

# The Measurement of State Work Engagement

## A Multilevel Factor Analytic Study

Kimberley Breevaart<sup>1</sup>, Arnold B. Bakker<sup>1</sup>, Evangelia Demerouti<sup>2</sup>, and Jørn Hetland<sup>3</sup>

<sup>1</sup>Department of Work and Organizational Psychology, Erasmus University Rotterdam, The Netherlands, <sup>2</sup>Department of Industrial Engineering and Innovation Sciences Human Performance Management Group, Eindhoven University of Technology, The Netherlands, <sup>3</sup>Department of Psychosocial Science, Faculty of Psychology, University of Bergen, Norway

**Abstract.** While diary studies have gained in popularity, the validity of the measures utilized in such studies remains an underresearched issue. This study examines the factor structure of the Utrecht Work Engagement Scale (UWES) on both between-person (trait) and within-person (state) levels. A multilevel confirmatory factor analysis was performed to confirm that the between-level factor structure also operates on the within-level. Data from 271 employees who filled in a state version of the UWES on five consecutive days were used to perform the analysis. Results showed that the UWES can be used to measure both trait and state work engagement. The three-factor multilevel model appeared to best fit the data. Implications for future research on engagement are discussed.

**Keywords:** employee engagement, multilevel factor analysis, state work engagement, trait work engagement

According to Bakker and Leiter (2010), contemporary organizations need employees who are psychologically connected to their work. The information and service economy of the 21st century requires employees who are willing and able to invest themselves fully in their roles. Organizations need employees who are energetic and dedicated, i.e., who are engaged with their work. It is therefore not surprising that, since the turn of the century, work engagement has gained significant popularity in the management (e.g., Macey, Schneider, Barbera, & Young, 2009) and scientific literature (e.g., Bakker & Leiter, 2010).

Most scholars use Schaufeli and Bakker's (2010; Schaufeli, Salanova, González-Romá, & Bakker, 2002) definition of work engagement. Accordingly, work engagement is a positive, fulfilling, work-related state characterized by vigor, dedication, and absorption. Vigor means that employees have high energy levels and great mental resilience. Dedication means being strongly involved in work and being enthusiastic about and proud of one's work. Finally, absorption means being fully concentrated on the work tasks and having the feeling that time flies. Schaufeli and Bakker (2003; Schaufeli et al., 2002) developed the Utrecht Work Engagement Scale (UWES) to enable the assessment of enduring work engagement. Some researchers have adapted the UWES to measure work engagement on a daily basis (e.g., Sonnentag, 2003; Xanthopoulou, Bakker, Demerouti, & Schaufeli, 2009a), and

their research has confirmed that there is substantial variation in work engagement within persons. However, the UWES was originally developed to measure work engagement in general and not to measure work engagement on a daily basis. Although most diary studies show the good internal consistencies of the UWES subscales, the factor structure of the state version of the UWES has never been established using multilevel analysis.

Recently, Sonnentag, Dormann, and Demerouti (2010) raised some concerns about the factor structure of the UWES when adapted to be used on a daily basis. Therefore, the present study examines both the between-person (trait) and the within-person (state) factor structure of the state version of the UWES. The three-factor structure of the 9-item version of the UWES has been confirmed in many different countries and occupations, although the three factors appear to be very closely related (Schaufeli & Bakker, 2010; Schaufeli, Bakker, & Salanova, 2006). Therefore, this article tests both the three-factor model and the one-factor model. Further, we show how to examine the factor structure of multilevel data by performing a multilevel confirmatory factor analysis (MCFA; Muthén, 1994). This may justify future research on daily work engagement using this adapted version of the UWES and guide future research on the psychometric properties of organizational psychological phenomena with a multilevel structure.

