Journal of Occupational and Organizational Psychology (2000), **73**, 53–66 *Printed in Great Britain*

The factorial validity of the Maslach Burnout Inventory–General Survey (MBI–GS) across occupational groups and nations

Nico Schutte

Utrecht University, The Netherlands

Salla Toppinen and **Raija Kalimo**

Finnish Institute of Occupational Health, Helsinki, Finland

Wilmar Schaufeli*

Utrecht University, The Netherlands

The factorial validity of the Maslach Burnout Inventory–General Survey (MBI–GS) was investigated among employees of a multinational company in the forest industry The present study includes data from Finnish, Swedish and Dutch employees (total $N = 9055$). The hypothesized three-factor model of the MBI–GS (Exhaustion, Cynicism, Professional Efficacy) was replicated across occupational groups (i.e managers, clerks, foremen technicians, blue-collar workers) and nations The fit of this model to the data was superior to alternative one- and two-factor models in all samples under investigation. In addition, the three-factor structure of the MBI–GS proved invariant across all occupational groups The internal consistencies of the three subscales are satisfactory, except for the cynicism scale in some subsamples Therefore, it is suggested to exclude one—ambiguous—cynicism item. Finally, some differences in levels of burnout are found between nations and occupational groups that are consistent with earlier findings.

Soon after its introduction in the early 1980s (Maslach & Jackson, 1981), the Maslach Burnout Inventory (MBI) became the almost universally accepted 'gold standard' to assess burnout. This popular psychological phenomenon was defined by the test authors as '. . . a syndrome of emotional exhaustion, depersonalisation, and reduced personal accomplishment that can occur among individuals who do "people work" of some kind' (Maslach & Jackson, 1986, p. 1). Emotional exhaustion is the key aspect of the syndrome and refers to feelings of being over-extended and drained from one's emotional resources. Depersonalization refers to negative, cynical, detached, and impersonal attitudes and feelings towards

*Requests for reprints should be addressed to Wilmar B Schaufeli, Utrecht University, Department of Psychology, P.O Box 80.140, 3508 TC Utrecht, The Netherlands (e-mail W SCHAUFELI@FSS UU NL)

other people. Finally, reduced personal accomplishment refers to a decline in one's feelings of competence and to a tendency to evaluate oneself negatively in particular regarding one's work with other people.

From the outset it was assumed by most influential authors that the specific emotional demands that are associated with dealing with difficult, troubled, suffering, or unmotivated recipients lie at the core of the syndrome (e.g. Cherniss, 1980; Freudenberger, 1974; Maslach, 1982). Therefore, they reasoned that burnout develops particularly among those professionals who interact intensively and frequently with such recipients, like nurses, physicians, teachers, lawyers, child-care workers, counsellors, probation officers, social-workers, and prison officers. Thus, traditionally, the burnout concept and its measurement are closely linked to the human services where professionals do 'people work' (Maslach & Schaufeli, 1993).

However, attempts were also made to measure burnout in occupational groups outside the human services (e.g. Goldembiewski & Munzenrider, 1988; Pines & Aronson, 1988). It turned out however that the original MBI could not be applied indiscriminately outside the human services, the setting for which it was originally designed (Maslach & Jackson, 1986). For instance, Golbembiewski and Munzenrider (1988) substituted 'co-worker' for 'recipient' in various MBI items. They were severely criticized by Garden (1987), who argued that this substitution wrongly assumes the equivalent of two relationships: that of the service provider with his or her recipient, and that of the worker with other co-workers.

Recently, the apparent need for a scale that measures burnout in occupational groups outside the human services was met by the introduction of the Maslach Burnout Inventory–General Survey (MBI–GS) (Schaufeli, Leiter, Maslach, & Jackson, 1996). The MBI–GS comprises three subscales that parallel those of the original MBI: Exhaustion (Ex), Cynicism (Cy) and Professional Efficacy (PE). The exhaustion component in the MBI–GS is generic; i.e. without direct reference to people as the source of those feelings as in the MBI-emotional exhaustion subscale. In a similar vein, cynicism reflects indifference or a distant attitude towards one's work in general, rather than towards recipients of one's services. Finally, professional efficacy covers a broader scope than personal accomplishment because social *and* non-social aspects of occupational accomplishments are included. Essentially, it assesses the employee's expectations of continued effectiveness at work. In a manner similar to the original MBI, a high degree of burnout is reflected in high Ex and Cy scores and in low PE scores. At the other end of the continuum, 'engagement' is being defined as an energetic state in which the employee is dedicated to excellent performance at work and is confident of his or her effectiveness. Accordingly, the MBI–GS can be used in non-human services settings to investigate levels and correlates of 'burnout'—a mental condition that is similar but not identical to the classical definition of the syndrome (see Maslach & Jackson, 1986, p. 1).

