Conceptualizing and Measuring Engagement: An Analysis of the Utrecht Work Engagement Scale

Maura J. Mills · Satoris S. Culbertson · Clive J. Fullagar

Abstract This article analyzes the Utrecht Work Engagement Scale (UWES; Schaufeli et al. in J Happiness Stud 3:71–92, 2002b) on a variety of levels. Study 1 critiques the method by which the original scale was developed, and analyzes a similar sample using both exploratory and, subsequently, confirmatory factor analyses. Study 2 uses three samples to explore the 17-item UWES-17, and the recent shorter version of the scale, the 9-item UWES-9. Factor structures and reliabilities of scores for both scale versions were examined for each sample. Although some cautions are warranted when using the UWES, this research leans toward supporting a multifactorial conceptualization of the construct. Preliminary construct validation of the use of the measures was also established via correlations with other pertinent constructs. Although research on the measure remains sparse, the UWES-9 holds promise as a parsimonious version of the UWES-17 that appears valid in use, appears to yield reliable scores in the samples herein, and also appears to capture the purported three-factor dimensionality of the engagement construct better than does the original UWES-17 version.

Keywords Engagement · Utrecht work engagement scale · Construct validity · Factor analysis · Parallel analysis · Scale development

Engagement is a construct naturally subsumed within the increasingly popular domains of positive psychology (Seligman and Csikszentmihalyi 2000) and positive organizational behavior (Luthans 2002), which aim to enhance employees' positive experiences of work. Although engagement has received increased interest in both the literature and practice, as

M. J. Mills (✉)
Department of Psychology, Hofstra University, 135 Hofstra University, Hempstead, NY 11549, USA
e-mail: Maura.Mills@hofstra.edu

S. S. Culbertson · C. J. Fullagar
Kansas State University, 492 Bluemont Hall, Mid-Campus Drive, Manhattan, KS 66506, USA

with many popular constructs, neither a ubiquitously-accepted measurement instrument nor a common conceptualization have emerged (Saks 2006). Nevertheless, there are commonalities in the diverse definitions of engagement that have been proposed by both scientists and practitioners. As Macey and Schneider (2008) noted, "Common to these definitions is the notion that employee engagement is a desirable condition, has an organizational purpose, and connotes involvement, commitment, passion, enthusiasm, focused effort, and energy so it has both attitudinal and behavioral components" (p. 4).

Schaufeli et al. (2002b) proposed that engagement is a multidimensional construct consisting of three dimensions: *vigor*, characterized by high levels of energy, effort, resilience, persistence, and motivation to invest in the work; *dedication*, characterized by involvement in work, enthusiasm, and a sense of pride and inspiration; and *absorption*, characterized by immersion in one's work and the sense of time passing quickly. Accordingly, Schaufeli et al. (2002b) developed and provided initial validity support for a measure of engagement that incorporated this multidimensionality. While their measure, the Utrecht Work Engagement Scale (UWES), is widely used in measuring engagement, the present study critiques the methodology of the original scale development, arguing that inappropriate factor extraction criteria, incorrect rotation method, and the use of PCA rather than EFA compromised the integrity and appropriateness of the scale from the outset.

Nevertheless, since its development, the psychometric properties of the UWES have received considerable support. For example, internal consistencies have been found to be acceptable, generally ranging from .80 to .90 (Montgomery et al. 2003; Salanova et al. 2001; Schaufeli and Bakker 2004). Additionally, a majority of studies have confirmed that the three-factor structure of the UWES is superior to a one-factor, or unidimensional, conceptualization of engagement (e.g., Montgomery et al. 2003; Schaufeli and Bakker 2004).

However, there have also been various concerns with the UWES. For example, Shirom (2003) expressed concern with the high intercorrelations among the three dimensions (typically exceeding .65, particularly between vigor and absorption), suggesting a certain amount of redundancy. In response to this criticism, Schaufeli and colleagues (Salanova et al. 2001; Schaufeli et al. 2006; Schaufeli et al. 2002a; b) argued that it is theoretically justifiable that vigor and absorption would be correlated, since such a relationship is indicative of the fact that being immersed in one's activities corresponds with high energy levels, and vice versa. Furthermore, they noted that there is no evidence of multicollinearity, arguing that the dimensions represent distinct constructs and should be measured as such.

Recognizing these high intercorrelations, Schaufeli et al. (2002a, b) explored a two-factor dimensionality of engagement by collapsing vigor and absorption into a single dimension. The two-factor solution provided a small, but statistically significant, deterioration in the goodness-of-fit indices when compared to the three-factor conceptualization. They concluded, therefore, that a three-factor conceptualization is most appropriate, and the high intercorrelations between vigor and absorption should not be surprising given the nature of the relationship.

Nevertheless, problems still exist. For example, in conducting a principal components analysis on the UWES measure, Sonnentag (2003) failed to support the three-factor structure and instead successfully used the UWES as an overall measure of engagement, citing the high reliability of resulting scores. Furthermore, in a meta-analysis examining the correlates of engagement, Christian and Slaughter (2007) found that vigor, dedication, and absorption were strongly intercorrelated and subsequently suggested that Schaufeli

et al.'s (2002b) measure be scored as a unidimensional composite rather than as the three separate scales. They also noted, however, that the three dimensions differentially related to such outcomes as health and commitment in a way that supported the use of a three-factor measure.

The fact that the three-factor solution has yet to be consistently confirmed in independent samples may be due to both substantive and methodological reasons (Van Prooijen and Van der Kloot 2001). Substantively, the original factor solution may not generalize to different samples or populations, given that it was developed based solely upon a population of students but is now being used in both academic and corporate settings. Methodological reasons for why exploratory factor analytic solutions such as those mentioned above cannot be replicated include inappropriate factor extractions criteria, incorrect rotation method, and the use of principal components analysis instead of exploratory factor analysis (Fabrigar et al. 1999). It may be that the original and/or updated factor models of engagement may be unreliable because the initial model was itself not empirically sound (Van Prooijen and Van der Kloot 2001). Thus, a key purpose of the present research was to examine Schaufeli et al.'s (2002b) development of the initial measure and related factor model of engagement, and to do so within the context of research that has outlined appropriate steps for such factor analyses (Study 1). A subsequent purpose of this research was to examine the dimensionality (Study 1, Study 2), reliability of scores (Study 1, Study 2), and construct validity (Study 2) of both versions of the UWES using four separate samples in order to determine whether the three-factor structure is indeed the most appropriate conceptualization of the UWES.

An additional contribution of this research is a thorough examination of a shorter version of the UWES (Study 1, Study 2). In an attempt to make the scale more parsimonious, Schaufeli et al. (2006) proposed a shortened version of the UWES, using only nine of the original 17 items. Nevertheless, as of yet research has failed to offer many investigations of this shorter version. In fact, to date, the authors found only two studies (Schaufeli et al. 2006; Seppälä et al. 2009) that have investigated the psychometric properties of this shorter version, one of which is the paper in which the nine-item version was originally proposed (Schaufeli et al. 2006). Thus, in addition to examining the methodological approach to the development of the UWES-17, we also sought to examine the psychometric properties of both versions of the UWES (the original 17-item version, UWES-17, and the shorter nine-item version, UWES-9). We evaluated the latter version to determine whether the short version is a feasible alternative for the longer, more established form, and did so while taking into consideration Smith et al.'s (2000) recommendations and notes of caution regarding short form development.

Finally, it should be noted that while various past studies have previously attempted to provide support for the external construct validity of inferences drawn from the measurement of the engagement construct, the sample- and usage-specific nature of validity cannot necessarily be generalized across studies (Messick 1995), and therefore it also warrants some initial investigation herein (Study 2). In this way, the present research contributes toward both securing and enriching the nomological network in which engagement lies. As various researchers (e.g., Cronbach and Meehl 1955) have argued, the establishment of such nomological networks is crucial if we are to make appropriate and meaningful inferences about both the usage and the implications of our construct measurements, and, subsequently, frame them within appropriate theoretical bounds and generate suitable hypotheses from such theory. Indeed, as Cronbach and Meehl (1955) stated, "a necessary condition for a construct to be scientifically admissible is that it occur in a nomological net" (p. 290).