## State Work Engagement

State work engagement (SWE) has been defined as a transient, work-related experience that fluctuates within individuals over a short period of time (Sonnentag et al., 2010). While trait work engagement focuses on interindividual differences, SWE focuses on intraindividual differences in work engagement. Recent diary studies investigating SWE have found that 30–40% of the variance in work engagement is explained within persons over the working week (e.g., Sonnentag, 2003; Xanthopoulou et al., 2009a, 2009b). Therefore, examining SWE and its predictors can explain why even generally engaged people sometimes have an off-day. In other words, the diary methodology uncovers the dynamic part of work engagement and allows for examining more proximal predictors of work engagement. For example, Xanthopoulou et al. (2009b) examined the relationship between job resources, personal resources, and work engagement, and found that coaching on a specific day had a positive effect on work engagement the next day through its effect on next-day optimism. Further, SWE data are more reliable, since SWE is captured much closer to real life than trait work engagement. Furthermore, studying both trait and state work engagement is important, since their predictors appear to be different. For example, research shows that job demands are positively related to SWE (Bakker, van Emmerik, Geurts, & Demerouti, 2008), while they drain energy in the long term and are therefore negatively associated with enduring work engagement.

However, SWE cannot be measured with the original UWES, because this instrument averages experiences of work engagement over a longer period of time, thereby ignoring possible short-term, within-person differences in engagement. Thus, researchers have adjusted the UWES items and the timeframe of the scale anchors (see Table 1). Sonnentag et al. (2010) recently stated that “there is a strong need to theoretically and empirically investigate the structural similarities of state- and trait-engagement models” (p. 28). They argue that more research is needed to examine whether the quality and configuration of SWE is identical to the quality and configuration of trait work engagement. One of their main concerns is whether state engagement can be assessed with a modified scale for the assessment of enduring work engagement.

It is unclear whether the experience of SWE is really similar to the experience of trait work engagement. When they are not similar, this reduces the factorial invariance and consequently the factorial validity of the adapted scale to measure SWE. Factorial invariance refers to the degree to which a construct is measured similarly across levels, whereas factorial validity refers to the degree to which the measurement of a construct conforms to the theoretical definition of that construct (Hoyle & Smith, 1994[not in refs]). For example, it is conceivable that the enduring work engagement items include feelings and attitudes that cannot change or develop from one day to another, or that

some features of engagement are more important on some days than on others. When feelings and attitudes cannot change from day to day, responses would be stable across days. This is not desirable from a methodological point of view, because it would minimize the within-person variance and maximize between-person variance, so that it would measure trait work engagement instead of SWE. Related to this is the more general problem with measuring state variables with diary studies. Participants may want to appear consistent in their behavior and therefore do not report behavior inconsistent with their previous responses (Visser, Krosnick, & Lavrakas, 2000[not in refs]). This desire to respond consistently may reduce real changes in behavior over time. However, this does not seem to be a problem with work engagement, because 30–40% of the variance can be explained by the within-person level.

This could also mean that some of the items cannot be answered on a daily basis, which could lead to fewer extracted factors in factor analysis, resulting in a conceptually different measure of SWE compared to trait work engagement. Finally, one of the vigor items refers specifically to the morning (“This morning, I felt like going to work”), which may not necessarily reflect the amount of vigor experienced on that day, but which may also be a reflection of the vigor experienced before going to work. Therefore, it is important to look at both the between-person and within-person factor structure of the state version of the UWES before using it to measure SWE.

Such research should be conducted using software appropriate for multilevel data since SWE scores are nested within people, which means that the scores of the same person on different days are dependent. When observations are dependent, but the data are analyzed on a single-level, the independence assumption is violated. This can affect parameter estimates such as factor loadings, which may lead to inaccurate statistical inferences. According to Chan (1998), we often do not know whether a construct has an identical structure across different levels of analysis, or whether it varies across levels. This is important because some statistical analyses assume that constructs have an identical structure at each level of analysis. According to Mumford (1998), any effect found in single-level analysis may reflect only a methodological artifact when aggregate and disaggregate measures do not have similar reliabilities and a similar pattern of item loadings.

## Multilevel Confirmatory Factor Analysis

In the present study, a multilevel factor analysis was conducted using the state version of the UWES to examine the factor structure of work engagement at both the within-person and between-person level of analysis simultaneously. The between-person factor structure refers to differences between individuals, which is work engagement aggregated across days. Compared to work engagement measured at a single timepoint, aggregating the work en-

agement scores over 5 days reduces measurement error (Shiffman, 2007), thereby providing a statistically more reliable and powerful measure of trait work engagement. The within-person factor structure refers to differences within individuals. This is the deviation of a persons' mean on each day and is referred to as state or daily work engagement. The analysis here is confirmatory, since there is some evidence suggesting that SWE is not substantially different from the trait work engagement. This evidence is discussed below.

