

SHORT NOTE

MEASURING PROFESSIONAL BURNOUT IN DUTCH-SPEAKING REGIONS: AN EVALUATION OF THE FACTORIAL VALIDITY OF THE MASLACH BURNOUT INVENTORY

Stijn VANHEULE, Yves ROSSEEL
Ghent University

and Stefan BOGAERTS
Tilburg University, The Netherlands

Professional burnout among professionals in human services has usually been studied by making use of the Maslach Burnout Inventory (MBI). Using confirmatory factor analysis, the present study evaluates the factorial validity of the three versions of this scale that are available for Dutch-speaking regions (the full version with 22 items, a version with 20 items and a version with 16 items). The sample consisted of 995 Flemish special educators. The version of the MBI with 16 items proved to be slightly better than the others, but the fit of the model was still unsatisfactory. Implications for further research are discussed.

Introduction

In the mid-1970s, the concept of professional burnout was first used in psychological thinking on excessive negative job stress. Herbert Freudenberger (1974), an American psychoanalyst working in alternative American health-care, observed that many caregivers (himself included) working with demanding and severely ill patients gradually became emotionally exhausted and lost their motivation. This expressed itself in several mental (e.g., feelings of frustration) and physical (e.g., fatigue) symptoms. He classified the state of these caregivers with the term 'burnout'. Shortly after, Maslach (1976) started to study burnout academically from a social-

Vanheule: Department of Psychoanalysis and Clinical Consulting, Ghent University; Rosseel: Department of Data Analysis, Ghent University; Bogaerts: TIAS Business School, The Netherlands.

Correspondence concerning this article should be addressed to Stijn Vanheule, Department of Psychoanalysis and Clinical Consulting, Ghent University, Henri Dunantlaan 2, 9000 Gent, e-mail: Stijn.Vanheule@UGent.be; to Yves Rosseel, Department of Data Analysis, Ghent University, Henri Dunantlaan 2, 9000 Gent, e-mail: Yves.Rosseel@UGent.be; or to Stefan Bogaerts, TIAS Business School, O.O. Box 90163, 5000 LE Tilburg, The Netherlands, e-mail: s.bogaerts@tias.edu

psychological perspective. From then on, the study of the concept became increasingly popular (Maslach, Schaufeli & Leiter, 2001).

Much attention has been paid to the study of symptoms and complaints associated with the construct – e.g., Burisch (1993) concluded that there were some 130. Based on factor-analysis, Maslach and Jackson (1981) reduced the multitude of symptoms to a three-dimensional structure. They defined burnout via the following state-description (Maslach & Jackson, 1986, p. 1): “Burnout is a syndrome of emotional exhaustion, depersonalisation and reduced personal accomplishment that can occur among individuals who do ‘people work’ of some kind.” Emotional exhaustion can be understood as a dysphoric feeling of being down; depersonalisation refers to assuming an impersonal attitude toward the people with whom one works, and reduced personal accomplishment indicates a reduction in the feeling of being competent (Maslach, 1993). According to this definition, burnout is a typical phenomenon in the context of people-oriented professions, in which interpersonal relations are pivotal (such as teachers, nurses, psychologists, etc.).¹ This description is the most widely accepted definition of burnout (Schaufeli & Enzmann, 1998). Based on this definition a questionnaire has been developed: the Maslach Burnout Inventory (MBI), which is internationally considered to be the pre-eminent instrument for assessing burnout (Schaufeli & Enzmann, 1998).²