1 Study 1

In our first study, we examined the original factor analytic study by Schaufeli et al. (2002) that proposed the UWES-17. As such, Study 1 fills a gap in the literature that is important to consider prior to further analyzing the structure and psychometric properties of the UWES. The only other independent evaluation of the UWES (Seppälä et al. 2009) failed to consider this literature gap and therefore did not conduct this important first step regarding the original scale development. In the measurement of engagement, we agree with Schaufeli and colleagues' contention that engagement should be measured via its own scale, as opposed to, for example, the antipode of burnout as measured via the Maslach Burnout Inventory (Maslach and Leiter 1997). However, we question the appropriateness of their sole use of confirmatory factor analytic (CFA) techniques in scale development. Schaufeli and colleagues contended that the items themselves were developed in regard to both (a) consideration of engagement's (negative) relationship to burnout, and (b) interview responses (Schaufeli et al. 2001). Although it appears that they used appropriate item analytic techniques to remove items and improve score reliability while simultaneously considering parsimony, we contend that factor identification would have been better conducted via an initial exploratory factor analysis (EFA) using appropriate factor extraction criteria, to be followed *subsequently* by a CFA. This approach is supported by the preponderance of research, including Fabrigar et al. (1999), who noted that EFA should be used when "there is insufficient basis to specify an a priori model" (p. 283). Fabrigar et al. (1999) emphasized that, in cases such as the present one, EFA is the most appropriate form of factor analysis because of the inordinate number of alternative models that would necessitate factor analyzing if CFAs were the method used. Byrne (2001) offered the same warning—albeit in a much firmer tone—stating that, "the application of CFA procedures to assessment instruments that are still in the initial stages of development represents a serious misuse of this analytic strategy" (p. 99).

Likewise, Van Prooijen and Van der Kloot (2001) highlighted the fact that CFA must be theory-driven. It is now acknowledged that the theoretical underpinnings of the concept of engagement are somewhat vague and indicate considerable redundancy with various other constructs (Macey and Schneider 2008; Saks 2006). Researchers agree that when insufficient theory is present, EFA and CFA should be used in conjunction with one another in that an EFA should be conducted initially, to be followed by a CFA for cross-validation purposes (Costello and Osborne 2005; Fabrigar et al. 1999; Gerbing and Hamilton 1996; Van Prooijen and Van der Kloot 2001; Worthington and Whittaker 2006). The subsequent CFAs should then be conducted on both the original data, for purposes of within-sample cross-validation (Van Prooijen and Van der Kloot 2001), and also on new data, "in order to answer questions such as 'does an instrument have the same structure across certain population subgroups?'" (Costello and Osborne 2005, p. 8). Therefore, one of the main contributions of the present paper, and the thrust of this Study 1, is its adherence to a more rigorous protocol for establishing and subsequently factor analyzing the UWES, thus making the present research more methodologically rigorous than prior studies (Schaufeli et al. 2002).

1.1 Method

1.1.1 Participants and Procedure

Study 1 was comprised of 344 students in the architectural professional degree program at a large public university in the Midwestern United States. These 344 were from an eligible

sample pool of 690, thus resulting in a response rate of 50%. Approximately half (51%) of participants were male, and their mean age was 21.32 years (SD = 1.05). Participants were in their first (18% of participants), second (16%), third (13%), forth (30%), and fifth (22%) years of the architecture program, respectively. Participation was entirely voluntary and anonymous. Surveys were administered using an online survey platform.

1.1.2 Measure

This EFA consisted of the UWES-17 (Schaufeli et al. 2002), although this scale also necessarily includes the UWES-9 (Schaufeli et al. 2006). The UWES-17 is comprised of 17 items (Table 2) measured on a seven-point Likert scale anchored by the response options 1 = *never* and 7 = *always*. Six items comprised the vigor sub-scale (e.g., "When studying I feel strong and vigorous."). Dedication was measured with five items (e.g., "I find my studies to be full of meaning and purpose."). Finally, the remaining six items comprised the absorption sub-scale (e.g., "I get carried away when I am studying."). The UWES-9 maintains the same factor structure and response scale, but is comprised of nine of the original 17 items (Table 3).

1.2 Analysis

Using a sample similar to the one utilized by Schaufeli et al. (2002) in their original scale development, we conducted an EFA on the 17 (of 24 original) items that they found to be appropriate and found to yield reliable scores ($\alpha = .92$). It should be noted here that our usage of an EFA as a first step prior to a CFA has the additional benefit of allowing for usage of appropriate factor extraction criteria. Schaufeli et al.'s (2002) failure to conduct an initial EFA necessarily precluded the application of appropriate factor extraction criteria. That is, not only did they force factors into an otherwise unknown measure, but they also failed to appropriately explore any other possible factorial structure that the measure might have. Conducting an initial EFA allowed us to employ a variety of techniques for determining the most appropriate factor structure of the UWES. Conducting these EFAs for both the original UWES-17 as well as the more recent UWES-9 yielded further information on how the two might differ, and which might be a more appropriate measure of the engagement construct.

We used Parallel Analysis as a first step in identifying the number of factors yielded by the measure. Originally proposed by Horn (1965), Parallel Analysis corrects for potential sampling error of other factor retention analytical methods by comparing the correlation matrices of the raw data to iterative correlation matrices of fabricated random datasets. The researcher must subsequently examine the eigenvalues of the raw data as compared to the mean and 95th percentile eigenvalues of the random data, retaining only factors for which the former is larger than the latter. The preponderance of research (e.g., Hayton et al. 2004; O'Connor 2000; Zwick and Velicer 1986) is increasingly recognizing Parallel Analysis as the most appropriate analytical tool for extracting factors. For instance, Zwick and Velicer (1986) compared five potential methods for extracting factors, and determined that Parallel Analysis outperforms all other methods in identifying the number of components to retain. Nevertheless, despite its superior performance, Parallel Analysis is (unfortunately) severely underutilized in factor analytic research (Hayton et al. 2004; O'Connor 2000). O'Connor (2000) speculates that this may be due to the advanced complexity of the procedure, and the inability of popular statistical analytic packages to perform the necessary calculations as part of their typical analytic repertoire.

Nevertheless, despite Parallel Analysis' strengths, further factor extraction procedures were subsequently used, for two very important reasons. First, as noted by Buja and Eyuboglu (1992), Parallel Analysis can be subject to errors of overextraction, thereby identifying trivial components as primary factors. Therefore, these researchers, along with O'Connor (2000), recommend that Parallel Analysis should be used as a first step in identifying factors that are likely to exist beyond mere chance, but that it should be followed by subsequent procedures for the purposes of trimming negligible factors. Second, while Parallel Analysis is superior at identifying the sheer number of factors, unlike other factor analysis programs, it is unable to identify the particular factor into which each individual item falls. Therefore, we followed up the Parallel Analysis with a more traditional factor analytic procedure which was able to accomplish this identification and placement of individual items. In conducting this phase of the EFA, we rotated the EFA solution with direct oblimin rotation to allow for correlation between the factors. Although varimax rotation is more commonly utilized than direct oblimin, the latter is more appropriate in this case given the high intercorrelations between the factors that have been found both in past research (e.g., Shirom 2003; Schaufeli et al. 2002) and also in Study 2 of the present research ($r = .58$ to .78; see Tables 6, 7, 8). Moreover, it has been argued that using orthogonal rotations such as varimax results in loss of important information—or, worse, misinformation—when factors can reasonably be expected to be related to one another, as in this case (Costello and Osborne 2005; Fabrigar et al. 1999). We used maximum likelihood (ML) estimation as this was the extraction method used by Schaufeli et al. (2002) and has also been recommended as best practice in EFA by various researchers (Costello and Osborne 2005; Fabrigar et al. 1999).

1.3 Results and Discussion

1.3.1 UWES-17

A Parallel Analysis for the UWES-17 (see Table 1) revealed five factors meeting the component extraction criteria of raw eigenvalues exceeding the mean and 95th percentile eigenvalues of the random datasets. However, upon closer comparison of the values, it is clear that only four factors have raw eigenvalues substantially exceeding their random counterparts (specifically, the raw data eigenvalue for the fifth proposed factor drops substantially, bringing it very closely in line with the 95th percentile random eigenvalue; .26 [raw] versus .24 [random], see Table 1). Given that Parallel Analysis can be susceptible to overextraction of minor components—and particularly considering Silverstein's (1987) finding that, when Parallel Analysis does overfactor, it tends to do so by 1 component—we consider the possibility that at least the proposed fifth factor found by the Parallel Analysis is negligible.

