BURNOUT AMONG DUTCH TEACHERS
AN MBI-VALIDITY STUDY

WILMAR B. SCHAUFELI
University of Nijmegen, Netherlands

JOHN DAAMEN AND HANNEKE VAN MIERLO
Occupational Health Service, Eindhoven, Netherlands

The study investigates the internal consistency as well as the factorial and discriminant validity of the Maslach Burnout Inventory (MBI) in a sample of 326 Dutch secondary teachers Compared to other teacher samples, the internal consistency of the three subscales (Emotional Exhaustion-EEX, Depersonalization-DEP, Reduced Personal Accomplishment-PAC) is relatively high Using confirmatory maximum likelihood factor analysis with LISREL, it is shown that the three-factor oblique model fits relatively well compared to various alternative models As in some other studies, one ambiguous item was identified ("I feel very energetic") that loads on two factors simultaneously (i e , EEX and PAC) Moreover, burnout can be distinguished successfully from self-reported psychological strain as well as from somatic complaints Nevertheless, EEX is substantially related to these more general symptoms Finally, levels of EEX, psychological strain, and somatic complaints are significantly higher compared to Dutch reference groups, whereas levels of DEP and reduced PAC are significantly lower It is concluded that the MBI is a valid and reliable self-report instrument to assess burnout in a non-English-speaking country like the Netherlands

Burnout is a major problem in education. Recently, Farber (1991) concluded from a review of the U.S. literature that, "Depending on the type of community and the time of assessment, between 5% and 20% of all teachers in this country are burned out at any given time" (p. 42). In the Netherlands,

60% of the teachers who had to quit their jobs and now receive disability pensions suffer from stress-related disorders, notably burnout (Knepper, 1991) This rate is nearly twice as high as the average across all other occupations

Although different definitions of burnout exist, this long-term stress reaction is most commonly described as

> a psychological syndrome of emotional exhaustion, depersonalization, and reduced personal accomplishment that can occur among individuals who work with other people in some capacity Emotional exhaustion refers to feelings of being emotionally overextended and depleted of one's emotional resources. Depersonalization refers to a negative, callous, or excessively detached response to other people, who are usually the recipients of one's services or care Reduced personal accomplishment refers to a decline in one's feelings of competence and successful achievement in one's work (Maslach, 1993, pp 20-21)

The most widely used self-report questionnaire, the Maslach Burnout Inventory (MBI, Maslach & Jackson, 1986), is based on this definition.

The MBI includes three subscales, which correspond with the dimensions of the syndrome (a) Emotional Exhaustion (EEX—9 items); (b) Depersonalization (DEP—5 items); (c) Personal Accomplishment (PAC—8 items). A high score on the first two subscales and a low score on the latter subscale are indicative of a high level of burnout. Because the original MBI was designed to measure burnout in the helping professions, a particular version for teachers has been developed (i.e., MBI Form Ed; Maslach & Jackson, 1986). Essentially, in this version the term substituted for recipients is *students* (e g., "I don't really care what happens to some students").

It is the purpose of the current study to evaluate three major psychometric features of the Dutch version of the MBI for teachers. (a) factorial validity, (b) internal consistency of its subscales, and (c) discriminant validity. The results are compared with similar studies that employed the original English version of the MBI.

Factorial Validity

Several studies have investigated the factorial validity of the MBI in teacher samples using exploratory factor analysis. In most cases, the three-factor structure as postulated in the test manual was more or less obtained (e.g., Gold, Bachelor, & Michael, 1989; Gold, Roth, Wright, Michael, & Chen, 1992). It is interesting to note that, in some of these studies, similar deviations from the original factor structure were observed: Item 6 ("Working with people all day is really a strain for me") and Item 16 ("Working with people directly puts too much stress on me") shifted from the EEX factor to the DEP factor. Moreover, Item 12 ("I feel very energetic"—PAC) appears to be somewhat ambiguous because it also loads on the EEX factor

Iwanicki and Schwab (1981) observed a basically different factorial structure of the MBI among Massachusetts teachers. In their sample, DEP separated into two factors: (a) depersonalization affected by teaching (e g , "I worry this job is hardening me emotionally") and (b) depersonalization affected by students (e.g., "I feel I treat some students as if they were impersonal objects"). However, Green and Walkey (1988) reanalyzed the data of Iwanicki and Schwab (1981) using a different criterion for the extraction of principal factors and failed to replicate their results. Instead the original three-factor solution was observed.

