SA Journal of Industrial Psychology, 2007, 33 (2), 7-15
SA Tydskrif vir Bedryfsielkunde, 2007, 33 (2), 7-15

A PSYCHOMETRIC EVALUATION OF MEASURES OF AFFECTIVE WELL-BEING IN AN INSURANCE COMPANY

WJ COETZER
S ROTHMANN
Ian.Rothmann@nwu.ac.za
WorkWell: Research Unit for People, Policy and Performance
North-West University

ABSTRACT

The objectives of this study were to validate two measures of affective well-being, namely the Maslach Burnout Inventory (MBI) and the Utrecht Work Engagement Scale (UWES) for employees in an insurance company, to assess their construct equivalence for different language groups and to determine the relationship between burnout and work engagement. A cross-sectional survey design with an availability sample (N = 613) was used. The MBI, UWES and a biographical questionnaire were administered. Structural equation modelling confirmed a three-factor model of burnout, consisting of Exhaustion, Cynicism and Professional Efficacy and a three-factor model of work engagement consisting of Vigour, Dedication and Absorption. Acceptable construct equivalence of the three-factor model of burnout and work engagement for different language groups was confirmed. A second-order factor analysis of the scales resulted in two factors, namely burnout and work engagement.

Key words:
Burnout, work engagement, construct equivalence

The insurance industry expanded considerably in the late 19th century (Chan, 2002), inducing acute competitiveness and rivalry between companies and employees (Lai, Chan, Ko & Boey, 2000). This, along with the increased demands from current operating and economic conditions world-wide, force organisations to make rapid changes to their workforce. Everywhere organisations are downsizing, outsourcing and restructuring, ultimately impacting on employees' work demands and obligations (Kickul & Posig, 2001) and leaving them with feelings of stress, insecurity, misunderstanding, undervaluation and alienation. These rapid changes in organisations, along with the changes in information technology, make the situation more complex for employees (Lindstršm, Leino, Seitsamo & Tordtila, 1997). They are faced not only with increased workloads and pressures but also with decreased job control (Chan, 2002; Lai et al., 2000). When the employee can no longer tolerate occupational pressures and feels totally overwhelmed by stress, he or she is likely to reach breaking point and experience burnout (Weisberg, 1994).

Burnout has been defined as "a persistent, negative, work-related state of mind in 'normal' individuals that is primarily characterised by exhaustion, which is accompanied by distress, a sense of reduced effectiveness, decreased motivation, and the development of dysfunctional attitudes and behaviors at work" (Schaufeli & Enzmann, 1998, p. 36). Maslach, Jackson and Leiter (1996) and Maslach, Schaufeli and Leiter (2001) describe burnout as a syndrome consisting of three dimensions, namely feelings of emotional exhaustion, depersonalisation (cynicism) and reduced personal accomplishment (professional efficacy). It is important to use a reliable and valid instrument to measure burnout. The development of the Maslach Burnout Inventory – General Survey (MBI-GS) (Schaufeli, Leiter, Maslach & Jackson, 1996) made it possible to make comparisons among different occupational groups. Thus the concept of burnout and its measurement were broadened to include all employees and not only those who do "people work" (Maslach & Leiter, 1997). The MBI-GS comprises three subscales: Exhaustion, Cynicism and Professional Efficacy.

Empirical studies revealed that some employees, regardless of high job demands and long working hours, do not develop burnout in comparison to others, but seemed to find pleasure in hard work and dealing with job demands (Schaufeli & Bakker, 2001). Consequently, theoretical and empirical studies commenced on the concept of work engagement, theoretically viewed as the antithesis of the burnout construct. Work engagement is defined as a positive, fulfilling, work-related state of mind that is characterised by vigour, dedication and absorption. Furthermore, it is not a momentary and specific state, but a more persistent and pervasive affective-cognitive state which is not focused on a particular object, event, individual or behaviour (Schaufeli, Salanova, González-Romá & Bakker, 2002).

Burnout and work engagement are important components of affective work-related well-being. A lack of research in terms of burnout and work engagement of employees in the insurance industry within the South African context necessitates the current research. It is not only important to obtain valid and reliable measurements of burnout and work engagement in South Africa from an empirical point of view, but also to enable the individual measurement of burnout and work engagement in a valid and reliable manner in the insurance industry context in South Africa. According to Van de Vijver and Leung (1997), measurement equivalence should be computed for measuring instruments in any multicultural setting where groups from different cultural groups are compared in terms of a specific construct.

The objectives of this study were to validate the Maslach Burnout Inventory – General Survey (MBI-GS) and the Utrecht Work Engagement Scale (UWES) for employees in an insurance company in South Africa, to assess their construct equivalence for different language groups and to determine the relationship between burnout and work engagement.

The Maslach Burnout Inventory – General Survey (MBI – GS)
The Maslach Burnout Inventory (MBI) (Maslach et al., 1996) is currently the most widely used research instrument to measure burnout. The psychometric properties of the MBI-GS are satisfactory. Internal consistencies varying from 0,73 (Cynicism) to 0,91 (Exhaustion) are reported by Leiter and Schaufeli (1996). Reliability analyses done by Schutte, Toppinen, Kalimo and Schaufeli (2000) showed that the Exhaustion and Professional Efficacy subscales were sufficiently internally consistent, but that one Cynicism item should be removed in order to increase the internal consistency beyond the criterion of 0,70. According to them, this might be caused by the ambivalence of the particular item: "I just want to do my job and not be bothered". In their studies, Schaufeli, Leiter and Kalimo (1995) and Leiter and Schaufeli (1996) also found that this item had the lowest factor loadings of the three subscales.

Four studies were found that used the MBI-GS on South African samples. In a sample of senior managers in a manufacturing industry, Rothmann and Jansen van Vuuren (2002) found

satisfactory alpha coefficients: Exhaustion = 0,79; Cynicism = 0,84 (after item 13 had been omitted); and Professional Efficacy = 0,84. Rothmann and Malan (2003) found higher alphas (Exhaustion = 0,89; Cynicism = 0,76; Professional Efficacy = 0,85) while Kruger, Veldman, Rothmann and Jackson (2002) found lower alphas for Cynicism (0,72) and Professional Efficacy (0,69). Storm and Rothmann (2003a) found alpha coefficients of 0,88 (Exhaustion), 0,78 (Cynicism) and 0,79 (Professional Efficacy) in a sample of police officers in South Africa.