To further investigate these findings, subsequent to the Parallel Analysis, an EFA with ML extraction and direct oblimin rotation determined that the UWES-17 could be most appropriately conceptualized as being comprised of four factors (Table 2). Recognizing the appropriateness of using multiple decision rules to determine factor structure, these four factors were indicated both by examination of a scree plot and also by eigenvalues greater than 1.00. Factors with eigenvalues less than 1.00 were not retained, as per widely-accepted factor analytic rules. The four retained factors are absorption and dedication as they currently stand (six items and five items, respectively), and the vigor factor, which appears to represent two separate factors, each of which consists of three items. While three-item factors may not be ideal, it is also not uncommon, and in this case each of the

Table 1 Study 1: UWES-17 and UWES-9—exploratory factor analyses (parallel analysis factor extraction, 100 random datasets)

Factor/ root	Raw data eigenvalue	Random data eigenvalue mean	Random data eigenvalue 95th percentile
UWES-17			
1	7.22	.46	.56
2	1.37	.38	.43
3	.74	.31	.36
4	.68	.25	.29
5	.26	.20	.24
6	.16	.16	.20
7	.06	.11	.14
8	.05	.07	.10
9	.01	.02	.05
10	−.02	−.01	.02
UWES-9			
1	4.27	.29	.37
2	.66	.19	.24
3	.49	.13	.17
4	.03	.06	.11
5	−.07	.01	.04
6	−.09	−.04	−.01
7	−.12	−.09	−.06
8	−.14	−.14	−.11
9	−.20	−.20	−.15

For UWES-17, values are reported up to a root of ten to allow for the reader to see past the cut-off rule for what constitutes a factor. Values for roots beyond ten are not reported for parsimony's sake, but are available from the first author upon request. For UWES-9, all available roots are reported

three items do indeed appear to better represent the factors identified here than would one factor consisting of six items (as proposed by Schaufeli et al. 2002a, b). The splitting of the vigor factor into two separate factors is further supported when we reflect on the item wording of the items in each of these two factors, one of which appears to represent vigor (as consistent with Schaufeli and colleagues' original factor delineation, and as indicated by items such as, "I feel strong and vigorous" and "I feel like going to class"), and the other appears to represent perseverance (as indicated by items such as, "I can continue studying for very long periods of time" and "I always persevere, even when things do not go well").

Nevertheless, following the progression of factor analyses supported by various researchers (Costello and Osborne 2005; Fabrigar et al. 1999; Van Prooijen and Van Der Kloot 2001) and outlined earlier, after having examined the scale via an EFA, we then conducted a CFA forcing four factors in order to further investigate the factor structure and determine whether it was consistent with the EFA. Unfortunately, the CFA model forcing four factors was found to be 'not positive definite,' preventing model admissibility. Therefore, we followed Wothke's (1993) recommendation that the unweighted least squares (ULS) estimation method be used in such a situation, since this method does not require the covariance matrix to be positive definite. It is worthwhile noting that another solution discussed by Wothke (1993) is smoothing the matrix. However, Wothke (1993) cautions that smoothing methods a) have not received widespread acceptance in analyses beyond regression, and b) substantially alter the covariance matrix and thus yield distorted

Table 2 Study 1: UWES-17—exploratory factor analysis factor structure matrix (maximum likelihood estimation, direct oblimin rotation)

	Factor 1 absorption	Factor 2 dedication	Factor 3 perseverance	Factor 4 vigor	Communalities	
					Initial	Extraction
"I get carried away when I am studying"	**.87**	−.36		.40	.70	.76
"Time flies when I am studying"	**.81**	−.38	.39		.65	.67
"It is difficult to detach myself from my studies"	**.78**	−.41		.35	.64	.62
"I am immersed in my studies"	**.75**	−.46	.36		.62	.58
"When I am studying I forget everything else around me"	**.75**	−.46	.41	.43	.58	.59
"I feel happy when I am studying intensely"	**.73**	−.51	.39	.55	.61	.61
"Study inspires me"	.51	**−.83**	.35	.49	.67	.72
"I am proud of my studies"	.37	**−.79**	.45		.59	.63
"My studies are full of meaning and purpose"	.47	**−.85**	.50	.43	.68	.73
"I am enthusiastic about my studies"	.53	**−.83**	.38	.59	.71	.77
"My studies are challenging"		**−.48**			.27	.25
"I am very resilient, mentally"		−.39	**.83**		.45	.68
"I can continue studying for very long periods of time"	.53	−.41	**.70**	.43	.53	.57
"I always persevere, even when things do not go well"		−.40	**.54**		.32	.32
"I feel bursting with energy"	.48	−.39		**.81**	.53	.67
"I feel like going to class"				**.69**	.39	.48
"I feel strong and vigorous"	.51	−.50	.53	**.68**	.55	.60
Eigenvalues of factors:	7.63	1.75	1.24	1.21		
Subsequent five eigenvalues: .76, .71, .51, .47, .45 percentage of variance explained by the factor	44.85	10.31	7.31	7.09		
Cumulative variance explained	44.85	55.16	62.47	69.58		

Factor loadings falling below the standard .35 were suppressed to allow for clearer factor identification

Coefficients for items within their corresponding factor column are bolded for easier reader identification of the items within each factor

χ^2 fit statistics. Therefore, smoothing was deemed an inappropriate solution for use in this context, and the ULS solution was utilized instead. The resulting fit statistics lend preliminary support to the four-factor structure (Table 4), although they should be taken with caution considering the method of estimation.

Nevertheless, due to the initial nonpositive definite problems encountered when attempting to estimate this four-factor model using ML estimation, we also conducted a CFA (via ML estimation) forcing the three factors that are theorized in Schaufeli et al.'s (2002) original UWES measure. This model was found to be a (arguably) tolerable, although certainly not ideal, fit, $\chi^2/df = 4.49$, RMSEA = .09, CFI = .88 (Table 4).

Table 3 Study 1: UWES-9—exploratory factor analysis factor structure matrix (maximum likelihood estimation, direct oblimin rotation)

	Factor 1 dedication	Factor 2 absorption	Factor 3 vigor	Communalities	
				Initial	Extraction
"Study inspires me"	**.86**	.54	.49	.63	.73
"I am enthusiastic about my studies"	**.89**	.57	.58	.69	.80
"I am proud of my studies"	**.73**	.42	.37	.47	.54
"I get carried away when I am studying"	.43	**.78**	.41	.47	.61
"I am immersed in my studies"	.49	**.75**	.33	.47	.57
"I feel happy when I am studying intensely"	.58	**.81**	.54	.59	.68
"I feel like going to class"	.38	.30	**.73**	.38	.53
"I feel bursting with energy"	.51	.53	**.79**	.51	.65
"I feel strong and vigorous"	.55	.56	**.68**	.50	.54
Eigenvalue	4.71	1.06	1.02		
Subsequent five eigenvalues: .51, .40, .40, .36, .33					
Percentage of variance explained by the factor	52.37	11.75	11.31		
Cumulative variance explained	52.37	64.12	75.43		

Factor loadings falling below the standard .35 were suppressed to allow for clearer factor identification

Coefficients for items within their corresponding factor column are bolded for easier reader identification of the items within each factor

Table 4 Study 1: confirmatory factor analyses

UWES version	χ^2	df	χ^2/df	RMSEA	NFI	CFI
UWES-17[†] (4 factors)	507.11**	113	4.49[a]	.09	.86	.88
UWES-17[‡] (4 factors)	696.72**	113	6.16	–	.98	–
UWES-17 (3 factors)	528.59**	116	4.55[a]	.09	.85	.88
UWES-9 (3 factors)	75.90**	24	3.16[a]	.08	.95	.97

For χ^2 coefficient, ** $p < .01$

For χ^2/df ratio, [a] <5 (as per Wheaton et al. 1977; [aa] <3 (as per Carmines and McIver 1981)

[†] Not positive definite [‡] ULS estimation

1.3.2 UWES-9

In keeping with the above analytical progression for the UWES-17, we examined the more recently proposed UWES-9 by first conducting EFAs on the data using both Parallel Analysis extraction criteria and, subsequently, ML estimation with direct oblimin rotation. For reasons previously explicated, we followed this with a CFA on the same data ($\alpha = .88$).

A Parallel Analysis for UWES-9 (see Table 1) revealed three factors meeting the component extraction criteria of raw eigenvalues exceeding the mean and 95th percentile eigenvalues of the random datasets, thus confirming the intended three-factor structure of the measure. Each of these three factors had raw eigenvalues considerably greater than their random counterparts, and the drop-off at the fourth root was considerable (see Table 1),

thereby indicating that the three found factors are strong and are unlikely to be susceptible to Parallel Analysis' occasional susceptibility toward overextraction. Subsequent to the Parallel Analysis, the EFA with ML estimation and direct oblimin rotation also found three factors (Table 3), again consistent with the measure's proposed factor structure. The subsequent CFA found that the three-factor model fits the data very well, and is a statistically significantly better fit than it is with the UWES-17, $\Delta\chi^2$ (92) = 452.69, $p < .001$ (Table 4).

It is interesting to note that none of the three items in the fourth (perseverance) factor found in the rotated EFA of the UWES-17 are present in the UWES-9. This allows for cleaner factor loadings in factor analyses of the latter scale than of the former. This is further supported by the consistency of the UWES-9 factor structure across various factor analytic procedures, whereas proposed factor structures of the UWES-17 were more variable across analytic methods. All of these considerations go toward supporting the use of the UWES-9 as more representative of the UWES's original factor structure than is the UWES-17. It is expected that Study 2 will support the findings from Study 1 in that it will provide further evidence that the UWES-9 is perhaps a better measure of Schaufeli et al.'s (2002) originally intended factor structure of the engagement construct than is the original, longer UWES-17.

