European Journal of Psychological Assessment

A Cross-National study on the Psychometric Quality of the Italian Version of the Dutch Work Addiction Scale (DUWAS)

--Manuscript Draft--

Manuscript Number:	EJPA-D-14-00029R2
Full Title:	A Cross-National study on the Psychometric Quality of the Italian Version of the Dutch Work Addiction Scale (DUWAS)
Short Title:	Validation of the Italian version of the DUWAS
Article Type:	Original Article
Keywords:	DUWAS; workaholism; work addiction; validation study; measurement equivalence
Corresponding Author:	Cristian Balducci, PhD University of Bologna Bologna, ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Bologna
Corresponding Author's Secondary Institution:	
First Author:	Cristian Balducci, PhD
First Author Secondary Information:	
Order of Authors:	Cristian Balducci, PhD
	Lorenzo Avanzi
	Chiara Consiglio
	Franco Fraccaroli
	Wilmar Schaufeli
Order of Authors Secondary Information:	
Abstract:	We investigated the psychometric properties of the Italian version of the Dutch Work Addiction Scale (DUWAS) by using an Italian sample (N = 1027) and a comparable Dutch sample (N = 7523). We first conducted multi-group confirmatory factor analysis and reliability analysis. We then examined the correlations of the DUWAS with job demands, work-family conflict, job-related affective well-being and psychological and physical strain. Results showed that a two-factor solution of the DUWAS was equivalent across the Italian and Dutch samples, with data supporting factor variance and covariance equivalence in addition to metric equivalence. Internal consistencies of the DUWAS and its subscales were adequate, and test-retest reliability showed a strong stability of the measure at one year. The DUWAS and its subscales showed a pattern of correlations in the expected direction with the adopted criterion measures. Overall, the Italian version of the DUWAS shows adequate validity and reliability.
Additional Information:	
Question	**Response**
Does the length of your paper (including abstract, references, tables, and figures) fall within the word limit?	Yes
Has the manuscript or any component of it already been published or is currently under consideration by another journal?	No
Has this manuscript or another version of it been submitted to this journal	No

previously?	
If the paper has been written by more than one person, can the corresponding author attest that each author has studied the manuscript in the form submitted, agreed to be cited as a coauthor, and has accepted the order of authorship?	Yes
Is there any conflict of interests concerning the content of this manuscript?	No
Have you ensured that all references to author names and affiliations have been removed from the manuscript and all related documents to ensure an anonymous review process?	Yes
Has the manuscript been prepared according to the Publication Manual of the American Psychological Association (6th ed.)?	Yes
Does the length of your paper (including abstract, references, tables, and figures) fall within the word limit?	Yes
State the word count. as follow-up to "Does the length of your paper (including abstract, references, tables, and figures) fall within the word limit?"	5000
Does your submission contain electronic supplementary material (ESM)?	No
Response to Reviewers:	Dear Professor Ziegler, We greatly appreciate your invitation to revise and resubmit the paper, and have closely considered the reviewer' suggestions and critique. Having incorporated many of the suggestions to take the criticism expressed into account we believe that our manuscript is now further improved. Below we provide detailed information on how we have dealt with each of the suggestions made by the reviewer. We hope that the paper now meets the requirements for publication in EJPA. Sincerely, AUTHORS Reviewer #1: Review EJPA-D-14-0029R1 COMMENT: Thank you for your thorough revision. Most of my concerns have been dealt with adequately. There are only a few minor issues remaining that I am sure can be dealt with easily by the authors. RESPONSE: Thanks a lot for these positive comments. Follow we tried to answer at each questions. COMMENT: 1) Table 1: I appreciate that the authors added factor loadings to Table 1. However, I think all factor loadings should be shown, including cross-loadings (e.g., loadings of WE items on WC). Considering the high correlation between the two subscales it would be important to know how well the items can be attributed to a particular factor. RESPONSE: We included only the factor loadings of the hypothesized model as derived through CFA (as – we believe – can be understood by reading the footnote reported in Table 1). So we do not have available cross-loadings for this model. To add the cross-loadings, we should test a different model by freeing additional patterns. However, if cross-loadings were a problem of the tested two-factor solution, we believe that the solution wouldn't have fit the data well and better than a one-factor solution.

To reassure you that the cross-loadings are not of great concern for the DUWAS, we add in the following table the cross-loadings of an EFA (PCA) carried out on the DUWAS - Italian sample only. As you can see, the highest cross-loading was .33 (ranging from .07 to .33). Thus, also from an exploratory point of view, we think that items are clearly better attributed to the hypothesized two-factor solution than to a one-factor solution. We do not report this table in the text, however, because we took a confirmatory factor analysis approach in our study and we believe it is not methodologically sound to present exploratory and confirmatory results by using the same sample. We hope you can agree with this.

SEE ADDITIONAL MATERIAL TO CHECK THE TABLE WITH FACTOR LOADINGS

COMMENT: 2) P.4 line 58: "The study sample was made by women in 61% of the cases…". This sentence sounds strange. You probably mean "made up of women". Since information on the % of women is already included in the description of each of the samples, I don't think it is necessary to provide this number.

RESPONSE: Thank YOU. Now, we have deleted the sentence, and we added in parenthesis in the previous sentence the female percentage, as follow:
“Overall, it was available a sample of 1027 participants (61% women).”

COMMENT: 3) Discussion: The discussion should address the limitation of only having metric invariance as opposed to scalar invariance. I agree with the authors that if 5 out of 10 intercepts have to be freed in order to achieve partial scalar invariance, then the resulting model is not very useful for further analyses. However, it should be noted that this limits the value of showing the invariance of factor variances and covariances. Furthermore, it limits the informativeness of group comparisons since without scalar invariance, we cannot be sure that differences in responses to the DUWAS items between the Italian and Dutch samples reflect real differences in the underlying traits. Instead, the differences in item responses could just be due to differential item functioning.