Sonnentag et al. (2010) argue that it is important to integrate theoretical models of trait and SWE. A challenging question that remains to be answered in order to integrate these models is whether trait and SWE have the same predictors and outcomes. According to Chen, Bliese, and Mathieu (2005), homologous models are a logical first step to considering multilevel relationships, because they tell us something about the breath and possible boundaries of theories. An important assumption of homologous models is that the predictors and outcomes of similar variables have to be similar across levels. Research indicates that trait and state work engagement are both affected by the availability of resources and affect similar outcomes (Bakker & Leiter, 2010).

Although not conclusive, there is some evidence indicating that the relationship between job resources and work engagement is highly consistent across levels. A meta-analysis by Halbesleben (2010) showed that trait work engagement was positively related to the job resources social support ( $\rho = .37$ ) and autonomy ( $\rho = .27$ ), to the personal resources self-efficacy ( $\rho = .59$ ) and optimism ( $\rho = .44$ ), as well as to job performance ( $\rho = .36$ ). Comparing these results to the studies that examined these relationships on a daily basis, it appears that the correlations are highly similar. Based on the research of Xanthopoulou, Bakker, Heuven, Demerouti, and Schaufeli (2008) as well as Xanthopoulou et al. (2009 [a or b?]), it appears that SWE is positively related to the job resources social support ( $r = .37$ ) and autonomy ( $r = .25$ ), to the personal resources self-efficacy ( $r = .28/.52$ ) and optimism ( $r = .42$ ) as well as to job performance ( $r = .34/.39$ ).

A second reason to expect the factor structure to be invariant across levels of analysis is that research shows a positive relationship between trait work engagement and SWE. Sonnentag (2003) reported a correlation of .66 between trait work engagement and SWE. Similarly, Xanthopoulou et al. (2009b) found a correlation of .76 between trait and SWE. Although this positive relationship is not perfect, these results indicate that employees who are most engaged in general are also most likely to be most engaged on a specific day. Therefore, daily measures of work engagement have only minor influence on interitem correlations, since it is unlikely that a generally engaged person would not be engaged at all at one out of 5 days. This means that SWE is not conceptually different from trait work engagement, but how engaged some-

one feels (the mean level) may still differ depending on the day. Despite the concerns raised by Sonnentag et al. (2010), there is some evidence that the factor structure of the state UWES may be similar on both the between-person level and within-person level. Therefore, we decided to conduct a confirmatory factor analysis.

Most techniques used to analyze diary data assume compound symmetry across days, i.e., the work engagement scores of all days (day 1, 2, 3, . . . k) are correlated equally strong with each other. However, it is conceivable that the effects of events affecting work engagement on one day do not decay immediately, but also influence work engagement on the following day(s). In the current article, we argue that the effect of these events on next days' work engagement is minimal so that we can assume compound symmetry. For example, in routine jobs, the events affecting work engagement are more or less the same every day. In our view, the work engagement of an employee with a nonroutine job is also unlikely to be influenced by the work engagement experienced on previous days. This is because every day different circumstances determine the work engagement experienced on that day. Also, as discussed before, work engagement fluctuates from day to day as a function of the available job resources on the specific day. This implies that it is likely that there will be no systematic pattern of correlations between work engagement scores over time. For example, an employee may be generally very engaged throughout a working week with the exception of Tuesday, because on that day social support was low due to the sick leave of several colleagues. Consequently, we assume that correlations between work engagement scores on different days will be very similar for individuals in different kind of jobs.

## Method

### Participants and Procedure

Data from three different studies using SWE as one of the variables were combined and used in the present study. The data from these three studies were gathered by students working on their master thesis at a Dutch University. They examined the relationship between leisure activities and work engagement, leadership and follower work engagement, and the crossover of work engagement between colleagues. All participants were asked to fill in the diary on 5 consecutive days. They filled in the questions about daily work engagement at the end of each workday.