For the Dutch-speaking regions, two adaptations of the original MBI have been made: one based on a Flemish translation (Vlaamse Maslach Burnout Inventory, MBI-VL) (Vlerick, 1993, 1995), and one based on a Dutch translation (Utrechtse Burnout Schaal, UBOS) (Schaufeli & Van Dierendonck, 2000). In the process of adapting these translations of the MBI, validity and reliability studies each time have resulted in the omission of some items from the original MBI. The items removed differ for each version: MBI-VL has 16 items and the corresponding version of the UBOS contains 20 items (see *infra*). In addition to both adaptations, the Flemish translation of the complete MBI is also available (Vlerick, 1995). As a result, three different versions of the same scale are available for measuring burnout in the Dutch-speaking regions. For those interested in measuring professional burnout, this is of course a confusing situation. Since the number of indicators for each subscale differs among the three versions, the summated scores are computed differently in each version. In some extreme cases, the three versions might contradict one another in their assessment of the severity of

¹ It was only more recently that researchers started studying how burnout can occur in other professions (see Schaufeli & Van Dierendonck, 1994; Maslach, Jackson & Leiter, 1996).

² Another often used instrument is the Burnout Measure (see Schaufeli & Enzmann, 1998, pp. 48-50).

burnout in an individual subject. The question remains: which of the three available versions is the best for measuring burnout in Flanders?

In this paper, we aim partially to answer this question by evaluating the *factorial validity* of these available versions based on an assessment of a substantial sample of special educators in Flanders (n = 995). In particular, confirmatory factor analysis (Jöreskog & Sörbom, 1999) is used to study the factorial validity for each of the three versions of the MBI scale. Good factorial validity is essential for a multi-dimensional scale like the MBI, since the summated scores on the three subscales (and not the individual items) are used as the set of indicators for professional burnout. In this paper we study the three available versions of the MBI, report on how well they fit our data, and evaluate which instrument fits best with this data.

Three versions of the MBI

The original Maslach Burnout Inventory (Maslach & Jackson, 1986; Maslach, Jackson & Leiter, 1996) consists of 22 items and it describes burnout as a three-dimensional syndrome characterised by emotional exhaustion (9 items), depersonalisation (5 items) and reduced personal accomplishment (8 items) (see Table 1 for some sample items).

Table 1.

Sample Items from the Maslach Burnout Inventory (source: Maslach & Jackson, 1982, p. 229).

<ul style="list-style-type: none">• Emotional Exhaustion subscale:<ul style="list-style-type: none">◦ “I feel emotionally drained from my work”◦ “Working with people all day is really a strain for me”• Depersonalisation subscale:<ul style="list-style-type: none">◦ “I’ve become more callous toward people since I took this job”◦ “I worry that this job is hardening me emotionally”• Personal Accomplishment subscale:<ul style="list-style-type: none">◦ “In my work, I deal with emotional problems very calmly”◦ “I feel I’m positively influencing other people’s lives through my work”
--

All items measure frequencies and are scored on a seven-point Likert-scale with fixed anchors that range from “never” to “every day”. Curiously, the three dimensions of the scale have not been deduced theoretically, but were labelled after a (explorative) factor analysis of an initial set of 47 variables (Maslach & Jackson, 1981). The current (third) edition of this questionnaire (Maslach, Jackson & Leiter, 1996) contains three versions of the MBI: one for human services professions (MBI-HSS) (i.e., professions where there is both frequent and intensive contact with people, e.g., in welfare, social service and (mental) health professions), one for teachers (MBI-ES), and a general version for non-social service workers (MBI-GS). In this

third edition, the MBI-HSS corresponds to the former MBI (Maslach & Jackson, 1986); the two other versions are new. In view of the subjects studied in our sample, we only take into consideration the MBI-HSS.

In line with the current version of the MBI, the Dutch adaptation (UBOS) consists of three versions: one for human services workers (UBOS-C), one for teachers (UBOS-L), and a general version for all professions (UBOS-A) (Schaufeli & Van Dierendonck, 2000). Before the publication of the UBOS-C (Schaufeli & Van Dierendonck, 2000), this scale was already known as the MBI-NL (Schaufeli & Van Dierendonck, 1994). After studying the construct validity of the MBI in several (large) samples of Dutch human services professionals, it was decided to remove two items from the original set of 22 items (1 from the emotional exhaustion subscale (item 16) and one item from the personal accomplishment subscale (item 12). Therefore, the UBOS-C has two items less than the original MBI.