Confirmatory factor analysis done by Schutte et al. (2000) showed that the three-factor model of the MBI-GS was clearly superior to alternative one-factor and two-factor models. Leiter and Schaufeli (1996) employed confirmatory factor analysis using linear structural equation modelling and also confirmed a three-factor structure. Taris, Schreurs and Schaufeli (1999) obtained similar results. However, in a sample of Spanish workers who used computer-aided technologies at their jobs, Salanova and Schaufeli (2000) found a four-factor model of burnout where the Efficacy subscale split into two factors that were labelled "goal attainment" and "self-confidence". According to the authors, this divergent result might have been caused by translation problems or by the specific sample being studied. It seems reasonable to expect that a three-factor structure will be obtained in this study. Confirmatory factor analyses by Rothmann and Jansen van Vuuren (2002), Rothmann and Malan (2003), Kruger et al. (2002) and Storm and Rothmann (2003a) consistently showed low loadings of item 13 on Cynicism. Despite various studies regarding the factor structure of the MBI-GS as described above, the equivalence of the factor structure for different language groups has not been studied to date.

The first hypothesis of this study is as follows:

H1: Burnout, as measured by the MBI-GS, can be defined as a three-dimensional construct, with an equivalent structure for different language groups and acceptable levels of internal consistency for each of its subscales, namely Exhaustion, Cynicism and Professional Efficacy.

The Utrecht Work Engagement Scale (UWES)

Schaufeli, Salanova et al. (2002) developed the Utrecht Work Engagement Scale (UWES). The UWES measures three dimensions of work engagement: Vigour, Dedication and Absorption. Schaufeli et al. (2002) reported acceptable internal consistencies ($\alpha > 0,70$) for the UWES. Recent confirmatory factor-analytical studies confirmed the factorial validity of the UWES (Schaufeli, Bakker, Hoogduin, Schaap & Kladler, 2001; Schaufeli, Martinez, Pinto, Salanova & Bakker, 2002; Schaufeli, Salanova et al., 2002). The findings showed internally consistent results for the three scales of the UWES.

In a cross-cultural study regarding the UWES for students in Spain, Portugal and the Netherlands, the factorial validity of the UWES was confirmed and the internal consistency of the scales was found to be satisfactory (Schaufeli, Martinez et al., 2002). Factor loadings of Absorption were found to be invariant across all samples, while factor loadings of Vigour were invariant for only two of the three groups. The three-factor model fit to the data was found to be superior in all three samples after removing three items, namely items 17, 16 and 11. Internally consistent Cronbach alphas ranged from 0,65 to 0,79 for Vigour (5 items); 0,77 to 0,85 for Dedication (5 items); and from 0,65 to 0,73 for Absorption (4 items).

Two studies regarding the internal consistency and factorial validity in South Africa were found. In their study, Storm and Rothmann (2003b) found that a re-specified one-factor model (after deleting items 3, 11, 15 and 16) fitted the data the best in their random, stratified sample of police members in South Africa. Although a three-factor model was initially tested and satisfactory results obtained, the fit with the data was superior for a one-factor model. Internal consistencies of the three subscales were determined at 0,78 (Vigour), 0,89 (Dedication) and 0,78 (Absorption). In a sample of emergency workers, Naudé (2003) found that a re-specified three-factor model (after deleting items 15 and 16) fitted the data best. Internal consistencies of the three subscales were determined at 0,70 (Vigour); 0,83 (Dedication) and 0,67 (Absorption). Although various studies regarding the factor structure of the UWES have been conducted, the equivalence of the factor structure for different language groups has not been studied to date.

The second hypothesis of this study is as follows:

H2: Work engagement, as measured by the UWES, can be defined as a three-dimensional construct, with an equivalent structure for different language groups and acceptable levels of internal consistency for each of its subscales, namely Vigour, Absorption and Dedication.

The relationship between burnout and work engagement

Burnout and work engagement are both components of affective well-being at work. Maslach and Leiter (1997) described burnout as an "erosion of engagement with the job". Work engagement, according to these authors, is characterised by energy, involvement and efficacy, the direct opposites of burnout. However, Schaufeli, Salanova et al. (2002) found that although burnout is related to work engagement, it is not the direct opposite of burnout. Schaufeli and Bakker (2004) found that the engagement and burnout scales were moderately negatively correlated. Schaufeli, Salanova et al. (2002) and Schaufeli and Bakker (2004) observed that a core burnout factor – consisting of (emotional) exhaustion and mental distance (cynicism) – and an extended engagement factor – including professional efficacy in addition to the three engagement scales – fitted the data best. Green, Walkey and Taylor (1991) also referred to exhaustion and mental distance (cynicism) as the core of burnout.

The third hypothesis of this study is as follows:

H3: Burnout and work engagement are two different but related components of work-related well-being.

RESEARCH DESIGN

Research approach

A cross-sectional survey design was used in this study. Questionnaires were used to gather data in a no-random field survey.

Research method

Participants

The total population of 1 100 employees in an insurance company was targeted. A response rate of 56,5% was achieved, of which 613 responses (98,55%) could be utilised. Descriptive information of the sample is given in Table 1.

Table 1
Characteristics of the participants

Item	Category	Frequency (Percentage)
Education	Grade 10 (Standard 8)	48 (8,35)
	Grade 12	303 (52,70)
	Grade 12 + Diploma	133 (23,13)
	Grade 12 + Higher Diploma or Degree	68 (11,83)
	Grade 12 + Higher Diploma or Degree (Honours)	16 (2,78)
	Grade 12 + Higher Diploma or Degree (Master's)	7 (1,22)
Gender	Male	262 (42,74)
	Female	351 (57,26)
Race	Black	67 (10,95)
	White	281 (45,92)
	Coloured	236 (38,56)
	Asian	28 (4,58)
Home Language	Afrikaans	208 (34,04)
	English	339 (55,48)
	Other	64 (10,47)

The sample consisted mainly of English-speaking (55,48%), married females (57,26%) with a Grade 12 school qualification (52,7%). The mean age of the participants was 35,5 years while the average length of service was 7,55 years.

Measuring Instruments

The *Maslach Burnout Inventory – General Survey* (MBI-GS) (Maslach et al., 1996) measures respondents' perceived experience of burnout. The MBI-GS has three subscales: Exhaustion (five items, e.g. "I feel used up at the end of the workday"), Cynicism (five items, e.g. "I have become less enthusiastic about my work") and Professional Efficacy (six items, e.g. "In my opinion, I am good at my job"). Cronbach alpha coefficients reported by Maslach et al. (1996) varied from 0,87 to 0,89 for Exhaustion, 0,73 to 0,84 for Cynicism and 0,76 to 0,84 for Professional Efficacy. Test-retest reliabilities after one year were 0,65 (Exhaustion), 0,60 (Cynicism) and 0,67 (Professional Efficacy) (Maslach et al., 1996). All items are scored on a seven-point frequency-rating scale ranging from 0 (*never*), to 6 (*daily*). Storm (2002) confirmed the three-factor structure of the MBI-GS. The following Cronbach alpha coefficients were obtained for the MBI-GS: Exhaustion: 0,88; Cynicism: 0,79; Professional Efficacy: 0,78 (Storm, 2002).