The sample consisted of 271 employees; 159 men and 112 women. The youngest employee was 20 and the oldest 64 ( $M = 36.75$ ,  $SD = 10.49$ ). Most employees were married (62.4%), highly educated (72.3%), and working in business services (23.2%), government (26.6%), or another sector, like education, health and welfare, or transport.

Table 1. Means, standard deviations, as well as between-level and within-level (between brackets) intercorrelations for the nine SWE items

Items	$\bar{x}$	$SD$	1	2	3	4	5	6	7	8
1. Today, I felt bursting with energy. ( <i>Vi</i> )	5.01	1.22								
2. Today, I felt strong and vigorous at my job. ( <i>Vi</i> )	5.09	1.23	.94 (.77)							
3. When I got up this morning, I felt like going to work. ( <i>Vi</i> )	4.93	1.42	.82 (.51)	.83 (.51)						
4. Today, I was enthusiastic about my job ( <i>De</i> )	5.13	1.26	.84 (.57)	.88 (.59)	.86 (.41)					
5. Today, my job inspired me ( <i>De</i> )	4.88	1.32	.80 (.56)	.87 (.55)	.83 (.40)	.93 (.63)				
6. Today, I was proud of the work that I do ( <i>De</i> )	5.09	1.29	.75 (.45)	.77 (.47)	.71 (.38)	.89 (.56)	.85 (.51)			
7. Today, I felt happy when I was working intensely ( <i>Ab</i> )	5.10	1.23	.79 (.49)	.82 (.52)	.81 (.46)	.92 (.55)	.87 (.54)	.87 (.59)		
8. Today, I was immersed in my work ( <i>Ab</i> )	4.72	1.41	.62 (.53)	.66 (.51)	.64 (.39)	.81 (.55)	.74 (.54)	.82 (.61)	.76 (.53)	
9. Today, I got carried away when I was working ( <i>Ab</i> )	3.91	1.48	.59 (.36)	.56 (.36)	.60 (.24)	.65 (.40)	.69 (.38)	.70 (.42)	.66 (.39)	.69 (.53)

Notes.  $N = 271$ . The original items are in Dutch. The items presented here have been translated to English with the translation-backtranslation procedure. However, the English items have not yet been tested empirically. *Vi* = vigor, *De* = dedication, *Ab* = absorption.

## Measure

### Work Engagement

The Dutch version of the 9-item UWES (Schaufeli et al., 2006) was used to measure work engagement. The questions were adapted to measure work engagement on a daily basis, thereby measuring SWE (see Table 1). The questionnaire used in the different studies were measured on a 7-point scale (0 = *strongly disagree*, 6 = *strongly agree*). Sample items are: "Today, my job inspired me" and "Today, I was very enthusiastic about my job." The internal consistency of the total scale is .93 and ranges from .81 to .89 for the three dimensions separately.

### Strategy of Analysis

Researchers usually perform either a factor analysis based on the total covariance matrix of the entire sample or a factor analysis based on the sample between-person matrix when looking at the factor structure of a person level construct (Dyer, Hanges, & Hall, 2005). However, according to Muthén (1994), both procedures are problematic. To counter the problems associated with these procedures, Muthén developed the multilevel confirmatory factor analysis (MCFA) procedure, which uses the within- and between-person covariance matrices simultaneously. Mplus (Muthén & Muthén, 2002) was used to perform the analysis.

The normative  $\chi^2$  statistic, the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR) were used to assess the fit of the models. CFI values of .90 to .95 are acceptable, and values of .95 to 1.00 are high; RMSEA

values less than .08 are acceptable, and less than .05 are high; and SRMR values up to .10 show acceptable fit (Schweizer, 2010). Since the sample in this study is large (i.e., 271 participants  $\times$  5 days = 1,355 cases for the within-person level) and the  $\chi^2$  statistic is affected by sample size, the main focus was on the other fit indices. Bayesian information criterion (BIC) was used to compare the fit of the different models, because it adjusts for the number of parameters that are estimated (Schwartz, 1978). Therefore, unlike the  $\chi^2$  test, the BIC can be used to compare models that are not nested. Lower BICs indicate a better fit.

## Results

### Descriptive Statistics

Table 1 shows the means, standard deviations, and intercorrelations for the 9 items at the between-level of analysis and the within-level of analysis.

