Educational and Psychological Measurement
Volume 66 Number 4
August 2006 701-716

10.1177/0013164405282471
http://epm.sagepub.com
hosted at
http://online.sagepub.com

The Measurement of Work Engagement With a Short Questionnaire

A Cross-National Study

Wilmar B. Schaufeli
Arnold B. Bakker
Utrecht University
Marisa Salanova
Jaume I University

This article reports on the development of a short questionnaire to measure work engagement—a positive work-related state of fulfillment that is characterized by vigor, dedication, and absorption. Data were collected in 10 different countries (N = 14,521), and results indicated that the original 17-item Utrecht Work Engagement Scale (UWES) can be shortened to 9 items (UWES-9). The factorial validity of the UWES-9 was demonstrated using confirmatory factor analyses, and the three scale scores have good internal consistency and test-retest reliability. Furthermore, a two-factor model with a reduced Burnout factor (including exhaustion and cynicism) and an expanded Engagement factor (including vigor, dedication, absorption, and professional efficacy) fit best to the data. These results confirm that work engagement may be conceived as the positive antipode of burnout. It is concluded that the UWES-9 scores has acceptable psychometric properties and that the instrument can be used in studies on positive organizational behavior.

Keywords: *work engagement; measurement; burnout; Utrecht Work Engagement Scale (UWES)*

Since the beginning of this century, increased attention has been paid to what has been coined *positive psychology*: the scientific study of human strength and optimal functioning (Seligman & Csikszentmihalyi, 2000). This approach is considered

Authors' Note: The authors express their gratitude to the following colleagues who have provided data (in alphabetical order): Sarah Jane Cotton (Australia), Edgar Bresó (Spain), Jan de Jonge (Netherlands), Maureen Dollard (Australia), Martin Euwema (Netherlands), Esther Greenglass (Canada), Asbjørn Grimsmo (Norway), Gabriele Haeslich (Germany), Jari Hakanen (Finland), Sandrine Hollet (France), Saar Langelaan (Netherlands), Pascale Le Blanc (Netherlands), Susana Llorens (Spain), Stig Berge Matthiesen (Norway), Maria Peeters (Netherlands), Astrid Richardsen (Norway), Peter Richter (Germany), Ian Rothmann (South Africa), Katariina Salmela-Aro (Finland), Sabine Sonnentag (Germany), Willem van

to supplement the traditional focus of psychology on disease, damage, disorder, and disability. The recent trend to concentrate on optimal functioning also has aroused attention in organizational psychology, as illustrated by Luthans's (2002) recent plea for "the study of positively oriented human resource strengths and psychological capacities that can be measured, developed, and effectively managed for performance improvement in today's workplace" (p. 698).

One of these positive states is work engagement, which is considered to be the antipode of burnout. The current article is about the development and psychometric evaluation of a short self-report questionnaire to measure work engagement. Contrary to those who suffer from burnout, engaged employees have a sense of energetic and effective connection with their work activities, and they see themselves as able to deal well with the demands of their jobs. Work engagement is defined as a positive, fulfilling work-related state of mind that is characterized by vigor, dedication, and absorption (Schaufeli & Salanova, in press; Schaufeli, Salanova, Gonzalez-Romá, & Bakker, 2002).

Rather than a momentary and specific state, engagement refers to a more persistent and pervasive affective-cognitive state that is not focused on any particular object, event, individual, or behavior. Vigor is characterized by high levels of energy and mental resilience while working, the willingness to invest effort in one's work, and persistence even in the face of difficulties. Dedication refers to being strongly involved in one's work and experiencing a sense of significance, enthusiasm, inspiration, pride, and challenge. Finally, absorption is characterized by being fully concentrated and happily engrossed in one's work, whereby time passes quickly and one has difficulties with detaching oneself from work. Accordingly, vigor and dedication are considered direct opposites of the core burnout dimensions of exhaustion and cynicism, respectively (Maslach, Schaufeli, & Leiter, 2001). Therefore, particularly the correlations between vigor and exhaustion and between dedication and cynicism are expected to be strongly negative. The remaining dimensions of burnout (i.e., professional efficacy) and of work engagement (i.e., absorption) are distinct aspects that are not considered as opposites.

Based on the above-mentioned definition, a self-report questionnaire—the Utrecht Work Engagement Scale (UWES)—has been developed that includes the three constituting dimensions of work engagement: vigor, dedication, and absorption. Originally, the UWES included 24 items, but after psychometric evaluation, 7 unsound items were eliminated so that three scales, totaling 17 items, remained (Schaufeli, Salanova, et al., 2002): Vigor (VI, 6 items), Dedication (DE, 5 items), and Absorption (AB, 6 items) scales (see the appendix). Using a large international database, the current article seeks to reduce the number of items of the UWES. The reason for shortening the UWES is basically pragmatic: Researchers strive to include as few items as possible

---

Rhenen (Netherlands), Toon Taris (Netherlands), Peter Vlerick (Belgium), Tony Winefield (Australia), Hans de Witte (Belgium), and Dieter Zapf (Germany). Please address correspondence to Wilmar B. Schaufeli, PhD, Utrecht University, Department of Psychology, P.O. Box 80140, 3508 TC Utrecht, the Netherlands.

for measuring a particular construct because respondents should not be unnecessarily bothered. Besides, long questionnaires increase the likelihood of attrition.