Abu-Hilal and Salameh (1992) reported some indirect evidence for the existence of two factors underlying the DEP dimension. They found in Jordanian teachers that only the items reflecting depersonalization as affected by teaching loaded on the DEP factor. The three remaining items that refer to depersonalization affected by students were not included in any factor. The authors speculated that Jordanian teachers deny that students could be a reason for their burnout because this would strongly conflict with the culturally based positive attitudes they hold toward students.

Walkey and Green (1992) reanalyzed the correlation matrices of six studies, some of which included teachers, and found, without exception, a near fusion of the EEX and DEP factors. This outcome agrees with that of Green, Walkey, and Taylor (1991), who studied four nonteacher samples from several countries and proposed the Core of Burnout Scale, which is the composite EEX and DEP score.

Three studies confirmed the original three-factor structure of the MBI in teacher samples using confirmatory maximum likelihood factor analyses (Evans & Fischer, 1993; Gold et al., 1989; Gold et al., 1992). Although several alternative models were tested, such as a one-factor model and a two-factor model (including a combined DEP and EEX [Core of Burnout] factor and a separate PAC factor), the fit of the three-factor oblique model appeared to be superior. However, in two cases, highly specific samples were used (i.e., college students in a teacher training program [Gold et al., 1989] and beginning teachers with less than 3 years experience [Gold et al., 1992]).

Internal Consistency

The internal consistencies (Cronbach's α) of the MBI subscales reported for teacher samples generally agree with the values mentioned in the test manual: .90, .79, and .71 for EEX, DEP, and PAC, respectively (Maslach & Jackson, 1986). As a rule, EEX appears to be the most internally consistent subscale, followed by PAC and DEP, respectively (e.g., Belcastro, Gold, & Hays, 1983; Gold et al., 1989).

Discriminant Validity

Discriminant validity studies that assess the specificity of burnout as measured by the MBI are almost lacking as far as teacher samples are concerned. One notable exception is the study of Meier (1984) among faculty members, which showed that a considerable overlap exists between burnout and depression, by employing a multitrait-multimethod methodology. Correlations between burnout and depression ranged from .55 to .60. Unfortunately, Meier employed a composite MBI burnout score. In contrast, in studies among nurses that differentiated between the three MBI dimensions, correlations were found with depression that ranged from .43 to .52 (EEX), .23 to .32 (DEP), and −.17 to −.32 (PAC) (e.g., Glass, McKnight, & Valdimarsdottir, 1993). Accordingly, about 20% of the variance in EEX is shared with measures of depression, which is more than twice as much as the amount of variance that is shared with DEP and PAC.

Belcastro, Gold, and Grant (1982) found that correctional teachers could be classified as burned-out or non-burned-out on the basis of their somatic complaints. Burned-out was defined as scoring above the normative mean from the test manual on each of the three MBI subscales. The frequency as well as the intensity of 22 somatic complaints (e.g., sleeplessness, acid in the stomach, headaches, tachycardia) were found to be significantly higher in burned-out teachers than in non-burned-out teachers. Thus burnout in teachers appears to be associated with a wide variety of somatic complaints.

Research Issues

The current study addresses three issues. First, the factorial validity of the MBI is investigated. More particularly, four factor-analytic models were examined using confirmatory maximum likelihood factor analyses: (a) the one-factor model (M_1), which assumes that all MBI items load on a general composite burnout factor; (b) the two-factor model (M_2), in which the EEX and DEP items cluster in one factor and the PAC items constitute the second factor; (c) the original three-factor orthogonal model (M_3); and (d) the three-factor oblique model (M_4), in which the three factors of M_3 are assumed to be correlated. Second, internal consistencies of the MBI subscales were studied.

Finally, the discriminant validity of the MBI was assessed using indicators of psychological strain (including depression) and somatic complaints It is expected that burnout is a specific syndrome that can be distinguished empirically from psychological strain and from somatic complaints.

Method

Samples

The current study uses a composite sample ($N = 326$) that consists of two subsamples of teachers from eight secondary (vocational) schools located in Eindhoven (three schools; $n = 134$) and Apeldoorn (five schools; $n = 192$). Teachers' participation was voluntary. Response rates ranged from 68% to 87%. The mean age of the total sample is 45.9 years ($SD = 7.2$); 52.5% are men ($n = 171$) and 47 5% are women ($n = 155$). Their mean work experience is 14.4 years ($SD = 8.0$).