The main purpose of this study is to investigate the factorial validity of the MBI–GS. Preliminary confirmatory factor analysis using linear structural modelling revealed that the hypothesized three-factor solution fitted well to the data of a Dutch, Canadian, and Finnish sample (Schaufeli, Leiter, & Kalimo, 1995).

However, these analyses were part of the process of test construction; they were used for identifying and eliminating unsound items. The current study attempts to investigate the factorial validity of the final version of the MBI–GS (Schaufeli *et al.*, 1996) by contrasting it with three alternative models: (1) a one-factor model that agrees with the one-dimensional view on burnout (Pines & Aronson, 1988; Shirom, 1989); (2) a two-factor model that includes the so-called 'core of burnout' factor (Ex and Cy) along with a separate PE factor (Green, Walkey, & Taylor, 1991); (3) a three-factor model with correlated Ex and Cy factors and an independent PE factor, as it becomes increasingly clear from studies with the original MBI that personal accomplishment develops largely independent from the other two burnout dimensions (Cordes & Dougherty, 1993; Lee & Ashforth, 1996).

The second aim of the present study is to assess the factorial validity of the MBI–GS across occupational groups and countries. Recently, a Canadian study showed that the hypothesized three-factor structure of the MBI–GS, with Ex, Cy, and PE as independent but correlated factors, fitted to the data of various occupational groups in a tertiary hospital and a psychiatric facility (Leiter & Schaufeli, 1996). The present study goes beyond the traditional occupational field in which burnout has been investigated—the human services—and examines the factor structure of the MBI–GS in a multinational corporate sample, including—for the first time—blue-collar workers. As noted before, the factor structure of the MBI–GS has been studied in samples from three different nations, but this was done during its construction. So, for the first time, the cross-national validity of the *final* version of the MBI–GS is tested. Following Hofstede (1980), we used samples from the same company in different countries (Finland, Sweden, and The Netherlands) in order to minimize the likelihood that all kinds of idiosyncracies would influence the results.

Method

Participants and procedure

Participants were employees of a multinational corporation in the forest industry whose top management and main production facilities are in Finland The current study includes all employees from Finland and one production unit from Sweden and one from The Netherlands. The latter two units employ practically every employee of the company in these nations. Further, the core activities of the two units are similar to that of the main facilities in Finland. Overall, 10 383 questionnaires were returned (Finland, $N = 9607$; Sweden, $N = 446$; The Netherlands, $N = 330$), which corresponds with a response rate of 63% (Finland 62%, Sweden 87%, The Netherlands 59%) Listwise deletion of missing data yielded a final sample size of 9055 employees (Finland, $N = 8529$; Sweden, $N = 267$, The Netherlands $N = 259$).

Gender ($\chi^2(2) = 45.55$, $p < .001$), age ($\chi^2(6) = 148.74$, $p < .001$), and occupational group ($\chi^2(2) = 11\ 8$, $p < .005$), were differently distributed across subsamples. The Dutch sample was almost exclusively male and relatively young (i.e. almost 70% was less than 40 years of age), whereas relatively more female and older employees were included in both Scandinavian samples (i.e. almost 70% of the Finnish sample was aged over 40).

Questionnaires were sent to the business units in envelopes bearing the employee's name. Respondents could either use a pre-stamped return envelope or take the questionnaire—anonymously—to a central point from where they were collectively mailed. The questionnaire was

accompanied by a covering letter that explained the purpose of the study, emphasized voluntary participation, and guaranteed absolute confidentiality.

Measure

The MBI–GS was used to assess burnout (Schaufeli *et al.*, 1996). The MBI–GS includes three subscales: Exhaustion (Ex) (five items, e g. 'I feel used up at the end of the workday'), Cynicism (Cy) (five items, e.g. 'I have become less enthusiastic about my work') and Professional Efficacy (PE) (six items, e g. 'In my opinion, I am good at my job'). All items are scored on a 7-point frequency rating scale ranging from '0' (never) to '6' (daily). High scores on Ex and Cy, and low scores on PE are indicative for burnout (i.e. PE-items are reversibly scored). The MBI–GS was simultaneously developed in The Netherlands, Canada, and Finland (Schaufeli *et al.*, 1995) The items were initially formulated in Dutch and English and checked by a bilingual psychologist Next, the items were translated from English into Finnish and from Finnish into Swedish and then checked by bilingual psychologists The English version was also used for checking the Swedish translation