The *Utrecht Work Engagement Scale* (UWES) (Schaufeli, Salanova et al., 2002) measures levels of work engagement. The UWES includes three dimensions, namely vigour, dedication and absorption, which are conceptually regarded as the opposite of burnout and are scored on a seven-point frequency-rating scale, varying from 0 (*never*) to 6 (*every day*). The questionnaire consists of 17 questions and includes questions like "I am bursting with energy every day in my work"; "Time flies when I am at work" and "My job inspires me". The alpha coefficients for the three subscales varied between 0,68 and 0,91. Storm (2002) obtained the following alpha coefficients for the UWES in the South African Police Service: Vigour: 0,78; Dedication: 0,89; Absorption: 0,78. Naudé (2003) obtained the following alpha coefficients in a sample of 405 emergency workers in South Africa: Vigour: 0,70; Dedication: 0,83; and Absorption: 0,67.

Statistical analysis

The statistical analysis was carried out with the SAS program (SAS Institute, 2000). In the first step, means, standard deviations, skewness and kurtosis were determined to describe the data. The reliability and validity of the MBI-GS and UWES were also determined by means of Cronbach alpha coefficients, mean inter-item correlations and their distribution scales, as well as confirmatory factor analysis with the use of the AMOS program (Arbuckle, 1999).

In order to test the factorial validity and construct equivalence of the MBI-GS and UWES for different language groups, structural equation modeling (SEM) methods were used with the maximum likelihood method of the AMOS program (Arbuckle, 1999). According to Jšresk og (1971), all tests of invariance across groups should begin with a global test of the equality of their covariance structures. In testing for these equivalencies, sets of parameters are tested in a logical order and by increasing restrictions in every step. The sets of parameters that are of most interest regarding group variances are: (a) factor loading paths, (b) factor variances/covariances, and (c) structural regression paths, while, according to Bentler (1995) – contradicting the view of Jöreskog – equality of error variances and covariances is generally the least important hypothesis to test, due to the restrictive nature of these tests.

The general procedure for the testing of hypotheses related to group invariance starts with scrutiny of the measurement model. The pattern of factor loadings for each observed measure should be tested first for its equivalence across the groups. Once the group invariances have been identified, these parameters are equally constrained, while subsequent tests of the structural parameters are conducted. While testing each new set of parameters, those known to be group-invariant are equally constrained, thus testing a series of increasingly restrictive hypotheses in an orderly sequence of analytic steps (Byrne, 2001).

Before the factorial invariance can be tested as described above, it is important to consider a baseline model for each group separately, which best fits the data from the perspectives of both parsimony and substantive meaningfulness. Baseline models need not be completely identical across groups. The number of factors also need not be equivalent across groups (Byrne, 2001). In testing for invariance, however, equality constraints are imposed on particular parameters. Therefore, the data for all groups must be analysed simultaneously to obtain efficient estimates (Bentler, 1995; Jš reskog & Sš rbom, 1996).

Hypothesised relationships are tested empirically for goodness of fit with the sample data. The χ^2 and several other goodness-of-fit indices summarise the degree of correspondence between the implied and observed covariance matrices. However, because the χ^2 statistic equals (N – 1) Fmin, this value tends to be substantial when the model does not hold and the sample size is large (Byrne, 2001). The following goodness-of-fit-indices were used as adjuncts to the χ^2 statistics: a) The Goodness of fit Index (GFI); b) The Adjusted Goodness of Fit Index (AGFI); c) The Normed Fit Index (NFI); d) The Comparative Fit Index (CFI); e) The Tucker-Lewis Index (TLI), and f) The Root Mean Square Error of Approximation (RMSEA).

RESULTS

Construct validity and construct equivalence of the MBI-GS

Problems of construct equivalence could arise when participants have difficulty with communicating in English (the language in which the instruments were administered). Because of relatively small sample sizes in the African language groups, it was decided to conduct the analyses for two language groups, namely Afrikaans/African and English.

First, a one-factor model of the MBI-GS consisting of 16 items was tested for each language group separately. Next, a three-factor model of the MBI-GS consisting of 16 items was tested for each language group separately. Statistics of the fit between the theoretical model and the empirical data for both language groups are given in Table 2.

Table 2

Goodness-of-fit statistics for the hypothesised MBI-GS models employees in an insurance company in different language groups

Model	χ^2	χ^2/df	GFI	AGFI	NFI	TLI	CFI	RMSEA
One-factor Model – Afrikaans and African	644,01	6,19	0,71	0,62	0,56	0,54	0,60	0,14
One-factor Model – English	934,49	8,99	0,66	0,56	0,59	0,56	0,62	0,15
Three-factor Model 1 – Afrikaans and African	286,32	2,84	0,88	0,84	0,80	0,84	0,86	0,08
Three-factor Model 2 – Afrikaans and African	145,67	1,71	0,94	0,91	0,90	0,94	0,95	0,05
Three-factor Model 1 – English	334,17	3,31	0,89	0,85	0,85	0,87	0,89	0,08
Three-factor Model – English	223,05	2,62	0,92	0,89	0,90	0,92	0,94	0,07

Table 2 shows that the fit of the three-factor model was superior to the one-factor model for both language groups. Regarding the three-factor model, the χ^2 values of 286,32 (df = 101; p = 0,01) obtained for the Afrikaans and African language group and of 334,17 (df = 101; p = 0,01) for the

English language group are indicative of a poor overall fit to the theoretical three-factor model of the MBI-GS. The goodness-of-fit indices also support this finding by not reaching the recommended critical values. Values lower than 0,90 for GFI, AGFI, NFI, TLI and CFI were found. The RMSEA value is also higher than the recommended value of 0,05. In order to obtain a better fit between the theoretical three-factor model with the population data, modification of the model is needed. To pinpoint possible areas of misfit, modification indices were examined. Looking at the regression weights, one parameter, which represents the cross-loading of Item 13 on the Cynicism factor, stands apart from the rest and accounts for substantial misspecification of the hypothesised factor loading.

