothesis 5. Burnout and depression will be associated with less desirable correlate levels (i.e., lower subjective health and life satisfaction, and more sleep disturbances and somatic symptoms), although relations involving depression will be more pronounced than those involving burnout.

Considering the Multidimensionality of the Burnout Construct

Despite encompassing a variety of symptoms, depression is typically measured as a single dimension (Ruscio & Ruscio, 2000). In contrast, although a comprehensive measurement of burnout should tap into three types of manifestations (exhaustion, cynicism and inadequacy; Maslach et al., 1997), some have suggested that burnout might be experienced (like depression) as a single overarching phenomenon (Cheng et al., 2016). This perspective is supported by the high correlations reported among burnout dimensions (Demerouti et al., 2010), and by research supporting a higher-order representation of burnout (Sinval et al., 2019). However, research has also revealed differentiated relations between burnout components and outcomes, supporting their distinctive nature (Collie et al., 2018). Thus, a core question is whether enough specificity remains in ratings of exhaustion, cynicism and inadequacy once global burnout levels are considered (Sinval et al., 2019).

These considerations suggest that a third approach might be more relevant to burnout measurement: Burnout may exist as a global phenomenon reflecting commonalities among ratings of exhaustion, cynicism and inadequacy, which themselves may retain a meaningful level of specificity left unexplained by global burnout levels (e.g., Mészáros et al., 2014). Two psychometric approaches can be used to assess this possibility. The first involves higher-order models, where participants' ratings are utilized to estimate first-order factors, which are then used to estimate a higher-order burnout factor (Sinval et al., 2019). Higher-order models, however, suffer from a critical limitation, that of forcing the ratio of variance explained by the global factor versus that of the specific factors to be identical for all indicators associated with the same first-order factor (Gignac, 2016; Morin et al., 2016). The second approach involves bifactor models, which do not share this limitation (Chen et al., 2006; Gignac, 2016). Bifactor models involves the explicit estimation of a global (G-) factor reflecting participants' global levels of burnout defined by all items, while also taking into account the unique qualities associated with each specific dimension (i.e., exhaustion, cynicism, and inadequacy) in the form of orthogonal specific (S-) factors not explained by the G-factor. This approach has been found to match the structure of burnout (Bianchi, 2020; Doherty et al., 2019; Hawrot & Koniewski, 2018; Mészáros et al., 2014; Schonfeld et al., 2019a; Verkulien et al., 2020).

However, in and of itself, the confirmatory factor analytic (CFA) approach, together with its bifactor counterpart, also suffers from another key limitation. As noted by Morin and colleagues (Morin et al., 2016, 2020), this approach is overly restrictive when used with multidimensional constructs as it fails to take into account the inherently fallible nature of questionnaire items, which can typically be expected to demonstrate at least some degree of association with non-target constructs sharing conceptual similarities (i.e., cross-loadings). Recent statistical research has shown that forcing even negligible cross-loadings (i.e., λ = .100) to be zero was likely to result in biased estimates of factor correlations (Morin et al., 2016) and associations with other constructs (Mai et al., 2018), or of the variance attributed to the G-factor (Murray & Johnson, 2013), whereas unnecessary cross-loading do not carry any risk in terms of estimation biases (Asparouhov et al., 2015; Morin et al., 2020).

Exploratory structural equation modeling (ESEM; Morin et al., 2013) and bifactor-ESEM (Morin et al., 2016), together with the confirmatory target rotation (Morin et al., 2020), makes it possible to estimate bifactor models including cross-loadings where factors remained defined in a theoretically-driven manner. When relying on bifactor-ESEM, it remains critical to contrast all four alternative models (CFA, ESEM, bifactor-CFA, bifactor-ESEM) given that each of them is able to absorb unmodeled sources of multidimensionality (Morin et al., 2016, 2020). Previous research has already documented the value of relying on an ESEM (e.g., Trépanier et al., 2015) or bifactor-ESEM (Bianchi, 2020; Doherty et al., 2019; Schonfeld et al., 2019a; Verkulien et al., 2020) representation of burnout. To our knowledge however, none of these approaches have been used in longitudinal research seeking to understand the reciprocal associations between burnout and depression distinction, which could possibly explain some inconsistencies observed in previous results (i.e., biased estimation of the latent construct representing burnout is likely to lead to a biased estimate of its associations with depression).

Hypothesis 6. The results will support the superiority of the bifactor-ESEM representation of burnout relative to alternative CFA, ESEM, and bifactor-CFA representations.

Research Question. We leave as an open question whether the burnout S-factors will result in an incremental contribution over and above the burnout G-factor in the prediction of depression.

The Present Study

In an effort to extend past research, the current study was designed to document the directionality of the longitudinal associations between burnout and depression, while relying on an optimized representation of the multidimensionality of the burnout constructs and a rigorous assessment of the generalizability of these predictions across four distinct time of measurement spanning a total of eight years (i.e., predictive equilibrium). To further improve our understanding of the distinctive nature of these two psychological constructs, this study also investigates the differential associations between burnout and depression, and a range of physical and psychological health-related correlates.

Method

Participants and Procedure

This study relies on data from the Cohort B of the Finnish Educational Transitions project (Salmela-

Aro, 2003-2020). This cohort of participants (*N* = 614 at the initial time of measurement) was recruited in their second year of general upper secondary schooling (aged 17-18), and followed across 7 waves of measurement between 2003 and 2017, with the goal of studying the post-education transition. The present study focused on the 542 early career respondents (54.4% female) who participated in this study during any of the last four measurement waves and who reported being employed at that time. These participants were thus surveyed in 2008/2009 (Time 1: aged 22-23), in 2011 (Time 2: aged 24-25), in 2013/2014 (Time 3: aged 26-27), and again in 2016/2017 (Time 4: aged 29-30). These time lags allowed us to maximize our ability to detect associations as Ford et al. (2014) reported in their meta-analysis that longitudinal effects tend to be small but increasing in magnitude as time lags increase up to 2-3 years. Of these participants, 13.1% reported having a permanent employment and 10.9% working full-time at Time 1, 24.4% reported having a permanent employment and 31.4% working full time at Time 2, 35.1% reported having a permanent employment and 43% working full time at Time 3, and 45.4% reported having a permanent employment and 56.1% working full time at Time 4.

Measures

Depression. At all four waves, depression was assessed using the 9 items from the Finnish Depression Scale (DEPS-10; Salokangas et al., 1994), focusing on participants' mood during the previous month (e.g., "I have felt low in energy or slowed down"; α_{T1} = .878, α_{T2} = .888; α_{T3} = .888; α_{T4} = .902). Items were rated on a 4-point scale (1 = not at all, 4 = very much).

Burnout. At all four waves, burnout was measured with the Finnish 10-item School Burnout Inventory (SBI; Salmela-Aro, Kiuru, et al., 2009), which was modified to change the referent from school to the work context (Salmela-Aro et al., 2011). The SBI assesses three dimensions: exhaustion (4 items, e.g., "I feel I am drowning in work"; α_{T1} = .785, α_{T2} = .791; α_{T3} = .810; α_{T4} = .783), cynicism (3 items, e.g., "I feel I am losing interest in work"; α_{T1} = .874, α_{T2} = .843; α_{T3} = .856; α_{T4} = .886), and inadequacy (3 items, e.g., "I often have feelings of inadequacy at work"; α_{T1} = .798, α_{T2} = .754; α_{T3} = .750; α_{T4} = .780). Items were rated on a 6-point scale (1 = completely disagree, 6 = completely agree).

Correlates. At the last wave, four distinct correlates were measured. *Sleep disturbances* were assessed with 3 items (α = .572; National Institute for Health and Welfare, 2020). The first item referred to the overall quality of sleep (i.e., "How would you rate the overall quality of your sleep during last month?") with a 4-point scale (1 = very poor, 4 = very good). The second item referred to the frequency of taking sleep medication (i.e., "How many times have you taken medicine to fall asleep during last month?") with a 4-point scale (1 = not once during last month, 4 = three times or more in a week). The third item referred to the frequency of difficulties in falling asleep, or waking up at night, using a 4-point scale (1 = rarely or never, 4 = nearly every day). *Somatic symptoms* were measured with 4 items where respondents had to indicate, on a 4-point scale (1 = rarely or never, 4 = nearly every day), the frequency to which they experienced any of the listed symptoms during the past 6 months and (e.g., "neck or shoulder pain"; α = .662; National Institute for Health and Welfare, 2020). *Life satisfaction* was measured using the 5-item (e.g., "For the most part my life is near my ideal", α = .879) Satisfaction with Life Scale (Diener et al., 1985). Items were rated on a 7-point scale (1 = completely disagree, 7 = completely agree). Finally, *subjective health* was measured with a single item (e.g., "How would you rate your health comparing to others of your age?"; National Institute for Health and Welfare, 2020) that was rated on a 5-point scale (1 = poor, 3 = average, 5 = good).

Analyses

Model Estimation

All analyses were conducted using Mplus 8 (Muthén & Muthén, 2017). Models were estimated using the robust weighted least square estimator with mean- and variance-adjusted statistics (WLSMV in Mplus), which has been demonstrated to outperform maximum-likelihood-based estimation when using ordinal indicators (especially with five or less response categories) following asymmetric response thresholds such as those used in the present study (Finney & DiStefano, 2013). Models were estimated using all participants who completed one measurement point (N = 558) using missing data algorithms implemented in Mplus for WLSMV estimation (Asparouhov & Muthén, 2010). More details on missing responses and time points are provided in Appendix 1 of the online supplements.

Measurement Models

Depression was represented by a one-factor model in which one a priori correlated uniqueness (CU) was added between a pair of items to control for the methodological artefact associated with the parallel wording of these items (Morin et al., 2020). Burnout was represented using the bifactor-ESEM framework (Morin et al., 2016, 2020) based on recent empirical evidence showing the value of incorporating a bifactor (Mészáros et al., 2014), ESEM (Trépanier et al., 2015), or both (Bianchi, 2020; Doherty et al., 2019; Schonfeld et al., 2019a; Verkulien et al., 2020) components to achieve a more accurate representation of burnout. Additional details on measurement models specification and selection is provided in Appendix 1.

