Examining the Factorial Structure, Measurement Invariance, and Convergent and Discriminant Validity of a Novel Self-report Measure of Work Ability: Work Ability – Personal Radar

Ville Ilmarinen[a*]

Juhani Ilmarinen[b]

Pekka Huuhtanen[b]

Veikko Louhevaara[b]

Ove Näsman[c]

[a]*Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland*

[b]*School of Medicine, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland*

[c]*The Archipelago Academy for Wellbeing at Work, Kasnäs, Finland*

[*] Corresponding author. Email: ville-juhani.ilmarinen@helsinki.fi

Examining the Factorial Structure, Measurement Invariance, and Convergent and Discriminant Validity of a Novel Self-report Measure of Work Ability: Work Ability – Personal Radar

For identification of groups and domains for work ability promotion, brief self-report measure, Work Ability – Personal Radar (WA-PR), based on the 'the house of work ability'-model is presented and psychometrically evaluated in the structural equation framework using data from technological sector (N = 3754).

The house model had acceptable fit to the data. Additionally, factor loadings in the model were invariant across groups, demonstrating metric invariance of the WA-PR. Scalar invariance of WA-PR was fully demonstrated across men and women, and partially demonstrated across age and employee groups. Comparisons between groups revealed lower levels of health and functional capacity, but higher levels of four other WA-PR -dimensions in older employees. Additionally, all house-structures showed convergence with alternative work ability measures.

WA-PR demonstrated potential for effective measurement of multiple work ability dimensions from employees' perspective. It provides means for efficient location of relevant domains and focus groups for work ability promotion.

Keywords: work ability; measurement; self-report; invariance; validity

Practitioner Summary:

A novel approach to multidimensional work ability measurement was developed to tackle the challenges of work ability promotion. The properties of the instrument were psychometrically evaluated in structural equation modelling framework. Instrument demonstrated potential for locating relevant domains and focus groups for work ability promotions at workplaces and organizations.

1. Introduction

There is a noticeable need for maintaining and improving work ability due to ageing of the work-force and high rate of disability and early retirements. The proportion of people at working age is declining, causing a threat to the dependency ratio. Proposed solution for controlling this ratio is lengthening working careers, but that cannot be executed without taking the work ability of the aging population and those in risk for early retirement into account. To contribute to this demand, different factors affecting ability to work should be validly detected, recognized, and discriminated from each other. This study addresses the issue of effective multidimensional work ability assessment by presenting and evaluating of short self-assessment method *Work Ability – Personal Radar (WA-PR)* that aims to empower efficient and accurate work ability monitoring and promotion.

The initial definition for the concept of work ability is represented by the question: 'How good is the worker at present and in the near future, and how able is he/she to do his/her work with respect to the work demands, health, and mental resources?' (Ilmarinen and Tuomi 1993). Although this conceptualization that focuses on the state of an individual is widely agreed upon, it does not take an explicit stand on the dimensions underlying work ability. As a result, conclusions about how work ability should be promoted remain unclear. These challenges have recently brought up more detailed investigations about the underlying factors of work ability (Sturesson et al. 2013; Gould et al. 2008). Based on these studies, more holistic models reflecting various interrelated dimensions of work ability have been presented.

One recently established multidimensional model of work ability is *the house of work ability* (Figure 1; Gould et al. 2008; Ilmarinen, Tuomi, and Seitsamo 2005). In the house-model, four floors of the house and its nearby environment represent five interrelated dimensions that underlie work ability. These core structures are based on a broad study of work ability (Gould et al. 2008) and are visualized as a house, giving a straightforward and practical reminder of the relevant dimensions.