The rejection of the postulated theoretical model in the previous section initiated, by implication, a model development process, in other words, an exploratory factor analysis process. Given the high cross-loading levels of item 13, it was decided to re-specify the model by deleting this variable. Also, errors of two item pairs (namely CY14-CY15 and PE11-PE12) were allowed to correlate, given the comparatively high covariance associated with these errors. Subsequent analysis therefore includes only 15 items, labelled Model 2. The various fit statistics in Table 2 indicate an incremental improvement from the first model fit with the empirical data. All fit indices indicated a marginally acceptable fit at best with the data for the Afrikaans and African language group, with χ^2 = 145,67 (df = 85; p = 0,01). With the exception of the AGFI and RMSEA, all fit indices indicate a marginally acceptable fit at best with the data for the English language group, with χ^2 = 223,05 (df = 85; p = 0,01). A difference of $\Delta\chi^2_{(16)}$ = 140,65 was found between Model 1 and Model 2 for the Afrikaans and African language group and a difference of $\Delta\chi^2_{(16)}$ = 111,12 was found between Model 1 and Model 2 for the English language group. Both these differences are substantial. The other fit statistics seem to support an acceptable fit of Model 2 with the empirical data for both language groups. Because this model represented acceptable comparative evidence of fit for both language groups between the empirical data and a theoretical model in line with the theoretical premises of the MBI-GS, no further modification of the model was deemed necessary.

Next, tests of invariance in different language groups for the MBI-GS were determined, and these are indicated in Table 3.

Table 3
Goodness-of-fit statistics for tests of invariance of the MBI-GS for employees in an insurance company in different language groups

Model	Groups	Comparative model	χ^2	*df*	$\Delta\chi^2$	*df*	p
Hypothesised model (Model 1)	Afrikaans/ English/ African		368,70	170	-	-	-
Factor loadings, variances and covariances constrained to be equal	Afrikaans/ English/ African	Model 1	392,60	190	23,9	20	NS

The results in Table 3 shows that construct equivalence exists, with factor loadings, variances and covariances constrained to be equal among the various language groups. The equality of error covariances was not tested, due to the restrictive nature of the test on the data and the relative unimportance thereof (Byrne, 2001). These results provide support for part of Hypothesis 1 in that burnout, as measured by the MBI-GS, can be defined as a three-dimensional construct, with an equivalent structure for different language groups.

Construct validity and construct equivalence of the UWES

The full three-factor model of the UWES, consisting of 17 items, was tested for each language group separately. Statistics of the fit between the theoretical model and the empirical data are given in Table 4.

Table 4
Goodness-of-fit statistics for the hypothesised UWES model for employees in an insurance company in different language groups

Model	χ^2	χ^2/df	GFI	AGFI	NFI	TLI	CFI	RMSEA
Model 1 (one-factor) – Afrikaans and African	430,03	3,61	**0,83**	**0,78**	**0,80**	**0,82**	**0,85**	**0,10**
Model 2 (one-factor) – Afrikaans and African	302,78	3,44	0,86	0,82	0,85	0,86	0,89	0,10
Model 1 (one-factor) – English	499,45	4,20	0,84	0,79	0,81	0,83	0,85	0,10
Model 2 (one-factor) – English	324,73	3,69	0,88	0,84	0,87	0,88	0,90	0,09
Model 1 (three-factor) – Afrikaans and African	360,31	3,11	0,85	0,80	0,83	0,86	0,88	0,09
Model 2 (three-factor) – Afrikaans and African	250,87	2,92	0,88	0,84	0,87	0,89	0,91	0,08
Model 1 (three-factor) – English	433,04	3,73	0,86	0,81	0,84	0,85	0,88	0,09
Model 2 (three-factor) – English	319,35	3,71	0,88	0,83	0,87	0,88	0,90	0,09
Model 1 (adjusted) – Afrikaans and African	297,46	3,42	0,87	0,82	0,87	0,88	0,90	0,09
Model 2 (adjusted) – Afrikaans and African	165,37	2,76	0,91	0,87	0,91	0,93	0,94	0,08
Model 1 (adjusted) – English	314,97	3,62	0,88	0,84	0,87	0,88	0,90	0,09
Model 2 (adjusted) – English	127,97	2,13	0,95	0,92	0,94	0,96	0,97	0,06

First, a unidimensional model, which assumes that all 17 UWES items load on one single factor, was tested. Table 4 provides a summary of the fit statistics for the hypothesised one-factor model. This model, however, revealed very poor overall fit, as indicated by the statistically significant χ^2 value of 430,03 (df = 119; p = 0,01) for the Afrikaans and African language group and 499,45 (df = 119; p = 0,01) for the English language group. All the other fit indices confirmed a poor fit with the data. In order to obtain a better fit between the theoretical one-factor model with the population data, modification of the model is needed.

To pinpoint possible areas of misfit, modification indices were examined. Looking at the regression weights, items 16 and 17 demonstrated comparatively low values. The standardised residual covariances confirmed the problematic nature of items 16 and 17, with loadings higher than 2,58 (Byrne, 2001).

The rejection of the postulated theoretical model in the previous section initiated, by implication, an exploratory factor analysis process where the constructs of work engagement are studied specifically in the insurance company worker population. Given the high cross-loading levels of items 16 and 17, it was decided to re-specify the model by deleting these variables. Also, errors of two item pairs (i.e. 1 and 4 and 11 and

12) were allowed to correlate, given the comparatively high covariance associated with these errors. Subsequent analysis therefore includes only 15 items, labelled Model 2. Fit statistics for Model 2 are presented in Table 2. The various fit statistics in Table 2 indicate an incremental improvement from the first model fit with the empirical data. This model, however, still revealed very poor overall fit as indicated by the statistically significant χ^2 value of 302,78 (df = 88; p = 0,01) for the Afrikaans and African language group and 324,73 (df = 88; p = 0,01) for the English language group. All the other fit indices confirmed a poor fit with the data.

Subsequently, the hypothesised 17-item three-factor UWES model was fitted with the data. In Table 4 the fit statistics are provided for the fit between the original model and the empirical data for both language groups. Statistics of the fit between the theoretical three-factor model and the empirical data for both language groups is given in Table 4. The χ^2 value of 360,31 (df = 116; p = 0,01) obtained for the Afrikaans and African language group, and of 443,04 (df = 116; p = 0,01) for the English language group, is indicative of a poor overall fit to the theoretical three-factor model of the UWES. The goodness-of-fit indices also support this finding by not reaching the recommended critical values. Values lower than 0,90 for GFI, AGFI, NFI, TLI and CFI were found. The RMSEA value is also higher than the recommended value of 0,05.