REPONSE: We are not in complete agreement with you when you say: “However, it should be noted that this limits the value of showing the invariance of factor variances and covariances”. Factor variance and covariance invariance can be tested if metric invariance hold (Barbaranelli & Ingoglia, 2013). Furthermore, if metric and factor variance and covariance invariance hold, it is possible to compare correlations and betas across groups (Barbaranelli & Ingoglia, 2013). In this sense, some comparisons across groups are possible and meaningful for research even if scalar invariance does not hold.
However, we are aware of the limitations due to the lack of scalar invariance, and thus following your suggestion (for which we thank you), we have now added some considerations on this. We specifically wrote:
“Loosely speaking, the metric of the DUWAS did not change across the Italian and Dutch samples, which means that comparisons between the DUWAS and other variables at latent factors level (e.g. correlations between factors) are meaningful. To the contrary, the DUWAS observed scores in the two countries cannot be compared because scalar invariance was not supported, meaning that it is not sure that differences in responses to the DUWAS items between the Italian and Dutch samples reflect real differences in the underlying traits. This result could be due to a number of reasons (Little, 1997), including the fact that we used an Italian convenience sample which was entirely comparable to the Dutch sample, which could have particularly affected the more stringent invariance tests.”

Barbaranelli, C. (2013). The multi-group structural equation modeling and the examination of factorial invariance [I modelli di equazioni strutturali molti-gruppo e l’esame dell’invarianza fattoriale]. In C. Barbaranelli, & Ingoglia, S. (Ed.), Structural Equation Modeling. Themes and perspectives [I modelli di equazioni strutturali. Temi e prospettive] (pp. 153–200). Milano: Led Edizioni.

COMMENT: 4) P.10 line 46: "alpha values met the criterion of .70". Either a reference for the criterion should be provided or this part of the sentence should be deleted. Furthermore, considering that the assumptions of alpha are probably violated in this data, it would be better to report a different coefficient such as omega (e.g.,

	Dunn, Baguley, & Brunsden, 2014). RESPONSE: Thank you for this suggestion. We initially have provided a reference for alpha, i.e. Nunnaly & Bernstein (1994), but then – due to the fact that the manuscript was already longer than permitted (i.e. 5000 words) – we have decided to eliminate the entire sentence. Additionally, in the results we have now added the bootstrapped confidence interval for Cronbach's alpha (Iacobucci & Duhachek, 2003). As you can see below, the 95% confidence interval for both dimensions was fairly narrow (.72-.77), and in both cases the lower interval was higher than .70: WE: .74 CI: .717-.767 WC: .74 CI: .719-.769 We have also computed the omega index (Dunn e al., 2014) for both the WE and WC scales. Again, we found acceptable values, just larger than alpha (i.e. WE: .75; WC: .75) We have reported these results in a paragraph included in the results section and reported below: "For the WE and WC subscales we also computed the bootstrap 95% confidence interval around α (see Iacobucci & Duhachek, 2003), which was found to be the same in both cases (i.e. CI=.72-.77). Since Cronbach's alpha relies on assumptions that could be unrealistic in field studies (i.e., true scores for different items are assumed to have the same variance), some have proposed to use alternative indexes to estimate scale reliability, such as the omega index (see Dunn, Baguley, & Brunsden 2014). Such index was found to be acceptable for the WE and WC subscales in the present study (ω = .75 in both cases)." Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British Journal of Psychology, 105(3), 399-412. Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. Journal of Consumer Psychology, 13, 478–487.

"A Cross-National study on the Psychometric Quality of the Italian Version of the Dutch Work Addiction Scale (DUWAS)" EJPA-D-14-00029R1;

European Journal of Psychological Assessment

Dear Professor Ziegler,

We greatly appreciate your invitation to revise and resubmit the paper, and have closely considered the reviewer' suggestions and critique. Having incorporated many of the suggestions to take the criticism expressed into account we believe that our manuscript is now substantially improved.

Below we provide detailed information on how we have dealt with each of the suggestions made by the reviewer. In italics the modifications that we have made in response to each suggestion.

We hope that the paper now meets the requirements for publication in EJPA.

Sincerely,

Reviewer #1: Review EJPA-D-14-0029R1

Thank you for your thorough revision. Most of my concerns have been dealt with adequately. There are only a few minor issues remaining that I am sure can be dealt with easily by the authors.

Thanks a lot for these positive comments.

1) Table 1: I appreciate that the authors added factor loadings to Table 1. However, I think all factor loadings should be shown, including cross-loadings (e.g., loadings of WE items on WC). Considering the high correlation between the two subscales it would be important to know how well the items can be attributed to a particular factor.

We included only the factor loadings of the hypothesized model as derived through CFA (as – we believe – can be understood by reading the footnote reported in Table 1). So we do not have available cross-loadings for this model. To add the cross-loadings, we should test a different model by freeing additional patterns. However, if cross-loadings were a problem of the tested two-factor solution, we believe that the solution wouldn't have fit the data well and better than a one-factor solution.

To reassure you that the cross-loadings are not of great concern for the DUWAS, we add in the following table the cross-loadings of an EFA (PCA) carried out on the DUWAS - Italian sample only. As you can see, the highest cross-loading was .33 (ranging from .07 to .33). Thus, also from an exploratory point of view, we think that items are clearly better attributed to the hypothesized two-factor solution than to a one-factor solution. We do not report this table in the text, however, because we took a confirmatory factor analysis approach in our study and we believe it is not methodologically sound to present exploratory and confirmatory results by using the same sample. We hope you can agree with this.