The final retained measurement models for burnout and depression were then combined into a global model at each time point. To ascertain the distinctive nature of burnout and depression, these time-specific models were contrasted with a restricted alternative in which burnout and depression items were used together to estimate a single global factor, while retaining optimal measurement structure for the burnout model (bifactor, ESEM, bifactor-ESEM). More precisely, pending a bifactor representation of burnout, the depression items were only incorporated to the measurement of the global burnout factor, whereas the specific factors remained defined as before (with, or without, cross-loadings). For ESEM or bifactor-ESEM parameterizations, this restricted model involved the reliance on the ESEM-within-CFA approach previously described by Morin et al. (2013).

Tests of longitudinal measurement invariance were then conducted on the optimal global model across the four time waves (5 factors [depression, burnout G-factor, and three burnout S-factors] × 4 waves = 20 factors) to ascertain that the construct definition remained unchanged over time. These tests were performed in the following sequence (Millsap, 2011): (1) configural invariance (same factor structure), (2) weak invariance (invariance of factor loadings), (3) strong invariance (invariance of factor loadings and thresholds), (4) strict invariance (invariance of factor loadings, thresholds and uniquenesses); (5) invariance of the latent variance-covariance matrix (invariance of factor loadings, thresholds, uniquenesses, factor variances and factor covariances); and (6) latent means invariance (invariance of factor loadings, thresholds, uniquenesses, factor variances, factor covariances and factor means). A priori correlated uniquenesses were added between matching indicators over time to avoid obtaining inflated stability estimates (Marsh, 2007). For all models, we report composite reliability indices (ω; McDonald, 1970).

Predictive Models

To test the longitudinal relations between burnout and depression, a total of nine ARCL models were estimated (see Figure 1). The first four models were used to contrast the theoretical models outlined in the introduction and to test Hypothesis 3. In Model 1 (M1), each construct measured at Time *t* was allowed to predict itself at Time *t* + 1 (autoregressions) to control for the stability of each construct over time (dotted directional arrows in Figure 1). This model corresponds to the theoretical *stability* model. Then, in the second model (M2), predictive paths were added to M1, allowing burnout levels at Time *t* to predict depression levels at Time *t* + 1 (black directional arrows in Figure 1). This model corresponds to the theoretical *burnout-as-antecedent* model. The third model (M3) removed these additional paths, and replaced them by predictive paths between depression levels at Time *t* and burnout levels at Time *t* + 1 (greyscale directional arrows in Figure 1). This model corresponds to the theoretical *burnout-as-consequence* model. Then, the fourth model (M4) combined the previous models, including reciprocal predictive paths allowing all constructs measured at Time *t* were allowed to predict the other constructs measured at Time *t* + 1 (directional arrows in Figure 1). This model corresponds to the theoretical *reciprocal effects* model. Finally, in the fifth model (M5), the effects of the burnout S-factors were constrained to be zero to assess the need to maintain these paths in the model once the effects of the burnout G-factor were taken into account. This last model was used to address our final Research Question. Time-specific correlations between the specific constructs (bidirectional arrows) were also freely estimated in all models (Jöreskog, 1979).

Once the optimal predictive representation was selected, three additional models were estimated in order to assess the predictive equilibrium of this model (Cole & Maxwell, 2003) and to test Hypothesis 4. This verification was done via the progressive inclusion of equality constraints on the autoregressive paths (M6), the predictive paths (M7), and the time-specific correlations (M8). Finally, in the last model (M9), the correlates measured at the last wave were incorporated into the previously retained model and specified to be predicted by burnout and depression at the last time wave (dashed directional arrows in Figure 1). This model was used to test Hypothesis 5.

Model Evaluation

Given the oversensitivity of the chi-square test (χ^2) to sample size and minor misspecifications (Marsh et al., 2005), we report this indicator for the sake of transparency but rely on sample size independent indices to assess model fit (Marsh et al., 2005; Yu, 2002): the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the root mean square error of approximation (RMSEA). CFI and TLI can be interpreted as good or excellent when values are respectively higher than .90 and .95. RMSEA can be interpreted as good or excellent when values are respectively smaller than .08 and .06. In tests of measurement invariance and predictive equilibrium, we focus on changes (Δ) in fit indices: a decrease of .010 or less for CFI and TLI, and an increase of .015 or less for RMSEA suggests that the more restrictive model should be preferred (Chen, 2007; Cheung & Rensvold, 2002). However, it is important to keep in mind that these guidelines have mainly been validated for tests of measurement invariance, making it hard to assess the extent to which they can generalized to more complex forms of longitudinal model comparisons. Importantly, via the principle of error propagation, ARCL models are able to partially or completely absorb unmodelled pathways (such as the paths going from burnout to depression in the estimation of the burnout-as-consequence model) via a simple inflation of the autoregressive paths and time-specific correlations. Thus, for model comparisons not related to measurement invariance or predictive equilibrium, we consider any change in model fit to suggest possible differences, and combine the examination of fit indices with a consideration of parameter estimates, similar to that recommended by Morin et al. (2020) for the comparison of CFA, ESEM, bifactor-CFA, and bifactor-ESEM models (described in Appendix 1 of the online supplements).

Results

Measurement Models: Depression

Table 1 present the goodness-of-fit indices of the depression measurement models across time points, and parameter estimates are reported in Table S1 of the online supplements. Results support the adequacy of this one-factor model over time (CFI/TLI > .95; RMSEA < .06), which resulted in well-defined (λ = .407 to .958, M_λ = .784) and reliable (ω = .929 to .965) factors across time points.

Measurement Models: Burnout

The goodness-of fit of the alternative burnout measurement models are reported in Table 1. Parameter estimates are reported in Table S2 for Times 1-2, and in Table S3 for Times 3-4. At Time 1, the CFA and bifactor-CFA solutions achieved an acceptable fit according to the CFI and TLI, but not the RMSEA. The ESEM solution demonstrated excellent fit according to the CFI and TLI, and an acceptable or marginally acceptable fit according to the RMSEA. The bifactor-ESEM solution achieved a satisfactory level of fit at all time points, and a marked improvement in fit relative to the CFA (ΔCFI $\geq$ +.029; ΔTLI $\geq$ +.034; ΔRMSEA $\geq$ -.055), bifactor-CFA (ΔCFI $\geq$ +.035; ΔTLI $\geq$ +.055; ΔRMSEA $\geq$ -.079), and even ESEM (ΔCFI $\geq$ +.005; ΔTLI $\geq$ +.011; ΔRMSEA $\geq$ -.027) solutions. Based on goodness-of-fit alone, the bifactor-ESEM solution should be retained. However, as noted by Morin et al., (2020), model selection should also be guided by an examination of parameter estimates.

ESEM versus CFA. Parameter estimates for the CFA and ESEM solutions were highly similar as CFA (λ = .612 to .937, M_λ = .791, ω = .811 to .927) and ESEM (λ = -149. to .994, M_λ = .624, ω = .475 to .934) factors appeared to be well-defined by strong target loadings, leading to satisfactory levels of composite reliability. When looking more closely at the ESEM solution, the presence of multiple statistically significant cross-loadings becomes apparent, but most of these cross-loadings remained low enough not to undermine the definition of the factors (λ_{CL} = -.398 to .690, M_{CL} = .201). Although some items had higher than desirable cross-loadings, these higher cross-loadings do not seem to generalize across all time points (item 6 = .558 at Time 1 and .681 at Time 2; item 3 = .544 and .203 at Time 1, .443 and .494 at Time 2, .528 and .247 at Time 3; item 7 = .690 at Time 4), suggesting that these particular items might be more suitable for the assessment of global levels of burnout than for its specific components. In addition, factor correlations (see Table S4) decreased substantially in the ESEM (r = .246 to .775, M_r = .479) relative to CFA (r = .434 to .992, M_r = .738) solution, supporting the ESEM solution (Asparouhov et al., 2015; Morin et al., 2020). Still, the presence of multiple cross-loadings suggests that a global factor might also need to be incorporated.

ESEM versus bifactor-ESEM. Parameter estimates from the bifactor-ESEM solution reveal a well-defined burnout G-factor characterized by high and positive factor loadings (λ = .335 to .919, M_λ = .664) and satisfactory composite reliability (ω = .931 to .948). Beyond this G-factor, the exhaustion S-factor retained a high level of specificity (λ = -.221 to .767, M_λ = .492, ω = .447 to .814), while the cynicism (λ = -.131 to .650, M_λ = .385, ω = .560 to .828) and inadequacy (λ = -.230 to .421, M_λ = .270, ω = .303 to .585) S-factors retained a lower level of specificity. It is important to remember that bifactor factor loadings are typically lower than those in first-order solutions due to the item-level covariance being partitioned into two sources (G- and S-factors) instead of one. These relatively weaker S-factors can still be considered reliable as they are estimated using latent variables that are naturally corrected for measurement error. The superiority of the bifactor-ESEM solution is also reinforced by the reduced magnitude of cross-loadings (λ_{CL} = -.237 to .650, M_{CL} = .111) relative to the ESEM solution. This solution was thus retained as our optimal solution, thus supporting Hypothesis 6.

Measurement Models: Combined

As shown in Table 1, the combined measurement models including burnout and depression resulted in a satisfactory level of fit to the data across time points. In these models, even though the time-specific associations between the two constructs fluctuated over time, these remained moderate-to-large in magnitude (r_{T1} = .474, r_{T2} = .639, r_{T3} = .569, r_{T4} = .560, all *p*s < .001), but not so large that they would suggest conceptual redundancies. This conclusion is supported by the comparison of these models with the more restricted alternative models in which a single global factor was used to represent global burnout and depression levels, which systematically resulted in a small (Time 2: ΔCFI = -.005; ΔTLI = -.004; ΔRMSEA = +.003) to large (Time 1: ΔCFI = -.026; ΔTLI = -.030; ΔRMSEA = +.025; Time 3: ΔCFI = -.036; ΔTLI = -.043; ΔRMSEA = +.032; Time 4: ΔCFI = -.043; ΔTLI = -.053; ΔRMSEA = +.044) decrease in model fit across time points. These results support Hypothesis 1 and 2.

Next, tests of longitudinal measurement invariance were conducted on this global model, and the negligible decrease in model fit (ΔCFI and ΔTLI ≤ .010 and ΔRMSEA ≤ .015) supported the configural, weak, strong, and strict invariance, the invariance of the correlated uniquenesses, the latent variances-covariances and the invariance of latent means across time. The final parameter estimates associated with this model are reported in Table 2, and reveal that all latent factors were well-defined and reliable. Latent variable correlations from this final longitudinal model are reported in Table S5. In this latent mean invariant model, the correlation between global burnout and depression remained moderate (*r* = .566, *p* < .001), corroborating our prior time-specific findings and lending further support to Hypothesis 2. This most invariant model was retained as for the predictive analyses.