To pinpoint possible areas of misfit, modification indices were examined. Looking at the regression weights, items 16 and 17 demonstrated comparatively low values. The standardised residual covariances confirmed the problematic nature of items 16 and 17, with loadings higher than 2,58. Given the high cross-loading levels of items 16 and 17, it was decided to re-specify the model by deleting these variables. Also, errors of one item pair (i.e. VI1-VI4) were allowed to correlate, given the comparatively high covariance associated with these errors. Subsequent analysis therefore includes only 15 items, labelled Model 2. Fit statistics for Model 2 are presented in Table 4. The various fit statistics in Table 4 indicate an incremental improvement from the first model fit with the empirical data. The difference between Model 1 ($\Delta\chi^2_{(116)}$ = 360,31) and Model 2 ($\Delta\chi^2_{(86)}$ = 250,87) is $\Delta\chi^2_{(30)}$ = 109,44 for the Afrikaans and African language groups and the difference between Model 1 ($\Delta\chi^2_{(116)}$ = 433,04) and Model 2 ($\Delta\chi^2_{(86)}$ = 319,35) is $\Delta\chi^2_{(30)}$ = 113,69 for the English language group. These differences are substantial. The goodness-of-fit indices do not reach the recommended critical values. Except for the CFI, values lower than 0,90 for the fit indices were found. The RMSEA value is also higher than the recommended value of 0,05.

These findings could possibly be explained in terms of the possibility of semantic differences in terms of understanding the content of the items by the different language groups. It is possible that certain items were misunderstood by some of the language groups, which led to inconsistent responses by the different language groups in this sample. Therefore some items were replaced with items that were written in a more familiar South African vocabulary, in order to address the possible semantic problems. Item 4 (“I feel strong and vigorous in my job.”) was replaced with item 19 (“I feel strong and full of energy in my work.”). Item 9 (“I feel happy when I am engrossed in my work.”) was replaced with item 18 (“I feel happy when my attention is totally focused on my work.”). Item 11 (“I am immersed in my work.”) was replaced with item 21 (“I enjoy devoting all my attention and energy to my work.”). Item 15 (“I am very resilient, mentally, in my job.”) was replaced with item 20 (“In my job I can comfortably deal with stressful situations and I easily recover from such situations.”).

The adjusted three-factor model was fitted with the data. In Table 4 the fit statistics are provided for the fit between the adjusted model and the empirical data for both language groups. According to Table 4, it is evident that the SEM analysis yielded a marginal fit at most between the theoretical model and empirical data for both language groups. The statistically significant χ^2 value of 297,46 (df = 87; p = 0,01) for the Afrikaans and African language group and 314,97 (df = 87; p = 0,01) for the English language group, along with the relatively elevated RMSEA values, indicate possible existing misspecifications in the theoretical model. None of the fit indices, except the CFI, reached the recommended critical values. Since model fit was not acceptable, further modification of the model was deemed necessary.

Inspection of the standardised residual covariances led to the identification of item 14 with two loadings > 2,58 for the Afrikaans and African language groups and one loading > 2,58 for the English language group, and of item 20 with one loading > 2,58 for the English language group. Also, errors of two item pairs (i.e. DE7-DE13 and AB3-AB21) were allowed to correlate, given the comparatively high covariance associated with these errors. Having identified possible areas of misspecification in the model, modification of the adjusted model is needed.

The theoretical model was re-specified by deleting items 14 and 20. The various fit statistics in Table 4 indicate a marginal improvement from the adjusted model fit with the empirical data with a significant χ^2 value of 165,37 (df = 60; p = 0,01) for the Afrikaans and African language group, and of 127,97 (df = 60; p = 0,01) for the English language group. All of the indices, except the AGFI (for the Afrikaans and African language group), the RMSEA (for both language groups), reached the recommended critical values. A difference of $\Delta\chi^2_{(17)}$ = 132,09 for the Afrikaans and African language group and $\Delta\chi^2_{(17)}$ = 187,00 for the English language group was found between Model 1 and Model 2, which is significant. Because this model represented acceptable comparative evidence of fit for both language groups between the empirical data and a theoretical model in line with the theoretical premises of the UWES, no further modification of the model was deemed necessary.

Next, tests of invariance in different language groups for the UWES were determined, and these are indicated in Table 5.

A difference of $\Delta\chi^2_{(18)}$ = 61,00 was found between the hypothesised adjusted three-factor UWES model (Model 1) and the hypothesised model with factor loadings, variances and covariances constrained to be equal. This difference is statistically significant (p < 0,01). The different factor loadings were then separately constrained and tested against Model 1. Model 2 (Model 1 with factor loadings on Vigour constrained to be equal) displayed a difference of $\Delta\chi^2_{(3)}$ = 7,91 with Model 1, which is non-significant. Model 3 (Model 2 with factor loadings of items 2, 5 and 7 on Dedication constrained to be equal) displayed a difference of $\Delta\chi^2_{(3)}$ = 0,87 with Model 2, which is non-significant. Model 4 (Model 3 with factor loadings on Absorption constrained to be equal) displayed a difference of $\Delta\chi^2_{(3)}$ = 9,06 with Model 3, which is non-significant. Model 5 (Model 4 with error covariances constrained to be equal) displayed a difference of $\Delta\chi^2_{(2)}$ = 4,24 with Model 4, which is non-significant. A difference of $\Delta\chi^2_{(3)}$ = 20,05 were found between Model 5 and Model 5 with covariances constrained to be equal.

The results in Table 5 shows that construct equivalence exists between the various language groups. The error covariances for the different language groups were not equivalent. But Byrne (2001) indicates that the equality of error covariances test has a restrictive nature on the data and is relatively unimportant. These results provide support for part of Hypothesis 2 in that work engagement, as measured by the UWES, can be defined as a three-dimensional construct with an equivalent structure for different language groups.