The original UWES-17 has encouraging psychometric features for its scores. For instance, internal consistencies (Cronbach's alpha) typically range between .80 and .90 (Demerouti, Bakker, Janssen, & Schaufeli, 2001; Durán, Extremera, & Rey, 2004; Montgomery, Peeters, Schaufeli, & Den Ouden, 2003; Salanova, Schaufeli, Llorens, Peiró, & Grau, 2001; Schaufeli & Bakker, 2004). Thus, values of Cronbach's alpha exceed the value of .70 that is traditionally used as a rule of thumb (Nunnally & Bernstein, 1994), and even more so, in almost all cases, alpha satisfies the more stringent value of .80 that is now considered a generally accepted standard (Henson, 2001).

Furthermore, confirmatory factor analyses have shown that the hypothesized three-factor structure of the UWES is superior to the one-factor model (Schaufeli, Martínez, Marques-Pinto, Salanova, & Bakker, 2002; Schaufeli, Salanova, et al., 2002). However, in contrast, using explorative factor analyses, Sonnentag (2003) did not find a clear three-factor structure and decided to use the total score of the UWES as a measure for work engagement. Furthermore, the three-factor structure of the slightly adapted student version of the UWES is largely invariant across samples from Spain, the Netherlands, and Portugal (Schaufeli, Martínez, et al., 2002). In a similar vein, the UWES can be used as an unbiased instrument to measure work engagement because its equivalence is acceptable for different racial groups (Storm & Rothmann, 2003). Although confirmatory factor analyses have supported the three-dimensional structure of the UWES, the dimensions are very closely related. That is, correlations between the scales usually exceed .65 (e.g., Demerouti et al., 2001; Salanova et al., 2001; Schaufeli, Martínez, et al., 2002; Schaufeli, Salanova, et al., 2002), whereas correlations between latent variables of a covariance structure model range from about .80 to more than .90 (Salanova et al., 2001; Schaufeli, Martínez, et al., 2002; Schaufeli, Salanova, et al., 2002).

Because engagement has been defined as the opposite of burnout (Maslach et al., 2001), it is expected that both concepts are negatively related. Indeed, the three burnout dimensions—as measured with the Maslach Burnout Inventory (MBI; Maslach, Jackson, & Leiter, 1996)—correlate negatively with the three dimensions of work engagement (Demerouti et al., 2001; Montgomery et al., 2003; Salanova et al., 2001; Schaufeli & Bakker, 2004; Schaufeli, Martínez, et al., 2002; Schaufeli, Salanova, et al., 2002). However, the pattern of relationships slightly differs from what was expected. Namely, the negative correlations between vigor and exhaustion and between dedication and cynicism do not appear to be the strongest but, instead, the correlations between lack of professional efficacy and all three aspects of engagement. In the discussion, we will elaborate on this. As a consequence, a second-order factor model, in which the three engagement scales weighted together with professional efficacy constitute on one factor (Engagement) and exhaustion and cynicism on the other factor (Burnout), fits best to the data (Salanova et al., 2001; Schaufeli & Bakker, 2004; Schaufeli, Salanova, et al., 2002). A similar result was obtained by Demerouti et al. (2001) using discriminant analyses. In this study, the three engagement scales

Table 1
Countries Included in the Database

Country	*n*	%
1. Australia	473	3.3
2. Belgium[a]	767	5.3
3. Canada	267	1.8
4. Finland	3,651	25.1
5. France	221	1.5
6. Germany	465	3.2
7. The Netherlands[a]	2,163	14.9
8. Norway[a]	2,114	14.6
9. South Africa	2,547	17.5
10. Spain	1,832	12.6
Total	14,521	100.0

a. Burnout data are (partly) not available (see text).

plus professional efficacy weighted on one discriminant function, whereas both other burnout scales weighted on the second remaining function.

The aims of the current study are five-fold: (a) to shorten the UWES as much as possible; (b) to compare the fit of a one-factor model of the short version to that of the three-factor model and to evaluate the cross-national invariance of both models; (c) to study some psychometric features of the short version of the UWES (i.e., internal consistency, stability, and correlations with the original scales); (d) to analyze the relationship between engagement and burnout, whereby based on previous studies (see above), we expect that the two-factor model with a reduced Burnout factor (including exhaustion and cynicism) and an expanded Engagement factor (including vigor, dedication, absorption, and professional efficacy) fits the data best; and (e) to carry out some descriptive analyses with the short UWES version, evaluating its relationship with gender, age, and occupation.