Predictive Models

Goodness-of-fit statistics associated with the alternative ARCL models are reported in the bottom section of Table 1. All of these models result in an acceptable level of fit to the data. However, when contrasting the first four theoretical models (M1 to M4), the model resulting in the highest level of model fit appeared to be the theoretical reciprocal effects model (M4). This conclusion was supported by an examination of the parameter estimates from models M1 to M4, as well as of the covariance residuals and modification indices associated with these models, which are all consistent with the presence of reciprocal effects between burnout and depression, thus supporting Hypothesis 3.

In M5, we constrained the relations involving the burnout S-factors to be zero to assess their relative contribution. This model led to a negligible decrease in model fit, suggesting that the longitudinal associations between depression and burnout can be entirely captured by the burnout G-factor. This conclusion was supported by the examination of the parameter estimates obtained in M4 in which the main associations did indeed appear to be limited to global burnout levels. This more parsimonious model (M5) was thus retained, suggesting that the response to our Research Question was that associations involving the S-factors were negligible. In terms of predictive equilibrium, constraining the autoregressions (M6), the cross-lagged predictions (M7), and the time-specific correlations (M8) to equality did not result in a substantial decrease in model fit, thus supporting the predictive equilibrium of the model. M8 was thus retained for interpretation, supporting Hypothesis 4.

Parameter estimates from this model are reported in Table 3. First, autoregressive paths were positive and moderate-to-high in magnitude for both burnout and depression (with standardized regression coefficients [β] varying between .420 and .522[1]), attesting to the relative stability of the constructs over time. Turning out attention to the cross-lagged effects, our results show that depression and burnout were reciprocally related over time (depression → burnout: β = .120 to .138; burnout → depression: β = .101 and .111), Thus supporting the hypothesized reciprocal model.

Finally, the correlates at the last wave were incorporated into M8. The resulting model M9 resulted in an excellent level of model fit, and revealed well-defined and reliable factors associated with sleep disturbances (λ = -.727 to .856; ω = .746), somatic symptoms (λ = .498 to .895; ω = .762), and life

[1] All unstandardized coefficients are constrained to be equality, but slight differences can be observed at the level of the standardized coefficients.

satisfaction (λ = .663 to .968; ω = .912). The additional cross-sectional paths estimated in this model are reported in Table 4 and revealed statistically significant associations between depression and all correlates. Depression negatively predicted subjective health and life satisfaction, and positively predicted sleep disturbances and somatic symptoms. In contrast, burnout only predicted life satisfaction (negatively) and somatic symptoms (positively). It is also interesting to note that the strength of associations found between depression and the correlates were, on average, much larger than those found between burnout and the correlates, thus partially supporting Hypothesis 5. These results support the distinction of burnout and depression not just in terms of measurement models, but also in terms of associations with correlates.

Discussion

This study sought to examine the distinctiveness and longitudinal associations between burnout and depression. Whereas previous longitudinal studies have reported positive associations between burnout and depression (see the review of Bianchi et al., 2015b), the direction of the associations remained inconclusive, reinforcing the need for further research. To increase the precision of these analyses, we relied on a bifactor-ESEM representation of burnout, verified the predictive equilibrium of its longitudinal associations with depression, and considered the differential associations between these two constructs and various correlates measured within the last wave of the study. This approach provided several new insights.

The Structure of Burnout Ratings

Although a secondary objective of the present study, designed to allow us to achieve the best possible operationalization of burnout ratings to obtain a clearer picture of its associations with depression, our results supported the value of a bifactor-ESEM operationalization of burnout. This result replicates recent conclusions (Bianchi, 2020; Doherty et al., 2019; Schonfeld et al., 2019a; Verkulien et al., 2020), and supports Hypothesis 6. One important advantage of the bifactor-ESEM representation is that it provides a way to achieve a direct estimate of employees' global levels of burnout while accounting for the specific levels of exhaustion, cynicism, and inadequacy over and above the global levels. Our results revealed a well-defined and reliable burnout G-factor across all time points, suggesting that all items tapped into a global burnout construct. The results also showed that some of the S-factors retained meaningful specificity over and above employees' global levels of burnout. Another important advantage of this bifactor-ESEM representation is that it provides a more complete, and accurate, accounting of the multidimensionality of burnout measurement, leading to more accurate estimates of associations with other constructs occurring at the global and specific level (e.g., Asparouhov et al., 2015; Mai et al., 2018; Morin et al., 2016, 2020). Our results also provided evidence for the longitudinal measurement invariance of this representation across four distinct time points, attesting to its robustness. In practical terms, these results show that it is possible to simultaneously obtain a direct and explicit estimate of global levels of burnout along with estimates of the level of each co-existing burnout component.

Cross-Sectional Associations and Distinctions between Burnout and Depression

Looking at the time-specific cross-sectional relations between burnout and depression, our results first supported the presence of statistically significant positive associations. The moderate strength of these associations (r = .474 and .639) supported the idea that these two constructs shared conceptual similarities without suggesting the presence of conceptual redundancies, thus supporting Hypothesis 2. Results of more systematic tests of discriminant validity in which global burnout levels and depression were combined to reflect a single factor further supported the idea that these two constructs reflected distinct psychological phenomena, thus supporting Hypothesis 1. These results are generally aligned with previous research results (e.g., Epstein & Privitera, 2017; Koutsimani et al., 2019; Szigeti et al., 2017) supporting the distinctive nature of these two constructs. Thus, at least cross-sectionally, it does seem that burnout cannot be subsumed as a subtype of depression (Bianchi et al., 2015b, 2017).

Reciprocal Longitudinal Associations between Burnout and Depression

Going beyond simple correlational evidence, results from our longitudinal analyses provided unambiguous support for the reciprocal effects model proposed in Hypothesis 3. Our longitudinal results thus supported the idea that burnout and depression tended to present moderately high, and comparable, longitudinal stability (β = .431 to .519). Beyond this stability, our results also revealed statistically significant positive reciprocal longitudinal associations between these two constructs. Thus, levels of burnout were found to predict increases over time in depression, while levels of depression were also found to predict increases over time in burnout. These reciprocal associations were found to be comparable to one another (β = .094 to .125), and occurred while controlling for the stability of both constructs, suggesting a true incremental predictive contribution. This predictive system demonstrated equilibrium over time, thus supporting Hypothesis 4 and the generalizability of the relations over the three time periods considered. Despite the importance of predictive equilibrium to ensure that time-related variations reflect true changes rather than random sampling variations (Cole & Maxwell, 2003), the present study is the first to document this generalizability over time.

In theoretical terms, these results match the downward spiral phenomenon proposed by the conservation of resources theory (Hobfoll et al., 2018) and consistent with the job demands resource model (Demerouti et al., 2001; Hakanen & Schaufeli, 2012). More specifically, our results suggest that two mechanisms seem to simultaneously underlie associations between burnout and depression. The first suggests that burnout leads to an increase in the psychological resources allocated by employees to cope with their job demands, thus leading to a depletion of these resources, which can no longer be allocated to other life areas. This spillover effect of work exhaustion into one's personal life may lead to increases in depression stemming from the inability to pursue non-work-related activities and to efficiently recover from work. The second suggests that depression itself involves a negative world view that come to encompass all spheres of life, and in a depletion of individuals' global store of psychological resources. This phenomenon leads to a reduction in the resources one can allocate to meeting job demands, leading to additional feelings of burnout and exhaustion. In combination, these two mechanisms take the form of a downward spiral within which these two mechanisms mutually reinforce one another. This finding is congruent with previous evidence showing that burnout and depression are mutually interrelated over time (Ahola & Hakanen, 2007; Salmela-Aro et al., 2009).

Finally, our results answered our Research Question by showing that, although some burnout S-factors retained specificity in the measurement models, these S-factors did not have any meaningful contribution to the longitudinal associations between burnout and depression. This result is important and suggest that the core mechanism underpinning associations between burnout and depression is related to employees' global levels of burnout, with no residual predictive value associated with the specific burnout components. This result is important for at least two different reasons. First, it suggests that the discrepant results obtained in previous studies regarding the longitudinal associations between burnout and depression could possibly be explained by their failure to consider the full multidimensionality of the burnout construct. On the one hand, studies separately considering separate burnout components without also considering its common core (i.e., the G-factor) might have obtained discrepant results reflecting this lack of control (i.e., resulting in inflated correlations reflecting this unmodelled common core between burnout components). On the other hand, studies relying on a single global burnout score, without having previously extracted the specificity unique to each component, might have obtained biased results reflecting this confusing combination of global and specific components into a single score. Second, by demonstrating this global/specific nature of the burnout construct and showing that associations with depression only occurred at the global level, these results further demonstrated the distinctive nature of burnout (as a multidimensional construct) and depression (as a unidimensional construct). It is, however, important to keep in mind that this finding does in no way indicate that the burnout items used to assess the S-factors do not tap into key components of burnout. Rather, these results simply suggest that the associations between these two constructs tend to occur at the global, rather than specific, level. Likewise, these conclusions regarding the dominant role of global levels of burnout in prediction should, for the moment and pending additional investigations, be limited to associations between burnout and depression.

Differential Cross-Sectional Associations with Correlates

Finally, we examined the cross-sectional associations between a series of context-general correlates related to participants' physical and psychological health and their levels of burnout and depression at the last wave. We expected both burnout and depression to be associated with all correlates, although we expected associations with depression to be more pronounced than those involving burnout due to the context-general nature of the correlates considered. In this regard, our results provided partial support to Hypothesis 5. Indeed, we found depression to be related to all correlates in the expected direction (associated with lower levels of subjective health and life satisfaction, and with higher levels of sleep disturbances and somatic symptoms), a finding that matches conclusions from earlier studies on the detrimental consequences of depression (Busseri & Peck, 2015; Chang-Quan et al., 2010).