The Flemish adaptation (MBI-VL) is based on the second edition of the MBI and consequently only exists in one version (only applicable for human services professions) (Vlerick, 1993, 1995). Compared to the Dutch version (Schaufeli & Van Dierendonck, 2000), 8 items are translated slightly differently in the Flemish version to better suit the common parlance of the Flemish-speaking. Again, after studying the construct validity of the scale in several samples of Flemish nurses, several items were removed. This time, 6 items were omitted: 3 items from the emotional exhaustion subscale (items 1, 2, 14), 1 from the depersonalisation subscale (item 5), and 2 from the personal accomplishment subscale (items 12, 19) (Vlerick, 1993, 1995). Importantly, of the 6 items removed from the MBI-VL, only 1 overlaps (i.e., item 12) with the 2 items removed from the UBOS-C.

It is not clear why a different set of items had to be removed from the UBOS-C and the MBI-VL. Perhaps the slightly different translation of some of the items results in some subtle differences. Perhaps the differences are related to cultural differences, or differences between the professions of the samples involved. In addition, different selection criteria have been used by the Dutch and Flemish researchers to remove potentially weak or conflicting items. In Schaufeli & Van Dierendonck (1994), items were removed if they loaded significantly on more than one factor. In Vlerick (1995), an additional criterion was used: items were removed if they had a standardised factor loading below .40 (in each of the three sub-samples used). While in the former study only divergent validity criteria were used, both divergent and convergent validity criteria have been used in the latter as a methodology for improving the factorial validity of the MBI scale.

Method

Samples and procedure

A total of 1317 questionnaires were sent to a representative sample of residential special educators working in Flanders, Belgium. 995 questionnaires were returned to us (response rate: 75.6%) through letterboxes we installed in all participating institutions ($n = 47$). Two sectors of employment were studied: special youth care ($n = 241$) and mentally handicapped care ($n = 754$) – in each sector approximately 10% of the total population of special educators was studied (Vlaamse Overheid, 2003). The subgroup of educators working in residential special youth care had been working in the sector for 9.8 years on average ($SD = 7.9$) (average age of 33.2, $SD = 8.6$) and was composed of 70% women. The sample of educators working in residential mental handicap care worked in the sector for 12.2 years on average ($SD = 7.7$) (average age of 34.4, $SD = 8$) and was composed of 72% women.

Measures

The participants responded anonymously to a questionnaire, containing items on demographic variables and professional burnout. Burnout was assessed by making use of the Flemish translation of the Maslach Burnout Inventory items (Vlerick, 1993, 1995). All 22 items were used in the questionnaire.

Data analysis

Many studies have addressed the factorial structure of the MBI using confirmatory factor analysis (CFA) (e.g., Bakker, Demerouti & Schaufeli, 2002; Beckstead, 2002; Taris, Schreurs & Schaufeli, 1999). This research usually favoured a three-dimensional structure with factors corresponding to the original factors as proposed by Maslach & Jackson (1986). Another recurring finding is that only models in which the factors are allowed to correlate fit the data well (notice that in the original study [Maslach & Jackson, 1981], the factors were thought of as perpendicular). Since it is not the aim of this paper to discuss the dimensionality of the MBI, we will only fit the standard correlated three-factor model to our data.

Based on the translations of the original 22 items, we test the item-factor structure of the three available versions of the MBI. All versions consist of the same three subscales. Only the selection of items differs. To distinguish between the three versions, we will use the following notation: MBI-VL22 corresponds to the original, complete 22-item scale, MBI-VL20 uses the

same items as the UBOS-C, and MBI-VL16 uses the 16 items selected by Vlerick (1993). The only difference between the three models (MBI-VL22, MBI-VL20 and MB-VL16) lies in the number of items used as indicators for the subscales.