Together, three bottom floors of the house represent individual's resources affecting work ability. The first floor concerns health and functional capacity (HF). Its status as the grounding element of the house demonstrates its role as a basis for work ability; at least some health and functioning is required for a person to be able to work. The associations between work ability and health are widely documented (for review, see: van den Berg et al. 2009). The second floor of the house illustrates occupational competence (CO), which consists of work ability related expertise, knowledge, and skills (Gould et al. 2008). CO is thought as highly relevant personal resource; without any 'know how' or skills (acquired e.g. by experience, training and education) coping with the job would not be possible. Both, HF and CO are necessary, but not sufficient, for being able to work (Tengland 2011). The third floor of the house consists of attitudes and motivations (AM) possessed towards work, representing attitudinal factors affecting work ability (Gould et al. 2008). Central location of AM in the model describes its strong interrelations with the other constructs. Altogether, HF, CO, and AM -dimensions are conceived as 'the person' -factors of the work ability opposing 'the work' and 'the context of life' -factors that are described next (Sturesson et al. 2013).

Working conditions, organization of work, work community, and management are core features of the fourth floor of the house of work ability (WM). These are illustrative of physical, psychosocial and organizational working contexts employees encounter (Tengland 2011). Work ability is always defined in relation to the current occupation and tasks; therefore, the fourth floor has an inevitable effect on person's ability to work (Tuomi et al. 2004; Goedhard and Goedhard 2005; Gould et al. 2008).

As a matter of fact, one way to comprehend work ability within the house model is examining the balance between individual resources (first three floors) and demands of the work (fourth floor) (Ilmarinen, Tuomi, and Seitsamo 2005). If there is balance between resources and demands of work, work ability is good. If resources are not sufficient to deal with demands, ability to work will decline. Overall, for work ability to remain good, the demands of work should not exceed the personal resources that person is capable of providing in an occupation.[†] Similar conceptualization of work ability have also been presented elsewhere in occupational health literature: i.e., demand-control (Karasek 1979), effort-reward imbalance (Siegrist 1996), and job demands-resources (Demerouti et al. 2001) models each deal with equilibria of various work related phenomena.

The nearing environment and societal surroundings are also incorporated in the house-model of work ability. Two buildings on the yard represent relations between work, family and close community ('the context of life'; Sturesson et al. 2013). These structures are conceptually related to the domain of work-family interface; a two-way framework with potential to produce negative and positive spill-overs from work to home (and/or vice versa; Grzywacz and Marks 2000). Therefore, these factors also have an influence on work ability. Collectively, in the measurement model presented here, the buildings on the yard of the house form the dimension of work, family and spare-time activities (FS).

[†] It is possible to have more resources available than is demanded, but the concepts of overqualification (leftover of personal resources; Brynin 2002) or underemployment (lack of demands; McKee-Ryan and Harvey 2011), are not in the scope of this paper.

The house of work ability -framework depicts that its constructs are interrelated and dependent of one another (Gould et al. 2008). There is a feedback cycle and reciprocated causation between the structures. Also, none of the dimensions are emphasized over the others, despite the fact that in the past some have been studied more in detail (van den Berg et al. 2009). Correspondingly, the measurement method presented and examined in this paper does not take a stand on the causal order or the relative importance of the different structures of work ability. Instead, the purpose of WA-PR is to provide method for measuring subjective experiences of work ability according to the dimensions depicted in the house-model. The objective is to provide versatile assessment based on which more accurate allocations of interventions and promotions of work ability may be executed.

Traditionally, work ability measures, like work ability index, WAI (Ilmarinen and Tuomi 1993; Tuomi et al. 1998) have included assessment of individual's illnesses and impairments. Contrasting this viewpoint, in WA-PR, weight on impairments, symptoms, diagnoses and other health-related issues affecting work ability are relaxed. Reason for not overplaying health aspects is that these factors may have been overemphasized in past at the cost of other relevant factors. The current view is that health certainly affects, but does not entirely determine person's ability to work (Sturesson et al. 2013; Tengland 2011). Additionally in WAI, the construct of work ability is assumed to be one-dimensional, albeit this postulation has recently been called into question (Martus et al. 2010; Radkiewicz and Widerszal-Bazyl 2005). In WA-PR, multidimensionality of work ability is assumed, based on theoretical framework of the house of work ability. In addition, the WA-PR approach is based on self-assessment of subjective experiences of personal resources, working context and work-life interface, so it does not require examinations by occupational physician or other health care professionals. By subjective assessment-procedures the 'demand specific work ability' (Nabe-Nielsen et al. 2014) is better captured instead of overall work ability, by addressing the current occupation of each employee. Altogether, WA-PR was constructed to provide an efficient multifaceted instrument for work ability assessment with relatively low implementation costs.