In contrast, burnout was only related to three of the four correlates, albeit in the expected direction. These results are consistent with those from prior studies documenting the deleterious effects of burnout for employees' physical and psychological health (e.g., Pikó, 2006; Salmela-Aro & Upadyaya, 2014). However, burnout was not related to sleep disturbances, suggesting that with respect to this particular correlate, predictions were limited to depression. The magnitude of the associations between burnout and the correlates was similar to that of depression for somatic symptoms, but not for subjective health and life satisfaction whose association with depression were more pronounced. Results associated with three out of four correlates (i.e., subjective health, life satisfaction, and sleep disturbances) thus supported Hypothesis 5, showing that relations were more pronounced with depression than with burnout due to the context-general nature of these correlates. These results lend further support to the conceptual and empirical differentiation between burnout and depression.

Limitations and Future Directions

The present study relied on a robust methodological approach to test the cross-sectional and longitudinal associations between burnout and depression. Despite its strengths, this study is not without limitations. Indeed, this study relies on self-reported questionnaires, which could be influenced by various self-report biases. To address this issue, we encourage researchers to administer informant-reported measures to colleagues and supervisors and to consider more objective health-related indicators, which could, in turn, complement the self-reported data obtained from participants. While the present analytic approach, over four time points, clarified the directionality of the observed relations over time, causality still cannot be inferred. In addition, it would be informative to examine whether and how our results generalize to shorter (6 months, or even daily or weekly measurement for an experience sampling approach) or longer (e.g., 10 years, 15 years, or even longer) time intervals. While shorter time lags would facilitate the study of daily or weekly fluctuations in mood among people exposed to changing environments (e.g., stressful occupations, critical transitions), longer time lags would allow researchers to better understand longer term developmental processes at play over the lifespan (e.g., from young adulthood until retirement).

The inclusion of additional context-general and context-specific correlates may potentially help to enrich interpretations regarding the differential impact of burnout and depression on participants' personal and work lives. A related limitation is that the associations between burnout/depression and the correlates is cross-sectional in nature as all of these constructs were only assessed at the same time point. Future studies should strive to investigate longitudinal predictive associations between these variables. Likewise, it would be interesting for future studies to consider the differential role of work-specific and context-general predictors of burnout and depression (see, for instance, Hakanen & Schaufeli, 2012). This avenue would also provide support for or against the distinction of burnout and depression. More specifically, demonstrating differentiated associations between distinct set of predictors and burnout or depression would support their distinctive nature, whereas observing similar associations would support their similarity. Bearing in mind that this study relied on a sample of early career Finnish employees, generalizing our findings to other countries, cultures, or age groups should be made with caution. For this reason, future studies should be conducing using more diverse samples of employees from different nations, as well as from different career stages. This might be particularly important as burnout tends to be less prevalent among older employees (e.g., Brewer & Shapard, 2004; Drybye et al., 2013) who might put more emphasis on maintaining their health (Salmela-Aro & Upadyaya, 2018) and become more content with their lives in general (Marsh et al., 2013). Another potential future avenue would be the reliance on person-oriented strategies to obtain a different insight into how burnout and depression combine within employees (e.g., Ahola et al., 2014).

Practical Implications

Findings of the present study suggest that several attempts could be made to reduce burnout and depression at the workplace. As for burnout, relying on either an employee-focused or an organization-focused approach (Maslach et al., 2001), several intervention strategies have been proposed and tested already such as interventions aiming to adjust the discrepancy between employees' goals and their actual work situation (Te Brake et al., 2001) or to increase staff support (Le Blanc et al., 2007). A review of intervention programs (Awa et al., 2010) suggests that burnout interventions tend to be effective when enhanced with refresher and follow-up courses or with organization-directed approaches (Panagioti et al., 2017). Likewise, intervention programs have also been proposed for depression management, focusing on, for instance, cognitive-behavioral therapy (CBT; Calear & Christensen, 2010) or a combination of CBT, positive psychology, and emotion-focused techniques (Meyer et al., 2015). Given that the reciprocal model was supported, it appears that systems-oriented perspectives accounting for both depression and burnout might be beneficial. Additionally, based on our results, practitioners should consider intensifying invention efforts toward both burnout or depression, as both of them seem to have the possibility of spilling over to other life areas. For example, by helping to reduce burnout (by way of the aforementioned approaches) might help employees allocate more energy to other aspects of their lives, thus decreasing their general negative mood (i.e., depression) in the process. Similarly, by helping individuals to develop a more positive outlook toward life in general (i.e., by reducing depression), might in turn help these individuals to become more energetic, proactive, and satisfied in their workplaces (i.e., reducing burnout). The two processes might thus lead to an upward spiral, mutually reinforcing one another.

References

Adler, D.A., McLaughlin, T.J., Rogers, W.H., Chang, H., Lapitsky, L., & Lerner, D. (2006). Job performance deficits due to depression. *American Journal of Psychiatry, 163*, 1569–1576.

Ahola, K., & Hakanen, J. (2007). Job strain, burnout, and depressive symptoms: A prospective study among dentists. *Journal of Affective Disorders, 104*, 103-110.

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.

Ahola, K., Kivimäki, M., Honkonen, T., Virtanen, M., Koskinen, S., Vahtera, J., et al. (2008). Occupational burnout and medically certified sickness absence: A population-based study of Finnish employees. *Journal of Psychosomatic Research, 64*, 185–193.

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

Armon, G., Shirom, A., & Melamed, S. (2012). The big five personality factors as predictors of changes across time in burnout and its facets. *Journal of Personality*, *80*, 403-427.

Asparouhov, T., & Muthén, B.O. (2010). *Weighted least square estimation with missing data*. www.statmodel.com/download/GstrucMissingRevision.pdf

Asparouhov, T., Muthén, B., & Morin, A.J.S. (2015). Bayesian Structural equation modeling with cross-loadings and residual covariances. *Journal of Management*, *41*, 1561-1577.

Awa, W.L., Plaumann, M., & Walter, U. (2010). Burnout prevention: A review of intervention programs. *Patient Education and Counseling*, *78*, 184-190.

Bianchi, R. (2020). Do burnout and depressive symptoms form a single syndrome? Confirmatory factor analysis and exploratory structural equation modeling bifactor analysis. *Journal of Psychosomatic Research*, *131*.

Bianchi, R., & Brisson, R. (2019). Burnout and depression: Causal attributions and construct overlap. *Journal of Health Psychology, 24*, 1574-1580.

Bianchi, R., & Mirkovic, D. (2020). Is Machiavellianism associated with depression? A cluster-analytic study. *Personality and Individual Differences*, *152*, 109594.

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2015a). Burnout does not help predict depression among French school teachers. *Scandinavian Journal of Work, Environment & Health*, *41*, 565-568.

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2015b). Burnout–depression overlap: A review. *Clinical Psychology Review*, *36*, 28-41.

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2017). Is it time to characterize burnout as a depressive syndrome? A review of recent research. *European Psychiatry*, *41*, S141.

Bianchi, R., Schonfeld, I. S., & Verkuilen, J. (2020). A five-sample confirmatory factor analytic study of burnout-depression overlap. *Journal of Clinical Psychology*.

Brewer, E.W., & Shapard, L. (2004). Employee burnout: A meta-analysis of the relationship between age or years of experience. *Human Resource Development Review*, *3*, 102-123.

Busseri, M.A., & Peck, E. (2015). Do (even) depressed individuals believe that life gets better and better? The link between depression and subjective trajectories for life satisfaction. *Clinical Psychological Science*, *3*, 715-725.

Calear, A.L., & Christensen, H. (2010). Systematic review of school-based prevention and early intervention programs for depression. *Journal of Adolescence*, *33*, 429-438.

Chang-Quan, H., Xue-Mei, Z., Bi-Rong, D., Zhen-Chan, L., Ji-Rong, Y., & Qing-Xiu, L. (2010). Health status and risk for depression among the elderly: a meta-analysis of published literature. *Age and*

Ageing, *39*, 23-30.
Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement. *Structural Equation Modeling*, *14*, 464-504.
Chen, F.F., West, S.G., & Sousa, K.H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*, 189-225.
Cheng, C., Bartram, T., Karimi, L., & Leggat, S. (2016). Transformational leadership and social identity as predictors of team climate, perceived quality of care, burnout and turnover intention among nurses. *Personnel Review, 45*, 1200-1216.
Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255.
Cole, D.A., & Maxwell, S.E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558-577.
Collie, R.J., Granziera, H., & Martin, A.J. (2018). Teachers' perceived autonomy support and adaptability: An investigation employing the job demands-resources model as relevant to workplace exhaustion, disengagement, and commitment. *Teaching and Teacher Education, 74*, 125-136.
Cordeiro, P., Paixão, P., Lens, W., Lacante, M., & Luyckx, K. (2016). The Portuguese validation of the Basic Psychological Need Satisfaction and Frustration Scale: Concurrent and longitudinal relations to well-being and ill-being. *Psychologica Belgica*, *56*, 193-209.
De Lange, A.H., Taris, T.W., Kompier, M.A., Houtman, I.L., & Bongers, P.M. (2004). The relationships between work characteristics and mental health: Examining normal, reversed and reciprocal relationships in a 4-wave study. *Work & Stress*, *18*, 149-166.
Demerouti, E., Bakker, A.B., Nachreiner, F., & Schaufeli, W.B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*, 499-512.
Demerouti, E., Mostert, K., & Bakker, A.B. (2010). Burnout and work engagement: Investigation of the independency of both constructs. *Journal of Occupational Health Psychology, 15*, 209-222.
Diener, E.D., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49,* 71-75.
Doherty, A., Mallett, J., Leiter, M., & McFadden, P. (2019). Measuring burnout in social work: factorial validity of the Maslach Burnout Inventory-Human Services Survey. *European Journal of Psychological Assessment*, 1-22.
Dyrbye, L.N., Varkey, P., Boone, S.L., Satele, D.V., Sloan, J.A., & Shanafelt, T.D. (2013). Physician satisfaction and burnout at different career stages. *Mayo Clinic Proceedings, 88,* 1358-1367.
Epstein, R.M., & Privitera, M.R. (2017). Physician burnout is better conceptualized as depression–Authors' reply. *The Lancet*, *389,* 1398.
Eriksen, H. R., Svendsrod, R., Ursin, G., & Ursin, H. (1998). Prevalence of subjective health complaints in the Nordic European countries in 1993. *European Journal of Public Health*, *8*, 294-298.
Finney, S.J., & DiStefano, C. (2013). Non-normal and categorical data in structural equation modeling. In G.R. Hancock, & R.O. Mueller (Eds.). *Structural equation modeling: A second course* (pp. 439-492). (2nd ed.). Charlotte, NC: Information Age.
Finnish Centre for Pensions (2019). *Statistics on Pensioners in Finland 2018*. Official Statistics of Finland (OSF): Statistical Yearbook of Pensioners in Finland [e-publication]. https://www.etk.fi/wp-content/uploads/statistics-on-pensioners-in-finland-2018-figures.pdf
Ford, M.T., Matthews, R.A., Wooldridge, J.D., Mishra, V., Kakar, U.M., & Strahan, S.R. (2014). How do occupational stressor-strain effects vary with time? A review and meta-analysis of the relevance of time lags in longitudinal studies. *Work & Stress*, *28*, 9-30.
Gignac, G.E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence, 55,* 57-68.
Greenberg, P.E., Kessler, R.C., Birnbaum, H.G., Leong, S.A., Lowe, S.W., Berglund, P.A., & Corey-Lisle, P.K. (2003). The economic burden of depression in the United States: how did it change between 1990 and 2000? *Journal of Clinical Psychiatry*, *64*, 1465-1475.
Hakanen, J.J., & Schaufeli, W.B. (2012). Do burnout and work engagement predict depressive symptoms and life satisfaction? *Journal of Affective Disorders*, *141*, 415-424.
Hakanen, J.J., Schaufeli, W.B., & Ahola, K. (2008). The Job Demands-Resources model: A three-year cross-lagged study of burnout, depression, commitment, and engagement. *Work & Stress*, *22*, 224-241.
Han, B. (2002). Depressive symptoms and self-rated health in community-dwelling older adults: a

