Industrial Health 2015, **53**, 113–123

Original Article

Cross-national and longitudinal investigation of a short measure of workaholism

Johanna RANTANEN[1]*, Taru FELDT[2], Jari J HAKANEN[3, 4], Katja KOKKO[5], Mari HUHTALA[2], Lea PULKKINEN[2] and Wilmar SCHAUFELI[6, 7]

[1]Department of Teacher Education, University of Jyvaskyla, Finland
[2]Department of Psychology, University of Jyvaskyla, Finland
[3]Helsinki Collegium for Advanced Studies, Finland
[4]Finnish Institute of Occupational Health, Finland
[5]Gerontology Research Center and Department of Health Sciences, University of Jyvaskyla, Finland
[6]Department of Work and Organizational Psychology and Professional Learning, University of Leuven, Belgium
[7]Department of Social and Organizational Psychology, Utrecht University, The Netherlands

Received June 11, 2014 and accepted October 9, 2014
Published online in J-STAGE November 8, 2014

Abstract: The present study investigated the factor structure of the 10-item version of the Dutch Work Addiction Scale (DUWAS). The DUWAS-10 is intended to measure workaholism with two correlated factors: working excessively (WE) and working compulsively (WC). The factor structure of the DUWAS-10 was examined among multi-occupational samples from the Netherlands (*n*=9,010) and Finland (*n*=4,567) using confirmatory factor analysis (CFA). CFAs revealed that the expected correlated two-factor solution showed satisfactory fit to the data. However, a second-order factor solution, where WE comprised the first-order factors "working frantically" and "working long hours", and WC the first-order factors "obsessive work drive" and "unease if not working", showed significantly better fit to the data. The expectation of factorial group invariance of the second-order factor structure between the Dutch and Finnish samples was also supported. Moreover, factorial time invariance was observed across a two-year time lag in a sub-sample of Finnish managers (*n*=459). In conclusion, the DUWAS-10 was found to be a comprehensive measure of workaholism, meeting the criteria of factorial validity in multiple settings, and can thus be recommended for use in both research and practice.

Key words: Workaholism, Cross-national, Factorial validity, Longitudinal study, Confirmatory factor analysis

Introduction

Individuals' investment in work and, particularly, its excessive and compulsive characteristics have been studied with inventories. The earliest –rather lengthy– measures are the 25-item Workaholism Battery (WorkBAT[1)]) and the 25-item Work Addiction Risk Test (WART[2)]). It was by combining items from these two scales that the 20-item Dutch Work Addiction Scale (DUWAS[3)]) was developed. Recently, abbreviated 17-[4, 5)] and 10-item[6, 7)] versions of the DUWAS have been published, as both researchers and practitioners have a need for questionnaires that are not only valid but also brief. Of these the 10-item version, on

*To whom correspondence should be addressed.
E-mail: k.johanna.rantanen@jyu.fi

the basis of results from Dutch and Spanish samples, has been found to be more valid than the 17-item version[8]. The main aim of the present study was to further test the factorial validity of the DUWAS-10 for cross-national and longitudinal research purposes. This was accomplished by investigating workaholism among diverse occupational groups in two national settings.

Although no unanimous definition of workaholism exists, conceptual reviews have noted two recurring features in most definitions: working unreasonably hard and working out of an internal, obsessive drive[9, 10]. Working unreasonably hard refers to the behavioral tendency to take on excessively many work-related tasks and devote an excessive amount of time to working at the expense of social relationships and leisure time. Typically, this heavy investment in work goes beyond the formal job description and employer expectations and is not motivated by financial hardship. Working out of an internal, obsessive drive refers to a cognitive, mental state of feeling obliged to work, of reluctance to switch oneself off from work, and preoccupation with work-related matters even when not working. Accordingly, workaholism is defined by Schaufeli *et al.*[5] (p. 322) as “the tendency to work excessively hard (the behavioral dimension) and being obsessed with work (the cognitive dimension), which manifests itself in working compulsively”. In the DUWAS-10, both sub-dimensions of workaholism − working excessively (WE) and working compulsively (WC) − are measured with five items each (see appendix Schaufeli *et al.*[5], p. 343).

The factor structure of the DUWAS-10

Previous factor analytic studies of the DUWAS-10 are based on cross-sectional studies and, as yet, to samples from only a few countries. Schaufeli *et al.*[6] performed a confirmatory factor analysis in a sample of 2,115 Dutch medical residents. Their results supported the hypothesized correlated two-factor structure of the DUWAS-10, showing an intercorrelation between the WC and WE factors of 0.55. Schaufeli *et al.*[5] in turn used confirmatory factor analysis to study the factorial validity of the DUWAS-10, in large samples of Dutch (n=3,797) and Japanese (n=1,655) employees with diverse occupations (e.g., nurses, engineers, managers). Their results showed that the hypothesized correlated two-factor model, with WC and WE as the factors, fitted to the data of both countries equally well. The correlation between the WC and WE was 0.50 and 0.59 in the Dutch and Japanese samples, respectively. However, the two-factor structure was not fully invariant across the Dutch and Japanese samples: two WE items (#3 and #5) and two WC items (#2 and #4) were revealed to have non-invariant factor loadings, indicating that these factor loadings varied significantly across the samples.

Del Líbano *et al.*[8] also performed confirmatory factor analyses comparing samples of Dutch (n=2,714) and Spanish (n=550) employees from different occupational sectors (e.g., services, education, industry and commerce). The correlated two-factor model fitted to their cross-national data relatively well, particularly after allowing, in both samples, the same two pairs of item residuals within the WC factor to correlate (residuals for items #1 and #3 and items #4 and #5). These error covariances could be interpreted owing to overlapping item content; that is, items 1 and 3 both refer to working out of an inner obligation and without joy, and items 4 and 5 both refer to employees’ negative feelings while not working. The two-factor structure (with two error covariances) was similar across countries; the only significant cross-sample differences were the factor correlation between WE and WC (0.79 for Spanish and 0.53 for Dutch employees), and one factor loading (item #3).