2 Study 2

Considering the findings of Study 1, and in an effort to explore the originally specified factor structure of the UWES, we cross-validated the results of Study 1 by conducting CFAs on different samples on which we forced three-factor solutions. Due to the extensive exploration of both forms of the measure with various exploratory factor analytic methods in Study 1, EFAs were not necessary in Study 2, as confirmed by a preponderance of research that has suggested that further EFAs are unnecessary and that, once completed, should then be confirmed via CFAs on subsequent datasets (e.g., see Costello and Osborne 2005; Fabrigar et al. 1999; Van Prooijen and Van Der Kloot 2001). It should be noted that whereas Study 1 involved only a student sample for the reasons discussed previously, in order to test the stability of the factor structure across samples, and also in order to increase the findings' applicability and generalizability to full-time employees, Study 2 utilized one additional student sample in addition to two samples of employed individuals.

Finally, although the main premise of Study 2 was to investigate the UWES short and long versions as measures of engagement, it was warranted also to provide support for engagement's construct validity, because without substantiation of such validity of use, any further discussion regarding its measures is necessarily moot. Therefore, in all three samples we investigated the construct validity of engagement and the use of its measurement using both the UWES-17 and the UWES-9. Both convergent and divergent validity coefficients were computed to ensure a more complete assessment of construct validity. Brief descriptions of the constructs used for validation purposes along with the measures used are provided in the descriptions of the samples.

2.1 Method

We examined the psychometric properties of both UWES versions using three separate samples, described below. As in Study 1, in Study 2 we again used ML estimation and direct oblimin rotation. Participants in all three samples in Study 2 completed the full UWES-17 (Schaufeli et al. 2002), which contains the nine items that comprise the

UWES-9 (Schaufeli et al. 2006). Depending on the sample, the reference was either school-related (for students) or work-related (for employees). In addition to the UWES, other measures used to assess construct validity were also administered, and are described below.

2.1.1 Sample 1

2.1.1.1 Participants and Procedure The first sample consisted of 247 first-year engineering and architecture undergraduate students at a large Midwestern public university. These 247 were from an eligible sample pool of 311, thus resulting in a response rate of 79%. The majority of participants (68%) were male and their mean age was 19.10 years (SD = 1.72). Mass emails were sent to all engineering and architecture freshmen including explanation of the project and a link to the online survey.

2.1.1.2 Additional Measures In addition to the UWES-17 (Schaufeli et al. 2002) and the UWES-9 (Schaufeli et al. 2006), several measures were also used for the purposes of preliminary construct validation. Given past research, we attempted to establish convergent validity by predicting that engagement would be correlated with self-efficacy (Connell and Wellborn 1991; Schaufeli et al. 2002; Schaufeli and Salanova 2007; Schaufeli et al. 2006), need for achievement (Hallberg et al. 2007), intrinsic motivation (Connell and Wellborn 1991; Koestner and Losier 2002), and affective occupational commitment (Meyer et al. 1993; Snape and Redman 2003). Engagement in one's work should also necessarily be negatively associated with self-reported amotivation, characterized by an absence of motivation to do an activity and diminished sense of control over that work. Therefore, engagement's presumed negative correlation with amotivation may go toward establishing the construct's divergent validity.

We assessed self-efficacy (operationalized as confidence in one's own academic ability and aptitude) using the College Academic Self-Efficacy Scale (Owen and Fromen 1988). This measure is comprised of 32 items, and respondents answer on a Likert scale ranging from 1 (*Very Little Confidence*) to 5 (*A Lot of Confidence*). Need for achievement was measured using the construct's subscale on the Manifest Needs Questionnaire (Steers and Braunstein 1976). This measure consists of five items, and is scaled on a Likert scale ranging from 1 (*Never*) to 7 (*Always*). Affective commitment was measured using the construct's subscale of the Occupational Commitment Scale (Meyer et al. 1993). Participants respond to this six-item measure on a Likert scale ranging from 1 (*Strongly Disagree*) to 7 (*Very Strongly Disagree*). Finally, intrinsic motivation and amotivation were measured using the Academic Motivation Scale—College Version (Vallerand et al. 1992). This measure has 28 items, each scaled on a Likert scale of 1 (*Very Strongly Disagree*) to 7 (*Very Strongly Agree*). Coefficient alphas for these measures are presented in Table 5. Note that all alpha values are of acceptable magnitude, with the exception of the low (.57) reliability of need for achievement (NAch). Therefore, any validity findings resulting from need for achievement's correlation with engagement and/or engagement factors should be taken with caution.

2.1.2 Sample 2

2.1.2.1 Participants and Procedure The second sample consisted of 98 (34% male, 98% Caucasian) county extension agents who were employed full-time in the Midwestern United States. These 98 were from an eligible sample pool of 245, thus resulting in a

Table 5 Study 2: Samples 1–3—means (M), standard deviations (SD), Skew, Kurtosis, and Cronbach's reliability coefficients (α) for all Study 2 variables

	M	SD	Skew	Kurtosis	α
Sample 1					
UWES-17-Vigor	3.78	1.27	−.19	−.62	.90
UWES-17-Ded.	4.77	1.11	−.67	.46	.86
UWES-17-Abs	3.88	1.20	−.21	−.36	.86
UWES-17-Eng	4.11	1.10	−.37	−.15	.95
UWES-9-Vigor	4.33	1.27	−.43	−.20	.82
UWES-9-Ded.	4.43	1.09	−.36	.02	.83
UWES-9-Abs	4.24	1.27	−.26	−.05	.70
UWES-9-Eng	4.33	1.11	−.43	.08	.91
NAch	5.04	.74	−.12	−.56	.57
Intrinsic Mot.	4.93	.94	−.17	−.29	.91
Amotivation[†]	1.90	1.29	.91	−.54	.93
Self-efficacy	3.61	.56	−.41	.856	.94
Aff. Commit.	5.50	1.26	−.96	.76	.90
Sample 2					
UWES-17-Vigor	4.96	.72	.24	−.28	.80
UWES-17-Ded.	5.43	.74	.10	−.52	.87
UWES-17-Abs	4.82	.69	.31	.03	.76
UWES-17-Eng	5.06	.65	.09	.41	.91
UWES-9-Vigor	4.89	.86	.30	.11	.83
UWES-9-Ded.	5.42	.80	.04	−.15	.84
UWES-9-Abs	4.97	.77	.39	−.01	.70
UWES-9-Eng	5.09	.71	.71	−.47	.90
PsyCap	4.45	.37	.08	.11	.82
Positive affect	3.78	.56	.01	.12	.89
Negative affect	1.60	.42	.97	.96	.81
Sample 3					
UWES-17-Vigor	3.86	.50	−.23	.98	.70
UWES-17-Ded.	4.04	.62	−.66	.68	.73
UWES-17-Abs	3.40	.52	.28	.40	.73
UWES-17-Eng	3.77	.49	−.33	.75	.89
UWES-9-Vigor	4.01	.63	−.64	.62	.76
UWES-9-Ded.	3.98	.56	−.30	.14	.48
UWES-9-Abs	3.75	.53	−.12	.41	.49
UWES-9-Eng	3.91	.51	−.49	.93	.84
Job satisfaction	4.24	.76	−.92	.59	.86
POS	3.50	.72	−.53	.48	.91
Positive affect	4.09	.43	−.09	−.17	.84
Negative affect	1.59	.46	.83	.91	.86
Turnover intent.	2.13	1.14	.98	.62	.93

[†] Amotivation had initial skew of 1.53 and kurtosis of 1.43, thus was transformed using log10 prior to analyses

NAch, need for achievement; Intrinsic Mot., intrinsic motivation; Aff. Commit., affective commitment; PsyCap, psychological capital; POS, perceived organizational support; Turnover Intent., turnover intentions

response rate of 40%. Extension agents work to develop community activities and foster community involvement. Their job is largely self-directed and self-structured. Participant mean age was 41.06 years (SD = 12.18). Fifty-five percent of participants had earned a

Table 6 Study 2: Sample 1—pearson correlation coefficients

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. UWES-17-Vig	–												
2. UWES-17-Ded	.69**	–											
3. UWES-17-Abs	.83**	.71**	–										
4. UWES-17-Eng	.94**	.85**	.94**	–									
5. UWES-9-Vig	.83**	.77**	.85**	.90**	–								
6. UWES-9-Ded	.80**	.90**	.75**	.89**	.80**	–							
7. UWES-9-Abs	.80**	.76**	.85**	.81**	.73**	.75**	–						
8. UWES-9-Eng	.88**	.88**	.89**	.97**	.92**	.92**	.91**	–					
9. NAch	.33*	.50**	.35**	.42**	.44**	.43**	.39**	.46**	–				
10. Intrinsic Mot.	.49**	.57**	.46**	.55**	.46**	.55**	.48**	.54**	.42**	–			
11. Amotivation	-.16**	-.40**	-.17**	-.25**	-.27**	-.31**	-.27**	-.31**	-.47**	-.34**	–		
12. Self-efficacy	.35**	.50**	.35**	.43**	.45**	.43**	.38**	.46**	.48**	.34**	-.42**	–	
13. Aff. Commit.	.28**	.48**	.28**	.37**	.35**	.41**	.35**	.40**	.49**	.46**	-.53**	.32**	–