By making use of LISREL (Jöreskog & Sörbom, 1999) the models are confirmatively fitted to two different samples: (1) the sector of special youth care (effective sample size: $n=230$), and (2) the sector of mentally handicapped care (effective sample size: $n=694$). The two sectors have been considered as different samples, since preceding explorative Chi2-tests had indicated that both sectors indeed differed significantly ($p<.001$) with respect to sum-scores on the 3 MBI subscales for the 3 models (MBI-VL22, MBI-VL20 and MB-VL16). The analyses were based on the variance-covariance matrix. Maximum likelihood was used to estimate the model parameters. Model fit is assessed using the standard Chi2-test, as well as the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), and the Root Mean Squared Error of Approximation (RMSEA). The GFI is an index of absolute fit that indicates how much better a model fits as opposed to no model at all. The AGFI is based on the GFI, but is adjusted for model complexity (as reflected by the degrees of freedom). For both indices, a cut-off value of .90 or higher is considered to indicate a good fit (Jöreskog & Sörbom, 1999). The RMSEA is a badness-of-fit measure of the error approximation in the population that indicates the discrepancy per degree of freedom. We used a cut-off value of 0.05 or lower to indicate a good model fit (Browne & Cudeck, 1993), although Hu & Bentler (1999) suggested a slightly less conservative value of .06. Note that we only fitted 'standard' models. This, for example, means that no correlations between error terms were allowed and that no other undocumented manipulations that may help to improve the fit of the model were applied.

Results

The fits of the models are summarised in Tables 2 and 3 for the special youth care group and the mentally handicapped care group respectively.

Table 2.
Model fits for the sector of special youth care.

Models	Df	Chi2	P	GFI	AGFI	RMSEA
MBI-VL22	206	398	<.001	.86	.83	.065
MBI-VL20	167	287	<.001	.88	.85	.059
MBI-VL16	101	150	<.001	.92	.90	.047

Table 3.
Model fits for the sector of mentally handicapped care.

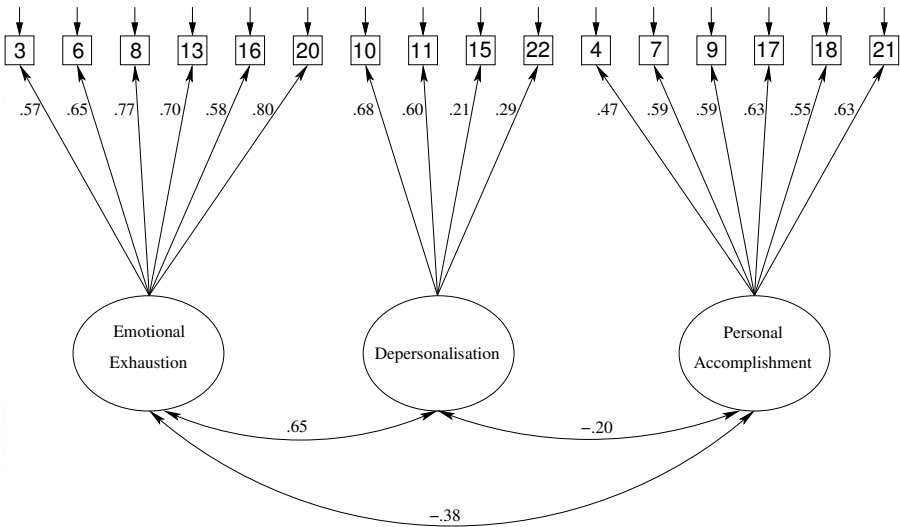
Models	Df	Chi2	P	GFI	AGFI	RMSEA
MBI-VL22	206	883	<.001	.89	.87	.071
MBI-VL20	167	583	<.001	.92	.90	.062