### Multilevel Confirmatory Factor Analysis

First, we used the intraclass correlation (ICC) to determine whether the use of multilevel analysis was justified. The ICC reflects the amount of between-person variability compared to the amount of total variability. Since the ICCs in this study range from .36 to .55, they can be classified as normal (James, 1982), justifying the use of multilevel analysis.

Next, we performed the MCFA. Table 2 shows that, for the multilevel model, the three-factor model shows a better



Table 2. Model fit for a priori multilevel models for 1-factor and 3-factor solutions

	$\chi^2$	df	Normative $\chi^2$	CFI	RMSEA	BIC	SRMR
Models							
1-factor model							
Multilevel	791.8	54	14.66	.90	.10	31117.6	W = .05, B = .05
3-factor model							
Multilevel	317.4	48	6.61	.96	.06	30686.5	W = .04, B = .04

Notes.  $N = 271$ .  $df$  = degrees of freedom, CFI = comparative fit index, RMSEA = root means square error of approximation, BIC = Bayesian information criterion, SRMR = standardized root mean square residual, W = within-person portion of the model, B = between-person portion of the model.

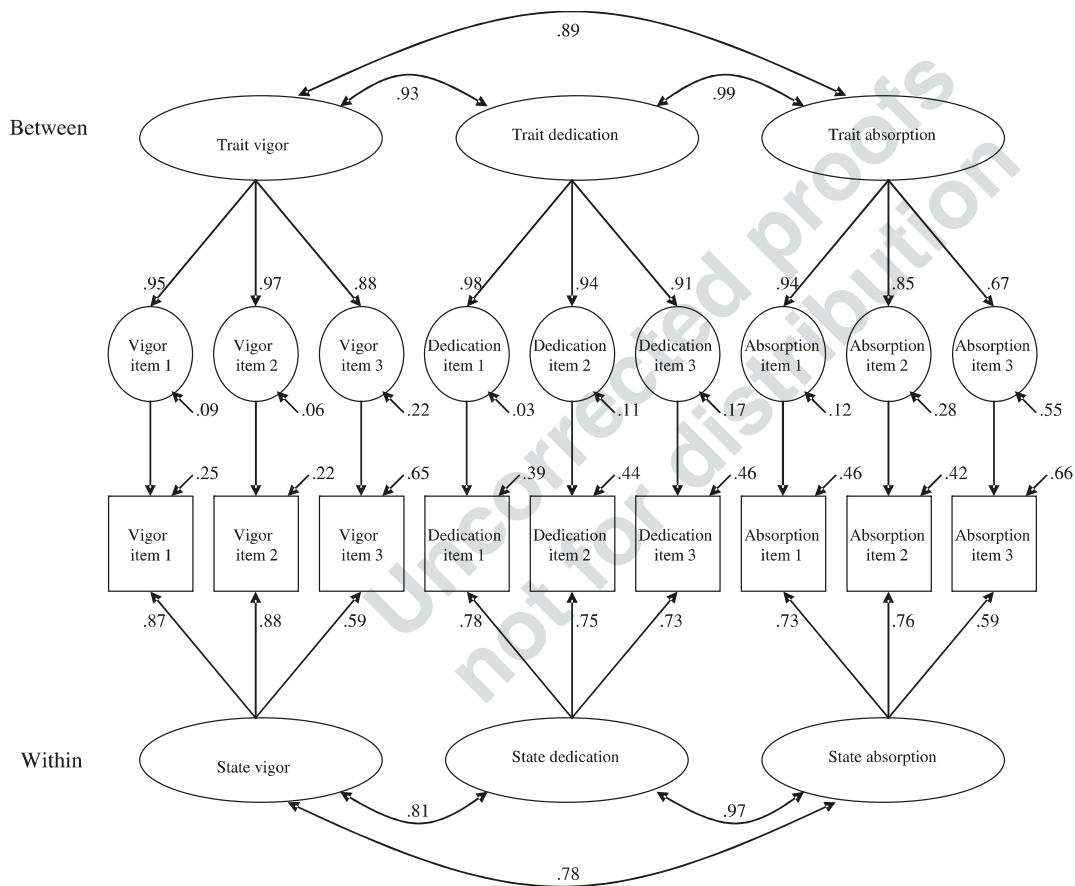


Figure 1. Path diagram of the final three-factor model.

fit to the data than the one-factor model, with BIC values being lower for the three-factor model. For the three-factor model, the CFI and SRMR values indicate good fit, and the RMSEA value indicates acceptable fit. For the one-factor model, only the SRMR value shows good fit, and the CFI shows acceptable fit.