In this study, psychometric properties of WA-PR are examined with structural equation modelling (SEM) techniques. SEM techniques are statistical methods that can be flexibly used for analysis of covariance structures (Kaplan, 2009). For example, SEM is often utilized for simultaneous examinations of latent structures (factor analysis) and associations between latent structures (regression analysis). In the present research, SEM is first used in investigating if the WA-PR formation reflects the underlying structure of work ability as depicted in the house of work ability. This is examined using confirmatory factor analysis (CFA) by testing whether the proposed house-model with five interrelated structures (factors) fits the WA-PR -data. Second, given that the house-model fits the data, the invariance of WA-PR is tested across groups. Testing of invariance hypothesis is essential step in measurement validation; only after conclusions about the similar composition of constructs and equivalent measurement across groups, the mean-level comparison between groups become meaningful (Schmitt and Kuljanin 2008). In this study, the invariance is tested across gender, age and employee groups separately. Finally, to understand if WA-PR is able to capture work ability in multidimensional manner, associations between its structures and other conceptually related measures are examined. This examination of convergent and discriminant validity also proposes locations for the WA-PR dimensions in relation to other work ability measures by sorting out how much overlap and separation there is between these supposedly related, but distinct conceptualizations.

2. Methods

2.1.Study design and setting

Study was conducted within the 'Good Work – Longer Career' -program in The Finnish Technology Industries [Teknologiateollisuus ry] that aims to maintain and promote work ability in the sector. The house of work ability was used as a common framework for the program, including the WA-PR measurement method. The framework was accepted by four trade unions and the employer association. In the program, WA-PR - measurements were used for promotion and maintenance of work ability in workplaces. The data collection was conducted by The Finnish Technology Industries. All measures used in this study were addressed to companies where participants filled out the questionnaire either with paper and pencil or online. Questionnaire was administered in Finnish. All participants were informed that their answers are used for research purposes, and were allowed not to participate or to quit filling the questionnaire at any time. Under the administration of Finnish Technology Industries, each participating company took care of the anonymity of the participants. No information based on which the participants may be identified was collected during the study.

2.2.Participants

Sample consisted of 3912 participants of whom complete data for all variables were available for 3754 (72.1% men). Participants were employees from 29 different companies under The Federation of Finnish Technology Industries (participants per organization ranging from 21 to 545). The companies represented predominantly branch of mechanical engineering which is the largest branch of technology industries, but companies from the branches of metal processing, electronic and electro-technical, and information technology industries were also represented. Participants were from all working-age groups (Under 35-year-olds: 1128; 35- to 44-year-olds: 1137; 45- to 54-year-olds: 991 and over 54-year-olds: 498), and blue-collar (45.7%) and white-collar (54.3%) employees.

2.3. Variables

2.3.1. Work Ability - Personal Radar

WA-PR consists of 18 items that are scored on five different subscales of work ability. Its five subscales are based on the theoretical framework of the work ability house (Gould et al. 2008; Ilmarinen, Tuomi, and Seitsamo 2005). Each subscale covers one element of the model and is measured with 3 to 5 items (see Appendix 1 for list of items). Development and selection of the items built on the previously reported core components within each of the work ability house structure (Gould et al. 2008). Items were designed to reflect individual's subjective experience on these components. All items were scored on an 11-point rating scale (range 0 to 10) where high score indicated positive experience.

2.3.2. Other measures

Three alternative work ability measures were chosen for examination of convergent and discriminant validity of WA-PR -dimensions. From the traditional WAI -items, Work Ability Score (WAS) was selected (Ilmarinen and Tuomi 1993; Tuomi et al. 1998). WAS-item states: 'Assume that your work ability at its best has a value of 10 points. How many points would you give your current work ability? (0 means that you cannot currently work at all)'. WAS is highly convergent with the rest of the WAI, and is also used as a single-item measure of work ability (El Fassi et al. 2013; Ahlstrom et al. 2010). The second measure selected represents another approach to work ability; the balance of personal resources and demands of work (RvD: 'Do the requirements of your work correspond to your resources?'). It was chosen as a general short representation of various equilibria concepts used in the literature of occupational health (Demerouti et al. 2001; Siegrist 1996; Karasek 1979). The selection of third measure: 'Are you able to perform in your current job until retirement?' (RET) was based on its practical utility. Since previous studies have associated early retirements (and intentions) to work ability (Van Den Berg, Elders, and Burdorf 2010), it was considered as relevant dependent measure for studying associations with WA-PR dimensions, and also for locating prominent work ability dimensions for lengthening working careers.