	Factor loading on WE	*Factor loading on WC*
1. I seem to be in a hurry and racing against the clock (WE1)	***.72***	*.13*
2. I find myself continuing to work after my co-workers have called it quits (WE2)	***.75***	*.11*
3. I stay busy and keep many irons in the fire (WE3)	***.64***	*.24*
4. I spend more time working than on socializing with friends, on hobbies, or on leisure activities (WE4)	***.55***	*.33*
5. I find myself doing two or three things at one time such as eating lunch and writing a memo, while taking on the	***.65***	*.20*
6. It's important to me to work hard even when I don't enjoy what I'm doing (WC1)	*.21*	***.66***
7. I feel that there's something inside me that drives me to work hard (WC2)	*.33*	***.65***
8. I feel obliged to work hard, even when it's not enjoyable (WC3)	*.31*	***.67***
9. I feel guilty when I take time off work (WC4)	*.07*	***.74***
10. It is hard for me to relax when I'm not working (WC5)	*.12*	***.60***

2) P.4 line 58: "The study sample was made by women in 61% of the cases…". This sentence sounds strange. You probably mean "made up of women". Since information on the % of women is already included in the description of each of the samples, I don't think it is necessary to provide this number.

Thank you. Now, we have deleted the sentence, and we have added in parenthesis in the previous sentence the female percentage, as follow:

"Overall, it was available a sample of 1027 participants (61% women)."

3) Discussion: The discussion should address the limitation of only having metric invariance as opposed to scalar invariance. I agree with the authors that if 5 out of 10 intercepts have to be freed in order to achieve partial scalar invariance, then the resulting model is not very useful for further analyses. However, it should be noted that this limits the value of showing the invariance of factor variances and covariances. Furthermore, it limits the informativeness of group comparisons since without scalar invariance, we cannot be sure that differences in responses to the DUWAS items between the Italian and Dutch samples reflect real differences in the underlying traits. Instead, the differences in item responses could just be due to differential item functioning.

We are not in complete agreement with you when you say: "However, it should be noted that this limits the value of showing the invariance of factor variances and covariances". Factor variance and covariance invariance can be tested if metric invariance hold (Barbaranelli & Ingoglia, 2013). Furthermore, if metric and factor variance and covariance invariance hold, it is possible to compare correlations and betas across groups (Barbaranelli & Ingoglia, 2013). In this sense, some comparisons across groups are possible and meaningful for research even if scalar invariance does not hold.

However, we are aware of the limitations due to the lack of scalar invariance, and thus following your suggestion (for which we thank you), we have now added some considerations on this. We specifically wrote:

"Loosely speaking, the metric of the DUWAS did not change across the Italian and Dutch samples, which means that comparisons between the DUWAS and other variables at latent factors level (e.g. correlations between factors) are meaningful. To the contrary, the DUWAS observed scores in the two countries cannot be compared because scalar invariance was not supported, meaning that it is not sure that differences in responses to the DUWAS items between the Italian and Dutch samples reflect real differences in the underlying traits. This result could be due to a number of reasons (Little, 1997), including the fact that we used an Italian convenience sample which was entirely comparable to the Dutch sample, which could have particularly affected the more stringent invariance tests."

Barbaranelli, C. (2013). The multi-group structural equation modeling and the examination of factorial invariance [I modelli di equazioni strutturali molti-gruppo e l'esame dell'invarianza fattoriale]. In C. Barbaranelli, & Ingoglia, S. (Ed.), Structural Equation Modeling. Themes and perspectives [I modelli di equazioni strutturali. Temi e prospettive] (pp. 153–200). Milano: Led Edizioni.

4) P.10 line 46: "alpha values met the criterion of .70". Either a reference for the criterion should be provided or this part of the sentence should be deleted. Furthermore, considering that the assumptions of alpha are probably violated in this data, it would be better to report a different coefficient such as omega (e.g., Dunn, Baguley, & Brunsden, 2014).

Thank you for this suggestion. We initially have provided a reference for alpha, i.e. Nunnaly & Bernstein (1994), but then – due to the fact that the manuscript was longer than permitted (i.e. 5000 words) – we have decided to eliminate the entire sentence. Additionally, in the results we have now added the bootstrapped confidence interval for Cronbach's alpha (Iacobucci & Duhachek, 2003). As you can see below, the 95% confidence intervals for both dimensions was fairly narrow (.72-.77), and in both cases the lower interval was higher than .70:

WE: .74 CI: .717-.767

WC: .74 CI: .719-.769

We have also computed the omega index (Dunn e al., 2014) for both the WE and WC scales. Again, we found acceptable values, just larger than alpha (i.e. WE: .75; WC: .75)

We reported these results in a paragraph included in the results section and reported below:

"For the WE and WC subscales we also computed the bootstrap 95% confidence interval around α (see Iacobucci & Duhachek, 2003), which was found to be the same in both cases (i.e. CI=.72-.77). Since Cronbach's alpha relies on assumptions that could be unrealistic in field studies (i.e., true scores for different items are assumed to have the same variance), some have proposed to use alternative indexes to estimate scale reliability, such as the omega index (see Dunn, Baguley, & Brunsden 2014). Such index was found to be acceptable for the WE and WC subscales in the present study (ω = .75 in both cases)."

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British Journal of Psychology, 105(3), 399-412.

Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. Journal of Consumer Psychology, 13, 478–487.

A Cross-National Study on the Psychometric Quality of the Italian Version of the Dutch Work Addiction Scale (DUWAS)

by

Cristian Balducci, Lorenzo Avanzi, Chiara Consiglio, Franco Fraccaroli, & Wilmar Schaufeli

~ Cristian Balducci – Department of Political Science, University of Bologna – Via dei Bersaglieri, 6/c – 40125 Bologna (BO) – Itaky; tel.: (+39) 051 2092722; E-mail: cristian.balducci3@unibo.it

~ Lorenzo Avanzi – Department of Psychology and Cognitive Sciences, University of Trento – Corso Bettini, 84, 38068 Rovereto (TN), Italy; tel.: (+39) 0464 808715. E-mail: lorenzo.avanzi@unitn.it

~ Chiara Consiglio – Department of Psychology – University of Rome "La Sapienza" – Via dei Marsi, 78, 00185 Roma, Italy; tel.: (+39) 06 49917723. E-mail: chiara.consiglio@uniroma1.it