longitudinal study. *Journal of the American Geriatrics Society*, *50*, 1549-1556.
Haslam, N., Holland, E., & Kuppens, P. (2012). Categories vs. dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine*, *42*, 903-920.
Hawrot, A., & Koniewski, M. (2018). Factor structure of the Maslach Burnout Inventory–educators survey in a Polish-speaking sample. *Journal of Career Assessment*, *26*, 515-530.
Hobfoll, S.E. (1989). Conservation of resources. *American Psychologist, 44,* 513-524.
Hobfoll, S.E. (1998). *Stress, Culture, and Community*. Plenum Press, NY.
Hobfoll, S.E., Halbesleben, J., Neveu, J.P., & Westman, M. (2018). Conservation of resources in the organizational context: The reality of resources and their consequences. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*, 103-128.
Honkonen, T., Ahola, K., Pertovaara, M., Isometsä, E., Kalimo, R., Nykyri, E., ... & Lönnqvist, J. (2006). The association between burnout and physical illness in the general population—results from the Finnish Health 2000 Study. *Journal of Psychosomatic Research*, *61*, 59-66.
Idris, M.A., & Dollard, M.F. (2014). Psychosocial safety climate, emotional demands, burnout, and depression: A longitudinal multilevel study in the Malaysian private sector. *Journal of Occupational Health Psychology*, *19*, 291-302.
Jöreskog, K.G. (1979). Statistical models and methods for the analysis of longitudinal data. In K.G. Jöreskog & D. Sörbom (Ed.), *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt Books.
Kantak, D.M., Futrell, C.M., & Sager, J.K. (1992). Job satisfaction and life satisfaction in a sales force. *Journal of Personal Selling & Sales Management*, *12*, 1-7.
Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*, 593-602.
Koutsimani, P., Anthony, M., & Georganta, K. (2019). The relationship between burnout, depression and anxiety: A systematic review and meta-analysis. *Frontiers in Psychology*, *10*, 284.
Kupferberg, A., Bicks, L., & Hasler, G. (2016). Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, *69*, 313-332.
Le Blanc, P.M., Hox, J.J., Schaufeli, W.B., Taris, T.W., & Peeters, M.C. (2007). Take care! The evaluation of a team-based burnout intervention program for oncology care providers. *Journal of Applied Psychology*, *92*, 213-227.
Mai, Y., Zhang, Z., & Wen, Z. (2018). Comparing exploratory structural equation modeling and existing approaches for multiple regression with latent variables. *Structural Equation Modeling*, *25*, 737-749.
Marsh, H.W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 774-798). New York, NY: Wiley.
Marsh, H.W., Hau, K., & Grayson, D. (2005). Goodness of fit in structural equation models. A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics* (pp. 275-340). Mahwah, NJ: Erlbaum.
Marsh, H.W., Nagengast, B., & Morin, A.J.S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology, 49,* 1194-1218.
Maslach, C., Jackson, S.E., & Leiter, M.P. (1997). Maslach Burnout Inventory: Third edition. In C. P. Zalaquett & R. J. Wood (Eds.), *Evaluating stress* (pp. 191–218). Scarecrow Education.
Maslach, C., Schaufeli, W., & Leiter, M. (2001). Job burnout. *Annual Review of Psychology, 52*, 397-422.
McDonald, R.P. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical & Statistical Psychology*, *23*, 1-21.
Mészáros, V., Ádám, S., Szabó, M., Szigeti, R., & Urbán, R. (2014). The Bifactor Model of the Maslach Burnout Inventory–Human Services Survey (MBI-HSS)—An Alternative Measurement Model of Burnout. *Stress and Health*, *30*, 82-88.
Meyer, B., Bierbrodt, J., Schröder, J., Berger, T., Beevers, C. G., Weiss, M., ... & Hautzinger, M. (2015). Effects of an internet intervention (Deprexis) on severe depression symptoms: randomized controlled trial. *Internet Interventions*, *2*, 48-59.
Millsap, R.E. (2011). *Statistical approaches to measurement invariance*. New York: Taylor & Francis.
Morin, A. J.S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric

multidimensionality. *Structural Equation Modeling, 23*, 116-139.
Morin, A.J.S., Marsh, H.W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: a second course* (pp. 395-436). Charlotte: Information Age.
Morin, A.J.S., Myers, N.D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM) and bifactor-ESEM. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (4th ed.). New York, NY: Wiley.
Murray, A.L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence, 41*, 407–422.
Muthén, L.K., & Muthén, B.O. (2017). Mplus user guide. Los Angeles, CA: Muthén & Muthén.
National Institute for Health and Welfare (2020). *School Health Promotion Study*. https://thl.fi/en/web/thlfi-en/research-and-expertwork/population-studies/school-health-promotion-study
Ozkan, A., & Ozdevecioğlu, M. (2013). The effects of occupational stress on burnout and life satisfaction: A study in accountants. *Quality & Quantity*, *47*, 2785-2798.
Panagioti, M., Panagopoulou, E., Bower, P., Lewith, G., Kontopantelis, E., Chew-Graham, C., ... & Esmail, A. (2017). Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Internal Medicine*, *177*, 195-205.
Penninx, B. W., Milaneschi, Y., Lamers, F., & Vogelzangs, N. (2013). Understanding the somatic consequences of depression. *BMC medicine*, *11*, 1-14.
Peterson, U., Demerouti, E., Bergström, G., Samuelsson, M., Åsberg, M., & Nygren, Å. (2008). Burnout and physical and mental health among Swedish healthcare workers. *Journal of Advanced Nursing, 62*, 84-95.
Piko, B.F. (2006). Burnout, role conflict, job satisfaction and psychosocial health among Hungarian health care staff: A questionnaire survey. *International Journal of Nursing Studies*, *43*, 311-318.
Raiziene, S., Pilkauskaite-Valickiene, R., & Zukauskiene, R. (2014). School burnout and subjective well-being: evidence from cross lagged relations in a 1-year longitudinal sample. *Procedia-Social and Behavioral Sciences*, *116*, 3254-3258.
Reiche, E.M.V., Nunes, S.O.V., & Morimoto, H.K. (2004). Stress, depression, the immune system, and cancer. *The Lancet Oncology, 5*, 617-625.
Rugulies, R. (2002). Depression as a predictor for coronary heart disease: a review and meta-analysis. *American Journal of Preventive Medicine, 23,* 51-61.
Ruotsalainen, J.H., Verbeek, J.H., Mariné, A., & Serra, C. (2014). Preventing occupational stress in healthcare workers. *Cochrane Database of Systematic Reviews*, *11*.
Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology*, *109*, 473-487.
Salmela-Aro, K. (2003-2020). *FinEdu – Finnish Educational Transitions Study*. Academy of Finland.
Salmela-Aro, K., Aunola, K., & Nurmi, J.E. (2008). Trajectories of depressive symptoms during emerging adulthood. *European Journal of Developmental Psychology, 5*, 439-465.
Salmela-Aro, K., Kiuru, N., Leskinen, E., & Nurmi, J.E. (2009). School burnout inventory (SBI) reliability and validity. *European Journal of Psychological Assessment*, *25,* 48-57.
Salmela-Aro, K., Rantanen, J., Hyvönen, K., Tilleman, K., & Feldt, T. (2011). Bergen Burnout Inventory: reliability and validity among Finnish and Estonian managers. *International Archives of Occupational and Environmental Health*, *84*, 635-645.
Salmela-Aro, K., Savolainen, H., & Holopainen, L. (2009). Depressive symptoms and school burnout during adolescence. *Journal of Youth and Adolescence, 38*, 1316-1327.
Salmela-Aro, K., & Upadyaya, K. (2014). School burnout and engagement in the context of demands–resources model. *British Journal of Educational Psychology*, *84,* 137-151.
Salmela-Aro, K., & Upadyaya, K. (2018). Role of demands-resources in work engagement and burnout in different career stages. *Journal of Vocational Behavior*, *108,* 190-200.
Salokangas, R.K.R., Stengård, E., & Poutanen, O. (1994). DEPS – Uusi väline depression Seulontaan [DEPS – An instrument for screening depression]. *Duodecim*, *110,* 1141–1148.
Schonfeld, I.S., & Bianchi, R. (2016). Burnout and depression: two entities or one?. *Journal of Clinical Psychology*, *72*, 22-37.
Schonfeld, I.S., Verkuilen, J., & Bianchi, R. (2019a). An exploratory structural equation modeling bi-factor analytic approach to uncovering what burnout, depression, and anxiety scales measure.