Table 5
Tests of invariance of the UWES in different language groups

Model	Groups	Comparative model	χ^2	*df*	$\Delta\chi^2$	Δdf	*P*
Hypothesised model: one-factor UWES (Model 1)	Afrikaans/English/African		627,53	176	-	-	-
Factor loadings constrained to be equal	Afrikaans/English/African	Model 1	674,89	191	47,36	15	$p < 0,01$
Model 1 with error covariances constrained to be equal	Afrikaans/English/African	Model 1	678,13	193	50,60	17	$p < 0,01$
Hypothesised model: three-factor UWES (Model 1)	Afrikaans/English/African		570,21	172	-	-	-
Factor loadings, variances and covariances constrained to be equal	Afrikaans/English/African	Model 1	632,12	191	61,91	19	$p < 0,01$
Factor loadings constrained to be equal	Afrikaans/English/African	Model 1	613,54	184	43,33	12	$p < 0,01$
Factor loadings on Vigour constrained to be equal	Afrikaans/English/African	Model 1	590,99	176	20,78	4	$p < 0,01$
Factor loading of item 1 on Vigour constrained to be equal	Afrikaans/English/African	Model 1	572,30	173	2,09	1	NS
Factor loadings of items 1 and 4 on Vigour constrained to be equal	Afrikaans/English/African	Model 1	575,64	174	5,43	2	NS
Factor loadings of items 1, 4 and 8 on Vigour constrained to be equal (Model 2)	Afrikaans/English/African	Model 1	577,09	175	6,88	3	NS
Factor loadings of items 1, 4, 8 and 12 on Vigour constrained to be equal	Afrikaans/English/African	Model 1	590,99	176	20,78	4	$p < 0,01$
Model 2 with Factor loadings on Dedication constrained to be equal	Afrikaans/English/African	Model 2	594,19	179	17,10	4	$p < 0,01$
Model 2 with Factor loadings of item 2 on Dedication constrained to be equal	Afrikaans/English/African	Model 2	577,15	176	0,06	1	NS
Model 2 with Factor loadings of items 2 and 5 on Dedication constrained to be equal	Afrikaans/English/African	Model 2	577,90	177	0,81	2	NS
Model 2 with Factor loadings of items 2,5 and 7 on Dedication constrained to be equal (Model 3)	Afrikaans/English/African	Model 2	578,40	178	1,31	3	NS
Model 2 with Factor loadings of items 2,5,7 and 10 on DE constrained to be equal	Afrikaans/English/African	Model 2	594,19	179	17,10	4	$p < 0,01$
Model 3 with factor loadings on AB constrained to be equal (Model 4)	Afrikaans/English/African	Model 3	584,92	182	6,52	4	NS
Model 4 with error covariances constrained to be equal	Afrikaans/English/African	Model 4	592,93	183	8,01	1	$p < 0,01$
Model 4 with covariances constrained to be equal	Afrikaans/English/African	Model 4	599,34	185	14,42	3	$p < 0,01$
Hypothesised model: three-factor adjusted UWES (Model 1)	Afrikaans/English/African		293,36	120	-	-	-
Factor loadings, variances and covariances constrained to be equal	Afrikaans/ English/African	Model 1	354,36	138	61,00	18	$p < 0,01$
Factor loadings constrained to be equal	Afrikaans/ English/African	Model 1	329,76	130	36,40	10	$p < 0,01$
Factor loadings on Vigour constrained to be equal (Model 2)	Afrikaans/English/African	Model 1	301,27	123	7,91	3	NS
Model 2 with factor loadings on Dedication constrained to be equal	Afrikaans/English/African	Model 2	320,81	127	19,54	4	$p < 0,01$
Model 2 with factor loading of item 2 on Dedication constrained to be equal	Afrikaans/English/African	Model 2	301,29	124	0,02	1	NS
Model 2 with factor loading of item 2 and 5 on Dedication constrained to be equal	Afrikaans/English/African	Model 2	301,70	125	0,43	2	NS
Model 2 with factor loading of item 2, 5 and 7 on Dedication constrained to be equal (Model 3)	Afrikaans/English/African	Model 2	302,14	126	0,87	3	NS
Model 2 with factor loading of item 2, 5, 7 and 10 on Dedication constrained to be equal	Afrikaans/English/African	Model 2	320,81	127	19,54	4	$p < 0,01$
Model 3 with factor loadings on Absorption constrained to be equal (Model 4)	Afrikaans/English/African	Model 3	311,20	129	9,06	3	NS
Model 4 with error covariances constrained to be equal (Model 5)	Afrikaans/English/African	Model 4	315,44	131	4,24	2	NS
Model 5 with covariances constrained to be equal	Afrikaans/English/African	Model 5	335,49	134	20,05	3	$p < 0,01$

Descriptive statistics and internal consistency of the scales

The descriptive statistics, alpha coefficients and inter-item correlations of the three factors of the MBI-GS and the three factors of the UWES are given in Table 6.

Table 6
Descriptive statistics, alpha coefficients and inter-item correlations of the MBI-GS and the UWES

Item	Mean	*SD*	Skewness	Kurtosis	*r*(Mean)	α
MBI-GS						
Exhaustion	15,11	0,05	-0,02	-0,58	0,54	0,86
Cynicism	9,16	0,11	0,20	-0,67	0,50	0,80
Professional Efficacy	28,68	0,09	-0,87	0,39	0,36	0,76
UWES						
Vigor	19,93	0,10	-0,58	-0,17	0,45	0,80
Dedication	15,63	0,06	-0,65	-0,28	0,62	0,87
Absorption	17,35	0,10	-0,93	0,94	0,38	0,69

The information in Table 6 indicates that the scores on the factors of the MBI-GS and the factors of the UWES are normally distributed. With regard to the internal consistency of the scales, Exhaustion, Cynicism, Professional Efficacy, Vigour and Dedication seem to demonstrate acceptable coefficient alphas above the 0,70 guideline provided by Nunnally and Bernstein (1994). Furthermore, except for Exhaustion (factor of the MBI-GS) and Dedication (factor of UWES), acceptable levels of inter-item correlations were obtained for all the rest of the factors, consistent with the guideline of $0,15 < r < 0,50$ suggested by Clark and Watson (1995). These results provide support for the aspect of internal consistency of Hypothesis 1 and 2.

The relationship between the MBI-GS and UWES

To assess the relationship between burnout and work engagement, a second-order principal component analysis was carried out on the three scales of the MBI-GS (Exhaustion, Cynicism and Professional Efficacy) and the UWES (Vigour, Dedication and Absorption). This resulted in two related factors ($r = -0,38$) with eigenvalues higher than 1, which explained 75% of the variance. Subsequently, a direct oblimin rotation was carried out on the scales of the MBI-GS and the UWES. The pattern matrix revealed that Exhaustion (loading = 0,95) and Cynicism (0,72) loaded on one factor (labelled *Burnout*), while Professional Efficacy (0,83), Vigour (0,77), Dedication (0,81) and Absorption (0,90) loaded on the second factor (labelled *Work Engagement*). Therefore, our third hypothesis, namely that burnout and work engagement are separate, but related aspects of work-related well-being, was accepted.