In line with del Líbano *et al.*[8], Littman-Ovadia *et al.*[11], using a sample of Hebrew white-collar employees (n=351), also observed that the correlated two-factor model of DUWAS-10 fitted their data better if the error covariances between WC items #1 and #3 and items #4 and #5 were allowed. In their study, the factor correlation between WE and WC was 0.76, a rather high similar to that found for their Spanish sample by del Líbano *et al.*[8]. Finally, Andreassen *et al.*[12], in their study of an occupationally heterogenous sample of Norwegian employees (n=661), concluded that although their data supported the correlated two-factor model of the DUWAS-10 to some extent, the model fit was not fully acceptable. Andreassen *et al.*[12] do not report any further analyses or model refinement; however, inspection of the factor loadings in their study shows that neither WE items #2 and #4 (loadings 0.44 and 0.41, respectively) nor WC item #5 (loading 0.46) loaded very strongly onto their respective latent factors of WE and WC.

The present study

As shown by the above review, the factorial validity of the DUWAS-10 has been investigated in several studies with reasonably large sample sizes, including two cross-national comparisons (the Netherlands vs. Japan, the Netherlands vs. Spain). However, despite the support obtained for the factorial validity of the DUWAS-10, not all the items[5)] or the factor correlation between WE and WC[8)] were invariant across the samples from different countries. Therefore more cross-national comparison research is needed before it can be concluded whether the DUWAS-10, originally developed in the Netherlands, meets both the criteria of factorial and measurement invariance across samples from different countries and the criteria of cross-national validity (i.e. the scale measures what it is intended to measure in different national contexts with different working life policies and culture), thereby allowing more reliable cross-national studies on workaholism in the future. In other words, cross-national invariance of the factor structure is a primary prerequisite that has be to thoroughly investigated and established for any scale (see for example Little[13)], Meredith and Teresi[14)]) before it can be used, in translation in their own countries, by researchers and practitioners with confidence that the translated items truly reflect the latent factors of the phenomenon in question similarly and as intended by the original scale.

In addition, model refinement by releasing error covariances between items that constitute latent factors was needed in previous factor analytic studies of the DUWAS-10 to obtain acceptable model fit[8, 11)]. In fact, model refinement by releasing error covariances that are contentually and conceptually well justified[8, 11)] as along with the low factor loadings observed for some items[12)] implies that the proposed correlated two-factor structure of DUWAS-10 may not cover all the dimensions underlying workaholism. In sum, the factorial validity of the DUWAS-10 would clearly benefit from further investigation, in order to confidently establish it as a valid and reliable short measure of workaholism in both scientific research and practice.

The present study resembles the earlier factor analytic studies of the DUWAS-10 by investigating and comparing the factorial validity of the scale in two countries: here, these are the Netherlands (*n*=9,010) and Finland (*n*=4,567). However, it goes beyond the earlier studies (1) by also investigating a longitudinal sub-sample (Finnish managers) with a two-year time lag; and (2) by examining distinct occupational groups within each country. To our knowledge, these two aspects have not yet been considered in factor analytic studies of the DUWAS-10. Thus, our datasets allow us to investigate the robustness of the factor structure of the DUWAS-10, not only across multisample occupational data from two countries (factorial group invariance) but also across time (factorial time invariance). Establishing both forms of factorial invariance is essential, as this would indicate that the DUWAS-10 can be used without having to worry about the possible structural instability of the scale, meaning, that the items do not function in conflicting ways but instead similarly reflect the latent factors of workaholism from one time point to another, and from one national context to another. Demonstrating that it has an invariant factor structure would be a fundamental indicator of the measurement validity of the scale.

The following hypotheses for our study were based on the theoretical definition of workaholism underpinning the DUWAS-10. Accordingly, workaholism consists of two dimensions of WE and WC that correlate but are independent constructs[5)]:

H1: Workaholism, as measured with the DUWAS-10, comprises two correlated factors (i.e., WE and WC) in both the Dutch and Finnish samples.

H2: The correlated two-factor structure of workaholism is replicated in occupational sub-samples from the Netherlands and Finland.

H3: The correlated two-factor structure of workaholism is invariant across the Netherlands and Finland.

H4: The correlated two-factor structure of workaholism is invariant across time in the sample of Finnish managers.

Subjects and Methods

Participants and procedure

The present study utilized two composite samples from the Netherlands (*n*=9,010) and Finland (*n*=4,567). In the Dutch composite sample, 50% were males and the mean sample age was 38 (range 16–77). This composite sample consisted of 15 subsamples that were investigated between 2006 and 2012. Typically, data were gathered within the larger framework of an employee satisfaction or engagement survey. Participation was voluntary and the anonymity of the respondents was guaranteed. The Dutch total sample consisted of the following occupational groups: medical residents (*n*=2,121, 61% female), managers (*n*=1,946, 71% male), white-collar workers (e.g., office clerks, *n*=1,326, 68% male), higher professionals (e.g., consultants, *n*=610, 59% male), executives (*n*=505, 62% male), teachers (*n*=445, 69% female), social workers (*n*=396, 80% female), paramedics (*n*=305, 65% female), blue-collar workers (e.g., production line workers, *n*=287, 54% male), lower professionals (e.g., technicians, *n*=257, 65% female), nurses (*n*=225, 68% female), physicians (*n*=166, 51% male), sales persons (*n*=163, 58% male), pink-collar workers (e.g., waitress, *n*=152, 74% female) and artists (*n*=106, 63% male).

In the Finnish composite sample, 60% were males and the mean age was 49 (range 24–79). This composite sample covered three occupationally homogenous sub-samples of dentists (n=2,785, 73% females) gathered in 2010 as part of the third phase of a longitudinal study conducted among Finnish dentists[15]; managers (n=898, 70% males) gathered in 2009[16, 17] and lawyers (n=702, 52% males) gathered in 2009 as part of a study among Finnish general law courts[18]. It also included one age-cohort representative sub-sample of Finnish 50-yr-old employees (n=182, 51% males) gathered in 2009 as part of the latest phase of an ongoing longitudinal study[19, 20]. All these data were utilized in the present study.