~ Franco Fraccaroli – Department of Psychology and Cognitive Sciences, University of Trento – Corso Bettini, 31, 38068 Rovereto (TN), Italy; tel.: (+39) 0464 808410. Email: franco.fraccaroli@unitn.it

~ Wilmar Schaufeli, Department of Social and Organizational Psychology, Utrecht University – P.O. Box 80.140, 3508 TC Utrecht, The Netherlands; tel.: (+31) 30 2539093. E-mail: W.B.Schaufeli@uu.nl

Summary

We investigated the psychometric properties of the Italian version of the Dutch Work Addiction Scale (DUWAS) by using an Italian sample (N = 1027) and a comparable Dutch sample (N = 7523). We first conducted multi-group confirmatory factor analysis and reliability analysis. We then examined the correlations of the DUWAS with job demands, work-family conflict, job-related affective well-being and psychological and physical strain. Results showed that a two-factor solution of the DUWAS was equivalent across the Italian and Dutch samples, with data supporting factor variance and covariance equivalence in addition to metric equivalence. Internal consistencies of the DUWAS and its subscales were adequate, and test-retest reliability showed a strong stability of the measure at one year. The DUWAS and its subscales showed a pattern of correlations in the expected direction with the adopted criterion measures. Overall, the Italian version of the DUWAS shows adequate validity and reliability.

Keywords: DUWAS, workaholism; work addiction; validation study; measurement equivalence.

A Cross-National study on the Psychometric Quality of the Italian Version of the Dutch Work Addiction Scale (DUWAS)

In recent years occupational health researchers have devoted increasing attention to the phenomenon of workaholism, especially with the aim of understanding its essential characteristics (Clark, Michel, Zhdanova, Pui, & Baltes, 2015). Although there are currently different definitions in the literature, a conceptualization of workaholism often used proposes that it consists of two different elements, namely working excessively hard and having a strong and irresistible inner drive to work (Schaufeli, Taris, & Bakker, 2008). The first is the behavioural and observable component and points to the exceptional amount of time that workaholics devote to the work activity. The second is the mental component and refers to a true obsession for work – i.e. the persistent focus of thoughts on work-related matters, even when the individual is not at work.

Based on the above definition, a scale has been developed – the Dutch Work Addiction Scale (DUWAS; Schaufeli, Shimazu, & Taris, 2009) – which builds upon the work of Robinson (1999) and Spence and Robbins (1992). The scale consists of 10 items, with the two components of workaholism (Working Excessively – WE, and Working Compulsively – WC) being assessed by five items each. The fit of the two-factor solution of workaholism was found to be adequate and superior to the one-factor solution, with the latent WE and WC factors correlating strongly (.50). Furthermore, the internal consistencies (i.e. Cronbach's alpha) of the two workaholism subscales were found to be adequate (i.e. .78 in both cases). The analysis of the nomological network of the DUWAS revealed, as expected, that high levels of workaholism (particularly as indicated by WE) correlated positively with a two-item measure of overwork (e.g. "How often do you take work home") and a measure of overtime (i.e. the actual working time relative to the contracted working time).

On the whole, there is at least sufficient evidence of reliability and validity of the DUWAS as a measure of workaholism. Even if the scale has also been successfully validated in other countries, such as Japan (Schaufeli et al., 2009) Spain (Del Libano, Llorens, Salanova, & Schaufeli,

2010) and Israel (Littman-Ovadia, Balducci, Ben-Moshe, 2014), cross-national validation studies are still scarce. Along this line of research, this study aims at exploring the psychometric properties of the Italian version of the DUWAS, which has been used in previous research on workaholism in Italy (e.g. Balducci, Cecchin, Fraccaroli, Schaufeli, 2012; Guglielmi, Simbula, Schaufeli, & Depolo, 2012) but for which a stringent psychometric evaluation is not yet available. Thus, the first objective of the present study is to examine the factor structure of the DUWAS by adopting a multiple sample (Italian and Dutch) confirmatory factor analysis (CFA) approach. The second objective is to explore the nomological network of the DUWAS by looking at the correlations of the scale with a number of criterion variables regarding working conditions (e.g. job demands) and health (e.g. job-related well-being) which are related with workaholism.

We advanced a number of hypotheses. First, we hypothesised that (*Hypothesis 1*) the 2-factor solution of the DUWAS would better fit the Italian data than a 1-factor solution and that the two factors would replicate, in terms of items composition, those emerged with Dutch data (Schaufeli et al, 2008) indicating that the Italian and Dutch versions of the tool are configurally equivalent and making it possible to test for even stronger forms of equivalence. Since people with workaholic tendencies may create their own job demands (Guglielmi et al., 2012) to the point that the work sphere becomes hypertrophic at the expense of other life spheres, we also hypothesized (*Hypothesis 2*) that the DUWAS and its subscales would be positively related with job demands – as conceptualized in a broad sense –, specifically workload and work-to-family conflict. As workaholism was also found to be related with negative work-related psychological states such as burnout (e.g. Van den Broeck et al., 2011) and considering that emotional reactions are crucial mediators of the stress process, we also hypothesized (*Hypothesis 3*) that the DUWAS and its subscales would be positively associated with job-related negative emotions (e.g. anger) and negatively associated with job-related positive emotions (e.g. satisfaction). Finally, given that workaholism has also been related with poorer perceived psychological and physical health (e.g.

Shimazu, Schaufeli, & Taris, 2010), we also hypothesized (*Hypothesis 4*) that the DUWAS would be positively related with psychological and physical strain symptoms.