* $p < .05$; ** $p < .01$; NAch, need for achievement; Intrinsic Mot., intrinsic motivation; Aff. Commit., affective commitment

Table 7 Study 2: Sample 2—pearson correlation coefficients

	1	2	3	4	5	6	7	8	9	10	11
1. UWES-17-Vig	—										
2. UWES-17-Ded	.77**	—									
3. UWES-17-Abs	.66**	.71**	—								
4. UWES-17-Eng	.91**	.90**	.88**	—							
5. UWES-9-Vig	.89**	.76**	.60**	.82**	—						
6. UWES-9-Ded	.79**	.96**	.63**	.86**	.78**	—					
7. UWES-9-Abs	.63**	.69**	.91**	.83**	.58**	.63**	—				
8. UWES-9-Eng	.88**	.91**	.80**	.95**	.90**	.91**	.82**	—			
9. PsyCap	.63**	.54**	.41**	.59**	.57**	.54**	.42**	.58**	—		
10. Positive Aff.	.70**	.75**	.51**	.71**	.73**	.75**	.55**	.77**	.66**	—	
11. Negative Aff.	−.31**	−.16	−.03	−.19	−.22*	−.23*	.03	−.16	−.38**	−.27**	—

* $p < .05$; ** $p < .01$; PsyCap, psychological capital; Positive Aff., positive affect; Negative Aff., negative affect

Table 8 Study 2: Sample 3—pearson correlation coefficients

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. UWES-17-Vig	–												
2. UWES-17-Ded	.73**	–											
3. UWES-17-Abs	.65**	.65**	–										
4. UWES-17-Eng	.89**	.91**	.86**	–									
5. UWES-9-Vig	.87**	.86**	.61**	.89**	–								
6. UWES-9-Ded	.71**	.81**	.61**	.83**	.72**	–							
7. UWES-9-Abs	.70**	.83**	.76**	.87**	.74**	.66**	–						
8. UWES-9-Eng	.88**	.93**	.73**	.96**	.92**	.88**	.88**	–					
9. Job Satisfaction	.67**	.81**	.54**	.77**	.82**	.60**	.69**	.79**	–				
10. POS	.36**	.44**	.33**	.43**	.45**	.38**	.36**	.44**	.49**	–			
11. Positive affect	.34**	.32**	.08	.28**	.28**	.24**	.23**	.28**	.17	.08	–		
12. Negative affect	-.10	-.16	-.13	-.15	-.13	-.15	-.13	-.15	-.05	-.02	-.22*	–	
13. Turnover intent.	-.32**	-.42**	-.31**	-.40**	-.46**	-.38**	-.31**	-.43**	-.53**	-.62**	.08	-.04	–

$* p < .05$; $** p < .01$; *POS* perceived organizational support; Turnover Intent., turnover intentions

bachelors degree, and 42% had completed a master's degree. Participants worked an average of 48.81 h per week (SD = 5.13) and had been employed in their current position an average of 11.06 years (SD = 9.64). After relevant meetings describing the research, mass e-mails were sent out including a link to the online survey.

2.1.2.2 Additional Measures In addition to measuring work engagement, we also measured psychological capital and positive and negative affectivity. Given past research (e.g., Avey et al. 2008), it was expected that psychological capital would correlate positively with work engagement, as would positive affectivity, and that negative affectivity would correlate negatively with the construct. Psychological capital was assessed using the Psychological Capital Questionnaire (PCQ; Luthans et al. 2007). This measure has 24 items, with participants responding to each item on a Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Positive and negative affectivity were assessed using the Positive and Negative Affectivity Schedule (PANAS; Watson et al. 1988). This measure consists of 20 items, ten each for positive and negative affectivity respectively. For each affective descriptor, respondents were asked to rate the extent to which they generally feel that way, ranging on a Likert scale from 1 (*very slightly/not at all*) to 5 (*extremely*). Coefficient alphas for these measures are presented in Table 5.

2.1.3 Sample 3

2.1.3.1 Participants and Procedure The third sample was comprised of individuals employed in a public sector organization in the Southern United States. Data were collected as part of a larger study examining attitude change of organizational newcomers. Engagement was assessed on the employees' three-month anniversary via an online survey emailed to participants. The other measures (for construct validation purposes) were assessed either 3 months earlier (at the start of employment), at the same time as, or 3 months later (on their six-month anniversary) than the engagement measure.

The number of individuals who completed the initial survey (at the start of employment) was 132 out of 141 available newcomers, thus resulting in a response rate of 94%. After 3 months, at the time engagement was measured, the number participating was 120 out of 128 (93%). Finally, the number participating at their six-month anniversary was 99 out of 113 (88%). The majority of participants were male (61%) and Caucasian (86%). Mean age was 39 years (SD = 10.7). Thirty percent of participants held administrative jobs, 28% held jobs as professional trainers, 17% held clerical jobs, and 11% were employed in technical jobs. These demographics were representative of the make-up of the organization as a whole.

2.1.3.2 Additional Measures For purposes of supporting convergent validity, we measured positive affectivity (e.g., Avey et al. 2008), overall job satisfaction (e.g., Saks 2006), and perceived organizational support (e.g., Kinnunen et al. 2008; Saks 2006), all of which should correlate positively with engagement as evidenced by, for instance, the research cited herein. For purposes of divergent validity, we measured negative affectivity (e.g., Avey et al. 2008) and turnover intentions (e.g., Kinnunen et al. 2008; Saks 2006), both of which we expected to correlate negatively with engagement as per the aforementioned references.

Positive and negative affectivity was measured using the 20-item PANAS (Watson et al. 1988), assessed at the start of the employee's tenure with the organization. At the same

time as engagement was assessed, we assessed overall job satisfaction with the Michigan Organizational Satisfaction Scale (Camman et al. 1979). Participants respond to this three-item measure on a Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Three months later we assessed turnover intentions using Boroff and Lewin's (1997) two-item measure, which asked participants to respond on a Likert Scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*. At this same time, we also assessed perceived organizational support using Eisenberger et al.'s (1986) Survey of Perceived Organizational Support. This six-item survey requires participants to respond on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Coefficient alphas for these measures are presented in Table 5.

2.2 Results and Discussion

2.2.1 Confirmatory Factor Analyses

For each sample, a series of CFAs were conducted, each of which is reported in Table 9. First, the proposed three-factor structure consisting of the vigor, dedication, and absorption dimensions was examined. Subsequently, in an effort to fully explore the effects of high correlations typically reported between absorption and vigor (Salanova et al. 2001; Schaufeli et al. 2006; Schaufeli et al. 2002), absorption and vigor were collapsed into a single factor (as per Schaufeli et al. 2002), termed "involvement," and a CFA was conducted to explore this possible two-factor structure of engagement. Finally, because a one-factor model of overall engagement that has at times been found to be a better fit than the three-factor model (e.g., Sonnentag 2003), we also conducted a CFA examining this one-factor model.

2.2.1.1 Sample 1 Consistent with past research as outlined earlier, the present sample supported the internal consistency of the UWES for all factor structures, with score reliabilities as shown in Table 5. Results for the UWES-17 (see Table 9) indicated that, although the correlation between involvement and dedication decreased, the fit of the two-factor model did not statistically significantly improve upon the fit of the three-factor model, $\Delta\chi^2$ (2) = 3.95, *ns*. Also, results revealed that the one-factor model of overall engagement was a very poor fitting model, $\chi^2/df = 10.47$, RMSEA = .20, and was inferior to both the three-factor model, $\Delta\chi^2$ (3) = 334.74, *p* < .001 and the two-factor model, $\Delta\chi^2$ (1) = 338.69, *p* < .001.[1]

CFA results for Sample 1 for the UWES-9 showed the two-factor structure was a statistically significantly worse fit than was the three-factor structure, $\Delta\chi^2$ (2) = 99.14, *p* < .001. Finally, the one-factor structure of the measure was found to be a statistically significantly worse fit than both the three-factor structure, $\Delta\chi^2$ (3) = 298.34, *p* < .001, and the two-factor structure, $\Delta\chi^2$ (1) = 199.20, *p* < .001. Therefore, results from the first sample revealed that the three-factor structure was found to be the best conceptualization of engagement for both versions of the UWES.

2.2.1.2 Sample 2 Consistent with past research and also with Sample 1, Sample 2 supported the internal consistency of all factor structures of the UWES, with reliabilities of scores shown in Table 5. For the UWES-17, as shown, contrary to Sample 1, here the

---

[1] Recognizing Chi-square's susceptibility to sample size, additional goodness-of-fit indices are provided in Table 8.