Method

Samples and Procedure

A database was constructed of 27 studies that have been carried out between 1999 and 2003 in 10 different countries (see Table 1). In all studies, the work engagement and burnout questionnaires were included as part of a larger employee well-being survey. In most cases, this survey was distributed at the work site; in 6 studies, the surveys were sent to the employees' home addresses. The surveys that were used in the various studies were distributed at different times of the year during the study period, and the order of both questionnaires in the survey differed.

Table 2
Occupational Groups Included in the Database

Occupational Group	*n*	%
Social work	822	5.8
Blue-collar	1,024	7.2
Health care	2,777	18.8
White-collar (profit)	1,374	9.7
White-collar (not for profit)	147	1.2
Teaching	3,041	21.4
Police	2,650	18.7
Management	871	6.1
Information missing	314	2.2
Total	14,521	100.0

Table 2 presents an overview of the occupational groups that are included in the database. The database includes slightly more men (n = 7,621 men [53.3%] vs. 6,684 women [46.7%]), and ages range from 16 to 68 years (M = 40.3; SD = 11.7).

Instruments

Engagement was assessed with the UWES (Schaufeli, Salanova, et al., 2002). A test manual of the UWES, as well as several language versions, may be downloaded from www.schaufeli.com. The items of the UWES (see the appendix) are grouped into three subscales that reflect the underlying dimensions of engagement: VI (6 items), DE (5 items), and AB (6 items). All items are scored on a 7-point frequency rating scale ranging from 0 (*never*) to 6 (*always*).

Burnout was assessed with the MBI–General Survey (MBI-GS; Schaufeli, Leiter, Maslach, & Jackson, 1996). The MBI-GS includes three subscales: Exhaustion (EX, 5 items; e.g., "feel used up at the end of a work day"), Cynicism (CY, 5 items; e.g., "doubt the significance of my work"), and Professional Efficacy (PE, 6 items; e.g., "I can effectively solve the problems that arise in my work"). The burnout items are scored in a similar manner as the items of the UWES. However, all PE items are reversibly scored so that high scores on EX, CY, and PE (i.e., lack of professional efficacy) are indicative of burnout. Internal consistencies (Cronbach's alpha) across the various countries range between .72 and .90, .73 and .86, and .73 and .83 for EX, CY, and PE, respectively. Previous studies have demonstrated the cross-national validity of the MBI-GS scores across samples of Finnish, Swedish, and Dutch employees (Schutte, Toppinnen, Kalimo, & Schaufeli, 2000) as well as their validity across various occupations (Bakker, Demerouti, & Schaufeli, 2002; Leiter & Schaufeli, 1996).

Analyses

Structural equation modeling as implemented by AMOS (Arbuckle, 1997) was used for data analysis. Model testing was carried out in all 10 national samples simultaneously by using multiple-group analyses. First, two factor-analytic models were tested: a model that assumes that all engagement items weight on one single factor (M1) and a model that assumes three correlated factors—Vigor, Dedication, and Absorption (M2). Because both models are nested, the χ^2 difference test can be used to assess the best-fitting model. Following Taris, Bok, and Meijer (1998), the factorial invariance of M1 and M2 was investigated by constraining the factor coefficients and the factor covariances to be equal across all national samples, respectively. When the fit of the constrained model to the data is not significantly statistically worse than the fit of the unconstrained model, invariance has been demonstrated. This means that the factor coefficients or the covariances between the factors do not differ significantly between countries.

Next, 3 second-order two-factor models were tested across samples, again using the multiple-group method: (a) the hypothesized model that assumes that EX and CY relate to a reduced Burnout factor and that all three engagement dimensions plus professional efficacy weight on an extended Engagement factor (M_{Hyp}), (b) the one-factor model that assumes one underlying factor (Well-Being) including all burnout and engagement dimensions (M_{Wel}), and (c) the alternative model that assumes that the three MBI dimensions weight on one factor, whereas the three UWES dimensions weight on another factor (M_{Alt}). Finally, the invariance across countries of the best-fitting model was assessed by comparing the unconstrained model with its constrained counterparts (see above).

Fit indices. Maximum likelihood estimation methods were used, and the input for each analysis was the covariance matrix of the items or the scale scores. A check on the normal distribution of the items revealed that the skewness and kurtosis of virtually all engagement items were within the acceptable range (±1.96), only in the French sample the distribution of three items (VI1, VI2, DE2) was slightly peaked. In five countries, the kurtosis of the burnout item "In my opinion, I am good at my job" (PE) exceeded the critical value, whereas in the German sample, another four burnout items were peaked as well.