First of all, it can be concluded that all versions of the MBI fit the data fairly well (see Table 2 & 3). This is not surprising, given the enormous support for the correlated three-factor model in previous research. However, there are some subtle differences between the three versions, and the fits of the models are by no means excellent. In both samples, the fit of the MBI-VL22 is worse than the other two. This suggests that removing some items benefited the fit of the model. Unfortunately, it is not obvious which version fits best with the structure of the data, the MBI-VL20 or the MBI-VL16. In the smaller sample, the MBI-VL16 seems to have some advantage over the MBI-VL20 (GFI is respectively .92 and .88 and RMSEA respectively .047 and .059). However, in the larger sample, the difference is barely noticeable and rather inverse: the GFI is quite similar (MBI-VL-16: .93 and MBI-VL20: .92) and the RMSEA measure for the MBI-VL16 is slightly higher (.070) than for the MBI-VL20 (.062). The RMSEA measure only drops under the recommended level of .05 for the MBI-VL16 model in the smaller sample.

A graphical representation of the best fitting model (MBI-VL16 in the sector of special youth care sample) together with the standardised coefficients is depicted in Figure 1. Even in this best fitting model, some items still appear to be problematic. Significant cross-loadings were not observed for this model (although items 10 and 4 come close due to high loadings on the Emotional Exhaustion factor). Inspection of the fit statistics indicated some degree of model misfit. A review of the modification indices suggests significant correlations between the error terms of item 6 and item 16 (.18), and between the error terms of item 4 and item 7 (.19). Similar problematic patterns (although not always involving the same items) were observed in all other models.

Figure 1.

Graphical representation of the MBI-VL16 for the special youth care sample.



Squares indicate the 22 items on the MBI, circles represent the 3 latent factors associated with the subscales. Standardised parameter estimates (factor loadings and correlations among factors) are indicated.

Discussion

This study was designed to compare and evaluate the three available versions of the Maslach Burnout Inventory, based on a test of the factorial validity using confirmatory factor analysis in a sample of Flemish special educators. We found that all three models fit the data fairly well and that the 3-factorial structure of the MBI was confirmed each time. In both samples, the MBI-VL22 had the lowest fit. The best fit was for the MBI-VL16, although the advantage of the MBI-VL16 over the MBI-VL20 was not clear-cut.

Perhaps the most disturbing finding was that none of the three versions of the MBI scale fit the data really well. There was always some degree of misfit, as illustrated by the fact that only in one case (MBI-VL16 in the smaller sample) did the RMSEA drop below the recommended .05 level. It is possible of course to improve the fit of the models by accepting correlated error terms, a method used, for example, by Beckstead (2002) and Byrne (1991). However, we do not agree with this solution since: (1) there are no theoretical reasons why correlated error terms should be included in the models, and

(2) including correlations between error terms corresponding to items from different subscales undermines the factorial structure of the scale. In our view, the only alternative solution is further to remove conflicting or weak items from the MBI. If one were to follow the criteria used in Vlerick (1993, 1995) on our dataset (e.g., standardised factor loadings should exceed 0.40 in both samples), four more items should be removed from the MBI-VL16. A more drastic solution is to use the criteria used by Kalliath et al. (2000). They examined the fit of the complete 22-item MBI scale and removed all items with low reliability (as measured by the squared multiple correlation between the item and the latent factor; .40 was taken as the minimal level). Using their approach, only seven items were retained.

Regardless of the particular criteria used, the consequence of adopting this 'remove problematic items' approach is that one would inevitably end up with a fourth version of the MBI for the Dutch-speaking regions. We think this is a step that will have to be made, but also suggest that this kind of test would best be performed on a large sample, representing several people-oriented professions (not only special educators, such as in our case) from different sectors (not only mentally handicapped care and special youth care), so that we can end up with a consensus version of the MBI for the Flemish region. While doing so, special attention should be paid to correlations between the factors. We observed substantial correlations between the emotional exhaustion factor and the depersonalisation factor (see Figure 1: correlation = .65). Correlations of this magnitude between both factors are not uncommon (see Schaufeli & Enzmann, 1998), but nevertheless these alert us that the added value of three-factor models over two-factor models should explicitly be put to the test in future research.