Table 9 Study 2: Samples 1–3—confirmatory factor analyses

UWES version/Sample	Model	χ^2	df	χ^2/df	RMSEA	NFI	CFI
UWES-17—Sample 1	3 Factors	911.85**	116	7.86	.17	.78	.80
	2 Factors	907.90**	118	7.69	.17	.77	.80
	1 Factor	1246.59**	119	10.48	.20	.72	.74
UWES-17—Sample 2	3 Factors	348.26**	117	2.98[b]	.14	.64	.71
	2 Factors	252.02**	118	2.14[b]	.11	.74	.83
	1 Factor	259.07**	119	2.18[b]	.11	.73	.83
UWES-17—Sample 3	3 Factors	286.89**	116	2.47[b]	.11	.73	.81
	2 Factors	293.21**	118	2.48[b]	.11	.73	.81
	1 Factor	297.88**	119	2.50[b]	.11	.72	.81
UWES-9—Sample 1	3 Factors	111.61**	24	4.65[a]	.12	.93	.94
	2 Factors	210.75**	26	8.11	.17	.88	.90
	1 Factor	409.95**	27	15.18	.24	.79	.80
UWES-9—Sample 2	3 Factors	46.32**	24	1.93[b]	.10	.91	.95
	2 Factors	59.43**	26	2.29[b]	.12	.88	.93
	1 Factor	64.98**	27	2.41[b]	.12	.87	.92
UWES-9—Sample 3	3 Factors	65.72**	24	2.74[b]	.12	.86	.90
	2 Factors	66.85**	26	2.57[b]	.11	.86	.91
	1 Factor	55.91**	27	2.07[b]	.09	.89	.94

For χ^2 coefficient, ** $p < .01$

For χ^2/df ratio, [a] ratio < 5 (as per Wheaton et al. 1977; [b] ratio < 3 (as per Carmines & McIver, 1981)

two-factor structure was a statistically significant improvement over the three-factor structure ($\Delta\chi^2$ (1) = 96.24, $p < .001$). Additionally, the one-factor structure was statistically significantly better than the three-factor structure ($\Delta\chi^2$ (2) = 89.19, $p < .001$) but statistically significantly worse than the two-factor structure ($\Delta\chi^2$ (1) = 7.05, $p < .01$). While unusual, this finding of the two-factor structure as the best fit a) supports the widely-accepted contention that the absorption and vigor dimensions are highly intercorrelated, and b) remains consistent with a multifactorial conceptualization of engagement as opposed to a unidimensional conceptualization as has been suggested (e.g., Christian and Slaughter 2007; Sonnentag 2003).

CFAs from Sample 2 for the UWES-9 indicated that the three-factor structure yielded a statistically significantly better fit than the two-factor structure ($\Delta\chi^2$ (2) = 13.11, $p < .01$). Furthermore, both the three-factor structure ($\Delta\chi^2$ (3) = 18.66, $p < .001$), and, to a lesser degree, the two-factor structure ($\Delta\chi^2$ (1) = 5.55, $p < .05$) were found to be statistically significantly better than the one-factor model. In Sample 2, as in Sample 1, for the UWES-9, the three-factor structure was found to be the most appropriate conceptualization of the construct. However, for the UWES-17, a two-factor solution was found to provide the best fit with the data. This is inconsistent with the analyses of Sample 1, although it still recognizes engagement as a multifactorial construct.

2.2.1.3 Sample 3 Consistent with results from the first two samples, and also consistent with past research, Sample 3 supported the internal consistency of the UWES-17 for all factor structures (Table 5). However, reliability results for scores resulting from the

UWES-9 were less desirable (Table 5), with scores on the one-factor structure being the only ones found to be sufficiently reliable ($\alpha = .84$). Therefore, results from Sample 3 should be interpreted with caution.

Results for the CFAs on the UWES-17 with this sample indicated the three-factor solution yielded a statistically significantly better fit than the two-factor solution ($\Delta\chi^2$ (2) = 6.32, $p < .05$), and the one-factor structure was statistically significantly inferior to both the three-factor structure ($\Delta\chi^2$ (3) = 10.99, $p < .05$) and the two-factor structure ($\Delta\chi^2$ (1) = 4.67, $p < .05$). CFA results for the UWES-9 with this sample indicated that the one-factor structure was found to be a statistically significantly better fit than were both the three-factor structure ($\Delta\chi^2$ (3) = 9.81, $p < .05$) and the two-factor structure ($\Delta\chi^2$ (1) = 10.94, $p < .001$). Note that these latter two factor structures did not statistically significantly differ from one another ($\Delta\chi^2$ (2) = 1.13, ns). Thus, results from the third sample indicate that, whereas the three-factor structure was found to be the best conceptualization of the construct of engagement in the UWES-17, this was not the case in the UWES-9, which yielded the one-factor structure as the best fit. This finding, however, should be taken with caution, since there were score reliability problems with two factors in Sample 3's UWES-9 (Table 5). Likewise, this may indicate that the UWES may be more applicable to the measurement of academic engagement in students than it is to the measurement of work engagement in full-time employees.

2.2.2 Construct Validation

2.2.2.1 Sample 1 Consistent with expectations, convergent validity was confirmed by engagement's statistically significant and strong positive correlations with self-efficacy, need for achievement, intrinsic motivation, and affective occupational commitment. This was the case for both versions of the UWES, with both the three-factor conceptualization and the one-factor overall engagement measure (see Table 6). Additionally, divergent validity was supported by engagement's strongly negative correlation with amotivation for both versions of the UWES and with both the three- and one-factor conceptualizations of the construct (see Table 6).

2.2.2.2 Sample 2 Convergent validity was supported via strongly significant positive associations between engagement and both psychological capital and positive affectivity. This held true for both versions of the UWES and for both factor conceptualizations of the construct (see Table 7). Evidence for divergent validity, however, was not so favorable. That is, whereas we would expect a statistically significant and negative correlation between engagement and negative affectivity, results for the UWES-17 revealed only non-statistically significant negative correlations with overall engagement ($r = -.19, p > .05$) and with the dedication ($r = -.16, p > .05$) and absorption ($r = -.03, p > .05$) factors. There was, however, a highly statistically significant correlation between negative affectivity and the vigor factor ($r = -.31, p < .01$). Results for the UWES-9 were comparable. That is, no statistically significant relationship was found between negative affectivity and engagement for either the one-factor measure ($r = -.16, p > .05$) or for the absorption component of the three-factor measure ($r = .03, p > .05$). Nevertheless, significant negative relationships with engagement were supported for both the vigor factor ($r = -.22, p < .05$) and the dedication factor ($r = -.23, p < .05$), thus lending only partial support to engagement's divergent validity (see Table 7).

2.2.2.3 Sample 3 Convergent validity was confirmed by engagement's statistically significant and strongly positive correlations with overall job satisfaction, perceived organizational support, and positive affectivity. For the most part, this held for both versions of the UWES and for both factor conceptualizations of engagement. The only exception to this was that in the UWES-17, positive affectivity was not significantly correlated with the absorption factor ($r = .08$, $p > .05$). This anomalous finding, however, was not replicated in the UWES-9, wherein the correlation was ($r = .23$, $p < .01$) (see Table 8).

Divergent validity was again examined via engagement's correlations with negative affectivity and turnover intentions. Unfortunately, with the former of these two, similar problems arose as were found in Sample 2 in that negative affectivity was not found to be significantly negatively associated with engagement in either version of the UWES or with either factor structure of engagement. Divergent validity was, however, supported by engagement's correlations with turnover intentions. For both the UWES-17 and the UWES-9 and for both factor conceptualizations of engagement, turnover intentions were found to be strongly statistically significantly negatively correlated with engagement (see Table 8).

3 General Discussion

This series of studies examining the Utrecht Work Engagement Scale(s) yielded some interesting findings. First, Study 1 filled a gap in the literature by focusing on the methodology behind the development of the Utrecht Work Engagement Scale, and by empirically reevaluating the appropriateness of the original scale development using more theoretically sound methodology and factor extraction techniques. Study 2 built upon the findings from Study 1, offering both an analysis of the psychometric properties of the UWES, while also speaking to how the shorter UWES-9 can best be evaluated and compared to the original longer version.

3.1 Construct Validity

Examinations of construct validity for all factor conceptualizations of both the UWES-17 and the UWES-9 in all three samples supported both the convergent and divergent validity of the use of an overall engagement construct and also of its three components. The only issue that arose was the lack of a statistically significantly negative correlation between engagement and negative affectivity in both Sample 2 and Sample 3 where the latter was measured. Nevertheless, while the lack of a statistically significant negative correlation here is undoubtedly noteworthy, it is not fatal, since the correlations were in fact negative (although not statistically so), and since other constructs correlated statistically significantly with engagement as expected. Moreover, this relationship was demonstrated by two distinct samples, thus lessening the possibility that it was a sample-specific finding, and lending credence to its likelihood.