To test for the equality of factor loadings across levels, additional models were tested. First, we tested a model constraining all factor loadings to be equal across the two levels. This model was compared to the freely estimated multilevel three-factor model. The constrained model showed a significant increase of  $\chi^2$  ( $\Delta\chi^2(6) = 47.7, p < .001$ ), indicating that the factor loadings are not equal across levels for all items. Since vigor item 3 is the only item referring

to the morning, we tested another model with all the factor loadings constrained except for the factor loading of vigor item 3. It appeared that there was no significant increase of  $\chi^2$  when this model was compared to the freely estimated model ( $\Delta\chi^2(5) = 7.4, ns$ ). This means that all factor loadings are equal across levels except for vigor item 3, which has a lower factor loading at the within level of analysis.

Figure 1 shows that, overall, the factor loadings at the between-level are higher (.67–.98) than at the within-level (.59–.88). Correlations between the three factors are also higher at the between-level (.89–.99) than at the within-level (.78–.97), indicating that discriminant validity is better for the subdimensions of work engagement at the within-level of analysis.

## Discussion

This study tests the factorial validity of the adapted UWES to measure SWE by using diary data and statistical methods for modeling both within-person variability (“state”) and reliable between-person variability (“trait”). This study is unique in that it is the first to test the validity of the measurement of SWE and – to our knowledge – one of the first to test the validity of a state measure in general. The UWES was originally developed to measure trait work engagement, but recently researchers have started to adapt the UWES to measure SWE. However, the use of the UWES on a daily basis has not yet been justified as the factorial multilevel structure, and invariance has not yet been examined. Therefore, in this study we performed MCFA (Muthén, 1994).

The results supported our hypothesis that the best-fitting model is the multilevel, three-factor model. This means that the adapted version of the UWES captures both trait and SWE, and there is no need for a conceptually different measure to assess SWE. Regarding the concerns expressed by Sonnentag et al. (2010), this study empirically tested the structural similarities of measures of trait and SWE. It appears that the adapted version of the UWES is satisfactory, since the three-factor structure holds both at the between- (trait work engagement) and the within-level (SWE) of analysis.

The relationships of the items to their corresponding factor were not different at between- and within-level of analysis, with the exception of one of the vigor items, thereby supporting the construct validity of almost all the items. Apparently, the third item of the vigor factor better reflects the construct at the between-level of analysis compared to the within-level of analysis. As outlined in the Introduction, this item is the only one that does not specifically refer to the day, but rather to the morning, which may be an explanation for the results. Feeling like going to work in the morning may also be a reflection of the vigor experienced the previous working day or the vigor experienced during breakfast. A solution may be to choose another item from the 17-item version of the UWES which can be adapted to refer to the specific day (e.g., “Today at work, I was very resilient, mentally”).

These results may also be an answer to another concern raised by Sonnentag et al. (2010) about capturing trait work engagement. They argued that SWE reflects a vivid experience, while trait work engagement is more likely to reflect an attitude. This is because, when measuring trait work engagement, people have to think back over a longer period of time, thereby introducing the risk of memory bias. This study shows that the adapted version of the UWES also captures trait work engagement. Therefore, “real” trait work engagement may be captured by measuring work engagement every day or every week over a longer period of time.

The results also showed that the three engagement fac-

tors are closely related on both between-level and within-level of analysis. This means that there is one general factor – work engagement – that consists of three different factors (vigor, dedication, and absorption) on both levels of analysis. This study thereby supports the original view that the three factors of work engagement can be distinguished, but also the more recent notion that the three factors can be combined into one single measure of work engagement (Schaufeli et al., 2006). Researchers could use the former approach when the three factors are theoretically expected to be differently related to a certain outcome variable, and use the total score when the three factors are not expected to be differently related to an outcome variable. Using the total score when expecting different relationships between the three factors and predictors and/or outcomes, may lead to loss of information.