Measures

The Maslach Burnout Inventory-Form Ed (Maslach & Jackson, 1986) is employed to assess the teachers' level of burnout. The scoring dimension ranges from *never* (0) to *every day* (6) Compared to a large reference group ($N = 2,951$) that includes various human services occupations (e g., nurses, physicians, correctional officers, and social workers), the teachers in the present sample ($N = 326$) showed significantly higher levels of EEX, $t(3,275) = 5 21, p < .001$ and PAC, $t(3,275) = 46 61, p < 001$, but lower levels of DEP, $t(3,275) = 12.88, p < 001$ (Schaufeli & van Dierendonck, in press)

Psychological strain and somatic complaints were measured (only in the Eindhoven subsample) with the corresponding subscales of a work stress questionnaire, the Vragenlijst Organisatie Stress-Doetinchem (VOS-D) (Bergers, Marcelissen, & de Wolff, 1986) Psychological Strain (11 items; $\alpha = .80$) includes anxiety (e g., feeling nervous, jittery), depression (e.g., feeling sad, blue, cheerful), and irritation (e g., feeling angry, annoyed). Somatic Complaints (14 items) include a variety of psychosomatic symptoms like sweating palms, upset stomach, trouble sleeping, and heart beating faster than usual. Both VOS-D subscales are scored similarly on 4-point scales, ranging from *never* (1) to *very often* (4) Compared to the norms from the VOS-D test manual that are based on a large representative sample of Dutch workers ($N = 2,800$), the teachers in the current sample ($N = 326$) report significantly higher levels of psychological strain, $t(2,933) = 17.15$, $p < .001$, and significantly more somatic complaints, $t(2,933) = 3 91, p < 001$ Of the normative group, 75% report less psychological strain and somatic symptoms than the teachers in the current sample

Results

First, the factorial validity of the MBI was studied Using the LISREL VII computer program, the fit of four plausible factor-analytic models were

compared. Unfortunately, the chi-square goodness-of-fit index provided by LISREL strongly depends on sample size. In large samples, the chi-square statistic is very powerful, and even quite good model fits will produce significant differences. Therefore, the Tucker-Lewis Index[1] (TLI) is presented, which is barely affected by sample size (Marsh, Balla, & McDonald, 1988). This index does not provide information about the absolute fit of a particular model; rather it assesses the fit relative to another (nested) factorial model in that particular sample. Moreover, this index can be used to compare the fit of a particular model with that of a similar model in other samples of different sizes.

Table 1 presents the results of the comparison of four factor-analytic models. The so-called null model (M_0), which is exclusively used for computing the TLI, is a baseline model of maximum independence between the items (i.e., a model without a factor structure).

The probability levels of all χ^2 statistics are less than .001, indicating a rather poor absolute fit. Most probably, this is caused by the large sample size. The best relative fit of the four models is found for the three-factor oblique model. This model fits significantly better than M_1, $\delta \chi^2_{(3)} = 857.56$, $p < .001$; M_2, $\delta \chi^2_{(4)} = 122.29$, $p < .001$; and M_3, $\delta \chi^2_{(3)} = 117.65$, $p < .001$.

The LISREL modification indexes (MO) provide information about whether or not the fit could be improved if a single constraint is relaxed (i.e., an item is allowed to load on another factor as well). Inspection of these indexes reveals that a considerable improvement of the fit would occur if M_4 is reestimated, with the PAC item "I feel very energetic" (Item 12; MO = 40.03) allowed to load on the EEX factor as well. Such an additional factor loading clearly makes sense psychologically, because not feeling energetic can be considered as just another symptom of exhaustion. The fit of the revised oblique model improved significantly, $\delta \chi^2_{(1)} = 43.93$, $p < .001$, TLI = .85. Accordingly, it appears that the model fit can be slightly improved when Item 12 is allowed to load on two factors simultaneously.

Table 2 shows that the fit of all models in the current sample is substantially better than that of both studies of Gold et al. (1989; Gold et al., 1992) In contrast, the fit of M_4 is poorer compared to the recent study of Evans and Fischer (1993). However, they used a different approach to test their models by including the 3 most indicative items per burnout dimension instead of all 22 items. This procedure is likely to have a positive impact on the model fit.

Internal consistencies of the MBI subscales in the present sample are somewhat higher than in most other teacher samples: α values of .93, .81, and 85 for EEX, DEP, and PAC, respectively These values of α could not be increased by deleting any items.