Statistical analysis

Three preliminary analyses were carried out. First, the frequency distribution of the items of the MBI–GS was studied in order to assess deviations from normality. Following West, Finch, and Curran (1995) it was assumed that the χ^2 goodness-of-fit statistics are not likely to be inflated if the skew and kurtosis for individual items does not exceed the critical values of 2.0 and 7 0, respectively. Next, the reliability of the subscales of the MBI–GS was assessed using Cronbach's α as an indicator of internal consistency. As a rule of thumb, a value for α of 70 is considered to be sufficient (Nunnaly, 1987). Finally, mean levels of burnout were compared across occupational groups and nations. Subsequently, the structural equation-modelling program implemented on LISREL version 8.12A (Joreskog & Sorbom, 1993), was used to assess the factor structure of the MBI–GS, using the maximum likelihood method. In response to contradictory claims made in the literature concerning the structure of burnout (see above), five increasingly less restricted models were fitted on the total sample ($N = 9055$) The null model (M0), in which all constructs were hypothesized to be uncorrelated and measured without error, served as a basis for model comparison with five alternative models Since the χ^2 goodness-of-fit indices vary with sample size (i e. in large samples the χ^2 statistic tends to be significant so that the null hypothesis that the model fits to the data is virtually always rejected) The Normed Fit Index (NFI) and the Non-Normed Fit Index (NNFI) are used to assess global model fit. The NFI represents the point at which the model being evaluated falls on a scale running from a null model to perfect fit. This index is normed to fall on a 0 to 1 continuum The NNFI could fall outside of the 0 to 1 range due to sampling fluctuations It is suggested by Marsh, Balla, and Hau (1996) that these indices are relatively insensitive to sample sizes. In addition, the NNFI takes model parsimony into account. It is more or less generally accepted that a value less than 90 for NFI and NNFI indicates that the fit of the model can be improved (Hoyle, 1995)

In addition, the Adjusted Goodness-of-Fit Index (AGFI), the Comparative Fit Index (CFI) and the Parsimony Normed Fit Index (PNFI) are given. The AGFI is a measure of the relative amount of variance accounted for by the model, corrected for the degrees of freedom in the model relative to the number of variables. Since the distribution of the AGFI is unknown, no statistical test or critical value is available (Joreskog & Sorbom, 1986). The CFI is the normed version of the RNI (Relative Noncentrality Index) The CFI is not systematically related to sample size and it is a pure measure of model fit in the sense that model parsimony is not taken into account (Marsh *et al.*, 1996) The PNFI explicitly penalizes model complexity, following the rationale that—all other things being equal—a good fit attained by a parsimonious model (with few unknowns and thus many degrees of freedom) is preferable to a complex model. It is obtained by multiplying the NFI by the ratio of the degrees of freedom of the model under consideration to the degrees of freedom of the null model

A one-factor model (M1) was assessed in which all items were hypothesized to load on a single factor. In the two-factor orthogonal model (M2a), a 'core of burnout' dimension was assumed

consisting of a joint Ex and Cy factor along with the PE factor. In the two-factor oblique model (M2b) both factors were allowed to correlate. Finally, two three-factor models were fitted: (1) a PE-independent model (M3a) in which Cy and Ex were allowed to correlate and PE was constrained to be independent, (2) an oblique model with Ex, Cy and PE items loading on their respective factors (M3b) The fit of these models was tested in the total sample as well as in each national sample separately.

Results

Frequency distribution

It was found that two items are—slightly—skewed: P1 (− 2.01) and P6 (− 2.21) and only item E1 has a peaked distribution, its kurtosis being 8.13. Thus, with these very few exceptions the items of the MBI–GS are normally distributed. It is therefore unlikely that the χ^2 goodness-of-fit statistic is inflated.

Internal consistency

Table 1 gives an overview of the mean values, standard deviations, and internal consistencies (Cronbach's α) for the three subscales of the MBI–GS for the total sample, the three national samples and two occupational groups. White-collar workers include managers, clerks (predominantly women), foremen, technical designers and laboratory personnel. The Ex and PE subscales are sufficiently internally consistent since Cronbach's α meets the criterion of .70 in virtually every (sub)sample. Only in the Dutch sample PE shows a value slightly below that criterion. As a rule, Ex is more internally consistent than PE. Item C3 (i.e. 'I just want to do my job and not be bothered') was found to fit poorly in the Cy subscale since α increased in all samples after this item was detected. Therefore, this item was excluded and a 4-item Cy subscale was used in the remaining analyses.