The goodness of fit of the models was evaluated using the following absolute goodness-of-fit indices (cf. Jöreskog & Sörbom, 1986): (a) the c2 goodness-of-fit statistic, (b) the root mean square error of approximation, (c) the goodness-of-fit index, and (d) the adjusted goodness-of-fit index. Because c2 is sensitive to sample size (i.e., the probability of rejecting a hypothesized model increases with sample size), the use of relative goodness-of-fit measures is strongly recommended (Bentler, 1990). Following Marsh, Balla, and Hau (1996), three relative goodness-of-fit measures were calculated: (a) normed fit index, (b) nonnormed fit index (NNFI), and (c) comparative fit index. Because the distribution of the goodness-of-fit index and the adjusted goodness-of-fit index is unknown, no statistical test or critical value is available (Jöreskog

& Sörbom, 1986). Values smaller than .08 for the root mean square error of approximation are indicative of an acceptable fit, and values greater than .1 should lead to model rejection (Cudeck & Browne, 1993). For all three relative fit indices, as a rule of thumb, values greater than .90 are considered as indicating a good fit (Hoyle, 1995). More recently, Hu and Bentler (1999) have recommended slightly higher cutoff values, such as .95 for the comparative fit index.

Results

Construction of the Short Version of the UWES

To reduce the number of items of the UWES as much as possible, an iterative process was carried out, whereby the samples of each country were analyzed separately. First, the most characteristic item of each scale was selected on the basis of face validity. Next, this item was regressed on the remaining items of that particular scale. The item with the highest β value was then added to the initial item. In the next step, the sum of these two items was regressed on the remaining items of the scale, and again the item with the highest β value was added to both items that were previously selected. Next, the sum of these three items was regressed on the remaining items of that scale, and so on. This iterative procedure was aborted either when no substantial variance was added by a subsequent item or when no similar additional item emerged across the 10 countries.

The most characteristic item for the VI scale was "At my work, I feel bursting with energy" (VI1). This item was supplemented in the next step by "At my job, I feel strong and vigorous" (VI2), as this item had the highest β values in all countries (ranging between .21 and .59; median = .42). Both items were summed and regressed on the remaining four VI items, whereby the item "When I get up in the morning, I feel like going to work" (VI3) showed the highest β values in almost all countries (ranging between .12 and .71; median = .39). Only in Spain and Finland did another item (VI6) show slightly higher β values. Regressing the sum of the three items (i.e., VI1, VI2, and VI3) on the remaining three VI items did not yield a particular additional item that could be included in the short version. Thus, the final short VI scale consists of VI1, VI2, and VI3.

The most characteristic item for the DE scale reads "I am enthusiastic about my job" (DE2). This item was supplemented by "My job inspires me" (DE3) because it had the highest β value in all but two countries (ranging between .20 and .49; median = .32). In Belgium and Canada, another DE item (DE4) had slightly higher β values. Next, DE2 and DE3 were summed and regressed on the remaining three DE items. The item "I am proud of the work that I do" (DE4) showed the highest β values in all countries (ranging between .12 and .48; median = .27). Regressing the sum of the three items (i.e., DE2, DE3, and DE4) on the remaining two DE scale items did not yield a particular additional item that could be included in the short version. Thus, the final short DE scale consists of DE2, DE3, and DE4.

Table 3
The Fit of the UWES-9 Models

Model	χ^2	*df*	GFI	AGFI	RMSEA	NFI	NNFI	CFI
One-factor model (M1)								
Freely estimated	6144.52	270	.89	.82	.04	.91	.89	.91
Constrained factor coefficients	7333.87	342	.88	.84	.04	.89	.89	.90
Three-factor model (M2)								
Freely estimated	3227.29	240	.95	.90	.03	.95	.93	.96
Constrained factor coefficients	4180.18	294	.93	.89	.03	.94	.93	.94
Constrained covariances	3504.17	267	.94	.90	.03	.95	.94	.95
Null model	63064.50	36	.33	.16	.35	—	—	—

Note: Multiple-group method employed (N = 14,521). UWES = Utrecht Work Engagement Scale; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; RMSEA = root mean square error of approximation; NFI = normed fit index; NNFI = nonnormed fit index; CFI = comparative fit index. For a description of the models, see text.

The most characteristic item for the AB scale was "I am immersed in my work" (AB4). This item was supplemented by "I get carried away when I'm working" (AB5) because it had the highest β values in all but two countries (ranging between .21 and .47; median = .27). In Belgium and Spain, another AB item (AB3) had slightly higher β values. Next, AB4 and AB5 are summed and regressed on the remaining four AB items. The item "I feel happy when I am working intensely" (AB3) showed the highest β values in almost all countries (ranging between .20 and .62; median = .37). Only in Canada and Spain did another item (AB6) show slightly higher β values. Regressing the sum of the three items (i.e., AB3, AB4, and AB5) on the remaining three AB scale items did not yield a particular additional item that could be included in the short version. Thus, the final short AB scale consists of AB3, AB4, and AB5.

Factorial Validity

The one-factor (M1) and three-factor (M2) models were fitted to all 10 national samples simultaneously. M2 fit quite well to the data with all fit indices meeting their corresponding minimums/maximums for acceptability (see Table 3). Moreover, in all national samples, all items had statistically significant coefficients on their latent factors. Although M1 also fit reasonably well to the data—with only NNFI not meeting (but approaching) its criterion of .90—the fit of M2 is superior to that of M1: $\Delta\chi^2(30) = 2917.23$, $p < .001$. The reason for this is that the latent factors of M2 were highly correlated in the various national samples: $.88 < r_{\text{vi}-\text{de}} < .98$ (median = .95), $.75 < r_{\text{de}-\text{ab}} < .97$ (median = .92), and $.75 < r_{\text{vi}-\text{ab}} < .96$ (median = .90).