To study the discriminant validity of the MBI, a second-order factor analysis of scale scores of the MBI and VOS-D scales was performed Table 3 shows the results of the principal components analysis after varimax rotation

Table 1

Comparison of LISREL Goodness-of-Fit Results of MBI Confirmatory Factor-Analytic Models (N = 326)

Model	χ^2	df	GFI	AGFI	RMSR	TLI
Null (M_0)	4,176 90*	231	251	179	360	—
One-factor (M_1)	1,451 51*	209	580	491	441	61
Two-factor oblique (M_2)	711 63*	208	816	776	088	81
Three-factor orthogonal (M_3)	877 53*	209	816	777	216	76
Three-factor oblique (M_4)	593 98*	206	853	820	078	84

Note For M_0-M_4, see text GFI = goodness-of-fit index, AGFI = adjusted goodness-of-fit index, RMSR = root mean square residual, TLI = Tucker-Lewis Index
*$p < 001$

Table 2

Comparison of MBI Confirmatory Factor-Analytic Models in Teacher Samples (Tucker-Lewis Index [TLI])

Model	Current study ($N = 326$)	Gold et al (1989)[a] ($N = 147$)	Gold et al (1992)[a] ($N = 133$)	Evans and Fischer (1993)[a] ($N = 351$)
One-factor (M_1)	61	48	33	57
Two-factor oblique (M_2)	81	—	46	82
Three-factor orthogonal (M_3)	76	62	—	—
Three-factor oblique (M_4)	84	66	60	94

Note For M_1-M_4, see text
a TLI computed by the authors

Clearly, two components emerge. (a) a relatively strong and specific MBI burnout component and (b) a somewhat weaker general strain component that includes less specific mental and somatic complaints as measured by the VOS-D. Quite interestingly, EEX loads substantially on the second factor, indicating that of all three burnout dimensions, emotional exhaustion is most closely related with general psychological and physical symptoms

Discussion

To date, virtually all studies on teacher burnout have been conducted in the United States, using the original English version of the MBI The present study addresses the factorial validity, reliability, and discriminant validity of the MBI in a sample of Dutch teachers

In the Netherlands, as in the United States, many teachers suffer from job stress However, the picture is somewhat more complex. Compared to a large

Table 3
Discriminant Validity: Factor Loadings (N = 134)

Scale	Factor I	Factor II
MBI EEX	.73	.48
MBI DEP	.83	13
MBI PAC	-.73	- 01
VOS-D Psychological strain	.12	.75
VOS-D Somatic complaints	11	.84
Explained variance (percentage)	46 5	20 0

Note MBI = Maslach Burnout Inventory (Maslach & Jackson, 1986), EEX = emotional exhaustion, DEP = depersonalization, PAC = personal accomplishment, VOS-D = Vragenlijst Organisatie Stress-Doetinchem (Bergers, Marcelissen, & de Wolff, 1986). Bold indicates factor loadings > 40

reference group consisting of a wide variety of Dutch human services professionals, the teachers in the current sample experienced significantly higher levels of emotional exhaustion. On the other hand, they reported significantly fewer symptoms of depersonalization and reported more feelings of personal accomplishment. In addition, they experienced significantly more psychological strain and have more somatic complaints than other Dutch workers. In short, the teachers showed physical and mental stress symptoms (e.g., EEX), but their attitudes toward pupils (DEP) and toward their work (PAC) were not seriously affected.

This study corroborates the three-factor oblique structure that has been found in other recent investigations that employed similar confirmatory factor analyses (Evans & Fischer, 1993; Gold et al., 1989; Gold et al., 1992). Compared to these studies, the three-factor model fits relatively well to the data of our study. In contrast to earlier findings, no ambiguous or problematic items were identified (e.g., Belcastro et al., 1983). There was only one exception to this rule, Item 12 ("I feel very energetic"). Reestimation of a modified model, in which this PAC item was allowed to load on EEX, slightly improved the fit of the model.

As in most other studies, EEX and DEP are closely related, $r = .59$, but they do not appear to constitute one Core of Burnout dimension (see Table 1), as has been suggested by Green et al. (1991).

The internal consistencies of the MBI subscales are relatively high compared to the α values found in other studies and follow the same pattern that is observed in the data for other occupational groups, EEX being the most internally consistent subscale, followed by PAC and DEP, respectively (cf Schaufeli, Enzmann, & Girault, 1993).

The present study shows that MBI burnout can be successfully discriminated from other symptoms like psychological strain (e.g., depression) and somatic complaints. However, one dimension of burnout (EEX) is substan-

tially related to more general symptoms. This finding agrees with a recent study among Dutch nurses that included, in addition to the variables of the present study, another burnout questionnaire (the Burnout Measure [BM], Schaufeli & van Dierendonck, 1993). Essentially, the BM assesses the level of mental and physical exhaustion (Pines & Aronson, 1988). In accordance to the findings of the present sample, a two-factor solution emerged consisting of (a) mental and physical symptoms (MBI EEX, BM, VOS-D scales) and (b) MBI burnout (EEX, DEP, PAC). As in the present study, EEX loaded on both dimensions. Accordingly, emotional exhaustion is the dimension of burnout that is most closely linked to psychological strain and somatic symptoms.