Rather, these findings force consideration of the possibility that the lack of a correlation between these variables does not go toward failing to support the construct validity of engagement, but rather, as Cronbach and Meehl (1955) suggest in their seminal paper, enriches the nomological framework under which we view engagement, as well as allowing for greater specificity of the engagement construct itself. Specifically, the lack of a statistically significant correlation between engagement and negative affectivity has important implications for the understanding of the engagement construct as a whole. That

is, while engagement has oftentimes been understood to be related to unpleasant experience as well as to positive experience, as evidenced, for example, by its negative relation with burnout (Maslach and Leiter 1997), the current findings suggest that engagement may be more cognitive than it is affective, or emotional. Further, these findings suggest that the affective component of engagement is more significantly related to positive affective experience than it is to negative affective experience. These results support Schaufeli and colleagues' (Schaufeli et al. 2002) earlier contention that engagement should *not* be measured as the *antipode* of burnout.

Additionally, it is interesting to consider that engagement may involve some dimensions of Csikszentmihalyi's 'flow' construct (1975, 1990), and may in fact go beyond affectivity and emotionality. In particular, flow specifies that the individual loses self-consciousness when he or she is completely absorbed in a task or activity. The individual may therefore be too involved in (enjoying) the task at hand to be concerned about any negative affectivity that he or she might be experiencing. These interesting findings should give rise to future research, which should further explore engagement's relationship with affectivity and, moreover, wellbeing.

3.2 Factor Structure

One of the primary goals of the present research was to examine the proposed factor structure of engagement as measured by the UWES. We offer a much-needed reexamination of Schaufeli and colleagues' (Schaufeli et al. 2002) development of the original UWES, including four initial EFAs (two per UWES version) prior to forcing factors in subsequent CFAs, and strict adherence to a factor analytic sequence suggested by various researchers (e.g., Fabrigar et al. 1999; O'Connor 2000; Van Prooijen and Van Der Kloot 2001).

Our initial series of EFAs conducted on the UWES-17 in Study 1 suggest the potential presence of fourth and/or fifth factors. While the fifth factor from the Parallel Analysis can be almost entirely discounted as a trivial factor, the fourth factor is a distinct possibility, given its comparatively strong presence in the Parallel Analysis EFA and also in the subsequent rotated EFA. The latter EFA indicated particular items falling within this fourth factor, the content of which allowed the researchers to determine that all four items hung together well and together could be conceptualized as a fourth factor of 'perseverance' that served to further subdivide the vigor factor. Although this finding was unable to be confirmed via ML estimation due to a nonpositive definite covariance matrix, a CFA using ULS estimation as recommended by Wothke (1993) provided some initial support to the four-factor construct. Nevertheless, considering the estimation method that had to be used in the present study due to a nonpositive definite matrix, future research should further explore the tenability of this fourth factor prior to its full acceptance. Furthermore, an additional CFA further confirmed that the three-factor model of the UWES-17 did not ideally fit the data. With the more parsimonious UWES-9, however, both EFAs and also the subsequent CFA were much cleaner and more supportive of the original three-factor structure of the UWES, in that the EFAs indicated three factors with items loading as suggested by Schaufeli et al. (2006), and the CFA supported this conceptualization of engagement with acceptable goodness-of-fit indices. These findings offer initial support of engagement as a three-factor construct, particularly in the UWES-9. The longer version UWES may be more likely to produce anomalies such as what we might call 'factor bursts' as indicated by the finding of a fourth (perseverance) factor.

In Study 2, results from the first sample support the three-factor conceptualization of engagement as statistically significantly better than both the two-factor and one-factor conceptualizations in the UWES-9, and as statistically significantly better than the one-factor conceptualization in the UWES-17. These findings should spur future research to further investigate the possibility of uni- or bi-dimensionality of the UWES in their samples. However, despite the fact that in the present study the three-factor structure was not found to be a statistically significantly better fit than the two-factor structure in the UWES-17, the authors would still recommend that the three-factor structure be retained, because it is consistent with the original structure of the UWES and also considering that some previous research has found the two-factor structure to be a statistically significantly worse fit than the three-factor structure (Schaufeli et al. 2002), a finding further suggested by the present research with the UWES-9. Nevertheless, it does warrant further research.

Likewise, as results from the subsequent studies were analyzed, support for the three-factor structure began to waver. Interestingly, in the second sample, factor analyses of the UWES-17 found that the two-factor structure provided the best fit with the data. Although this necessarily fails to support the three-factor structure, it continues to support engagement as best conceptualized multifactorially. In the 9-item version, however, the original three-factor structure was once again supported as the statistically significantly best fit. Furthermore, CFAs from the third sample arguably yielded the most inconsistent results. Whereas the three-factor structure was once again found to be the statistically significantly best fit in the longer measure, the one-factor structure was found to be the statistically significantly best fit as measured by the shorter measure. This inconsistent finding, however, may be readily explained by an examination of score reliabilities, which will be discussed below, and would lead us to interpret results from the third sample with caution.

3.3 UWES-17 versus UWES-9

Arguably, some of the most important practical implications of the present research are regarding the assessment of the two UWES versions in relation to one another, particularly whereas the UWES-9 has been the subject of very little empirical research (for the two exceptions, see Seppälä et al. 2009, and Schaufeli et al. 2006). Therefore, in addition to offering an assessment of Schaufeli and colleagues' development of the original UWES-17 (Schaufeli et al. 2002), the present research also analyzed this scale in comparison to its more recently proposed parsimonious alternative, the UWES-9 (Schaufeli et al. 2006).

It appears as though the UWES-9 could serve as a viable—and perhaps even preferable—alternative to the longer UWES-17. In Study 2's Sample 1 and Sample 2, the UWES-9 yielded much the same results as did the UWES-17, in addition to yielding a cleaner and more consistent factor structure over the UWES-17 in each analytic step in Study 1 (e.g., limited cross-loadings and the lack of the fourth factor). Therefore, despite the fact that the UWES-9 yielded some inconsistent results in Study 2's Sample 3, we believe that the Sample 3 results for that version may be sample-specific and potentially due to either the nature of the sample, and/or problems with Item 14 in the absorption factor ("I get carried away when I am working"), which should be further investigated by future research. Furthermore, Sample 3 results for the UWES-9 were the only results that failed to support engagement as a multifactorial construct.

Whereas reliabilities (see Table 5) of scores for all factor structures of the UWES-17 were supported in all samples, score reliabilities for all factor structures of the UWES-9 were supported only in the first two samples of Study 2. Unfortunately, score reliability results for the UWES-9 in Sample 3 were less desirable, with the one-factor structure being

the only one found to yield sufficiently reliable scores ($\alpha = .84$). For the three-factor structure, whereas scores on the vigor factor were reliable ($\alpha = .76$), the dedication factor yielded scores with very poor reliability ($\alpha = .48$), although an item analysis failed to identify any faulty items included in the factor. The absorption factor also yielded scores with poor reliability ($\alpha = .49$), although an item analysis identified Item 14 ("I get carried away when I am working") as a faulty item, without which the reliability of scores from the scale increases to $\alpha = .81$.

One explanation for the reliability problem with this third sample could be that these participants were in the third month of new jobs and therefore were still considered to be within the 'socialization period,' during which they may still have been exploring their new organization, job demands, and the skills necessary to meet such requirements. That is, we may not expect a measure of job engagement to be stable for individuals who are still learning their jobs. (Note, for instance, the lack of problematic reliabilities using the extension agent working sample, in which the mean tenure was 11.06 years (SD = 9.64).)

Nevertheless, future researchers would do well to further examine the viability of Item 14. However, this item was only found to be so problematic with the UWES-9, presumably because of the lesser number of items per factor in that version. Future researchers should examine the viability of the item and the possibility of replacing it in the UWES-9 with a different absorption item from the UWES-17. Nevertheless, we believe that after such additional research and possible replacements, this initial research on the UWES-9 appears to support the measure as a viable parsimonious alternative to the UWES-17, and that once such further research and adaptations are made, the UWES-9 may indeed replace the UWES-17 as a more parsimonious measure of the engagement construct that more reliably recognizes engagement's multifactorial structure than does the UWES-17.

3.4 Limitations

Despite the contributions of the current research, the present series of studies is not without its limitations. The sample in Study 1 as well as one of the samples in Study 2 both use a university student sample, often criticized as lacking generalizability. However, in this case such a sample was sought for purposes of maintaining sample similarity with Schaufeli and colleagues' (Schaufeli et al. 2002) sample upon which the original UWES-17 was developed. The student sample is also appropriate here since the engagement measure is often used with students to measure their engagement in their schoolwork. However, we also balanced these student samples with worker samples. We hope that the relative similarity of the results with those of the working samples herein may lend credibility to the use of student samples.