Psychological Assessment, *31*, 1073-1079.
Schonfeld, I.S., Verkuilen, J., & Bianchi, R. (2019b). Inquiry into the correlation between burnout and depression. *Journal of Occupational Health Psychology, 24*, 603-616.
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.
Shirom, A., Melamed, S., Toker, S., Berliner, S., & Shapira, I. (2005). Burnout and health review. In G.P. Hodgkinson & J.K. Ford (Eds.), *International review of industrial and organizational psychology Volume 20* (pp. 269-308). West Sussex, UK: Wiley.
Sinval, J., Queirós, C., Pasian, S., & Marôco, J. (2019). Transcultural adaptation of the Oldenburg Burnout Inventory (OLBI) for Brazil and Portugal. *Frontiers in Psychology*, *10*, 338.
Szigeti, R., Balázs, N., Bikfalvi, R., & Urbán, R. (2017). Burnout and depressive symptoms in teachers: Factor structure and construct validity of the Maslach Burnout inventory-educators survey among elementary and secondary school teachers in Hungary. *Stress and Health*, *33*, 530-539.
Te Brake, H., Gorter, R., Hoogstraten, J., & Eijkman, M. (2001). Burnout intervention among Dutch dentists: long-term effects. *European Journal of Oral Sciences*, *109*, 380-387.
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.
Trépanier, S.G., Fernet, C., Austin, S., & Ménard, J. (2015). Revisiting the interplay between burnout and work engagement: An exploratory structural equation modeling approach. *Burnout Research*, *2*, 51-59.
Upadyaya, K., Vartiainen, M., & Salmela-Aro, K. (2016). From job demands and resources to work engagement, burnout, life satisfaction, depressive symptoms, and occupational health. *Burnout Research*, *3*, 101-108.
Verkuilen, J., Bianchi, R., Schonfeld, I.S., & Laurent, E. (2020). Burnout–Depression Overlap: Exploratory Structural Equation Modeling Bifactor Analysis and Network Analysis. *Assessment*.
World Health Organization. (2017). *Depression and other common mental disorders: global health estimates* (No. WHO/MSD/MER/2017.2). World Health Organization.
Yu, C.Y. (2002) *Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes*. Los Angeles, CA: University of California.

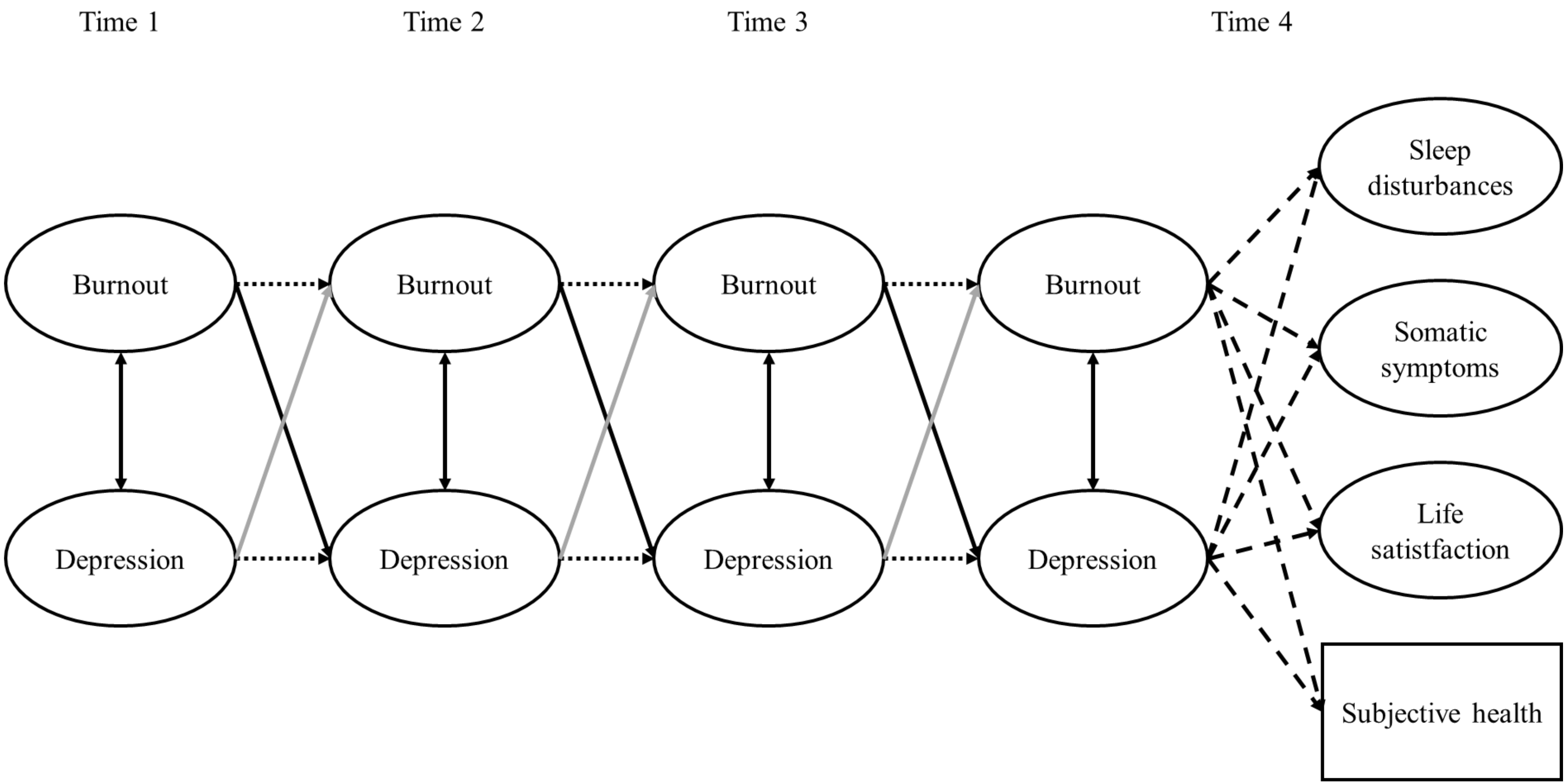

Figure 1

Autoregressive Cross-Lagged Model for Testing the Longitudinal Relationships between Burnout, Depression, and Correlates

Note. Ovals reflect latent variables (the measurement part of the model whereby observed items are used to define latent variables is not shown to maximize clarity), while rectangles reflect manifest variables. Dotted directional arrows represent autoregressive paths whereby each construct at Time t predicts itself at Time t + 1. Directional arrows represent reciprocal predictive paths between burnout and depression (black arrows refer to the burnout-as-antecedent model, whereas greyscale arrows refer to the burnout-as-consequence model). Bidirectional arrows represent time-specific correlations between burnout and depression. Dashed directional arrows represent predictive paths from burnout and depression to the correlates.

Table 1
Goodness-of-Fit Results from the Models Estimated in the Present Study

Models	χ^2	df	CFI	TLI	RMSEA	90% CI
Depression						
Time 1	34.993*	26	.998	.997	.032	.000, .057
Time 2	65.370*	26	.993	.990	.059	.041, .077
Time 3	51.612*	26	.995	.994	.047	.028, .065
Time 4	58.214*	26	.994	.992	.055	.036, .073
Burnout						
Time 1 CFA	175.949*	32	.970	.959	.116	.099, .133
Time 1 ESEM	47.840*	18	.994	.985	.070	.046, .095
Time 1 Bifactor CFA	212.971*	25	.961	.931	.149	.131, .168
Time 1 Bifactor ESEM	15.992*	11	.999	.996	.037	.000, .073
Time 2 CFA	283.272*	32	.958	.941	.138	.123, .153
Time 2 ESEM	66.558*	18	.992	.980	.081	.061, .102
Time 2 Bifactor CFA	276.556*	25	.958	.924	.156	.140, .173
Time 2 Bifactor ESEM	21.334*	11	.998	.993	.048	.014, .078
Time 3 CFA	242.636*	32	.965	.951	.125	.110, .139
Time 3 ESEM	89.732*	18	.988	.970	.097	.078, .117
Time 3 Bifactor CFA	261.780*	25	.961	.930	.149	.133, .166
Time 3 Bifactor ESEM	33.934*	11	.996	.985	.070	.044, .098
Time 4 CFA	294.545*	32	.958	.941	.140	.126, .155
Time 4 ESEM	79.797*	18	.990	.975	.091	.071, .111
Time 4 Bifactor CFA	327.452*	25	.952	.913	.170	.154, .187
Time 4 Bifactor ESEM	22.537*	11	.998	.992	.050	.019, .080
Complete Measurement Model (Including Burnout and Depression)						
Time 1: Distinct factors	320.364*	123	.974	.964	.068	.059, .077
Time 1: Single global factor	529.247*	133	.948	.934	.093	.085; .101
Time 2: Distinct factors	304.071*	123	.983	.976	.058	.050, .066
Time 2: Single global factor	365.380*	133	.978	.972	.063	.056; .071
Time 3: Distinct factors	367.065*	123	.974	.963	.066	.058, .074
Time 3: Single global factor	707.303*	133	.938	.920	.098	.091; .105
Time 4: Distinct factors	326.961*	123	.979	.971	.062	.054, .071
Time 4: Single global factor	764.807*	133	.936	.918	.106	.098; .113
Longitudinal Invariance of the Measurement Model						
Configural	2939.104*	2394	.979	.975	.020	.018, .023
Weak	3059.678*	2490	.978	.975	.021	.018, .023
Strong	3232.570*	2649	.977	.975	.020	.018, .023
Strict	3285.026*	2706	.977	.976	.020	.017, .022
Correlated uniquenesses	3300.701*	2709	.977	.976	.020	.018, .022
Variance-covariance	3307.044*	2754	.978	.978	.019	.017, .022
Latent means	3391.022*	2769	.976	.975	.020	.018, .022
Predictive Autoregressive Cross-lagged Models						
M1. Stability	3873.236*	2898	.962	.963	.025	.023, .027
M2. Burnout-as-antecedent	3857.433*	2886	.962	.962	.025	.023, .027
M3. Burnout-as-consequence	3801.252*	2886	.964	.965	.024	.022, .026
M4. Reciprocal effects	3660.228*	2874	.969	.970	.022	.020, .025
M5. M4 + Effects of S-factors at zero	3787.822*	2892	.965	.965	.024	.022, .026
M6. M5 + Autoregressions equal over time	3781.419*	2902	.966	.966	.024	.021, .026
M7. M6 + Predictions equal over time	3743.591*	2906	.967	.968	.023	.021, .025
M8. M7 + Equal time-specific correlations	3782.947*	2918	.966	.967	.023	.021, .025
M9. M8 + Correlates incorporated	5151.346*	3946	.958	.958	.024	.022, .026

Note. *$p < .01$; CFA: Confirmatory factor analyses; ESEM: exploratory structural equation model; χ^2: WLSMV chi-square; df: Degrees of freedom; CFI: Comparative fit index; TLI: Tucker-Lewis index; RMSEA: Root mean square error of approximation; 90% CI: RMSEA 90% confidence interval.