References

- Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2002). Validation of the Maslach Burnout Inventory – General Survey: an internet study. *Anxiety, Stress and Coping, 15*, 245-260.
- Beckstead, J. W. (2002). Confirmatory factor analysis of the Maslach Burnout Inventory among Florida nurses. *International Journal of Nursing Studies, 39*, 785-792.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Thousand Oaks, CA: Sage.
- Burisch, M. (1993). In search of theory: some ruminations on the nature and etiology of burnout. In: W. Schaufeli, C. Maslach & T. Marek (Eds.), *Professional Burnout – Recent Developments in Theory and Research* (pp. 75-94). Washington: Taylor & Francis.
- Byrne, B. M. (1991). The Maslach Burnout Inventory: Validating factorial structure and invariance across intermediate, secondary and university educators. *Multivariate Behavioral Research, 26*, 583-605.

- Freudenberger, H. J. (1974). Staff Burnout. *Journal of social Issues*, 30, 159-165.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-31.
- Jöreskog, K. G. & Sörbom, D. (1999). *Lisrel 8.3: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software Inc.
- Kalliath, T., Gillespie, D., O'Driscoll, M., & Bluedorn, A. (2000). A test of the Maslach Burnout Inventory in three samples of healthcare professionals. *Work and Stress*, 14, 35-50.
- Maslach, C. (1976). Burned-out. *Human Behaviour*, 5, 16-22.
- Maslach, C. (1993). Burnout: A multidimensional perspective. In: W. Schaufeli, C. Maslach & T. Marek (Eds.), *Professional Burnout – Recent Developments in Theory and Research* (pp. 19-32). Washington: Taylor & Francis.
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, 2, 99-113.
- Maslach, C., & Jackson, S. E. (1982). Burnout in Health Professions: A Social Psychological Analysis. In: G. S. Sanders & J. Suls (Eds.), *Social Psychology of Health and Illness* (pp. 227-252). Hillsdale: Lawrence Erlbaum.
- Maslach, C., & Jackson, S. E. (1986). *The Maslach Burnout Inventory. Manual*. Second edition. Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *The Maslach Burnout Inventory. Manual*. Third edition. Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job Burnout. *Annual Review of Psychology*, 52, 397-422.
- Schaufeli, W. B., & Enzmann, D. (1998). *The Burnout Companion to Study and Practice*. London: Taylor & Francis.
- Schaufeli, W. B., & Van Dierendonck, A. (1994). Burnout, een gemeten begrip. *Gedrag & Gezondheid*, 22, 153-173.
- Schaufeli, W. B., & Van Dierendonck, A. (2000). *UBOS, Utrechtse Burnout Schaal, Handleiding*. Lisse: Swets & Zeitlinger.
- Taris, T. W., Schreurs, P. J. G., & Schaufeli, W. B. (1999). Construct validity of the Maslach Burnout Inventory – General Survey: a two-sample examination of its factor structure and correlates. *Work & Stress*, 13, 223-237.
- Vlaamse Overheid (2003). *Welzijn & Volksgezondheid in Cijfers 2003*. Brussel: Ministerie van de Vlaamse Gemeenschap. Departement Welzijn, Volksgezondheid en Cultuur.
- Vlerick, P. (1993). Exploratieve analyse van de factoriële validiteit van de Maslach Burnout Inventory. *Tijdschrift Klinische Psychologie*, 23, 235-250.
- Vlerick, P. (1995). Confirmatorische analyse van de factorstructuur en de validiteit van de Maslach Burnout Inventory. *Gedrag & Gezondheid*, 32, 35-40.

Received December, 2003

Revision Received January, 2005

Accepted February, 2005