Factorial validity: total sample

Table 2 shows that the fit in the total sample of each of the less restrictive factor models was superior to that of the null model (M0). In addition, each of the subsequent nested factor models shows an improved fit over the prior more restrictive model as is indicated by the monotonously increasing values for NFI, NNFI, CFI and AGFI. In accordance with the assumption that burnout is a syndrome (i.e. a cluster of related symptoms) it was found that M2b fits better to the data than M2a and that the correlations between factors are significant ranging from |.35| to |.67|. When considering the (N)NFI, the fit of the three-factor correlated model (M3b), remains slightly below .90 which is generally considered to be the minimum value for an acceptable fit (Hoyle, 1995). Based on the so-called modification indices provided by LISREL and on theoretical considerations, model M3b was respecified. In this respecified model (M3c), errors of three item pairs (i.e. C1–C2, C4–C5, and P4–P5) were allowed to correlate. M3c shows an excellent fit for the data with NFI and NNFI values of .95 and .94, respectively. Since the model fit was satisfactory and the results agreed with the theoretical assumptions

Table 1. Means, standard deviations, and internal consistencies (Cronbach's α) of MBI–GS subscales

Sample	Exhaustion (5 items)			Professional efficacy (6 items)			Cynicism (5 items)			Cynicism (4 items)[a]		
	M	SD	α	*M*	SD	α	*M*	SD	α	*M*	SD	α
Total (*N*=9055)	1.48	1.41	0.86	4.66	1.69	0.83	1.48	1.69	0.75	1.21	1.53	0.82
Total White-C. (*N*=3378)	1.53	1.36	0.86	4.82	1.46	0.83	1.31	1.56	0.78	1.13	1.44	0.84
Total Blue-C. (*N*=5677)	1.46	1.43	0.86	4.55	1.80	0.82	1.58	1.75	0.74	1.25	1.58	0.81
Finland (*N*=8529)	1.51	1.40	0.86	4.65	1.68	0.83	1.49	1.68	0.76	1.22	1.52	0.83
Sweden (*N*=267)	1.45	1.55	0.76	4.78	1.74	0.70	1.42	1.86	0.65	1.15	1.69	0.66
Netherlands (*N*=259)	0.86	1.29	0.76	4.87	1.61	0.68	1.11	1.92	0.62	0.75	1.38	0.75

[a]Item C3 is excluded

Note White-C=White-collar workers, Blue-C =blue-collar workers

Table 2. Fit statistics of LISREL models[a]

Model	χ^2	d.f.	NFI	NNFI	CFI	AGFI	PNFI
M0	83 151.56	105					
M1	38 313.49	90	0.54	0.46	0.54	0.38	0.46
M2a	17 047.86	90	0.73	0.69	0.75	0.72	0.63
M2b	15 852.18	89	0.75	0.71	0.77	0.73	0.64
M3a	11 536.35	89	0.86	0.84	0.86	0.79	0.73
M3b	9 565.23	87	0.88	0.86	0.89	0.81	0.73
M3c	4 361.28	84	0.95	0.94	0.95	0.91	0.76

[a]Total sample; N=9055

Note. M0=null model; M1=one-dimensional model, M2a=two-dimensional uncorrelated model with (Ex+Cy) and PE, M2b=two-dimensional correlated model with (Ex+Cy) and PE; M3a=three-dimensional model with Ex and Cy correlated, M3b=three-dimensional correlated model with Ex, Cy and PE; M3c=respecified M3b (for explanation see text) All χ^2 values and all reported decrements in χ^2 for the increasingly less constrained models are significant at $p< 01$

underlying the structure of the MBI–GS no further modifications of the model were deemed necessary. Figure 1 displays a graphical representation of the three-factor respecified model (M3c). As can be seen, Ex and Cy show the highest correlation, followed by Cy and PE, and Ex and PE, respectively. The estimated factor loadings are fairly high ranging from .63 to .87.