The final conclusion from the present study is that the MBI is a reliable and valid instrument to study burnout among teachers in a non-English-speaking country. The three-dimensional structure of the MBI has been confirmed (factorial validity) and burnout can be distinguished from less specific psychological and physical symptoms (discriminant validity).

Of course, the nature of burnout cannot be determined exclusively in psychometric investigations such as the present one. Additional longitudinal, theory-driven research is needed in which a priori models are tested. The MBI seems to be a suitable tool for such investigations

Note

1 TLI = (n/df$_n$–t/df$_t$)/(n/df$_n$–1) with TLI = Tucker-Lewis Index, n = chi-square of null model, *df$_n$* = degree of freedom of null model, *t* = chi-square of target model, *df$_t$* = degree of freedom of target model

References

Abu-Hilal, M M., & Salameh, K M (1992) Validity and reliability of the Maslach Burnout Inventory for a sample of non-Western teachers *Educational and Psychological Measurement*, 52, 161-169

Belcastro, P A., Gold, R S., & Grant, J (1982) Stress and burnout Physiologic effects on correctional teachers *Criminal Justice and Behavior*, 9, 387-395

Belcastro, P A., Gold, R S., & Hays, L C (1983) Maslach Burnout Inventory Factor structures for samples of teachers *Psychological Reports*, 53, 364-366

Bergers, G P A., Marcelissen, F H G., & de Wolff, Ch J (1986) VOS-D Vragenlijst Organisatie Stress-D, handleiding (VOS-D test manual) Nijmegen University of Nijmegen, Department of Psychology

Evans, B K., & Fischer, D G (1993) The nature of burnout A study of the three-factor model of burnout in human service and non-human service samples *Journal of Occupational and Organizational Psychology*, 66, 29-38

Farber, B A (1991) *Crisis in education Stress and burnout in the American teacher* San Francisco Jossey-Bass

Glass, D C, McKnight, J D., & Valdimarsdottir H (1993) Depression, burnout and perceptions of control in hospital nurses *Journal of Consulting and Clinical Psychology*, 61, 147-155

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

Gold, Y, Roth, R A, Wright, C. R., Michael, W B, & Chen, C Y (1992) The factorial validity of a teacher burnout measure (educators survey) administered to a sample of beginning teachers in elementary and secondary schools in California *Educational and Psychological Measurement*, *52*, 761-768

Green, D E, & Walkey, F. H (1988). A confirmation of the three-factor structure of the Maslach Burnout Inventory *Educational and Psychological Measurement*, *48*, 579-585

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

Iwanicki, E F, & Schwab, R L. (1981) A cross validation study of the Maslach Burnout Inventory *Educational and Psychological Measurement*, *41*, 1174-1176

Knepper, S (1991) De epidemie van psychische arbeidsongeschiktheid Feiten en achtergronden [The epidemic of mental work incapacity Facts and backgrounds] *Gedrag & Gezondheid*, *19*, 315-319

Marsh, H W, Balla, J. R, & McDonald, R P (1988) Goodness-of-fit indexes in confirmatory factor analysis The effects of sample size. *Psychological Bulletin*, *103*, 391-410

Maslach, C (1993) Burnout A multidimensional perspective In W B Schaufeli, C Maslach, & T Marek (Eds), *Professional burnout Recent developments in theory and research* (pp 19-32) Washington Taylor and Francis

Maslach, C, & Jackson, S E (1986) *MBI Maslach Burnout Inventory Manual research edition* Palo Alto University of California, Consulting Psychologists Press

Meier, S T (1984) The construct validity of burnout *Journal of Occupational Psychology*, *57*, 211-219

Pines, A, & Aronson, E (1988) *Career burnout Causes and cures* New York Free Press

Schaufeli, W B, Enzmann, D, & Girault, N. (1993). Measurement of burnout A review In W B Schaufeli, C. Maslach, & T Marek (Eds), *Professional burnout Recent developments in theory and research* (pp. 199-215) Washington Taylor and Francis

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

Schaufeli, W B, & van Dierendonck, D (in press) De Nederlandse versie van de Maslach Burnout Inventory [The Dutch version of the Maslach Burnout Inventory] *Gedrag & Gezondheid*

Walkey, F H, & Green, D (1992) An exhaustive examination of the replication of the factor structure of the Maslach Burnout Inventory *Educational and Psychological Measurement*, *52*, 309-323