Finally, the dataset of the Finnish managers also included follow-up data on workaholism (n=459, 68% males and 51% of the Time 1 sample), with a two-year time lag from 2009 to 2011. The two-year follow-up period provided a good basis for study of the factorial invariance of the DUWAS-10 over time. If a scale meets the criteria of factorial time invariance over a long follow-up period, it is reasonable to conclude that the scale has good structural validity (i.e., the structure of the scale remains the same regardless of the long measurement time lag). Attrition analyses published earlier for the data of Finnish managers have shown that the follow-up respondents did not differ significantly from non-respondents in their total workaholism score at the baseline measurement in 2009 (see Mäkikangas *et al.*[21]). In addition, no significant differences were found between the follow-up respondents and non-respondents in background characteristics (gender, education, management level, working hours).

Measure

Workaholism was measured with the short 10-item DUWAS[5] in which five items measure working excessively (WE, e.g. “I seem to be in a hurry and racing against the clock” and “I find myself continuing to work after my coworkers have called it quits”) and five measure working compulsively (WC, e.g., “It’s important to me to work hard even when I don’t enjoy what I’m doing” and “I feel guilty when I take time off work”). The response scale for these statements ranges from 1 “(almost) never” to 4 “(almost) always”.

Data analyses

The data analyses proceeded in three stages. *In the first stage,* the correlated two-factor structure of the DUWAS-10 was investigated using confirmatory factor analysis (CFA). In the two-factor model, the five WE items were set to load on the latent factor of WE only, and the five WC items were set to load on the latent factor of WC only. Both the Dutch and Finnish composite samples were split into two equally sized random samples: an initial model development sample (i.e. 50% of the composite sample) and a cross-validating sample (i.e. the other 50% of the composite sample). If good factorial validity was observed for the initial model development samples and the factor structure of the DUWAS-10 was cross-validated using the cross-validating samples, CFAs were continued within each Dutch and Finnish sub-sample representing various occupational groups.

In the second stage, factorial group invariance was tested between the Dutch and Finnish composite samples. To do this, the Dutch and Finnish composite samples were combined into one data matrix and analyzed as a multisample data set. A freely estimated model (no constraints between the Dutch and Finnish sample) was compared to a constrained model in which the corresponding factor loadings were constrained to be equal across both samples.

In the third stage, the factorial time invariance of the DUWAS-10 was tested in the sub-sample of Finnish managers using two-wave longitudinal data with a two-year time lag. A freely estimated model (no invariance constraints between Time 1 and Time 2) was compared to a constrained model in which the corresponding factor loadings were set to be equal across Time 1 and Time 2.

The analyses were performed with the Mplus statistical package[22] using the missing data method (i.e., Mplus estimates models utilizing all the data that are available but without imputing data, for example, if a response is missing on one/some items of the scale) and robust maximum likelihood (MLR), a method of estimation that is robust to non-normality of the observed variables. The overall goodness-of-fit of the estimated models were evaluated using the following three absolute goodness-of fit indices: (1) χ^2 test, (2) Root Mean Square Error of Approximation (*RMSEA*) and (3) Standardized Root Mean Square Residual (*SRMR*). A non-significant χ^2 test indicates a good fit, as also do RMSEA and SRMR values of 0.05 or less, whereas values of 0.06–0.08 indicate a reasonable fit, and values $\geq$0.10 a poor fit[23, 24].

Because the χ^2 test is sensitive to sample size, the use of relative goodness-of fit indices is also strongly recommended in the case of large sample sizes[23]. Consequently, the following relative goodness-of-fit indices were also used to evaluate model fit: (1) Comparative Fit Index (*CFI*) and (2) Tucker-Lewis Index (*TLI*), with values $\geq$0.90 indicating a good fit. In addition to evaluating the overall goodness-of-fit, the factorial validity of each DUWAS

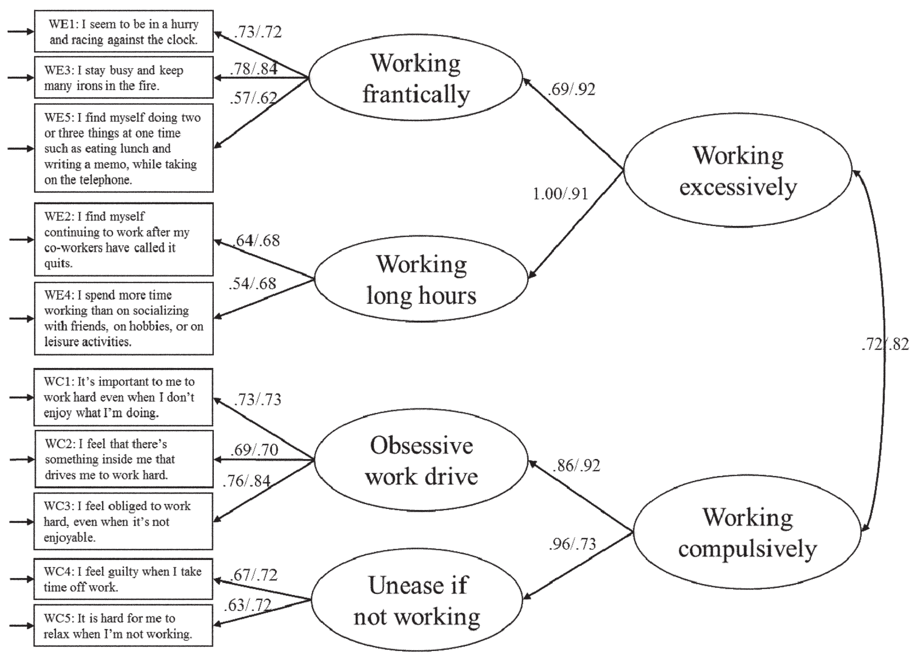

Fig. 1. The second-order factor model of the DUWAS-10: Standardized validity, that is, factor loadings for Dutch/Finnish composite samples. The items of working excessively (WE) and working compulsively (WC) are published by Schaufeli *et al* (2009).

item and item reliability was evaluated according to (1) standardized validity coefficients (i.e., standardized factor loadings), which indicate the direct structural relations between the true score and the item score[25] and (2) Cronbach's alpha coefficients for unweighted sum scores.