Based on the results of this study, future research using the adapted UWES to measure work engagement on a daily basis is justified. It would be interesting for future research to focus on more detailed measures of work engagement, since we focused on SWE measured once every day, for example, the factor structure when SWE is measured three times a day or every hour. The concerns raised by Sonnentag et al. (2010) could become prominent with these more detailed measures of SWE. Maybe employees do not feel very energetic in the morning, but this effect may be less apparent if they receive some very good news in the afternoon, which boosts their energy levels. Having multiple measures on one day may result in a conceptually different measure of SWE, because it also captures those feelings and thoughts that have a low day-level frequency.

Finally, future research should focus on the integration of models of trait work engagement and SWE. Although some research has been done on the same antecedents and consequences of trait work engagement and SWE, there is little integration thereof. This study shows that there is indeed reason to expect that models of trait work engagement and SWE can be integrated, and that there is no need to develop new models for SWE. One important proposition of the job demands-resources model (Bakker & Demerouti, 2007) is that job resources particularly influence work engagement when job demands are high. Although research has supported this interaction for trait work engagement (e.g., Hakanen, Bakker, & Demerouti, 2005), research on SWE has not yet paid much attention to this interaction effect on a daily basis. It would be interesting to know if daily job resources particularly boost SWE when daily job demands are high.

Of course, the present study also has some limitations. Although MCFA has many advantages, there is an important limitation to this method: There is no rule of thumb about the amount of observations needed to perform a MCFA. One study, described in Mok (1995), indicates that 800 observations or more are needed to perform a MCFA. Although more research is needed to explore this question, the present study clearly had more than 800 observations (1830). Another potential limitation in this study could be

the implicit assumption that the correlations between the work engagement items are equal for all days, because parameter estimates are biased when these correlations are not equal. However, this cannot be solved in any currently available statistical software package. Fortunately, as we argued in the Introduction, there appeared to be no systematic pattern of correlations between the work engagement scores across days. Another limitation may be that in this study, SWE was measured only once every day. As indicated earlier, it is important for future research to look at the factor structure of the adapted version of the UWES when SWE is measured multiple times every day.

Despite these limitations, a major strength of this paper is that it used an innovative statistical method to analyze daily diary data. Most researchers using daily diaries do not address the psychometric properties of their adjusted measures. This study is one of the first to analyze the factorial validity of an adjusted trait measure to measure a state construct. In the future, we recommend researchers using the method applied in this study to test the validity of their adjusted trait measures for measuring states.