In the next step, the factor coefficients in both models were constrained to be equal across all national samples. Because the fit of both constrained models deteriorated significantly—for M1: $\Delta\chi^2(72) = 1189.35$, $p < .001$; for M2: $\Delta\chi^2(54) = 952.89$, $p < .001$—it was concluded that the factor coefficients differed systematically across

countries. Finally, a model with the covariances of the factors constrained to be equal across samples was simultaneously fitted to the data of all national samples. Compared to the unconstrained three-factor model, the fit of this constrained model also deteriorated significantly: $\Delta\chi^2(27) = 267.88$, $p < .001$. Hence, the covariances of the three latent engagement factors differed significantly between countries.

In sum, the three-factor model fit slightly better to the data of the 10 countries than did the one-factor model. However, the three subscales of the short version of the UWES are highly interrelated. The one-factor and three-factor models are not invariant; that is, factor coefficients and covariances between factors differ across countries.

Additional Psychometric Analyses

Internal consistency. Cronbach's alpha of the three-item VI scale varied across countries between .60 and .88 (median = .77), with Finland ($\alpha = .65$) and France ($\alpha = .60$) as the only 2 countries with values lower than .70. Cronbach's alpha of the short DE scale varied between .75 and .90 (median = .85), whereas alpha values for the three-item AB scale varied between .66 and .86 (median = .78), with Spain ($\alpha = .66$) as the only country with a value lower than .70. Finally, Cronbach's alpha for the total nine-item scale varied between .85 and .92 (median = .92) across all 10 countries. In sum, with very few exceptions, the internal consistencies of the scores of the three-item scales seem to be satisfactory in the sense that they exceed the value of .70 (Nunnally & Bernstein, 1994), whereas values of Cronbach's alpha of the total nine-item scale are good across all national samples, satisfying the more stringent value of .80 (Henson, 2001).

Stability. In two countries—Australia ($n = 293$) and Norway ($n = 2{,}111$)—the UWES was administered twice with an interval of 1 year. The stability coefficients for VI, DE, and AB for Australia were .61, .56, and .60, respectively, and for Norway were .71, .66, and .68, respectively. The corresponding values of the total nine-item score for Australia and Norway were .64 and .73, respectively.

Correlations with the original scales. The correlations between the short three-item VI scale and the original six-item scale exceeded .90 in all countries except France ($r = .83$; median = .91). The correlations between the short three-item DE scale and the original five-item scale exceeded .95 in all countries (median = .96). The correlations between the short three-item AB scale and the original six-item scale exceeded .90 in all countries, except Belgium ($r = .85$) and Spain ($r = .89$; median = .92). In sum, with only few exceptions, the short UWES scales share more than 80% of their variances with the corresponding longer original versions.

Relationship With Burnout

For Norway and Belgium, as well as for a part of the Dutch sample ($n = 488$), no burnout data were available, thus leaving a total sample of $n = 11{,}152$ to be used for

Table 4
The Fit of the Second-Order Burnout and Engagement Models

Model	χ^2	*df*	GFI	AGFI	RMSEA	NFI	NNFI	CFI
One-factor model (M_{Wel})	4894.11	72	.88	.72	.08	.84	.74	.85
Alternative model (M_{Alt})	3325.19	64	.91	.77	.07	.89	.81	.90
Hypothesized model (M_{Hyp})								
Freely estimated	1545.33	64	.96	.90	.05	.95	.91	.95
Constrained factor coefficients	3291.90	92	.92	.85	.06	.90	.87	.90
Constrained covariance	1615.67	71	.96	.90	.04	.95	.92	.95
Null model	31509.94	120	.47	.26	.15	—	—	—

Note: Multiple-group method employed (total sample, n = 11,152). GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; RMSEA = root mean square error of approximation; NFI = normed fit index; NNFI= nonnormed fit index; CFI = comparative fit index. For a description of the models, see text.

investigating the relationship between work engagement and burnout. As expected, burnout and engagement were negatively correlated; this was especially true for vigor and exhaustion ($-.32 < r < -.44$; median = –.40) and for dedication and cynicism ($-.37 < r < -.54$; median = –.50). After all, these dimensions were meant to be each other's direct opposites. Interestingly, reduced professional efficacy was consistently and substantively correlated with all three engagement dimensions ($-.36 < r < -.73$; median = –.49), suggesting that M_{Hyp} would probably be the best-fitting model.

As shown in Table 4, the fit of the hypothesized model (M_{Hyp}) was indeed superior to that of the one-factor model (M_{Wel}), $\Delta\chi^2(8) = 3348.78$, $p < .001$, as well as to that of the alternative model (M_{Alt}). Moreover, the so-called modification indices of M_{Alt} indicated that the fit of the model could be improved significantly when instead of relations to burnout, professional efficacy would weight on engagement. All scales had statistically significant coefficients on the corresponding latent factor of M_{Hyp}, whereas the correlations between the latent Burnout and Engagement factors ranged between –.45 and –.64 (median = –.58).