When comparing different, nested models against each other, we used the Satorra-Bentler scaled χ^2 difference test, which is recommended when using MLR as a method of estimation[26]. A significant χ^2 difference test denotes that the model with fewer degrees of freedom (i.e., less constraints) fits better with the data, and, vice versa, a non-significant χ^2 difference test denotes that the model with greater degrees of freedom (i.e., more constraints) fits better with the data.

Results

The factor structure of the DUWAS-10

The correlated two-factor structure of the DUWAS-10 fitted the data quite well in both the Dutch and Finnish initial model development samples: $\chi^2(34)=927.36$ $p<0.001$, RMSEA=0.08, SRMR=0.05, CFI=0.91, TLI=0.88; and $\chi^2(34)=526.62$ $p<0.001$, RMSEA=0.08, SRMR=0.04, CFI=0.93, TLI=0.90, respectively. However, inspection of the modifications indices (MI) indicated that in both datasets the same four MIs would improve the model fit. These MIs referred to freeing the error covariances between WE items 1 and 3, and items 2 and 4, and between WC items 1 and 3 and items 4 and 5.

On closer inspection, we noticed that the content of WE items 1 and 3 differs from that of WE items 2 and 4, although they all tap an excessive style of working (see item contents in Fig. 1). WE items 1 and 3 refer to being in a rush and doing things hectically; that is, a frantic style of working. The content of WE item 5 also is suggestive of this. Instead, WE items 2 and 4 refer to devoting an excessive amount of time to work. Hence it seemed that WE items 1, 3 and 5 and WE items 2 and 4 might represent meaningful sub-dimensions of working excessively which we dubbed "working frantically" and "working long hours", respectively.

Similarly, the content of WC items 1 and 3 differs from WC items 4 and 5, although they all tap a compulsive style of working (see item contents in Fig. 1). WC items 1 and 3 refer to working without joy and because of obligation and an inner "must". The content of WC item 2 also refers to this. In turn, WC items 4 and 5 refer to feelings of unease when detaching oneself from work. Hence it seemed that WC items 1, 2 and 3 and WC items 4 and 5 might represent meaningful sub-dimensions of working compulsively, which we dubbed "obsessive work drive" and "unease if not working", respectively.

When this second-order factor model of the DUWAS-10 (Fig. 1) was estimated for the initial model development samples, the fit with the data seemed better than the original correlated two-factor structure of the DUWAS-10, in both the Dutch and Finnish samples: $\chi^2(30)=424.31$ $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.96, TLI=0.94; and $\chi^2(30)=218.81$, $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.97, TLI=0.96, respectively. The Satorra-Bentler scaled χ^2 difference tests further confirmed that in both the Dutch and the Finnish data, the second-order factor model of the DUWAS-10 fitted the data significantly better than the original correlated two-factor structure: $\chi^2(4)=509.13$, $p<0.001$; and $\chi^2(4)=294.15$, $p<0.001$, respectively.

When these same steps were repeated for the Dutch and the Finnish cross-validating samples, identical results were obtained[1]. Hence, our first hypothesis (H1) was *not* confirmed as such; instead, it was qualified by a more complex second-order factor structure. This included − as hypothesized − the two correlated factors of WE and WC, but with each sub-divided further into the finer-grained factors of working frantically and working long hours (WE) and obsessive work drive and unease if not working (WC). Therefore, this second-order factor model was chosen to be tested in the occupational groups of the national samples.

As shown in Table 1, the model fit was (very) good in the Dutch and Finnish sub-samples, thus lending further support for the second-order factor model of the DUWAS-10. Again these results did not confirm our second hypothesis (H2) as such, but qualified it further by a more complex second-order factor structure.

Figure 1 shows standardized factor loadings for the Dutch and Finnish composite samples and the model fit for these composite samples were $\chi^2(30)=807.57$ $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.96, TLI=0.94; and $\chi^2(30)=431.94$, $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.97, TLI=0.96, respectively. Overall these factor loadings were good for all DUWAS-10 items. In the Dutch composite sample, eight out of ten, and in the Finnish sample, all the first-order factor loadings, were above 0.60, and most were above 0.70. Only WE items 4 and 5 had slightly lower loadings (0.54 and 0.57) in the Dutch sample. The second-order factor loadings and the correlation between WE and WC were also high, which indicates that the second-order factor model of the DUWAS-10 effectively taps the distinctive features that underlie the construct of workaholism. The Cronbach's alpha coefficients were 0.72/0.76 (Dutch/Finnish sample) for working frantically, 0.51/0.63 for working long hours, 0.77/0.80 for obsessive work drive, and 0.59/0.68 for unease if not working. The Cronbach's alpha coefficients for WE were 0.72/0.80, for WC 0.80/0.80 and for whole scale of workaholism 0.82/0.86.

To obtain the latent correlations for the four sub-dimensions of the DUWAS-10, a four-correlated factor model of workaholism was estimated for both the Dutch and Finnish composite samples. The model fit was $\chi^2(29)=788.69$ $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.96, TLI=0.94 in the Dutch and $\chi^2(29)=424.46$, $p<0.001$, RMSEA=0.06, SRMR=0.03, CFI=0.97, TLI=0.96 in the Finnish composite sample. The factor loadings, shown in Fig. 1, are equal, and the correlations between the latent factors, that is, sub-dimensions of the DUWAS-10, are given in Table 2.

The factorial group invariance of the DUWAS-10

Next, the factorial invariance of the second-order factor model of the DUWAS-10, as shown in Fig. 1, was tested using the Dutch and Finnish composite samples with a multi-sample procedure, by comparing a freely estimated model to a constrained model in which the corresponding factor loadings were set to be equal between both samples. The freely estimated as well as the constrained second-order factor models showed a good fit with the data: $\chi^2(60)=1232.39$, $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.97, TLI=0.95; and $\chi^2(68)=1455.72$, $p<0.001$, RMSEA=0.06, SRMR=0.04, CFI=0.96, TLI=0.95, for the Dutch and Finnish composite samples, respectively.