Method

Sample and Procedure

Two different composite samples were involved in the study. An Italian sample, on which most of the analyses were based, and a Dutch sample, which was used to test for measurement equivalence of the Italian version of the DUWAS. The Italian sample has been built by using data coming from three different surveys conducted with an anonymous self-reported questionnaire. Survey 1 focused on self employed individuals or managers by using an online version of the questionnaire. Potential participants were contacted by email among acquaintances of the Italian researchers. They were asked to fill-in the survey and to forward its link to other potential participants with the characteristics specified above. A total of 171 individuals filled in the survey (40.1% women). Due to the sampling strategy, it was not possible to compute the response rate. Participants had different jobs, among which lawyer, engineer, architect, entrepreneur, university professor or researcher, manager, etc. In 76.7% of the cases participants had at least a university degree. Survey 2 was carried out in a national healthcare agency in Northern Italy. A total of 574 employees participated (76.5% women), with a response rate varying from 48.2% to 93.3% in the various organizational departments. Participants were medical doctors (6.5%), nurses (67.4%), administrative staff (24.1%), and others (e.g. personnel responsible for cleaning rooms) (2.0%). For a subgroup of 234 participants of Survey 2, we had available also 1-year follow-up data on workaholism. Survey 3 was conducted in a public environmental protection agency in Central Italy. Participants were 282 employees (44.7% women) in non-managerial position, playing an administrative (38.2%) or a technical (61.8%) role; the response rate was 54.2%. Overall, it was available a sample of 1027 participants (61% women). The age of participants was distributed as follows: 7.9% were aged under 29 years, 34.0 were aged between 30 and 39, 30.7% between 40 and 49, 23.5% between 50-59 and 4.0% were aged 60 or more.

The Dutch sample has been extracted from the DUWAS database, which contains data from more than 11.000 participants. We eliminated from the analyses participants working in the most dissimilar occupations as compared to those included in the Italian dataset. The final sample consisted of 7523 participants (42.4% women) with complete data on the DUWAS. The most represented occupations were medical doctor (30.2%), manager (21.4%), white collar employee (16.2%), higher professional (7.7%), executive (6.2%) and social worker (5.1%). Participants' age distribution was as follows: 21.0% were aged under 29; 40.0 between30 and 39; 23.1% between 40 and 49; 14.3% between 50 and 59 and 1.6% were aged 60 or more.

Measures

Workaholism was measured by using the DUWAS (Schaufeli et al., 2008), which investigates the experience of aspects of the construct by means of two five-item scales: Working excessively (WE) and Working compulsively (WC) (see Table 1). Responses are given on a frequency scale varying from 1 ("never or almost never") to 4 ("almost always or always"). The Italian version of the scale was derived from its English version by using the back translation method.[1]

Workload was measured by five items (e.g. "My job requires me to work very fast") forming the psychological demands scale of the Job Content Questionnaire (Karasek et al., 1998). Responses were given on a 4-point scale ranging from 1 ("strongly disagree") to 4 ("strongly agree"). Cronbach's alpha for this scale was .71.

Work to family conflict was investigated by means of two items (e.g., "In the last 6 months how often did your job or career interfere with your responsibilities at home, such as yard work, cooking, cleaning, repairs, shopping, paying the bills, or childcare?") taken from Grzywacz, Frone, Brewer, and Kovner (2006). Responses varied on a scale from 0 ("never") to 5 ("5 or more days per week"). The two items were strongly intercorrelated: r = .71.

Job-related affective well-being was assessed by using a shortened 8-item version of the

[1] The scale is freely available for scientific purposes on the internet at: www.wilmarschaufeli.nl

Job-related Affective Well-being Scale (JAWS; Van Katwyk, Fox, Spector, & Kelloway, 2000). The JAWS investigates the experience of positive and negative work-related affective states across the previous 30 days, with responses given on a 5-point scale ranging from 1 ("never") to 5 ("very often"). Based on a two-dimension (i.e. arousal and pleasure) model of work-related affect, specific subscales may be derived from the JAWS. We derived and used the following four 2-item subscales: high-arousal negative job-related affect (e.g. "Anger"; *r*: .55); low-arousal negative job-related affect (e.g. "Pessimism"; *r*: .70); high arousal positive job-related affect (e.g. "Energy"; *r*: .73), and low-arousal positive job-related affect (e.g. "Satisfaction"; *r*: .45). We also used the JAWS total score, which was derived by summing the eight scale items after reversing the score of the four negative affect items (alpha = .79).

While the measures described above were available in all the three surveys, those described below were available only in the second and third survey.

Psychological strain was explored by using the 12-item version of the General Health Questionnaire (GHQ-12; Goldberg, 1972). The GHQ-12 investigates the experience of a number of psychological symptoms, mainly of anxiety and depression (e.g. "You have been capable of making decisions"), with responses ranging from 0 ("no" or "more than usual", according to specific items) to 3 ("much more than usual" or "much less than usual"). Alpha for this scale was .86.

Physical strain was investigated by means of the following yes/no question: have you received a medical diagnosis for any of the following illnesses? The question was followed by a number of items, among which we selected "gastritis", "ulcer" and "colitis/irritable colon syndrome", which may be physical manifestation of stress.

Finally, a specific item which was available only in the first survey questioned participants on the average number of hours of work in a week.

Analysis

To use the full potential of the Italian sample as far as the DUWAS is concerned, we first replaced missing values at the DUWAS items by using the expectation-maximization – EM –

algorithm in SPSS. Cross-cultural equivalence of the DUWAS was explored by means of multi-group confirmatory factor analysis (CFA), which was implemented in LISREL 8.71 by using the maximum likelihood estimator. The fit of the one-factor model of workaholism was first compared to the fit of the two-factor model on the Italian and Dutch data separately. Next, a series of multiple-group CFA was run, by which more stringent forms of equivalence (i.e. configural, metric, scalar, factor variance and factor covariance) were tested (Cheung, 2008; Vanderberg & Lance, 2000).

CFA results were evaluated by using the χ^2 statistic and a variety of other more practical fit indices (cfr. Hu & Bentler, 1999; Meade, Johnson, & Braddy, 2008; Schweizer, 2010): the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), the comparative fit index (CFI), and the McDonald's non centrality index (NCI). Nested models in multiple group analysis were evaluated not only by using the $\Delta\chi^2$, which is sensitive to sample size, but also by using the ΔCFI and ΔNCI, which have been found to be the sensitive indexes to lack of invariance (Meade et al., 2008). ΔCFI values up to .002 indicate that the nested models are equivalent in terms of fit (Meade et al., 2008), while appropriate values for the ΔNCI have been computed for different number of factors and items of the investigated scale (Meade et al., 2008, p. 586). In the present case (i.e. 2 factors and 10 items) the suggested cutoff for the ΔNCI is .008.