Nevertheless, generalizability may also be limited by the recruitment procedure. Specifically, it is necessary to consider the possibility that mass e-mails distributed to potential participants by an authority figure may have coerced participation from otherwise hesitant individuals. However, it must be noted that in some cases participation rates were approximately only half of the eligible individuals. While this is still a respectable participation rate, it also begs consideration of the possibility that those individuals who opted to participate were somehow characteristically different on the constructs of interest than those who opted not to participate. This may seem particularly pertinent given that the primary construct of interest is engagement. However, its potential negative influence on the results is tempered by the consideration that the present research is designed primarily to investigate the structure of measures tapping this construct, rather than utilizing the construct in a nomological net.

Another potential limitation is that all data were collected via self-report measures, and that only one measure was used for each construct of interest. Therefore, mono-method bias or common method variance may problematic when the issue of validity is considered. Future research may consider examining the psychometric properties of these constructs and their relation with engagement with a multitrait-multimethod (MTMM) matrix (see Campbell and Fiske 1959), which would go toward overcoming this limitation. However, the present study does take a step in the right direction by evaluating various constructs tapping both convergent and discriminant construct validity, and utilizing multiple samples (evidencing what Campbell and Fiske would call heterotrait-monomethod triangles). Furthermore, self-report measures appear to be the best if not the only way to measure many of the constructs examined herein. Indeed, recent research has attested to the appropriateness of self-report measures, arguing that they may not produce such biases as are often attributed to them. Spector (2006) has questioned the problem of common method variance, while Goffin and Gellatly (2001) found that self- and peer-report measures may be largely redundant, and that self-reported responses are primarily driven by experiences as opposed to systematic bias resulting from defensive responding.

Finally, arguably the greatest limitation with this research is the score reliability problem found in Sample 3 with the UWES-9 for both the absorption and dedication factors. However, given that one of the purposes of this research is to evaluate the psychometric properties of the UWES versions, the unreliability of scores yielded by these two factors on Sample 3's UWES-9 is in itself contributory to the findings and the literature, and highlights an issue of possible concern for the UWES-9 and, more specifically, for the inclusion of Item 14.

3.5 Conclusion

In sum, the present research aims to analyze the two versions Utrecht Work Engagement Scale in addition to evaluating the initial scale development of the original UWES. The present research fills gaps left open by past research by using exploratory factor analysis as a first step in conceptualizing engagement for appropriate subsequent measurement. We apply best practices for scale development to critique the development of the Utrecht Work Engagement Scale, which we argue was methodologically flawed. As such, in Study 1 we describe the initial development of the UWES and contrast it with best practices in scale development, followed in Study 2 by a series of confirmatory studies further examining the exploratory findings from Study 1. Both studies aimed to explore how engagement might best be measured, and how the Utrecht Work Engagement Scale might be improved.

We found that conceptualizing engagement as a multifactorial construct can be beneficial and can aid in interpretation when compared to a one-factor conceptualization. In fact, the presence of a fourth factor, perseverance, may be more fully explored on different samples. The present research also contributes new evidence of convergent and divergent validity, analyzes the UWES factor structure, and critically compares the UWES-17 versus the UWES-9, the latter of which we believe holds great potential as a favored version of the UWES. Throughout this discussion, we have suggested various cautions that researchers would do well to consider prior to utilizing either of the UWES measures, and have also made various specific suggestions for how future research might best proceed.

References

Avey, J. B., Wernsing, T. S., & Luthans, F. (2008). Can positive employees help positive organizational change? Impact of psychological capital and emotions on relevant attitudes and behaviors. *Journal of Applied Behavioral Science, 44*, 48–70.

Boroff, K. E., & Lewin, D. (1997). Loyalty, voice, and intent to exit a union firm: A conceptual and empirical analysis. *Industrial and Labor Relations Review, 51*, 50–63.

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27*, 509–540.

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

Camman, C., Fichman, M., Jenkins, D., & Klesh, J. (1979). *The Michigan Organizational Assessment Questionnaire*. Unpublished Manuscript: University of Michigan.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables. In G. W. Bornstedt & E. F. Borgatta (Eds.), *Social measurement: Current issues* (pp. 65–115). Beverly Hills: Sage.

Christian, M. S., & Slaughter, J. E. (2007, August). *Work engagement: A meta-analytic review and directions for research in an emerging area*. Paper presented at the 67th annual meeting of the Academy of Management, Philadelphia, PA.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-esteem processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes in development: Minnesota symposium on child psychology* (Vol. 23, pp. 167–216). Hillsdale, NJ: Lawrence Erlbaum.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*, 1–9.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281–302.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety: Experiencing flow in work and play*. San Fransisco, CA: Jossey-Bass.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper & Row.

Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology, 71*, 500–507.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299.

Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling, 3*, 62–72.

Goffin, R. D., & Gellatly, I. R. (2001). A multi-rater assessment of organizational commitment: Are self-report measures biased? *Journal of Organizational Behavior, 22*, 437–451.

Hallberg, U. E., Johansson, G., & Schaufeli, W. B. (2007). Type A behavior and work situation: Associations with burnout and work engagement. *Scandinavian Journal of Psychology, 48*, 135–142.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*, 191–205.

Horn, J. L. (1965). A rationale and test for a number of factors in factor analysis. *Psychometrica, 30*, 179–185.

Kinnunen, U., Feldt, T., & Mäkikangas, A. (2008). Testing the effort-reward imbalance model among Finnish managers: The role of perceived organizational support. *Journal of Occupational Health Psychology, 13*, 114–127.

Koestner, R., & Losier, G. F. (2002). Distinguishing three ways of being internally motivated: A closer look at introjection, identification, and intrinsic motivation. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 101–121). Rochester, NY: University of Rochester Press.

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

Luthans, F., Avolio, B. J., & Youseff, C. (2007). *Psychological capital: Developing the human capital edge*. Oxford: Oxford University Press.

Macey, W. H., & Schneider, B. (2008). The meaning of employee engagement. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 3–30.

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

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741–749.

Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology, 78*, 538–551.

Montgomery, A. J., 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, and Coping, 16*, 195–211.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*, 396–402.

Owen, S. V., & Fromen, R. D. (1988). *Development of a college academic self-efficacy scale*. Paper presented at the Annual Meeting of the National Council on Measurement in Education. New Orleans, LA.

Saks, A. M. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology, 21*, 600–619.

Salanova, M., Schaufeli, W. B., Llorens, S., Peiró, J. M., & Grau, R. (2001). Desde el 'burnout' al 'engagement': ¿una nueva perspectiva? [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. (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., & Salanova, M. (2007). Efficacy or inefficacy, that's the question: Burnout and engagement, and their relationships with efficacy beliefs. *Anxiety, Stress, & Coping, 20*, 177–196.

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement, 66*, 701–716.

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

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

Schaufeli, W. B., Taris, T. W., Le Blanc, P. M., Peeters, M., Bakker, A. B., & De Jonge, J. (2001). Maakt arbeid gezond? Op zoek naar de bevolgen werknemer [Does work make people healthy? In search of the engaged worker]. *De Psycholoog, 36*, 422–428.

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

Seppälä, P., Mauno, S., Feldt, T., Hakanen, J., Kinnunen, U., Tolvanen, A., et al. (2009). The construct validity of the Utrecht Work Engagement Scale: Multisample and longitudinal evidence. *Journal of Happiness Studies, 10*, 459–481.

Shirom, A. (2003). Feeling vigorous at work? The construct of vigor and the study of positive affect in organizations. In D. Ganster & P. L. Perrewe (Eds.), *Research in organizational stress and well-being* (Vol. 3, pp. 135–165). Greenwich, CT: JAI Press.

Silverstein, A. B. (1987). Note on the parallel analysis criterion for determining the number of common factors or principal components. *Psychological Reports, 61*, 351–354.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*, 102–111.

Snape, E., & Redman, T. (2003). An evaluation of a three-component model of occupational commitment: Dimensionality and consequences among United Kingdom human resource management specialists. *Journal of Applied Psychology, 88*, 152–159.

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.

Spector, P. (2006). Method variance in organizational research: Truth or urban legend? *Organizational Research Methods, 9*, 221–232.

Steers, R. M., & Braunstein, D. N. (1976). A behaviorally based measure of manifest needs in work settings. *Journal of Vocational Behavior, 9*, 251–266.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senecal, C. B., & Vallieres, E. F. (1992). The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*, 159–172.

Van Prooijen, J. W., & Van Der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement, 61*, 777–792.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.

Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology* (pp. 84–136). San Francisco: Jossey Bass.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendation for best practices. *Counseling Psychologist, 34*, 806–838.

Wothke, W. (1993). Nonpositive definite matrices in structural equation modeling. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 256–293). Newbury Park, CA: Sage.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432–442.