The Satorra-Bentler scaled χ^2 difference test showed

[1] The original correlated two-factor structure of the DUWAS-10 showed reasonably good fit to the data in both the Dutch and Finnish cross-validating samples: $\chi^2(34)=977.24$ $p<0.001$, RMSEA=0.08, SRMR=0.05, CFI=0.91, TLI=0.88; $\chi^2(34)=605.22$ $p<0.001$, RMSEA=0.09, SRMR=0.05, CFI=0.92, TLI=0.89, respectively. Inspection of the modification indices (MI) indicated that in both datasets freeing the error covariances between WE items 1 and 3 and items 2 and 4, and between WC items 1 and 3 and items 4 and 5 would improve the model fit. When the second-order factor model of the DUWAS-10 presented in Fig. 1 was estimated using the cross-validating samples, the fit to the data seemed better than the original correlated two-factor structure of the DUWAS-10, in both the Dutch and Finnish samples: $\chi^2(30)=415.31$ $p<0.001$, RMSEA=0.05, SRMR=0.03, CFI=0.96, TLI=0.95; and $\chi^2(30)=249.12$, $p<0.001$, RMSEA=0.06, SRMR=0.03, CFI=0.97, TLI=0.95, respectively. The Satorra-Bentler scaled χ^2 difference tests further confirmed that both in the Dutch and the Finnish data, the second-order factor model of the DUWAS-10 fitted the data significantly better than the original correlated two-factor structure: $\chi^2(4)=555.13$, $p<0.001$; and $\chi^2(4)=336.41$, $p<0.001$, respectively.

Table 1. Factorial invariance of the DUWAS-10 across the Dutch and Finnish sub-samples: Goodness-of-fit criteria for the second-order factor model

Dutch Sub-samples	n	χ^2	df	p	RMSEA	SRMR	CFI	TLI
Medical residents	2,121	184.02	30	0.000	0.05	0.03	0.96	0.95
Managers	1,946	329.48	30	0.000	0.07	0.05	0.93	0.89
White-collar workers	1,326	152.22	30	0.000	0.06	0.04	0.96	0.94
Higher professionals	610	108.21	30	0.000	0.07	0.05	0.95	0.92
Executives	505	70.30	30	0.000	0.05	0.04	0.97	0.95
Teachers	445	49.92	30	0.013	0.04	0.03	0.98	0.97
Social workers	396	88.20	30	0.000	0.07	0.05	0.94	0.92
Paramedic	305	55.17	30	0.003	0.05	0.04	0.96	0.94
Blue-collar workers	287	75.71	30	0.000	0.07	0.04	0.94	0.91
Lower professionals	257	36.00	30	0.177	0.03	0.04	0.99	0.98
Nurses	225	45.40	30	0.035	0.05	0.04	0.97	0.95
Physicians	166	65.22	30	0.000	0.08	0.05	0.93	0.90
Sales persons	163	33.51	30	0.301	0.03	0.04	0.99	0.99
Pink-collar workers	152	42.83	30	0.061	0.05	0.04	0.98	0.96
Artists	106	32.55	30	0.342	0.03	0.04	0.99	0.99
Finnish Sub-samples								
Dentists	2,785	264.63	30	0.000	0.05	0.03	0.97	0.96
Managers	898	150.75	30	0.000	0.07	0.04	0.96	0.93
Lawyers	702	83.39	30	0.000	0.05	0.03	0.98	0.97
50-year-old employees	182	53.52	30	0.005	0.07	0.05	0.93	0.90

RMSEA: Root Mean Square Error of Approximation, SRMR: Standardized Root Mean Square Residual, CFI: Comparative Fit Index, TLI: Tucker-Lewis Index

Table 2. Latent correlations between the four sub-dimensions of DUWAS-10 for the Dutch (n=9,010, below diagonal) and Finnish (n=4,567, above diagonal) composite samples

Latent factor	1.	2.	3.	4.
1. Working frantically	-	0.84***	0.70***	0.53***
2. Working long hours	0.70***	-	0.68***	0.56***
3. Obsessive work drive	0.45***	0.61***	-	0.67***
4. Unease if not working	0.45***	0.71***	0.83***	-

*** p<0.001

that constraining the factor loadings to be equal between the Dutch and Finnish samples produced a significant loss in model fit: $\chi^2(8)$=228.96, p<0.001. However, MacCallum *et al.*[27] have argued that this is to be expected in very large samples (here n=13,577). They argue that with large samples it is better to test a null hypothesis of *small* (H_0: F_A*– F_B* ≤ δ*) rather zero differences between the nested models based on the comparison of RMSEA values. This is done based on a noncentral χ^2 distribution and a noncentrality parameter.

Assuming that the maximum difference between RMSEA fit index is 0.01, the critical value T for the χ^2 difference at a significance level of 0.05 is 1,414.89 in our case[2]. The observed χ^2 difference value of 228.96 is well below this critical value of 1,414.89, and the p-value for the null hypothesis of a small difference between the freely estimated and constrained second-order factor models

[2] This critical value T was obtained following the statistical theory and method for testing for a small difference in fit presented by MacCallum *et al.*[24] (pp. 27–28). *First*, we estimated the unconstrained [$\chi^2(60)$=1232.39, p<0.001, RMSEA=0.05, SRMR=0.03, CFI=0.97, TLI=0.95] and constrained [(68)=1455.72, p<0.001, RMSEA=0.06, SRMR= 0.04, CFI=0.96, TLI=0.95] second-order factor models and calculated the χ^2 difference; $\chi^2(8)$=228.96. *Second*, we calculated the delta value δ* with formula $\delta^*=df_{\text{constrained}} \times \varepsilon^2_{max} - df_{\text{unconstrained}} \times \varepsilon^2_{min}$, where ε_{max} and ε_{min} are theoretical RMSEA values (see MacCallum *et al.*[24]); $\delta^*=68 \times 0.06^2 - 60 \times 0.05^2 = 0.0948$. *Third*, we calculated the noncentrality parameter λ with the formula $\lambda=(N-1) \times \delta^*$; λ=(13,577–1) × 0.0948=1,287. *Fourth*, using SPSS code (cf. SAS code, designated Program C, provided on the Web at http://dx.doi.org/10.1037/1082-989X.11.1.19.supp; MacCallum *et al.*[24], p. 28), we calculated the critical value T for the χ^2 difference at the significance level of 0.05 based on a noncentral χ^2 distribution using the difference in degrees of freedom between the constrained and unconstrained model and the noncentrality parameter. This critical T value was 1,414.89. *Fifth*, we compared whether the χ^2 difference between the constrained and unconstrained model (228.96) was smaller than the critical value T (1,414.89). It was, and therefore the constrained model was accepted.