DISCUSSION

The aim of this study was to investigate the psychometric properties of the MBI-GS and the UWES for employees in an insurance company in South Africa and to assess the relationship between burnout and work engagement. Firstly, the results supported a three-dimensional factor structure of the MBI-GS, as has been consistently found across various samples, occupational groups and countries (Leiter & Schaufeli, 1996; Schaufeli, Salanova et al., 2002; Schutte et al., 2000; Storm, 2002; Taris et al., 1999). The three-dimensional factor structure of the UWES was also confirmed, a finding supported by research in different samples, groups and countries (Naudé, 2003; Schaufeli, Martinez et al., 2002; Schaufeli, Salanova et al., 2002; Storm & Rothmann, 2003b). Also, reliability analysis confirmed sufficient internal consistency of the subscales of the MBI-GS and the UWES. The construct equivalence of the scales for Afrikaans/African and English participants was confirmed.

Based on both conceptual and empirical grounds, item 13 ("I just want to do my job and not be bothered") was eliminated from the original MBI-GS, resulting in a 15-item scale. This is consistent with the study of Storm (2002), where item 13 was deleted to confirm the three-factor structure of the MBI-GS. Schutte et al. (2000) also excluded this item from a cross-national study of the factorial validity of the MBI-GS. According to these authors problems might be caused by the ambivalent nature of this item. On the one hand, a high score 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 work engagement: one concentrates on the task and does not want to be interrupted.

In examining the factor structure, some undesirable psychometric characteristics were found to be associated with several items of the UWES. Items 16 and 17 (in the initial three-factor model) and items 14 and 20 (in the adjusted three-factor model) showed high standardised residual errors. Additionally, these items had the highest modification indices. These findings suggest that the items may require either deletion or content modification, in which instance the latter must rather be considered. The particular items may be problematic because they do not correspond to the conceptual domain of the particular dimension. However, it is more likely that they are somewhat ambiguous, or that they are either sample- or country-specific. The deletion of items from the UWES for reasons of bias and model-fit improvement resulted in the sacrifice of model parsimony, in other words, relationships have been eliminated which could be viewed as an erosion in meaning of the work engagement construct. Also, it is possible, due to sampling procedure (subgroup representation), that these findings could have been obtained by pure chance. Also, the problems of some of the items may be related to difficult words that some of the participants could have found difficult to understand and/or interpret (e.g. vigorous, immersed and resilient). This resulted in the adjustment of the initial UWES questionnaire with the replacement of items 4, 9, 11 and 15 respectively with items that were written in a more familiar South African vocabulary.

The prominent correlated errors in this study presented another problem. In general, the specification of correlated error terms for the purpose of achieving a better-fitting model is not an acceptable practice. Correlated error terms in measurement models represent systematic, rather than random, measurement error in item responses. They may derive from characteristics specific either to the items or the respondents (Aish & Jšresk og, 1990). For example, if these parameters reflect item characteristics, they may represent a small omitted factor. However, as may be the case in this instance, correlated errors may represent respondent characteristics that reflect bias such as yea-/nay-saying, social desirability (Aish & Jšresk og, 1990), as well as a high degree of overlap in item content (Byrne, 2001).

Previous research with psychological constructs in general (e.g., Jöreskog, 1982; Newcomb & Bentler, 1988; Tanaka & Huba, 1984), and with measuring instruments in particular (Byrne, 1988, 2001), has demonstrated that the specification of correlated errors can often lead to substantially better fitting models. Bentler and Chou (1987) also argue that the specification of a model that forces these error parameters to be uncorrelated is rarely appropriate with real data. Therefore, it was considered more realistic to incorporate the correlated errors in this study, rather than to ignore their presence.

It is believed that this confusing state of affairs regarding the UWES does not reflect weaknesses inherent in the instrument, but is rather due to more general factors. First, the UWES is a recently constructed measuring instrument. Therefore, relatively few studies have critically reviewed its psychometric properties. Secondly, the UWES is an instrument that was originally constructed from data based on samples of individuals in the Netherlands (Schaufeli & Bakker, 2001). Despite a few studies of the UWES in South Africa (e.g. Naudé, 2003; Storm & Rothmann, 2003b), more research regarding work engagement in different occupational settings in South Africa is required. Schaufeli, Martinez et al. (2002) also found that the hypothesised three-factor model of work engagement was invariant across Spanish, Dutch and Portuguese samples.

The results of this study showed that burnout and work engagement are separate aspects of a larger factor, namely affective work-related well-being. This contradicts the view of Maslach and Leiter (1997) that work engagement is the direct opposite of burnout. The results confirmed findings elsewhere that exhaustion and cynicism form part of a core burnout factor, while vigour, dedication, absorption and professional efficacy form part of an extended work engagement factor (Schaufeli & Bakker, 2004).

In conclusion, the results of this study could serve as a standard for measuring burnout and work engagement levels of employees in an insurance company. The three-factor structure of the MBI-GS and the UWES is largely confirmed with acceptable internal consistency of its subscales of Exhaustion, Cynicism, Professional Efficacy, Vigour, Dedication and Absorption.

This study had several limitations. First, items of the measuring instruments were allowed to correlate in the model specification. This may impose interpretation problems because, as correlated error terms are added to the model, the correspondence between the posited construct of interest and the empirically defined factor becomes unclear (Gerbing & Anderson, 1984). Second, the sample size, and specifically the distribution of language groups and the sampling procedure in the present study were limitations. The study, which was conducted in the Western Cape, did not include sufficient numbers of African language speakers, which makes it difficult to assess the validity of the results for African language speakers. Future studies could benefit in terms of a stratified random-sample design, which would ensure sufficient representation of the different groups in the total population of employees in an insurance company. Third, the instruments were administered in English, which could have impacted on the scores of participants.

RECOMMENDATIONS

According to the results obtained in this study, the use of the MBI-GS is recommended to assess burnout and the UWES to assess work engagement in employees in an insurance company.

In the MBI-GS, item 13 should be omitted from the questionnaire and in the UWES, items 14, 16 and 17 should be omitted from the questionnaire in the multicultural context. Item 20 was an item that replaced item 15 due to semantic problems and may therefore need to be rewritten in a more acceptable South African language format.

It is suggested that future research focus on the reliability and validity of the MBI-GS and the UWES for other occupational settings. It is also important to determine norm levels for other occupations in South Africa for both questionnaires respectively. It is recommended that larger samples with a more powerful sampling method be utilised to enable generalisation of the findings to other similar groups. It is also necessary to translate the MBI-GS and the UWES into other languages used in South Africa.