Table 2

Standardized Parameter Estimates from the Latent Mean Invariant Bifactor-ESEM Measurement Model

	Burnout (λ)	Exhaustion (λ)	Cynicism (λ)	Inadequacy (λ)	Depression (λ)	δ
Exhaustion						
Item 1	**.371****	**.463****	.061	.341**		.528
Item 4	**.499****	**.589****	-.036	.016		.402
Item 8	**.512****	**.556****	-.177**	.077*		.391
Item 10	**.492****	**.647****	.004	.067		.335
Cynicism						
Item 2	**.748****	-.059**	**.467****	.048*		.216
Item 5	**.854****	-.036**	**.351****	-.054**		.144
Item 6	**.807****	-.061*	**.146****	-.116**		.311
Inadequacy						
Item 3	**.674****	.209**	.011	**.538****		.212
Item 7	**.878****	-.047**	.154**	**-.095****		.194
Item 9	**.823****	.045	-.176**	**.061**		.285
Depression						
Item 1					.583**	.660
Item 2					.833**	.307
Item 3					.811**	.342
Item 4					.746**	.444
Item 5					.751**	.435
Item 6					.871**	.241
Item 7					.862**	.256
Item 8					.915**	.163
Item 9					.846**	.285
ω	.936	.754	.581	.411	.943	

Note. *p < .05; **p < .01.; λ: factor loading (target loadings are in bold); δ: uniquenesses; ω: model-based omega composite reliability.

Table 3

Parameter Estimates from the Final Reciprocal Model (M8)

Predictor (t)	Outcome (t+1)	t → t+1 *b* (S.E.)	T1 → T2 β (S.E.)	T2 → T3 β (S.E.)	T3 → T4 β (S.E.)
		Autoregressive paths			
Burnout	Burnout	.496 (.051)**	.431 (.039)**	.475 (.047)**	.489 (.049)**
Depression	Depression	.528 (.044)**	.455 (.032)**	.504 (.039)**	.519 (.042)**
		Predictive cross-lagged paths			
Depression	Burnout	.125 (.047)**	.109 (.041)**	.121 (.046)**	.125 (.047)**
Burnout	Depression	.110 (.040)**	.094 (.035)**	.104 (.039)**	.106 (.040)**

Note. *p < .05; **p < .01.; The final model included invariant predictive paths, which explains why the non-standardized coefficients (b) are invariant across time periods. Conversely, the standardized coefficients (β) are a function of the variances of latent constructs on which no constraints were imposed, and thus differ slightly across time periods

Table 4

Parameter Estimates from the Final Predictive Model (M9)

Predictor	Correlate	*b* (S.E.)	β (S.E.)
Burnout	Subjective health	-.283 (.150)	-.281 (.140)*
Depression	Subjective health	-.380 (.107)**	-.385 (.109)**
Burnout	Sleep disturbances	-.114 (.153)	-.099 (.131)
Depression	Sleep disturbances	.865 (.154)**	.768 (.094)**
Burnout	Somatic symptoms	.358 (.131)**	.345 (.114)**
Depression	Somatic symptoms	.366 (.094)**	.359 (.094)**
Burnout	Life satisfaction	-.269 (.109)*	-.256 (.095)**
Depression	Life satisfaction	-.469 (.073)**	-.455 (.069)**

Note. *p < .05; **p < .01.; The final model included invariant predictive paths, which explains why the non-standardized coefficients (b) are invariant across time periods. Conversely, the standardized coefficients (β) are a function of the variances of latent constructs on which no constraints were imposed, and thus differ slightly across time periods.

Online Supplements for:

Reciprocal Associations between Burnout and Depression: An Eight-Year Longitudinal Study

These online supplements are to be posted on the journal website and hot-linked to the manuscript. If the journal does not offer this possibility, these materials can alternatively be posted on one of our personal websites (we will adjust the in-text reference upon acceptance).

We would also be happy to have some of these materials brought back into the main manuscript, or included as published appendices if you deem it useful. We developed these materials to provide additional technical information and to keep the main manuscript from becoming needlessly long.

Appendix 1

Missing Responses and Time Points

Models were estimated using all 542 participatns who completed at least one measurement point. More precisely, 342 participants participated at Time 1, 435 at Time 2, 453 at Time 3, 418 at Time 4. Of those, 60 completed one measurement point, 87 completed two measurement points, 153 completed three measurement points, and 242 at all four measurement points. Among participants who responded to each time of measurement, missing data at the item level was very low (Time 1: 0% to 2.05%, M = .97%, SD = .87%; Time 2: 0% to 5.29%, M = 2.72%, SD = 2.53%; Time 3: 0% to 6.62%, M = 3.53%, SD = 3.28%; Time 4: 0% to .72%, M = .36%, SD = .24%). Missing responses and time point were handled using algorithms implemented in Mplus for WLSMV estimation (Asparouhov & Muthén, 2010). This procedure is robust under the assumption that data is missing at random (MAR), thus allowing missingness to be conditioned on all latent and observed variables included in the model, which comprise the constructs themselves at the preceding time point (Enders, 2010).

Specification and Selection of the Burnout Measurement Model

In CFA, items were only associated with their a priori factors, all cross-loadings were constrained to zero, and factors were allowed to correlate. In ESEM, the factors were defined in the same way as in the CFA, but all cross-loadings were freely estimated and targeted to be as close to zero as possible through the application of a confirmatory oblique target rotation (Browne, 2001). In bifactor-CFA, all items were associated with one G-factor as well as with their a priori S-factor, cross-loadings were constrained to zero between the S-factors, and factors were specified as orthogonal as per typical bifactor specifications (Morin et al., 2020; Reise, 2012). In bifactor-ESEM, factors were defined as in bifactor-CFA, but cross-loadings were freely estimated between all S-factors and, once again, targeted to be close to zero via the orthogonal target rotation.

To select the optimal measurement model, we followed Morin et al.'s (2020) recommendations, and contrasted first-order and bifactor CFA and ESEM solutions. When contrasting first-order models, the ESEM solution should be preferred as long as factors remain equally well-defined, cross-loadings remain reasonable in magnitude, and estimates of factor correlations reduced (Morin et al., 2020). The retained first-order solution is then contrasted with its bifactor counterpart. Support for the bifactor solution would come from the similar or improved model fit and the observation of a well-defined G-factor together with at least a subset of well-defined S-factors (Morin et al., 2020).

References

Asparouhov, T., & Muthén, B.O. (2010). *Weighted least square estimation with missing data*. www.statmodel.com/download/GstrucMissingRevision.pdf

Browne, M. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*, 111-150.

Enders, C.K. (2010). *Applied missing data analysis*. New York: Guilford.

Morin, A.J.S., Myers, N.D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM) and bifactor-ESEM. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (4th ed.). New York, NY: Wiley.

Reise, S.P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*, 667-696.

Table S1

Standaridzed Parameter Estimates from the 1-Factor Depression Measurement Models at Each Time Point

	Time 1		Time 2		Time 3		Time 4	
	DE (λ)	δ	DE (λ)	δ	DE (λ)	δ	DE (λ)	δ
I1	.407**	.835	.463**	.786	.512**	.737	.481**	.768
I2	.875**	.234	.842**	.291	.861**	.258	.872**	.239
I3	.671**	.550	.722**	.479	.756**	.429	.803**	.356
I4	.600**	.640	.713**	.492	.626**	.608	.781**	.391
I5	.739**	.453	.782**	.389	.764**	.416	.782**	.388
I6	.882**	.223	.883**	.220	.894**	.200	.875**	.234
I7	.852**	.274	.835**	.302	.890**	.209	.812**	.341
I8	.951**	.096	.958**	.083	.927**	.140	.937**	.121
I9	.863**	.255	.853**	.272	.880**	.226	.890**	.209
ω	.929		.938		.965		.945	

Note. *p < .05; **p < .01; I1-I9: item 1-9; λ: Factor loading; δ: Item uniqueness; ω: model-based composite reliability.

Table S2

Standaridzed Parameter Estimates from the Alternative Burnout Measurement Models at Time 1 and Time 2