## References

- Bakker, A. B., & Demerouti, E. (2007). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology*, 22, 309–328. doi 10.1108/02683940710733115
- Bakker, A. B., & Leiter, M. P. (Eds.). (2010). *Work engagement: A handbook of essential theory and research*. New York: Psychology Press.
- Bakker, A. B., van Emmerik, I. H., Geurts, S. A. E., & Demerouti, E. (2008). *Recovery turns job demands into challenges: A diary study on work engagement and performance*. Working paper, Erasmus University Rotterdam.
- Chan, D. (1998). Functional relations among constructs in the same content domain in different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, 83, 234–246. doi 10.1037/0021-9010.83.2.234
- Chen, G., Bliese, P. D., & Mathieu, J. E. (2005). Conceptual framework and statistical procedures for delineating and testing multilevel theories of homology. *Organizational Research Methods*, 8, 375–409.
- Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *Leadership Quarterly*, 16, 149–167. doi 10.1016/j.leaqua.2004.09.009
- Hakanen, J. J., Bakker, A. B., & Demerouti, E. (2005). How dentists cope with their job demands and stay engaged: The moderating role of job resources. *European Journal of Oral Sciences*, 113, 479–487. doi 10.1111/j.1600-0722.2005.00250.x
- Halbesleben, J. R. B. (2010). A meta-analysis of work engagement: Relationships with burnout, demands, resources, and consequences. In M. P. Leiter & A. B. Bakker (Eds.), *Work engagement: A handbook of essential theory and research* [pages?]. New York: Psychology Press.
- [not in text]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.
- James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, 67, 219–229. doi 10.1037/0021-9010.67.2.219
- Macey, W. H., Schneider, B., Barbera, K. M., & Young, S. A. (Eds.). (2009). *Employee engagement: Tools for analysis, practice, and competitive advantage*. New York: Blackwell.
- Mok, M. (Ed.). (1995). *Sample size requirements for 2-level designs in educational research*. Sydney, Australia: Macquarie University [Press?].
- Mumford, M. D. (Ed.). (1998). Situations, interactions, and followers: The case of individualized leadership. *Leadership: The multiple-level approaches, classical and new wave*. Stamford, CT: JAI.
- Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research*, 22, 376–398. doi 10.1177/0049124194022003006
- Muthén, L. K., & Muthén, B. O. (2002). Mplus VERSION 2.12: Addendum to the Mplus users's guide. Los Angeles, CA: Muthén & Muthén.
- Schaufeli, W. B., & Bakker, A. B. (2003). *Utrecht Work Engagement Scale: Preliminary manual*. Department of Psychology, Utrecht University, The Netherlands. Retrieved from <http://www.schaufeli.com>
- Schaufeli, W. B., & Bakker, A. B. (2010). Defining and measuring work engagement: Bringing clarity to the concept. In M. P. Leiter & A. B. Bakker (Eds.), *Work engagement: A handbook of essential theory and research* [pages?]. New York: Psychology Press.
- Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A crossnational study. *Educational and Psychological Measurement*, 66, 701–716. doi 10.1177/0013164405282471
- Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of burnout and engagement: A confirmatory factor analytic approach. *Journal of Happiness Studies*, 3, 71–92. doi 10.1023/A:1015630930326
- [not in text]Schaufeli, W. B., Taris, T. W., & Van Rhenen, W. (2008). Workaholicism, burnout and engagement: Three of a kind or three different kinds of employee well-being? *Applied Psychology: An International Review*, 57, 173–203. doi 10.1111/j.1464-0597.2007.00285.x
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464. doi 10.1214/aos/1176344136
- Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment*, 26, 1–2. doi 10.1027/1015-5759/a000001
- [not in text]Seppälä, P., Mauno, S., Feldt, T., Hakanen, J., Kin-nunen, U., Tolvanen, A., & Schaufeli, W. B. (2009). The construct validity of the Utrecht Work Engagement Scale: Multi-sample and longitudinal evidence. *Journal of Happiness Studies*, 10, 459–481. doi 10.1007/s10902-008-9100-y
- Shiffman, S. (2007). Designing protocols for ecological momentary assessment. In A. A. Stone, S. Shiffman, A. A. Atienza, & L. Nebeling (Eds.), *The sciences of real-time data capture: Self-reports in health research* (pp. 27–53). New York: Oxford University Press.
- Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: A new look at the interface between nonwork and

- work. *Journal of Applied Psychology*, 88, 518–528. doi 10.1037/0021-9010.88.3.518
- Sonnentag, S., Dormann, C., & Demerouti, E. (2010). Not all days are equal: The concept of SWE. In M. P. Leiter & A. B. Bakker (Eds.), *Work engagement: A handbook of essential theory and research* [pages?]. New York: Psychology Press.
- Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2009a). Reciprocal relationships between job resources, personal resources, and work engagement. *Journal of Vocational Behavior*, 74, 235–244. doi 10.1016/j.jvb.2008.11.003
- Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2009b). Work engagement and financial returns: A diary study on the role of job and personal resources. *Journal of Occupational and Organizational Psychology*, 82, 183–200. doi 10.1348/096317908X285633
- Xanthopoulou, D., Bakker, A. B., Heuven, E., Demerouti, E., & Schaufeli, W. B. (2008). Working in the sky: A diary study on the role of job resources and personal resources. *Journal of Occupational and Organizational Psychology*, 82, 183–200. doi 10.1037/1076-8998.13.4.345

Kimberley Breevaart

---

Erasmus University Rotterdam  
Institute of Psychology  
Woudestein, T11-11  
P. O. Box 1738  
3000 DR Rotterdam  
The Netherlands  
Tel. +31 10 408 9034  
Fax +31 10 408 9009  
E-mail breevaart@fsw.eur.nl

Uncorrected proofs  
not for distribution