Factorial validity: occupational groups

In order to test the invariability of the three-factor (respecified) model across occupational groups, models were fitted independently in five groups of Finnish employees (i.e. managers, clerks, foremen, technical designers/laboratory personnel, and blue-collar workers). Table 3 shows the fit statistics of M3c in each of the professional groups. Only among technical designers and laboratory personnel (group 4), three additional pairs of error-terms were allowed to correlate: E1–E2, E4–E5 and P1–P2, these correlating error-terms referred to items within the same subscale. When we consider the three-factor respecified model, the NNFI ranged between .90 and .94 for the five subsamples which indicates good global model fit.

Factorial validity: national subsamples

In the next step the various models were tested in the three national samples. Table 4 shows the fit statistics of M3c for Finland, Sweden and The Netherlands. As a logical consequence of its dominant contribution to the total sample, similar model fits are found in the Finnish subsample. However, overall model fit was not as good in the Swedish and the Dutch subsamples as is expressed by lower fit indices. Nevertheless, in both samples a similar pattern of monotonously increasing NFI, NNFI, CFI and AGFI values is observed as in the total sample and in the Finnish subsample. The fit of M3c in both subsamples could still be further improved as is indicated by (N)NFI values that are lower than .90. Inspection of

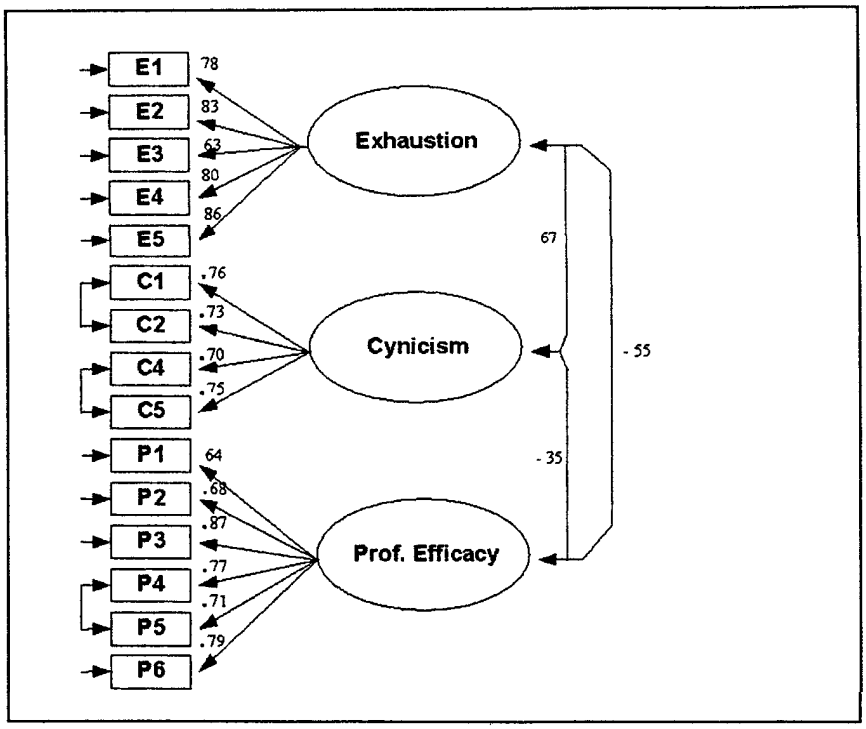

Figure 1. The three-factor re-specified model.

the modification indices suggested that in the Swedish sample five additional pairs of error-terms should be allowed to correlate (i.e. E3–E4, C1–C4, C2–C4, P1–P2 and P2–P5) and in the Dutch sample six—different—item pairs (i.e. E1–E2, E1–E5, E2–E5, C1–C2, C1–C5 and P4–P5) were released. Again all correlations refer to error-terms within the same subscale.

After these modifications the overall model fit as measured by the NFI and the NNFI remained only slightly below the criterion of .90 with a (N)NFI of .85 for Sweden and .86 for The Netherlands. For the same model CFI scores of .89 and .90 were found for Sweden and The Netherlands, respectively.

Further inspection of the modification indices revealed that the model fit could be enhanced by allowing some items to load on more than one factor simultaneously. Since such modifications could not be theoretically accounted for within our model, these constraints were maintained, and no further modifications were attempted.

When considering the factor loadings in the national samples, the Finnish sample shows relatively high loadings, which are evenly distributed; ranging from .64 to .86. Loadings for The Netherlands are on an average somewhat lower and the variability is larger with scores ranging from .48 to .89. This effect is seen even more clearly for the Swedish sample where loadings are found in the range .37 to .89.