was 1.00, thus confirming the null hypothesis. This means that the measurement properties of the second-order factor model of the DUWAS-10 proved to be similar (i.e., a small difference based on the noncentral χ^2 distribution) but not fully identical (i.e., zero difference based on the Satorra-Bentler scaled χ^2 difference test) between the Dutch and Finnish composite samples. In sum, the results did *not* confirm our third hypothesis (H3) as such, but qualified it by showing that the more complex second-order factor structure of the DUWAS-10 showed reasonable measurement invariance and stability of factor structure across the national samples.

The factorial time invariance of the DUWAS-10

To further investigate the factorial validity of the DUWAS-10, we tested whether the measurement properties of the second-order factor model of the DUWAS-10 remained invariant across time in the Finnish sub-sample of managers (n=459) with two measurement points two years apart. The model fit was good for both the freely estimated and the time-constrained model: $\chi^2(148)=314.37$ $p<0.001$, RMSEA=0.05, SRMR=0.05, CFI=0.95, TLI=0.94; and $\chi^2(156)=323.71$ $p<0.001$, RMSEA=0.05, SRMR=0.05, CFI=0.95, TLI=0.94, respectively. In addition, the Satorra-Bentler scaled χ^2 difference test showed that the factor loadings could be constrained to be equal between Time 1 and Time 2 without significant loss of model fit: $\chi^2(8)=8.42$, $p=0.393$. This means that also our fourth hypothesis (H4) was *not* confirmed as such, but was qualified by showing that the more complex second-order factor structure of the DUWAS-10 showed reasonable measurement invariance and stability of factor structure across time.

In the second-order factor model, the stability coefficient indicating rank-order stability across the two-year time lag was 0.66 (R^2=0.44) for WE and 0.76 (R^2=0.57) for WC. We also estimated a longitudinal correlated four-factor model [model fit: $\chi^2(150)=269.56$, $p<0.001$, RMSEA=0.04, SRMR=0.05, CFI=0.97, TLI=0.96] in order to obtain the stability coefficients for working frantically (β=0.60, R^2=0.36), working long hours (β=0.65, R^2=0.42), obsessive work drive (β=0.60, R^2=0.35), and unease if not working (β=0.71, R^2=0.50). Lastly, we estimated a longitudinal third-order factor model of workaholism, consisting of the second-order factors of WE and WC, which in turn consisted of the first-order factors of working frantically, working long hours, obsessive work drive, and unease if not working, in order to obtain the stability coefficient for total workaholism. The stability coefficient for total workaholism was 0.78 (R^2=0.61) and the model fit for this longitudinal third-order factor model was: $\chi^2(156)=359.69$, $p<0.001$, RMSEA=0.05, SRMR=0.05, CFI=0.94, TLI=0.93.

Discussion

The present study investigated the factorial validity of the DUWAS-10, a short measure for workaholism[5]. As expected, it was found that the ten items of the DUWAS tapped both the sub-dimensions of working excessively (WE) and working compulsively (WC), which correlated positively with each other. However, a more fine-grained structure was uncovered in our large-scale dataset of Dutch and Finnish employees. Namely, WE (the behavioral component of workaholism) split further into the sub-dimensions of working frantically and working long hours, and WC (the cognitive component of workaholism) split further into the sub-dimensions of obsessive work drive and unease if not working. Hence, the DUWAS with only ten items appears to be a somewhat more complex, multi-dimensional measure of workaholism than initially expected.

Our results strongly support this second-order factor model of DUWAS-10, where WE and WC correlate positively but each divides into two sub-dimensions. This does not contradict the expected correlated two-factor structure, but instead amplifies it. Not only was the second-order factor model of the DUWAS-10 corroborated by the results for the composite samples from the Netherlands and Finland, but it also fitted the data of the diverse occupational sub-samples in both countries. Furthermore, this more complex factor structure of workaholism was invariant across the Dutch and Finnish samples, and also across a two-year period among Finnish managers.

Our longitudinal study also revealed an interesting secondary finding. The longitudinal confirmatory factor analysis indicated that not only was the structure of the scale stable over time, but the relative order of individuals in their workaholism scores was also rather stable. The rank-order stabilities (stability coefficients) for the sub-dimensions of the DUWAS-10 ranged between 0.60 and 0.76. On the one hand, this result indicates that workaholism seems to represent a relatively stable working style of an individual rather than a fluctuating experience or action pattern. On the other hand, the range in the explanation rates, from 0.36 to 0.57, indicates that change is possible: the first measurement of workaholism did not explain all the variance observed between our participants in the second measurement.

It should be also acknowledged that the Cronbach's alphas for the sub-dimensions of working long hours and unease if not working in the Dutch data were somewhat low; however, this was not surprising, as each of these sub-dimensions consisted of only two items. It is not uncommon for scales with a smaller number of items to have lower Cronbach's alphas than scales with a larger number of items[28]. In addition, these two dimensions had acceptable Cronbach's alphas in the Finnish data, and more importantly, the factor loadings for the items of these dimensions were appropriate in both the Dutch and Finnish data. Standardized factor loadings can also be interpreted as validity coefficients that indicate the direct structural relations between the true score and the item score[25].

From a theoretical point of view, our results emphasize that workaholism − as assessed with the DUWAS-10 − is a multifaceted phenomenon and that the underlying structure would appear to be more complex than earlier presented. WE is more than just devoting excessive time to work, that is, working long hours in comparison to others and spending less off-job time with family, friends, and/or hobbies. The sub-dimension of working frantically taps the behavioral tendency of workaholics to take on too many work tasks[29]. Interestingly, and in line with the present results, Molino *et al.*[30] also found that WE items #2 and #4 share variance (Fig. 1) over and above the latent WE factor of the DUWAS-10, and in the study by Andreassen *et al.*[12] these two items loaded more weakly than items #1, #3, and #5 on the latent factor of WE.