Results

Table 1 reports the descriptive statistics at the item level of the DUWAS including factor loadings as emerged in the CFA analysis described below.

[Table 1 about here]

Table 2 reports the χ^2 and other fit indices of a series of CFA conducted on the Italian and Dutch DUWAS data. Model 1a in comparison to Model 1 and Model 2a in comparison to Model 2 suggested that in both the Italian and Dutch samples a 2-factor (Working excessively–WE and Working compulsively–WC) solution fitted the DUWAS data sufficiently well and significantly better than a 1-factor (Workaholism) solution ($\Delta\chi^2$ (1) = 202.425, $p < .001$, in Italy; $\Delta\chi^2$ (1) =

3063.103, $p < .001$, in the Netherlands). The suboptimal fit of the two-factor solution in the Dutch sample could be improved (i.e. $\Delta\chi^2$ (1) =398.37, p <.001) by freeing the covariance between a couple of items' errors (i.e. WE1 and WE3 – see Table 1). However, since this is not good practice (Schweizer, 2010) and considering that the large size of the Dutch sample could have contributed to such suboptimal fit, we didn't adopt such solution. The first multi-group analysis thus tested a model (Model 3) of configural equivalence by simultaneously evaluating the fit of Model 1a and Model 2a. The practical fit indices of Model 3 (RMSEA = .078; CFI = .958; NCI = .906) all indicated at least an acceptable fit, suggesting an invariant two-factor solution of the DUWAS in both Italy and the Netherlands. Model 4 tested for metric equivalence (i.e., equal factor loadings) and it also achieved an acceptable fit; while the χ^2 of this model deteriorated significantly with respect to its predecessor model ($\Delta\chi^2$ (8) = 70.674, $p < .05$), a ΔCFI = .002 and a ΔNCI of .003 suggested that metric equivalence was supported. By applying the same logic we rejected scalar equivalence (Model 5, see Table 2). To identify which items contributed the most to the lack of scalar equivalence and to look for whether partial scalar equivalence could hold, we ran additional analyses[2]. By freeing five items' intercepts (WE1, WE2, WC1, WC2, and WC3 – see Table 1) we obtained a model with a comparable fit to the metric invariance model – ΔCFI = .002; ΔNCI = .005. However, since the lack of scalar equivalence was due to a substantial number of items, the subsequent models were nested within the metric invariance model.

Thus Model 6, testing for factors' variance equivalence, was nested within Model 4 (metric equivalence). While the χ^2 of Model 6 slightly worsened in comparison to that of Model 4 ($\Delta\chi^2$ (2) = 16.708, $p < .05$), both a null ΔCFI and a ΔNCI of .001 suggested that factor variance equivalence was tenable. Finally, in Model 7 we constrained factor covariance to be the same across countries and again, despite a deterioration in the χ^2 of the model in comparison to that of Model 6 ($\Delta\chi^2$ (1) = 20.878, $p < .05$), a null change of the ΔCFI and a ΔNCI of .001 suggested that there was evidence

[2] We thank an anonymous reviewer for this suggestion.

for factor covariance equivalence in addition to factor variance equivalence. Overall these findings supported *Hypothesis 1*.

[Table 2 about here]

Additional analyses were run on the Italian dataset only. Internal consistency reliability (Cronbach's α) of the DUWAS overall scale was good (.82), while the internal consistency of the WE and WC subscales was adequate (.74 in both cases). For the WE and WC subscales we also computed the bootstrap 95% confidence interval around α (see Iacobucci & Duhachek, 2003), which was found to be the same in both cases (i.e. CI=.72-.77). Since Cronbach's alpha relies on assumptions that could be unrealistic in field studies (i.e., true scores for different items are assumed to have the same variance), some have proposed to use alternative indexes to estimate scale reliability, such as the omega index (see Dunn, Baguley, & Brunsden 2014). Such index was found to be acceptable for the WE and WC subscales in the present study (ω = .75 in both cases). Finally, the one-year test-retest reliability was r = .57 for the overall scale, while of r = .62 and r = .54 for the WE and WC subscales, respectively.

The DUWAS correlated positively and quite strongly with workload (see Table 3). In the subsample of self-employed workers (Survey 1, see above), the DUWAS (r = .35) and particularly WE (r = .43) showed a positive and significant correlation with the number of hours worked in a week. Furthermore, as expected, there was a positive correlation (r = .41) between the DUWAS and work-to-family conflict; quite understandably, this high correlation was mainly due to the WE subscale (r with criterion = .44). Hence we found support for *Hypothesis 2*.

[Table 3 about here]

Correlations between the study variables indicated that workaholism was positively and significantly related with both high and low arousal job-related negative affect (r varying from .21 to .25). Furthermore, and quite surprisingly, the DUWAS measures correlated positively (r varying from .12 to .15) also with high arousal job-related positive affect (e.g., enthusiasm, energy), hence *Hypothesis 3* was not supported. However, the pattern of correlations between the DUWAS

measures and overall job-related affect (i.e. a measure combining both positive and negative job-related affective states) indicates that higher workaholic tendencies go hand in hand with lower job-related affective well-being. The DUWAS and its subscales showed also positive and significant correlations with psychological strain as investigated by the GHQ-12, and with participants' self-report of a doctor's diagnosis of colitis and (especially) gastritis. Therefore *Hypothesis 4* was supported.

Finally, we didn't find any significant difference at the DUWAS according to either gender or age class. However, when we compared self employed individuals or managers vs. employees, we found that the first group reported higher levels of workaholism at the DUWAS [$F(1, 1025) = 88.50$, $p < .001$; $M(SD)$: 2.53 (0.44) vs. 2.10 (0.55)] as well as at the WE and WC subscales.