Table 3. Fit statistics SEM models, across occupational groups in Finnish sample

Model	χ^2	d.f.	NFI	NNFI	CFI	AGFI	PNFI
Group 1[a] (*N*=859)							
M0	7 779.35	105					
M3c	474.11	84	0.94	0.94	0.95	0.89	0.75
Group 2[b] (*N*=877)							
M0	9 448.53	105					
M3c	612.00	84	0.94	0.93	0.94	0.75	0.75
Group 3[c] (*N*=1242)							
M0	11 966.20	105					
M3c	821.65	84	0.93	0.92	0.94	0.88	0.75
Group 4[d] (*N*=252)							
M0	2 660.00	105					
M3c	274.21	81	0.90	0.90	0.92	0.81	0.69
Group 5[e] (N=5294)							
M0	50 517.63	105					
M3c	2 528.67	84	0.95	0.94	0.95	0.91	0.76

[a]Managers
[b]Clerks (predominantly women)
[c]Foremen.
[d]Technical designers and laboratory personnel
[e]Blue-collar workers
Note. All χ^2 values and all reported decrements in χ^2 for the less constrained models are significant at $p< 01$

Table 4. Fit statistics of LISREL models in the Finnish, Swedish and Dutch samples

Model	χ^2	d.f.	NFI	NNFI	CFI	AGFI	PNFI
Finland (*N*=8529)							
M0	81 290.97	105					
M3c	4 047.96	84	0.95	0.94	0.95	0.91	0.76
Sweden (*N*=267)							
M0	1 879.07	105					
M3c	280.24	79	0.85	0 85	0.89	0.82	0.64
Netherlands (*N*=259)							
M0	1 920.43	105					
M3c	270.38	81	0.86	0.86	0.90	0.83	0.66

Note. All χ^2 values and all reported decrements in χ^2 for the increasingly less constrained models are significant at $p<.01$.

Subsequently the measurement models of the national samples were compared simultaneously in LISREL, the results of which are given in Table 5. For the comparison the respecified models derived in the previous section were used. When comparing the national samples, the factor loadings, correlations between the

Table 5. Simultaneous comparison SEM models, across national samples

Model	χ^2	d.f.	NFI	NNFI	CFI	GFI	PNFI
Finland[a] and The Netherlands[c]							
M0	4477.43	210					
M3c	408.77	168	0.91	0.93	0.94	0.94	0.73
Finland[a] and Sweden[b]							
M0	4376.52	210					
M3c	404.38	166	0.91	0.93	0.94	0.94	0.72
Sweden[b] and The Netherlands[c]							
M0	3799.50	210					
M3c	562.19	163	0.86	0.86	0.89	0.88	0.66

[a] N=263.
[b] N=267.
[c] N=259
Note All χ^2 values and all reported decrements in χ^2 for the increasingly less constrained models are significant at $p< 01$. AGFI is not available for simultaneous model comparison for which reason the GFI is given.

factors, error-terms, and correlation between error-terms were allowed to vary between the national samples.

In addition, a random sample of 263 observations was drawn from the Finnish data, in order not to let these dominate the analyses. As can be seen from Table 5, the (N)NFI ranged between .86 and .93, closely approaching or reaching the .90 criterion; CFI ranged between .89 and .94.

Mean levels of burnout

Significant differences in burnout scores were found between blue- and white-collar workers, with the latter scoring lower on Cy (F(1,9578) = 41.67, p<.001) and higher on PE (F(1,9431) = 140.89, p<.001). No differences were found for Ex (F(1,9999) = .93, n.s.).

Significant differences in burnout scores were found between nations for each subscale, after controlling for age, gender, and type of job (blue- or white-collar): Ex (F(2,9992) = 37.80, p<.001), Cy (F(2,9570) = 13.68, p<.001), PE (F(2,9423) = 3.29, p<.05). Subsequent analyses revealed no significant differences in mean scores between Finnish and Swedish respondents for any of the MBI–GS subscales. However, compared with Finland and Sweden, respondents from The Netherlands scored significantly lower on Ex and on Cy. Furthermore, the Dutch respondents had significantly higher scores on PE than the Finnish respondents. Thus, overall the Dutch sample had more favourable burnout levels.

Discussion

The present study confirms that burnout—as measured by the MBI–GS—is a three-dimensional concept. Like the original MBI that has been designed for use

in the human services (Maslach & Jackson, 1981, 1986), the MBI–GS consists of three distinctive—yet related ($|.35| < r < |.67|$)—dimensions (Exhaustion, Cynicism, and reduced Professional Efficacy). Confirmatory factor analysis showed that the fit of the expected three-factor model was clearly superior to alternative one-factor and two-factor models. Basically, this result was replicated in five occupational groups (managers, clerks, foremen, technical professionals, and blue-collar workers) as well as in samples from three nations (Finland, Sweden and The Netherlands).