In the next and final step, all factor coefficients and the covariance of the latent burnout and engagement factors of M_{Hyp} were constrained to be equal across all national samples, respectively. Compared to the freely estimated M_{Hyp}, the fit of both constrained models deteriorated significantly: For the constrained factor coefficients, $\Delta\chi^2(28) = 1746.57$, $p < .001$, and for the constrained covariance, $\Delta\chi^2(7) = 70.34$, $p < .001$. Hence, it is concluded that the factor coefficients as well as the covariance between the latent Burnout and Engagement factors differed systematically across the eight countries involved.

Descriptive Analyses: Relationships With Age, Gender, and Occupational Group

Engagement was weakly positively related with age; that is, correlations with the short versions of VI, DE, and AB ranged from .00 to .28 (median = .08), 02 to .28

(median = .09), and .00 to .27 (median = .12), respectively. Although in many cases statistically significant, these correlations generally lacked practical significance. For instance, with the exception of the Canadian sample, correlations in all other samples were .15 or less.

Relationships between work engagement and gender were weak but equivocal. In the Australian, Canadian, and French samples, no gender differences were observed. On the other hand, in the Belgian, German, Finnish, and Norwegian samples, men scored slightly higher on the three engagement dimensions than did women, whereas the reverse was true for the South African (only VI), Spanish (only DE and AB), and Dutch samples. However, the gender differences also lack practical significance, which is illustrated by the fact that in all cases, Cohen's *d*—a statistic for effect size that is independent of sample size—was lower than .20 (Cohen, 1969). In the present study, such low effects were not considered meaningful.

To explore the relationship of engagement with occupational group, a pooled sample was used because not all occupational groups were represented in each country. Systematic differences were found between occupational groups in levels of vigor, $F(7, 13644) = 78.30$, $p < .001$, dedication, $F(7, 13630) = 84.24$, $p < .001$, and absorption, $F(7, 13635) = 90.38$, $p < .001$. The highest levels of vigor were found among educators ($M = 4.41$), managers ($M = 4.40$), and police officers ($M = 4.14$), whereas the lowest scores were observed for blue-collar workers ($M = 3.47$), social workers and counselors ($M = 3.89$), and health care workers ($M = 3.94$). The highest levels of dedication were found among police officers ($M = 4.55$), managers ($M = 4.48$), and educators ($M = 4.40$), whereas the lowest scores were observed for blue-collar workers ($M = 3.40$), white-collar workers in the nonprofit sector ($M = 4.14$), and social workers and counselors ($M = 4.17$). The highest levels of absorption were found among police officers ($M = 4.05$), managers ($M = 3.78$), and educators ($M = 3.70$), whereas the lowest scores were observed for blue-collar workers ($M = 2.74$), white-collar workers in the nonprofit sector ($M = 3.49$), and health care workers ($M = 3.55$). Post hoc tests revealed that all differences between the high-scoring occupations and the low-scoring occupations were statistically significant ($p < .001$). The Cohen's *d* effect sizes of the differences between blue-collar workers and the three highest-scoring occupational groups (i.e., police officers, managers, and educators) exceeded .80 and have to be qualified as "strong," whereas the differences between the remaining low-scoring groups with the three highest-scoring groups have to be qualified as "small" ($.20 < d < .30$). In other words, particularly blue-collar workers are less engaged in their work than are police officers, managers, and educators.

Discussion

Using a large international database, the current study set out to develop a short questionnaire to measure work engagement and validate its scores. The point of departure was a longer scale that was recently introduced—the 17-item UWES (Schaufeli, Salanova, et al., 2002). Scale construction was successful because after an iterative

process, the three original scales of the UWES could be reduced to 3 items each. The shortened versions of the scales correlated highly with their original longer counterparts, sharing more than 80% of their variances. Furthermore, internal consistencies of the scores from the three short scales were sufficient in almost all 10 countries that were included in the database. In fact, in only 3 of 30 cases (10%), Cronbach's alpha was slightly lower than .70; in 7 cases (23%), values of alpha ranged between .70 and .80; whereas in the remaining 27 cases (67%), alpha exceeded .80.