Discussion

The first aim of the study was to test the factorial validity of the Italian version of the Dutch Work Addiction Scale (DUWAS). Consistent with previous studies (e.g. Schaufeli et al., 2009), model fitting results confirmed the robustness of the two-factor solution (i.e. working excessively, WE, and working compulsively, WC). Moreover, results fully supported configural equivalence (i.e. equivalence of the number of constructs and observed variables) of the DUWAS, corroborating previous validation studies (e.g. Del Libano et al, 2010), thus attesting the stability of its factorial structure irrespective of the national context. In our study, also the equivalence in factor loadings, factor variances and covariances were also confirmed. Loosely speaking, the metric of the DUWAS did not change across the Italian and Dutch samples, which means that comparisons between the DUWAS and other variables at latent factors level (e.g. correlations between factors) are meaningful. To the contrary, the DUWAS observed scores in the two countries cannot be compared because scalar invariance was not supported, meaning that it is not sure that differences in responses to the DUWAS items between the Italian and Dutch samples reflect real differences in the underlying traits. This result could be due to a number of reasons, including the fact that we used an Italian convenience sample which was not entirely comparable to the Dutch sample.

The second aim of the study was to explore the nomological net between workaholism and other related concepts. Correlations confirmed that workaholism, and particularly the WE subscale, is strongly and positively related with job demands (i.e. workload and work-family conflict) supporting the hypothesis that workaholics to some extent may contribute to increase their job demands, for example by accepting new tasks before completing the previous ones (see e.g. Guglielmi et al, 2012). Moreover, and not surprisingly, among managers and self-employed workers (*N*=171) the DUWAS was also strongly and positively associated with the number of hours worked in a week. As expected, the DUWAS was also positively linked to psychological and physical strain reactions (GHQ and physical symptoms). This is in line with the previous literature attesting the association among workaholism, stress or burnout and reduced psychological and physical well-being (e.g. Schaufeli, Bakker, van der Heijden, & Prins, 2009). Therefore, it is likely that workaholics may constitute a high-risk group for ill health. As expected, the DUWAS showed also negative associations with overall job-related affective well-being (JAWS) and, concerning the JAWS component scales, with low and high arousal negative job-related affective states (e.g. pessimism and anger, respectively). Interestingly, the DUWAS was positively related with high arousal job-related positive affective states (e.g. energy). This suggests that individuals with workaholic tendencies do experience, to a certain extent, positive job-related affect, but this affect is mainly limited to states of high arousal type.

The present study has a number of limitations, some of which have been already noted above (e.g. the availability of convenience samples). Additionally, the variables were self-reported, hence common method variance may have inflated correlations. However, despite these limitations we believe that the present study attests that the Italian version of the DUWAS is a sound scale for assessing workaholism in organizational research conducted in Italy.

References

Balducci, C., Cecchin, M., Fraccaroli, F., & Schaufeli, W. B. (2012). Exploring the relationship between workaholism and workplace aggressive behaviour: The role of job-related emotion. *Personality and Individual Differences*, *53*, 629-634.

Burke, R. J., Matthiesen, S. B., & Pallesen, S. (2006). Personality correlates of workaholism. *Personality and Individual Differences*, *40*, 1223-1233.

Cheung, G.W. (2008). Testing equivalence in the structure, means, and variances of higher order constructs with structural equation modeling. *Organizational Research Methods*, *11*, 593–613.

Clark, M. A., Michel, J. S., Zhdanova, L., Pui, S. Y., & Baltes, B. B. (2015). All work and no play? A meta-analytic examination of the correlates and outcomes of workaholism. *Journal of Management*. Advance online publication. doi: 10.1177/0149206314522301

del Líbano, M., Llorens, S., Salanova M., & Schaufeli, W. (2010). Validity of a brief workaholism scale. *Psicothema*, *22*(1), 143-150.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399-412.

Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire*. *Maudsley monograph N. 21*. London: Oxford University Press.

Grzywacz, J. G., Frone, M. R., Brewer, C. S., & Kovner, C. T. (2006). Quantifying work-family conflict among registered nurses. *Research in Nursing and Health*, *29*(5), 414-426.

Guglielmi, D., Simbula, S., Schaufeli, W., & Depolo, M. (2012). Self-efficacy and workaholism as initiators of the job demands-resources model. *Career Development International*, *17*(4), 375-389.

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

Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology, 13,* 478–487.

Karasek, R., Brisson, C., Kawakami, N., Houtman, I., Bongers, P., & Amick, B. (1998). The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of Occupational Health Psychology, 3*(4), 322-355.

Littman-Ovadia, H., Balducci, C., & Ben-Moshe, T. (2014). Psychometric properties of the Hebrew version of the Dutch Work Addiction Scale (DUWAS-10). *The Journal of Psychology, 148*(3), 327-346.

Meade, A.W., Johnson, E. C., & Braddy, P.W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93,* 568–592.

Robinson, B. E. (1999). The work addiction risk test: Development of a tentative measure of workaholism. *Perceptual and Motor Skills, 88,* 199-210.

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

Schaufeli, W. B., Bakker, A. B., van der Heijden, F. M. M. A., & Prins, J. T. (2009). Workaholism, burnout, and well-being among junior doctors: The mediating role of role conflict. *Work & Stress, 23,* 155-172.

Schaufeli, W. B., Shimazu, A., & Taris, T. W. (2009). Being driven to work excessively hard: the evaluation of a two-factor measure of workaholism in the Netherlands and Japan. *Cross-Cultural Research, 43,* 320–348.

Schaufeli, W. B., Taris, T. W., & Bakker, A. B. (2008). It takes two to tango: Workaholism is working excessively and working compulsively. In R. J. Burke & C. L. Cooper (Eds.), *The*

Long Work Hours Culture: Causes, consequences and Choices (pp. 203-225). Bingley, UK: Emerald Group Publishing Limited.

Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment, 26*, 1–2.

Shimazu, A., Schaufeli, W. B., & Taris, T. (2010). How does workaholism affect worker health and performance? The mediating role of coping. *International Journal of Behavioral Medicine, 17*, 154–160.

Spence, J. T., & Robbins, A. S. (1992). Workaholics: Definition, measurement, and preliminary results. *Journal of Personality Assessment, 58,* 160-178.

Van den Broeck, A., Schreurs, B., De Witte, H., Vansteenkiste, M., Germeys, F., Schaufeli, W.B. (2011). Understanding Workaholics' Motivations: A Self-Determination Perspective, *Applied Psychology: An International Review, 60* (4), 600–621.

Vanderberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–69.

Van Katwyk, P. T., Fox, S., Spector, P. E., & Kelloway, E. K. (2000). Using the Job-related Affective Well-being Scale (JAWS) to investigate affective responses to work stressors. *Journal of Occupational Health Psychology*, *5*, 219–230.

Table 1. Descriptive statistics and factor loadings on the Italian sample of the DUWAS items (N = 1027)

	M (*SD*)	*Factor loading on WE*	*Factor loading on WC*
1. I seem to be in a hurry and racing against the clock (WE1)	2.63 (0.89)	.60***	–
2. I find myself continuing to work after my co-workers have called it quits (WE2)	2.22 (0.88)	.62***	–
3. I stay busy and keep many irons in the fire (WE3)	2.35 (0.88)	.62***	–
4. I spend more time working than on socializing with friends, on hobbies, or on leisure activities (WE4)	2.27 (0.99)	.60***	–
5. I find myself doing two or three things at one time such as eating lunch and writing a memo, while taking on the telephone (WE5)	2.24 (0.95)	.60***	–
6. It's important to me to work hard even when I don't enjoy what I'm doing (WC1)	2.27 (0.91)	–	.63***
7. I feel that there's something inside me that drives me to work hard (WC2)	2.41 (0.95)	–	.72***
8. I feel obliged to work hard, even when it's not enjoyable (WC3)	2.05 (0.91)	–	.68***
9. I feel guilty when I take time off work (WC4)	1.64 (0.85)	–	.55***
10. It is hard for me to relax when I'm not working (WC5)	1.66 (0.85)	–	.44***

Note. WE = Working excessively; WC = Working compulsively. Standardized factor loadings of the DUWAS two-factor model (see Table 2 – Italian sample) are reported. ***p < .001.

Table 2. Goodness of fit statistics for tests of cross-cultural equivalence of the DUWAS-10 (IT, *N* = 1027; NL, *N* = 7523)

	χ^2	df	SRMR	RMSEA	CFI	NCI
Model 1 (1-factor, IT)	390.654	35	.061	.108 (.099-.117)	.924	.845
Model 1a (2-factor, IT)	188.229	34	.041	.068 (.058-.077)	.967	.928
Model 2 (1-factor, NL)	4,639.739	35	.083	.146 (.143-.149)	.870	.736
Model 2a (2-factor, NL)	1,576.636	34	.055	.079 (.076-.082)	.957	.903
Model 3 (m.g.[a] – configural equivalence)	1,756.580	68	n.a.	.078 (.075-.081)	.958	.906
Model 4 (m.g. – metric equivalence)	1,827.254	76	n.a.	.075 (.072-.078)	.956	.903
Model 5 (m.g. – scalar equivalence)[b]	2,932.166	84	n.a.	.094 (.092-.097)	.929	.846
Model 6 (m.g. – factor variance equivalence)	1,843.962	78	n.a.	.074 (.071-.077)	.956	.902
Model 7 (m.g. – factor variance and covariance equivalence)	1,864.838	79	n.a.	.074 (.071-.077)	.956	.901

Note. IT = Italy; NL = The Netherlands. [a]m.g. = multiple-group. n.a. = not available. [b]To obtain an identified model, the means of the latent factors were constrained to “0” in the Italian sample.

Table 3. Properties and Pearson's product moment correlations of study variables (767 < *N* < 1027).

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1- Workaholism	2.17	0.56	-												
2- Working excessively	2.34	0.64	.88**	-											
3- Working compulsively	2.01	0.63	.87**	.54**	-										
4- Workload	2.77	0.49	.40**	.45**	.26**	-									
5- Work-to-family conflict	1.86	1.89	.41**	.44**	.28**	.35**	-								
6- High-arousal negative job-related affect	2.24	1.02	.25**	.21**	22**	.21**	.17**	-							
7- Low-arousal negative job-related affect	2.40	1.05	.24**	.21**	.21**	.20**	.15**	.61**	-						
8- High-arousal positive job-related affect	3.09	1.05	.15**	.14**	.12**	.12**	.06	-.12**	-.14**	-					
9- Low-arousal positive job-related affect	3.10	0.97	-.07*	-.05	-.08*	-.10**	-.08*	-.28**	-.31**	.60**	-				
10- Overall job-related affective well-being	3.39	0.73	-.16	-.13**	-.15**	-.15**	-.13**	-.71**	-.73**	.66**	.76**	-			
11- Psychological strain (GHQ-12)	2.06	2.73	.31**	.28**	.26**	.23**	.22**	.37**	.44**	-.05	-.18	-.37**	-		
12- Gastritis (0 = no; 1 = yes)	0.25	0.43	.18**	.16**	.16**	.07	.09**	.16**	.17**	-.06	-.13	-.18**	.14**	-	
13- Ulcer (0 = no; 1 = yes)	0.03	0.17	.05	.06	.03	.00	.04	.10**	.09**	-.04	-.05	-.10**	.06	.26**	-
14- Colitis (0 = no; 1 = yes)	0.20	0.40	.11**	.11**	.09**	.05	.06	.10**	.11**	-.01	-.10	-.11**	.12**	.37**	.13**

Note. **$p < .01$; *$p < .05$