2.4.Statistical analyses

2.4.1. Factorial structure

CFA was performed to examine the fit of the house-model to the data consisting of 18 WA-PR -variables. Before specifying the model, pair-wise scatterplot visualization with local regression lines was conducted to investigate if the linearity assumptions regarding the associations between variables were met. For the most part, variables were linearly associated. However, there were serious violations of linearity in the associations between CO2 and other variables. Although the associations seemed linear at lower tail and in the middle of joint distributions, linearity was sharply reversed towards the higher tail. In other words, those who reported to have the best possible professional competence (10) rated themselves lower in other items than those who reported "9" in professional competence. Because of this violation, CO2 was excluded from the subsequent modelling.

Based on the house of work ability, a model with five factors was specified in which WA-PR -variables loaded on the latent structural components of the house.

Indicator variables for the floors and nearing environment were as follows. Work related health (HF1), capability (HF2) and physical work ability (HF3) loaded on the latent variable of Health and Functional capacities. Sufficiency of training (CO1) and opportunities to learn (CO3) loaded on the latent variable of Competence. Received appreciation (AM1), trustworthiness of the employer (AM2), commitment (AM3), motivation (AM4) and fair treatment (AM5) loaded on the latent variable of Attitudes and Motivation. Organization of work (WM1), supervisor's support (WM2), supervisor's feedback (WM3) and colleague's support (WM4) loaded on the latent variable of Work. Finally, reconciliation of work and family (FS1), time and resources for friends and hobbies (FS2) and flexibility of working times (FS3) loaded on the latent variable of Work and Spare-time activities. Loadings from first indicator variables of each latent factor were fixed at one. All the latent factors were permitted to correlate with each other, reflecting that all structures are thought as interdependent dimensions of work ability. To compare if the house-model with five factors would provide a better fit to the data than alternative modelling, competing one-factor model was also specified. In this model, all indicator variables loaded on one factor, General Work Ability, and the loading from the HF1 was fixed at one.

The sample covariance matrix was analysed using Lavaan package (Rosseel 2012) with R software (R Core Team 2014). Preliminary analyses indicated that there is a substantial multivariate kurtosis in the data (Mardia's multivariate kurtosis coefficient: 552.70, $p < .001$), which violates the multi-normality assumptions of maximum likelihood (ML) minimization function widely used in CFA. Since high kurtosis affects the test of covariances (Curran, West, and Finch 1996), Satorra-Bentler -rescaling for the chi-square statistic (later referred as: S-Bχ^2) was used for the analyses (Satorra and Bentler 1994). Besides S-Bχ^2, alternative fit indexes (AFIs) were used for model testing and comparison, because of the sample size sensitiveness of chi square statistics (Bentler and Bonett 1980). AFIs chosen were root-mean-square error of approximation (RMSEA; Browne and Cudeck 1993), standardized root-mean-square residual (SRMR; Bentler 1995; Joreskog and Sorbom 1984) comparative fit index (CFI; Bentler 1990), Non-normed fit index (TLI; Tucker and Lewis 1973) and Akaike's information criterion (AIC; Akaike 1987). AFIs depending on chi-square statistic (all but SRMR) were calculated using S-Bχ^2 -scores.

2.4.2. Measurement invariance

Demonstrating measurement invariance is an important step, especially if one's interest is to compare group standings with the measurement instrument (Schmitt and Kuljanin 2008). Measurement can be concluded as invariant (or equivalent) when members of different groups (e.g. men and women) with identical standing on a measured construct (e.g. health and functional capacity) are given the same result with the instrument in use (e.g. WA-PR HF-subscale).