Analogously, WC is more than just working hard out of obligation and without joy, that is, motivated by a compulsive inner drive. The sub-dimension of feeling unease if not working taps the difficulty of relaxing and disengaging oneself from work when it is time to do so. In support of this, del Líbano *et al.*[8] and Littman-Ovidia *et al.*[11] also found that WC items #1 and #3 as well as #4 and #5 share some additional variance over and above the latent WC factor (Fig. 1). All in all, the present study suggests that the DUWAS-10 not only taps the latent constructs of WE and WC, but also the more specific dimensions that constitute WE and WC.

Strengths, limitations, and avenues for future research

The strengths of the present study are the use of large, cross-national samples of employees with heterogeneous occupational backgrounds. The study contributes in two important ways to the literature. The first is the finding of factorial invariance in these multi-sample datasets, meaning that the structure of the scale did not vary across either the countries or occupational groups investigated. Second, our longitudinal study enabled us to confirm the structural stability of the DUWAS-10 empirically over time. The factorial time invariance that was observed across a relatively long time span (2 yr) indicates that the DUWAS-10 can be used in long-term follow-up studies as well as in intervention studies or organizational development projects without any fears that possible mean level changes in the workaholism dimensions would be biased by possible structural instability of the scale over time. For example, Mäkikangas *et al*.[21] have observed that workaholism is sensitive to changes in the individual's work situation such as quitting the current job for a new one.

However, despite these strengths, our study also has its limitations. Although our participants had heterogeneous occupational backgrounds, in quantity they were more representative of white-collar than blue- and pink-collar workers, which may limit the generalization of the results. Some of our sub-samples (e.g., artists) were rather small in relation to the statistical power required, although most were adequate. Also, our longitudinal data included only one professional group, namely Finnish managers. Hence, future research should include more balanced samples in terms of occupational status, while longitudinal data sets from different cultures with occupationally heterogeneous sub-samples are also needed. In addition, the present study was based on self-reports. Future research would benefit from using co-workers and spouses/partners as observers of the focal person's workaholism to reduce possible self-report bias. This kind of scale development of DUWAS is already in progress[31, 32].

Finally, the criterion validity (e.g., associations with antecedents and/or consequences of workaholism) of the DUWAS-10 was beyond the scope of the present study; this is an aspect that should be thoroughly investigated in future studies. Evidence already exists that WE is positively and more strongly associated not only with overwork, workload and high weekly working hours but also with autonomous motivation, intrinsic job satisfaction and work engagement than WC, which in turn, is positively and more strongly associated with controlled motivation and job exhaustion than WE[5, 6, 11, 33]. Hence, and most importantly, to obtain further confirmation and theoretical understanding of the four sub-dimensions of workaholism detected here for the first time, they should first be replicated in another study and their criterion validity verified. For example, although it is plausible that all the sub-dimensions of workaholism are related to the following phenomena to some extent, it might be that working frantically is most strongly related to poor time- and self-management skills at work, working long hours to experience of work demands interfering with family and home life[34], obsessive work drive to high overcommitment at work[35] and unease if not working to low psychological detachment as a specific mechanism of recovery from work[36].

Conclusions

Our results show that the DUWAS-10 is a brief measure with good factorial validity across occupations, nations, and measurements over time. Due to the fact that the subdimensions of the DUWAS-10 correlate highly, it is possible to use the DUWAS-10 as one-dimensional measure of workaholism. However, at the same time our results indicated that the DUWAS-10 encompasses a wide range of different features of workaholism despite comprising only 10 items. Therefore, it is also recommended that WE and WC be assessed separately, as the items measuring these on the DUWAS-10 load on different factors, and thus they refer to different aspects of workaholism. For example, by crossing high/low on WE and WC, a fourfold table emerges representing four groups of workers: excessive workers (high on WE), compulsive workers (high on WC), workaholics (high on WE and WC) and non-workaholics (low on WE and WC). Meaningful differences between these groups of workers have been found in work characteristics, personality and subjective well-being[37].

Hence, we conclude that in practise, for example in occupational health care or human resource management it makes sense to discriminate at least between WE and WC dimensions of workaholism because these dimensions are stable, reliable, and invariant. Understanding of the underlying dimensionality of workaholism may also help occupational health practitioners in identifying workaholics and distinguishing them from, for example, those who are burned out because of heavy extrinsic job demands. We further conclude that in scientific research the DUWAS-10 could also be used as a four-dimensional scale because we uncovered a more fine-grained structure of the DUWAS-10 consisting of two first-order factors loading on WE ("working frantically" and "working long hours") and two on WC ("obsessive work drive" and "unease if not working"). If these four dimensions are replicated and confirmed in future studies, an interesting question might be for example to investigate whether the mean levels of these dimensions increase in individuals at the same time or whether there are some sequential steps between them that might describe the development of workaholism from mild to severe level. This kind of knowledge could be applied in building early screening tools to prevent further development of workaholism.

Acknowledgements

This research was supported by grants from the Academy of Finland (118316, 127125, 135347, 138369) and the Finnish Work Environment Fund (108124).