The factorial validity of the MBI–GS was clearly demonstrated across occupational groups. The three-factor model fit quite well in each group separately, only in the group of technical professionals additional error-terms were allowed to correlate in order to improve model fit. Similar results were obtained by Leiter and Schaufeli (1996), who compared four occupational groups of tertiary care hospital (i.e. clerical and maintenance personnel, technical professionals, nurses, and managers).

Global model fit for the slightly adjusted three-factor model—in which some error-terms within subscales were allowed to correlate—was good in the Finnish sample. Fit indices for the three-factor adjusted model were somewhat lower in Sweden and The Netherlands. This is also observed in the simultaneous comparison in which Sweden and The Netherlands both have lower fit indices compared to Finland. This may be accounted for by some poorly performing items as the average factor loadings are smaller (.63 and .67, against .76). It may also reflect the fact that the samples are relatively small and that—within the limits of our theoretical models—some additional error terms were allowed to correlate. This speculation is supported by the relatively high fit indices on the CFI, a pure measure of model fit in which model parsimony is not taken into account and which is not systematically related to sample size (Marsh *et al.*, 1996).

Reliability analyses revealed that the Ex and PE subscales are sufficiently internally consistent, but that one Cy item should be removed in order to increase the internal consistency beyond the criterion of .70. This might be caused by the ambivalence of the particular item (C3; 'I just want to do my job and not be bothered'). On the one hand, a high score on this item may indicate disengagement and social isolation by closing oneself off from contacts with others at work. On the other hand, a higher score may indicate strong motivation and engagement: one concentrates on the task and does not want to be interrupted. The ambivalent nature of this item is substantiated by the fact that it has the highest standard deviation of all items (2.22) as well as the highest rate of missing values: 6.3% against an average of 3.4% for the other items. Furthermore, Schaufeli *et al.* (1995) found that this item had the lowest factor loading in three independent samples. A similar result was also obtained by Leiter and Schaufeli (1996). Accordingly, it seems that C3 is an ambivalent, hence unsound, item.

In addition to the previous psychometric analyses, differences in burnout levels were investigated. Significant differences were found in the pattern of burnout scores of white-collar and blue-collar workers, the former scoring higher on PE and

lower on Cy. It can be argued that working conditions are more favourable for managers than for workers, offering more autonomy, higher job complexity, meaningful work, and more respect from co-workers (Kalimo & Toppinen, 1999), resulting in a higher levels of PE. The relatively high scores on Cy for blue-collar workers reflect indifference and a more distant attitude towards their jobs. This might be explained by the culture on the shopfloor where distrust, resentment, and scepticism towards management and the organization traditionally prevail. In a similar vein, Leiter and Schaufeli (1996) observed that clerical and maintenance staff of a hospital scored significantly higher on Cy than did technologists, whereas management scored highest on PE.

Finally, significant differences in burnout scores across national samples were observed, with the Dutch employees scoring consistently more favourably than either their Swedish or their Finnish colleagues (after controlling for age, gender and type of job). In a similar vein, a recent study found that Dutch teachers have significantly lower levels of emotional exhaustion and depersonalization compared with Canadian teachers, after controlling for demographic variables (i.e. gender and age) and factors related to work (i.e. experience in teaching, number of hours employed, and type of school) (Van Horn, Schaufeli, Greenglass, & Burke, 1997).

As in the Hofstede's classic study at IBM (Hofstede, 1980), the current study took place within a single multinational company. Not only are all sites production plants that employ workers with similar occupational levels and duties but many managers in other nations are Finnish who communicate intensively with each other in worldwide networks. Consequently, the sites in Sweden and The Netherlands are expected to share important aspects of the corporate culture with those in Finland—the corporation's home country. The observed differences in burnout levels are therefore consistent with the hypothesis of differences between countries.

On the other hand, since for both Sweden and The Netherlands analyses were based on data from a single site, the data would equally support alternative explanations, e.g. the differences found could reflect idiosyncratic differences of the sites. The hypothesis of population differences therefore needs further investigation, using representative samples of workers from different countries (cf. Golembiewski, Munzenrider, & Luo, 1996).