Although the three-factor model including vigor, dedication, and absorption fit significantly better to the data than did the one-factor model that assumed that all items weighted on one underlying Engagement factor, this result was not unequivocal. First, the one-factor model also fit reasonably well to the data, with three of four fit indices meeting their criterion and the remaining index (NNFI) approaching its criterion of .90. Second, correlations between the latent Vigor, Dedication, and Absorption factors were very high with medians > .90 across the national samples. Finally, without exception, the internal consistency of the scores of the total nine-item version appeared to be very high in all national samples. So, practically speaking, rather than computing three different scores for VI, DE, and AB, researchers might consider using the total nine-item score as an indicator of work engagement. In doing so, for instance, problems with multicollinearity are avoided when VI, DE, and AB are entered simultaneously as independent predictors in a regression equation. Alternatively, researchers could consider structure coefficients when using all three scales as predictors (cf. Courville & Thompson, 2001). On the other hand, when using structural equation modeling, the three aspects may be used as indicators of the latent engagement construct. For the time being, it seems that the total UWES-9 score can be used as an overall measure of work engagement. However, a final conclusion as to using a single composite engagement score versus three scale scores still stands out. Future research should uncover whether VI, DE, and AB have different causes and consequences so that instead of a single score, a differentiation between the three aspects would be preferred.

As far as the 1-year stability of engagement is concerned, this is of the same magnitude as burnout (Schaufeli & Enzmann, 1998, pp. 96-97): Between 31% and 53% of the variance in (aspects of) engagement of the second measurement is explained by the first measurement. This means that, like burnout, engagement is a chronic rather than a transient state.

As expected, engagement as measured with the UWES-9 is negatively related to burnout. This is particularly the case for the direct opposites vigor and exhaustion, as well as for dedication and absorption. In addition, it appeared that professional efficacy is rather strongly related to all three engagement dimensions. Hence, our hypothesized two-factor model was confirmed by the data: exhaustion and cynicism weighted on a factor representing the "core of burnout" (cf. Green, Walkey, & Taylor, 1991), whereas vigor, dedication, absorption, and professional efficacy weighted on an extended Engagement factor. Two explanations might be given for this result. First, during the past decade, evidence has accumulated on the divergent role that lack of professional efficacy plays in burnout as compared to exhaustion and cynicism (e.g.,

Lee & Ashforth, 1996; Leiter, 1993), whereas conceptually speaking, it makes sense that engaged workers feel efficacious in their jobs (Maslach & Leiter, 1997). Future longitudinal research should uncover whether professional efficacy might be considered a consequence (or an antecedent) of engagement rather than a constituting element. Based on the cross-sectional analyses of the current study, it can only be concluded that work engagement is related to professional efficacy, but no conclusion can be drawn about any causal order. An alternative explanation might be that lack of professional efficacy is measured with items that are positively formulated and that are subsequently reversed to constitute a "negative" score that is supposed to be indicative of lack of professional efficacy. Recently, Bouman, Te Brake, and Hoogstraten (2002) showed that the notoriously low negative correlations between professional efficacy and both other burnout dimensions change dramatically in much higher positive correlations when, instead of reversing positively formulated items, negative items are used to measure lack of efficacy. Future research that includes negatively worded inefficacy items instead of positive PE items should demonstrate whether a two-factor model in which inefficacy positively weights on burnout fits better to the data than a model in which inefficacy negatively weights on engagement.

As a rule, no indications were found for factorial invariance across samples from the various countries. This means that the structure of the relationships between the items (and scales) is similar across countries but that the size of the estimates (i.e., factor coefficients and covariances of the factors) differs systematically. Earlier studies using the original 17-item UWES showed its invariance across countries (Schaufeli, Martínez, et al., 2002) and racial groups (Storm & Rothmann, 2003). However, in these cases, university students from various countries or police officers from different racial groups were included, respectively. In other words, factorial invariance was demonstrated for members of the same group originating from different countries or racial backgrounds. In contrast, the present study includes different occupational groups from different countries. As a result, factorial invariance is less likely to be observed. Therefore, future research on factorial invariance of the UWES should include similar occupational groups from different countries.

Contrary to the idea that burnout decreases with age (Schaufeli & Enzmann, 1998, p. 76), it seems that work engagement slightly increases with age. However, the relationship with age is so weak that it can hardly be considered meaningful. Except for cynicism—men are usually more cynical than are women—no systematic gender differences were observed for burnout (Schaufeli & Enzmann, 1998, p. 76). Also, levels of engagement did not seem to differ systematically between both genders. As far as occupational groups are concerned, it appeared that blue-collar workers were less engaged compared to managers, educators, and police officers. A possible explanation might be that compared to the latter, the former might draw less on job resources that are known to be positively related to work engagement (Schaufeli & Bakker, 2004). However, like the relationships with age and gender, the relationships between engagement and occupational group should be interpreted with caution because instead of using representative national samples, convenience samples have been used. For instance, the Finnish sample consisted largely of educators, whereas the

South African sample consisted largely of police officers. In conclusion, we hope that the introduction of this short questionnaire to measure engagement, which seems to have encouraging psychometric features, stimulates further research on positive organizational psychology.

Appendix
Work and Well-Being Survey (UWES)

The following 17 statements are about how you feel at work. Please read each statement carefully and decide if you ever feel this way about your job. If you have never had this feeling, cross the "0" (zero) in the space after the statement. If you have had this feeling, indicate how often you felt it by crossing the number (from 1 to 6) that best describes how frequently you feel that way.