References

1) Spence JT, Robbins AS (1992) Workaholism: definition, measurement, and preliminary results. J Pers Assess **58**, 160–78.
2) Robinson BE (1999) The Work Addiction Risk Test: development of a tentative measure of workaholism. Percept Mot Skills **88**, 199–210.
3) Schaufeli WB, Taris TW (2004) ©The Dutch Work Addiction Scale (DUWAS) www.wilmarschaufeli.nl. Accessed June 10, 2014.
4) Schaufeli WB, Taris TW, Bakker AB (2006) Dr. Jekyll and Mr. Hyde: On the differences between work engagement and workaholism. In: Research companion to working time and work addiction, Burke RJ (Ed.), 193–217, Edward Elgar, Cheltenham.
5) Schaufeli WB, Shimazu A, Taris TW (2009) Being driven to work excessively hard: the evaluation of a two-factor measure of workaholism in the Netherlands and Japan. Cross-cult Res **43**, 320–48.
6) Schaufeli WB, Bakker AB, van der Heijden FMMA, Prins JT (2009) Workaholism among medical residents: it is the combination of working excessively and compulsively that counts. Int J Stress Manag **16**, 249–72.
7) Shimazu A, Schaufeli WB, Taris TW (2010) How does workaholism affect worker health and performance? The mediating role of coping. Int J Behav Med **17**, 154–60.
8) Del Líbano M, Llorens S, Salanova M, Schaufeli W (2010) Validity of a brief workaholism scale. Psicothema **22**, 143–50.
9) McMillan LHW, O'Driscoll MP (2006) Exploring new frontiers to generate an integrated definition of workaholism. In: Research companion to working time and work addiction, Burke R (Ed.), 89–107, Edward Elgar, Cheltenham.
10) Ng TWH, Sorensen KL, Feldman DC (2007) Dimensions, antecedents, and consequences of workaholism: a conceptual integration and extension. J Organ Behav **28**, 111–36.
11) Littman-Ovadia H, Balducci C, Ben-Moshe T (2014) Psychometric properties of the Hebrew version of the Dutch Work Addiction Scale (DUWAS-10). J Psychol **148**,

327–46.
12) Andreassen CS, Hetland J, Pallesen S (2014) Psychometric assessment of workaholism measures. J Manag Psychol **29**, 7–24.
13) Little TD (1997) Mean and covariance structures (MACS) analyses of cross-cultural data: practical and theoretical issues. Multivariate Behav Res **32**, 53–76.
14) Meredith W, Teresi JA (2006) An essay on measurement and factorial invariance. Med Care **44** Suppl 3, S69–77.
15) Hakanen JJ, Perhoniemi R (2011) Hammaslääkäreiden työhyvinvointi Suomessa—seitsemän vuoden seurantatutkimus [Work-related well-being among Finnish dentists—A seven-year follow-up study]. Hammaslaakarilehti **11**, 1–15.
16) Huhtala M (2013) Virtues that work: Ethical organisational culture as a context for occupational well-being and personal work goals. Jyväskylä Studies in Education, Psychology and Social Research 479, University of Jyväskylä, Finland. https://jyx.jyu.fi/dspace/handle/123456789/42178. Accessed June 10, 2014.
17) Huhtala M, Feldt T, Lämsä AM, Mauno S, Kinnunen U (2011) Does the ethical culture of organisations promote managers' occupational well-being? Investigating indirect links via ethical strain. J Bus Ethics **101**, 231–47.
18) Hakanen JJ, Perhoniemi R, Rodríguez-Sánchez AM (2012) Too good to be true? Similarities and differences between engagement and workaholism among Finnish judges. Spec issue on work engagem. Cienc Trabajo **14**, 72–80.
19) Pulkkinen L (2006) The Jyväskylä Longitudinal Study of Personality and Social Development (JYLS). In: Socioemotional development and health from adolescence to adulthood, Pulkkinen L, Kaprio J and Rose RJ (Eds.), 29–55, Cambridge University Press, New York.
20) Pulkkinen L, Kokko K (2012) Foundational issues in longitudinal data collection. In: Handbook of developmental research methods, Laursen B, Little TD and Card NA (Eds.), 129–147, Guilford Press, New York.
21) Mäkikangas A, Schaufeli W, Tolvanen A, Feldt T (2013) Engaged managers are not workaholics: evidence from a longitudinal person-centered analysis. Eur J Work Organ Psychol **29**, 135–43.
22) Muthén LK, Muthén BO (1998–2009) Mplus User's guide, 5th Ed., Muthén & Muthén, Los Angeles.
23) Hu L, Bentler PM (1999) Cut-off criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling **6**, 1–55.
24) Kline RB (2005) Principles and practice of structural equation modelling, 2nd Ed., Guilford Press, New York.
25) Bollen KA (1989) Structural equations with latent variables, John Wiley & Sons, New York.
26) Satorra A, Bentler PM (2001) A scaled difference chi-square test statistic for moment structure analysis. Psychometrika **66**, 507–14.
27) MacCallum RC, Browne MW, Cai L (2006) Testing differences between nested covariance structure models: power analysis and null hypotheses. Psychol Methods **11**, 19–35.
28) Steiner DL, Norman GR (2008) Health measurement scales: A practical guide to their development and use, 4th Ed, Oxford University Press, Oxford.
29) Machlowitz M (1980) Workaholics: Living with them, working with them, Simon & Schuster, New York.
30) Molino M, Ghislieri C, Colombo L (2012) Working excessively: theoretical and methodological considerations. G Ital Med Lav Ergon **34**, Suppl Psicologia A5–A10.
31) Falco A, Kravina L, Girardi D, Dal Corso L, Di Sipio A, De Carlo NA (2012) The convergence between self and observer ratings of workaholism: a comparison between couples. TPM Test Psychom Methodol Appl Psychol **19**, 311–24.
32) Mazetti G, Schaufeli WB, Guglielmi D (submitted) Are workaholism and work engagement in the eye of the beholder? A multirater perspective on opposite forms of working hard.
33) Broeck AV, Schreurs B, Witte HD, Vansteenkiste M, Germeys F, Schaufeli W (2011) Understanding workaholics motivations: a self-determination perspective. Appl Psychol **60**, 600–21.
34) Bellavia GM, Frone MR (2005) Work-family conflict. In: Handbook of work stress, Barling J, Kelloway EK and Frone MR (Eds.), 113–147, Sage, California.
35) Siegrist J, Starke D, Chandola T, Godin I, Marmot M, Niedhammer I, Peter R (2004) The measurement of effort-reward imbalance at work: European comparisons. Soc Sci Med **58**, 1483–99.
36) Sonnentag S (2012) Psychological detachment from work during leisure time: the benefits of mentally disengaging from work. Curr Dir Psychol **21**, 114–8.
37) Schaufeli WB, Van Wijhe C, Peeters M, Taris TW (2011) Werkverslaving, een begrip gemeten [Workaholism: the measurement of a concept]. Gedrag Organ **24**, 43–63.