Author's Note
This material is based upon work supported by the National Research Foundation under Grant number 2053344.

REFERENCES

Aish, A.M. & Jöreskog, K.G. (1990). A panel model for political efficacy responsiveness: An application of LISREL 7 with weighted least squares. *Quality and Quantity, 19*, 716– 723.

Arbuckle, J.L. (1999). *Amos 4.0*. Chicago, IL: Smallwaters.

Bentler, P.M. (1995). *EQS: Structural equations program manual.* Encino, CA: Multivariate Software, Inc.

Bentler, P.M. & Chou, C.P. (1987). Practical issues in structural modeling. *Sociological Methods and Research, 16,* 78–117.

Byrne, B.M. (1988). The Self Description Questionnaire III: Testing for equivalent factorial validity across ability. *Educational and Psychological Measurement, 48,* 397–406.

Byrne, B.M. (2001). *Structural equation modeling with AMOS : Basic concepts, applications and programming*. Mahwah, NJ: Erlbaum.

Chan, K.B. (2002). Coping with work stress, work satisfaction, and social support: An interpretive study of life insurance agents. *Journal of Social Science, 30,* 657–685.

Clark, L.A. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309– 319.

Gerbing, D.W. & Anderson, J.C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research, 11,* 572–580.

Green, D.E., Walkey, F.H. & Taylor, A.J.W. (1991). The three-factor structure of the Maslach Burnout Inventory. *Journal of Science Behavior and Personality, 6,* 453–652.

Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409–426.

Jöreskog, K.G. (1982). Analysis of covariance structures. In C. Fornell (Ed.), *A second generation of multivariate analysis Volume l: Methods* (pp. 200–242). New York: Praeger.

Jöreskog, K.G. & Sörbom, D. (1996). *LISREL 8: User's reference guide.* Chicago, IL: Scientific Software International.

Kickul, J. & Posig, M. (2001). Supervisory emotional support and burnout: An explanation of reverse buffering effects. *Journal of Managerial Issues, 13,* 328–345.

Kruger, M.M., Veldman, H., Rothmann, S. & Jackson, L.T.B. (2002, July). *Burnout, job satisfaction, sense of coherence and job stress in a local government.* Paper presented at the 25th International Conference of Applied Psychology, Singapore.

Lai, G., Chan, K.B., Ko, Y.C. & Boey, K.W. (2000). Institutional context and stress appraisal: The experience of life insurance agents in Singapore. *Journal of Asian and African Studies, 35,* 209–228.

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

Lindstršm, K., Leino, T., Seitsamo, J. & Tordtila, L. (1997). A longitudinal study of work characteristics and health complaints among insurance employees in VTD work. *International Journal of Human-Computer Interaction, 9,* 343–368.

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

Maslach, C. & Leiter, M.P. (1997). *The truth about burnout.* San Francisco, CA: Van Nostrand.

Maslach, C., Schaufeli, W.B. & Leiter, M. (2001). Job burnout. *Annual Review of Psychology, 52,* 397–422.

Naudé, J.L.P. (2003). *Occupational stress, coping, burnout and work engagement of emergency workers in Gauteng.* Unpublished doctoral thesis, PU for CHE, Potchefstroom, South Africa.

Newcomb, M.D. & Bentler, P.M. (1988). *Consequences of adolescent drug use: Impact on the lives of young adults.* Newbury Park, CA: Sage.

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Rothmann, S. & Jansen van Vuuren, J.A. (2002, March). *Sense of coherence, self-efficacy, locus of control, coping and burnout of senior managers in a manufacturing industry.* Poster session presented at the 1st South African Burnout Conference, Potchefstroom.

Rothmann, S. & Malan, H. (2003). Koherensiesin, selfdoeltreffendheid, lokus van beheer en uitbranding by maatskaplike werkers. *South African Journal of Industrial Psychology, 29*(4), 43–51.

Salanova, M. & Schaufeli, W.B. (2000). Exposure to information technology and its relation to burnout. *Behaviour and Information Technology, 19*, 385–392.

SAS Institute. (2000). *The SAS System for Windows: Release 8.01*. Cary, NC: SAS Institute Inc.

Schaufeli, W.B. & Bakker, A.B. (2001). Werk en welbevinden: Naar een positieve benadering in de Arbeids- en Gezondheidspsychologie [Work and wellbeing: Towards a positive occupational health psychology]. *Gedrag en Organizatie, 14*, 229–253.

Schaufeli, W.B. & Bakker, A.B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behaviour, 25*, 1–23.

Schaufeli, W.B., Bakker, A.B, Hoogduin, K., Schaap, C. & Kladler, A. (2001). On the clinical validity of the Maslach Burnout Inventory and the burnout measure. *Psychology and Health, 16*, 565–583.

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

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* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.

Schaufeli, W.B., Martinez, I., Pinto, A.M., Salanova, M. & Bakker, A.B. (2002). Burnout and engagement in university students: A cross national study. *Journal of Cross-Cultural Psychology, 33*, 464–481.

Schaufeli, W.B., Salanova, M., González-Romá, V. & Bakker, A.B. (2002). The measurement of engagement and burnout: A confirmation analytic approach. *Journal of Happiness Studies, 3,* 71–92.

Schutte, N., Toppinen, S., Kalimo, R. & Schaufeli, W.B. (2000). The factorial validity of the Maslach Burnout Inventory – General Survey (MBI-GS) across occupational groups and nations. *Journal of Occupational and Organizational Psychology, 73*, 53–66.

Storm, K. (2002). *Burnout and work engagement in the South African Police Service.* Unpublished doctoral thesis, PU for CHE, Potchefstroom.

Storm, K. & Rothmann, S. (2003a, June). *A psychometric analysis of the Maslach Burnout Inventory in the South African Police Service.* Paper presented at the 6th Annual Conference of the Society of Industrial and Organisational Psychology of South Africa, Johannesburg.

Storm, K. & Rothmann, S. (2003b). The validation of the Utrecht Work Engagement Scale in the South African Police Services. *South African Journal of Industrial Psychology, 29*(4), 62–70.

Tanaka, J.S. & Huba, G.J. (1984). Confirmatory hierarchical factor analysis of psychological distress measures. *Journal of Personality and Social Psychology, 46,* 621–635.

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 and Stress, 13*, 223–237.

Van de Vijver, F. & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.

Weisberg, J. (1994). Measuring workers' burnout and intention to leave. *International Journal of Manpower, 15*(1), 4–14.