Never	Almost Never	Rarely	Sometimes	Often	Very Often	Always
0	1	2	3	4	5	6
Never	A few times a year or less	Once a month or less	A few times a month	Once a week	A few times a week	Every day

1. At my work, I feel bursting with energy.[a] (VI1)
2. I find the work that I do full of meaning and purpose. (DE1)
3. Time flies when I am working. (AB1)
4. At my job, I feel strong and vigorous.[a] (VI2)
5. I am enthusiastic about my job.[a] (DE2)
6. When I am working, I forget everything else around me. (AB2)
7. My job inspires me.[a] (DE3)
8. When I get up in the morning, I feel like going to work.[a] (VI3)
9. I feel happy when I am working intensely.[a] (AB3)
10. I am proud of the work that I do.[a] (DE4)
11. I am immersed in my work.[a] (AB4)
12. I can continue working for very long periods at a time. (VI4)
13. To me, my job is challenging. (DE5)
14. I get carried away when I am working.[a] (AB5)
15. At my job, I am very resilient, mentally. (VI5)
16. It is difficult to detach myself from my job. (AB6)
17. At my work, I always persevere, even when things do not go well. (VI6)

Source: Schaufeli and Bakker (2003).
Note: VI = Vigor scale; DE = Dedication scale; AB = Absorption scale.
a. Shortened version (Utrecht Work Engagement Scale–9 [UWES-9]).

References

Arbuckle, J. L. (1997). *Amos users' guide* (Version 4.0). Chicago: Smallwaters.

Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2002). Validation of the Maslach Burnout Inventory–General Survey: An Internet study. *Anxiety, Stress & Coping*, *15*, 245-260.

Bentler, P. M. (1990). Comparative fit indexes in structural equation models. *Psychological Bulletin*, *107*, 238-246.

Bouman, A. M., Te Brake, H., & Hoogstraten, J. (2002). Significant effects due to rephrasing the Maslach Burnout Inventory's personal accomplishment items. *Psychological Reports*, *91*, 825-826.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement*, *61*, 229-248.

Cudeck, R., & Browne, M. W. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. Scott Long (Eds.), *Testing structural equation models* (pp. 1-9). Newbury Park, CA: Sage.

Demerouti, E., Bakker, A. B., Janssen, P. P. M., & Schaufeli, W. B. (2001). Burnout and engagement at work as a function of demands and control. *Scandinavian Journal of Work and Environment and Heath*, *27*, 279-286.

Durán, A., Extremera, N., & Rey, L. (2004). Engagement and burnout: Analysing their association patterns. *Psychological Reports*, *94*, 1048-1050.

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*, 453-472.

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development*, *34*, 177-189.

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.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

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

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. (1993). Burnout as a developmental process: Consideration of models. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 237-250). Washington, DC: Taylor & Francis.

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

Luthans, F. (2002). The need for and meaning of positive organizational behavior. *Journal of Organizational Behavior, 23*, 695-706.

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: Lawrence Erlbaum.

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

Maslach, C., & Leiter, M. P. (1997). *The truth about burnout: How organizations cause personal stress and what to do about it*. San Francisco: Jossey-Bass.

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

Montgomery, A., Peeters, M. C. W., Schaufeli, W. B., & Den Ouden, M. (2003). Work-home interference among newspaper managers: Its relationship with burnout and engagement. *Anxiety, Stress & Coping*, *16*, 195-211.

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

Salanova, M., Schaufeli, W. B., Llorens, S., Peiró, J. M., & Grau, R. (2001). Desde el "burnout" al "engamement": una nueva perspective [From "burnout" to "engagement": A new perspective]. *Revista de Psicología del Trabajo y de las Organizaciones*, *16*, 117-134.

Schaufeli, W. B., & Bakker, A. B. (2003). *Test manual for the Utrecht Work Engagement Scale*. Unpublished manuscript, Utrecht University, the Netherlands. Retrieved from http://www.schaufeli.com

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 Behavior, 25*, 293-315.

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

Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). Maslach Burnout Inventory—General Survey. In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *Maslach Burnout Inventory—Test manual* (3rd ed., pp. 22-26). Palo Alto, CA: Consulting Psychologists Press.

Schaufeli, W. B., Martínez, I., Marques-Pinto, A., 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. (in press). Work engagement: An emerging psychological concept and its implications for organizations. In S. W. Gilliland, D. D. Steiner, & D. P. Skarlicki (Eds.), *Research in social issues in management: Vol. 5. Managing social and ethical issues in organizations*. Greenwich, CT: Information Age Publishers.

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

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

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, *55*, 5-14.

Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: A new look at the interface between non-work and work. *Journal of Applied Psychology*, *88*, 518-528.

Storm, K., & Rothmann, I. (2003). A psychometric analysis of the Utrecht Work Engagement Scale in the South African police service. *South African Journal of Industrial Psychology*, *29*, 62-70.

Taris, T. W., Bok, I. A., & Meijer, Z. Y. (1998). Assessing stability and change of psychometric properties of multi-item concepts across different situation: A general approach. *Journal of Psychology*, *123*, 301-316.