In conclusion, our findings with the MBI–GS are encouraging: its factorial validity across nations and occupational groups is largely confirmed, its subscales are internally consistent and most differences in levels of burnout are in line either with previous research or theorizing. Based on the present results we argue that the MBI–GS is a suitable instrument for measuring burnout across various occupational groups and nations. Hence, the MBI–GS opens up further possibilities for burnout research outside the traditionally researched North American human services settings.

Acknowledgements

The authors wish to express their gratitude towards Pertti Mutanen and Jukka Vuori for their assistance in data analysis using linear structural equation modelling.

References

Cherniss, C. (1980). *Staff burnout· Job stress in the human services.* Beverly Hills, CA: Sage.

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

Freudenberger, H J. (1974) Staff burnout. *Journal of Social Issues, 30,* 159–165

Garden, A. M. (1987). Depersonalization· A valid dimension of burnout? *Human Relations, 40,* 545–560.

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

Golembiewski, R T., & Munzenrider, R. F (1988) *Phases of burnout: Developments in concept and applications.* New York: Preager

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

Green, D. E , Walkey, F. H , & Taylor, A. J W. (1991). The three-factor structure of the Maslach Burnout Inventory· A multicultural, multinational confirmatory study. *Journal of Social Behavior and Personality, 6,* 453–472.

Hofstede, G (1980). *Culture's consequences: International differences in work-related values* Beverly Hills, CA. Sage.

Hoyle, R. H. (1995). The Structural Equation Modeling approach: Basic concepts and fundamental issues. In R H. Hoyle (Ed.), *Structural Equation Modeling· Concepts, issues, and applications* (pp 1–15). Thousand Oaks, CA: Sage

Joreskog, K. G., & Sorbom, D. (1986) *LISREL user guide version VI* (4th ed.) Mooresville, IL: Scientific Software International

Joreskog, K G., & Sorbom, D. (1993). *LISREL version 8 12A.* Chicago, IL: Scientific Software International.

Kalimo, R., & Toppinen, S. (1999) Organizational well-being—ten years of research and development in a forest industry corporation In C. L Cooper & M Kompier (Eds.), *Improving productivity through stress prevention* (pp. 52–85). Chichester· Wiley.

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

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

Marsh, H. W., Balla, J. R., & Hau, K T (1996). An evaluation of Incremental Fit Indices A clarification of mathematical and empirical properties. In G. A Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: issues and techniques* (pp. 315–353). Mahwah, NJ: Erlbaum

Maslach, C. (1982). *Burnout The cost of caring* Englewood Cliffs, NJ: Prentice-Hall.

Maslach, C., & Jackson, S E. (1981). The measurement of experienced burnout. *Journal of Occupational Behaviour, 2,* 99–113.

Maslach, C., & Jackson, S E. (1986). *MBI: Maslach Burnout Inventory manual (research edition)* (1st ed 1981). Palo Alto, CA: Consulting Psychologists Press.

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

Maslach, C , & Schaufeli, W. B (1993). Historical and conceptual development of burnout. In W. B Schaufeli, C. Maslach, & T Marek (Eds), *Professional burnout· Recent developments in theory and research* (pp. 1–16). Washington, DC: Taylor & Francis.

Nunnaly, N. C (1987). *Psychometric theory.* New York: McGraw-Hill.

Pines, A. M., & Aronson, E (1988) *Career burnout· Causes and cures.* New York: The Free Press.

Schaufeli, W. B., Leiter, M. P., & Kalimo, R (1995, September) *The General Burnout Questionnaire. Cross-national development and validation.* Paper presented at the APA/NIOSH Congress Work, Stress and Health, Creating Healthier Workplaces, Washington, DC.

Schaufeli, W. B., Leiter, M. P , Maslach, C , & Jackson, S. E. (1996). MBI–General Survey In C. Maslach, S E. Jackson, & M. P. Leiter (Eds), *Maslach Burnout Inventory manual* (3rd ed.) Palo Alto, CA. Consulting Pyschologists Press

Shirom, A. (1989) Burnout in work organizations. In C. L. Cooper & I. Robertson (Eds), *International review of industrial and organizational psychology* (pp 25–48) Chichester: Wiley.

Van Horn, J E, Schaufeli, W. B., Greenglass, E. R., & Burke, R J. (1997). A Canadian–Dutch comparison of teachers' burnout *Psychological Reports, 81,* 371–382.

West, S G, Finch, J. F, & Curran, P. J (1995) Structural equation models with nonnormal variables: Problems and remedies In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications* (pp 65–75). Thousand Oaks, CA: Sage.

Received 2 October 1997; revised version received 22 December 1998