	CFA				ESEM				Bifactor CFA					Bifactor-ESEM				
	E (λ)	C (λ)	I (λ)	δ	E (λ)	C (λ)	I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ
Time 1																		
Exhaustion																		
I1	.699**			.512	**.633****	-.166	.256**	.513	.424**	.528**			.541	**.354****	**.569****	.023	.282**	.471
I4	.754**			.431	**.769****	.327**	-.257**	.353	.446**	.607**			.432	**.424****	**.682****	.129*	-.162*	.313
I8	.757**			.427	**.744****	-.130	.176*	.395	.726**	.645**			.395	**.440****	**.652****	-.224**	-.001	.331
I10	.701**			.508	**.636****	-.075	.170	.527	.774**	.561**			.512	**.383****	**.550****	-.011	.132*	.533
Cynicism																		
I2		.866**		.249	.007	**.700****	.182	.274	.797**		.453**		.268	**.779****	-.064*	**.395****	.070*	.227
I5		.937**		.123	.003	**.994****	.001	.009	.756**		.580**		.029	**.867****	-.068*	**.418****	-.044	.067
I6		.829**		.313	-.059	**.331****	.558**	.329	.968**		.231**		.375	**.846****	-.119**	**-.023**	-.115*	.256
Inadequacy																		
I3			.719**	.483	.394**	-.012	**.525****	.417	.434**			.629	.004	**.650****	.305**	.018	**.316****	.384
I7			.905**	.181	-.023	.253**	**.732****	.130	.794**			-.249	.002	**.919****	-.075*	.015	**.058**	.147
I9			.777**	.396	.197**	.101	**.618****	.362	.417**			.060	.367	**.762****	.105**	.018	**.282****	.329
ω	.819	.927	.845		.812	.870	.795		.936	.749	.704	.702		.931	.785	.560	.334	
Time 2																		
	CFA				ESEM				Bifactor CFA					Bifactor-ESEM				
	E (λ)	C (λ)	I (λ)	δ	E (λ)	C (λ)	I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ
Exhaustion																		
I1	.612**			.625	**.677****	.219**	-.321**	.500	.423**	.445**			.623	**.335****	**.431***	.154*	.366**	.544
I4	.771**			.406	**.750****	-.049	.084	.422	.534**	.566**			.395	**.511****	**.536****	.005	.193	.413
I8	.759**			.424	**.814****	-.179**	.147*	.379	.510**	.599**			.381	**.505****	**.610****	-.108**	.164	.334
I10	.798**			.363	**.701****	.058	.053	.424	.573**	.511**			.411	**.548****	**.493****	.043	.196	.416
Cynicism																		
I2		.805**		.352	-.013	**.834****	.022	.290	.697**		.472**		.291	**.745****	-.057	**.330****	.011	.333
I5		.921**		.152	.046	**.839****	.107	.115	.832**		.455**		.102	**.874****	-.002	**.433****	-.042	.047
I6		.829**		.313	.082	**.205****	.681**	.237	.789**		.106		.367	**.863****	-.106	**-.131**	-.055	.224
Inadequacy																		
I3			.670**	.551	.544**	.203**	**.075**	.485	.729**			.669	.020	**.610****	.060	-.019	**.761***	.045
I7			.886**	.214	.073	.463**	**.503****	.147	.959**			-.264	.011	**.915****	-.070	.055	**-.043**	.153
I9			.736**	.458	.348**	.181*	**.395****	.416	.760**			.064	.418	**.742****	.102	-.038	**.131****	.421
ω	.826	.889	.811		.934	.846	.475		.939	.697	.584	.689		.938	.715	.570	.585	

Note. **p* < .05; ***p* < .01; CFA: confirmatory factor analysis; ESEM: exploratory structural equation modeling; I1-I10: item 1-10; E: exhaustion; C: cynicism; I: inadequacy; G-B: global burnout factor as part of a bifactor model; S-E: exhaustion specific factor as part of a bifactor model; S-C: cynicism specific factor as part of a bifactor model; S-I: inadequacy specific factor as part of a bifactor model; λ: Factor loading; δ: Item uniqueness; ω: model-based composite reliability; Target factor loadings are in bold.

Table S3

Standaridzed Parameter Estimates from the Alternative Burnout Measurement Models at Time 3 and Time 4

	Time 3																	
	CFA				ESEM				Bifactor CFA					Bifactor-ESEM				
	E (λ)	C (λ)	I (λ)	δ	E (λ)	C (λ)	I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ
Exhaustion																		
I1	.654**			.572	**.732****	.263**	-.398**	.427	.427**	.487**			.581	**.429****	**.454****	-.110	.426**	.416
I4	.781**			.390	**.720****	.016	.052	.435	.509**	.586**			.398	**.440****	**.767****	.174**	-.088	.181
I8	.876**			.232	**.901****	-.162**	.151**	.198	.755**	.670**			.231	**.579****	**.617****	-.172	.036	.253
I10	.748**			.441	**.765****	-.099	.079	.432	.675**	.596**			.424	**.518****	**.511****	-.237**	.066	.411
Cynicism																		
I2		.828**		.314	-.035	**.919****	-.030	.220	.782**		.538**		.255	**.715****	-.032	**.503****	.097**	.225
I5		.926**		.142	.003	**.793****	.180**	.153	.751**		.517**		.121	**.804****	-.004	**.466****	-.050	.134
I6		.834**		.304	.059	**.451****	.419**	.328	.919**		.272**		.363	**.771****	-.011	**.221****	-.167**	.328
Inadequacy																		
I3			.681**	.536	.443**	.494**	**-.149***	.476	.566**			.623	.043	**.658****	.183*	.043	**.364****	.399
I7			.850**	.278	.070*	.321**	**.627****	.182	.798**			-.329	.047	**.871****	-.077**	.089	**-.230***	.175
I9			.775**	.399	.336**	.240**	**.382****	.381	.470**			-.018	.363	**.801****	.100	-.057	**-.038**	.343
ω	.851	.898	.814		.867	.870	.563		.940	.774	.704	.675		.938	.814	.673	.303	
	Time 4																	
	CFA				ESEM				Bifactor CFA					Bifactor-ESEM				
	E (λ)	C (λ)	I (λ)	δ	E (λ)	C (λ)	I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ	G-B (λ)	S-E (λ)	S-C (λ)	S-I (λ)	δ
Exhaustion																		
I1	.701**			.508	**.912****	.107	-.280	.139	.513**	.406**			.572	**.666****	**.668**	-.064	-.102	.095
I4	.707**			.501	**.572****	-.135*	.395*	.498	.483**	.516**			.501	**.677****	**-.032**	-.146	-.061	.516
I8	.802**			.357	**.602****	-.171*	.532**	.331	.555**	.551**			.388	**.820****	**-.221**	-.140	-.231	.205
I10	.728**			.469	**.719****	-.205**	.316	.413	.464**	.655**			.355	**.707****	**.083**	-.230**	-.087*	.432
Cynicism																		
I2		.871**		.241	.129*	**.812****	.038	.202	.743**		.458**		.237	**.640****	.091	**.611****	.108	.197
I5		.931**		.133	.051	**.822****	.175**	.140	.803**		.529**		.076	**.675****	-.066	**.650****	.147	.095
I6		.848**		.280	.018	**.654****	.330**	.287	.770**		.271**		.333	**.629****	-.088	**.439****	.305*	.310
Inadequacy																		
I3			.728**	.470	.528**	.247**	**.188**	.423	.797**			-.584	.023	**.731****	.180	.040	**.172***	.401
I7			.868**	.246	-.004	.690**	**.364****	.205	.934**			.321	.024	**.640****	-.060	.413**	**.594****	.064
I9			.769**	.409	.272**	.359**	**.410****	.375	.786**			.002	.382	**.725****	-.071	.172	**.253**	.376
ω	.825	.915	.833		.851	.893	.480		.942	.691	.710	.657		.947	.447	.828	.553	

Note. *p < .05; **p < .01; CFA: confirmatory factor analysis; ESEM: exploratory structural equation modeling; I1-I10: item 1-10; E: exhaustion; C: cynicism; I: inadequacy; G-B: global burnout factor as part of a bifactor model; S-E: exhaustion specific factor as part of a bifactor model; S-C: cynicism specific factor as part of a bifactor model; S-I: inadequacy specific factor as part of a bifactor model; λ: Factor loading; δ: Item uniqueness; ω: model-based composite reliability; Target factor loadings are in bold.

Table S4

Latent Factor Correlations from the Three-Factor CFA (below the diagonal) and ESEM (above the diagonal) Solutions for the Burnout Inventory Across the Four Waves

	Exhaustion	Cynicism	Inadequacy
Time 1			
Exhaustion	—	.343**	.400**
Cynicism	.434**	—	.775**
Inadequacy	.683**	.935**	—
Time 2			
Exhaustion	—	.540**	.402**
Cynicism	.586**	—	.673**
Inadequacy	.785**	.992**	—
Time 3			
Exhaustion	—	.482**	.417**
Cynicism	.504**	—	.643**
Inadequacy	.729**	.933**	—
Time 4			
Exhaustion	—	.455**	.246**
Cynicism	.556**	—	.377**
Inadequacy	.774**	.946**	—

Note. * $p < .05$; ** $p < .01$; CFA: Confirmatory factor analysis; ESEM: Exploratory structural equation modeling.

Table S5

Latent Correlations

	1	2	3	4	5	6	7	8	9	10
1. Global Burnout T1	—									
2. Exhaustion T1	0	—								
3. Cynicism T1	0	0	—							
4. Inadequacy T1	0	0	0	—						
5. Depression T1	.566**	.273**	-.022	.082*	—					
6. Global Burnout T2	.302**	.072	.193*	.077	.271**	—				
7. Exhaustion T2	.027	.312**	-.045	-.074	.099	0	—			
8. Cynicism T2	.038	.047	.162	-.198	.012	0	0	—		
9. Inadequacy T2	.111	.205*	-.024	.480**	.194*	0	0	0	—	
10. Depression T2	.281**	.180**	.219**	-.041	.369**	.566**	.273**	-.022	.082*	—
11. Global Burnout T3	.308**	.051	.021	.087	.219**	.492**	.166**	.079	-.129	.426**
12. Exhaustion T3	.083	.342**	.036	-.131	.165*	.106	.247**	-.248**	-.160	.168**
13. Cynicism T3	-.272**	.114	.249	.094	-.192*	-.179*	-.299**	.201	.039	-.170
14. Inadequacy T3	-.075	.135	.129	.216	.035	-.073	.133	.078	.377**	.002
15. Depression T3	.324**	.214*	.091	.145*	.418**	.408**	.224**	-.057	.056	.515**
16. Global Burnout T4	.101	.179*	.165	-.006	.172**	.448**	-.068	-.072	.176*	.291**
17. Exhaustion T4	.062	.269**	.093	.069	.182**	.072	.386**	-.294**	-.029	.229**
18. Cynicism T4	-.129	.127	.034	-.079	-.038	-.065	.171	.113	-.091	-.046
19. Inadequacy T4	.090	-.053	-.173	.403**	.071	.016	-.168	.209	.419**	-.188**
20. Depression T4	.316**	.093	-.075	-.043	.338**	.304**	.086	-.168	.166	.348**

Table S5 (continued)

	11	12	13	14	15	16	17	18	19
11. Global Burnout T3	—								
12. Exhaustion T3	0	—							
13. Cynicism T3	0	0	—						
14. Inadequacy T3	0	0	0	—					
15. Depression T3	.566**	.273**	-.022	.082*	—				
16. Global Burnout T4	.409**	-.037	-.083	.115	.471**	—			
17. Exhaustion T4	.136*	.498**	-.133	.002	.266**	0	—		
18. Cynicism T4	-.133	.135	.716**	-.058	-.145*	0	0	—	
19. Inadequacy T4	-.035	.034	-.051	.600**	.084	0	0	0	—
20. Depression T4	.262**	.045	-.254**	.220*	.552**	.566**	.273**	-.022	.082*

Note. *p < .05; **p < .01.; T: Time.