Invariance of WA-PR was tested across gender, age groups (under 35, 35 to 44, 45 to 54, and over 54-year-olds), and employee groups (blue-collar vs. white-collar) separately. The grounds for selecting all these subpopulations from the sample was to gather understanding if WA-PR is invariant for the same data divided by different demographics, and furthermore, to understand if the observed mean-differences between these groups are in fact meaningful.

Invariance testing was conducted in a stepwise manner. First, *metric measurement invariance* was tested. Metric invariance examines if the factor loadings in the model are equivalent for different groups (Cheung and Rensvold 2002). If the loadings are similar, it is supporting the notion that latent constructs of model are manifested and composed in same way for people with different demographics; in other words, meanings of the specified factors are same across groups (Gregorich 2006). Obtaining metric invariance warrants the test of *scalar invariance*, which is performed by constraining item intercepts to be invariant across groups (Schmitt and Kuljanin 2008). Scalar invariance stands for equal offsets for item responses (the response when standings on the measured construct are identical). Establishing scalar invariance is crucial step for measures that are used for group comparisons. Only after concluding invariance across groups for item intercepts (and factor loadings before that), becomes the comparison of groups meaningful (Gregorich 2006).

Before proceeding to invariance testing of factor loadings and item intercepts, *configural invariance* of the model was examined to ensure that the proposed WA-PR item clusters are identical for different groups, and to provide a baseline model against which subsequent models are tested.

Because the chi-square difference-test for invariance is overly sensitive for detecting lack of invariance for large sample sizes, two alternative fit-index difference-test for invariance (ΔAFIs) were used: difference in comparative fit indexes, ΔCFI (Bentler 1990) and difference in root-mean-square error of approximation, ΔRMSEA (Browne and Cudeck 1993). The cut-off criteria for ΔAFIs were chosen as: -0.010 for ΔCFI and 0.015 for ΔRMSEA, as suggested by (Chen 2007). To be concluded as invariant, it was necessitated that both of these criteria are met.

2.4.3. Convergent and discriminant validity

Associations between WA-PR and other work ability measures were examined by extending the confirmatory measurement model from the research question 1 to structural equation model. This was done by adding three observed variables (WAS, RvD, and RET) to the model as dependent variables regressed by all latent structures. The associations between the measures were investigated from the path coefficients between WA-PR dimensions and dependent variables.

3. Results

3.1. Descriptive analyses

The variable means, standard deviations and correlations between variables are presented in the Table 1. All WA-PR variables were inter-correlated (Mean = .39; SD = 0.13). Gender and age were not strongly associated with other variables. Correlation between gender and employee group indicated that women were more likely to be in white-collar job. This was further concluded by cross-tabulation which also revealed that women were overrepresented (and men underrepresented) in the in white-collar occupations ($\chi^2 = 197.69$; df = 1; $p < 0.01$). In the sub-population of blue-collar employees, men were overrepresented and women underrepresented in 35 to 44 age group ($\chi^2 = 9.08$; df = 3; $p = 0.03$). In addition, WA-PR variables correlated positively and moderately strongly with alternative work ability measures (Mean = .41, SD = 0.09, range .20-.72). Also inter-correlations between WAS, RvD and RET were moderately strong (range .50-.57).

3.2. Factorial structure

Results from the CFA for WA-PR -data are presented in Table 2. For house-model with five factors (Model 1), SRMR and RMSEA indicated acceptable fit (< .08), but scaled chi-square test as well as TLI and CFI (< .90) were indicative of poor fit, demonstrating that the overall fit of Model 1 was unsatisfactory.

After investigating the sources of ill-fit from residual covariance -matrix, two additional covariance parameters were added to house-model (Model 2). First error covariance was added between residuals of commitment (AM3) and motivation (AM4).

This covariance indicates that there was some variation common to commitment and motivation that the third floor of the house did not account for. This may result from the global overlap between the concepts of commitment and motivation. Second error covariance was added between superior's support (WM2) and superior's feedback (WM3). The ground for adding this error covariance was the same object presented only in these items: superior. Responses for these indicators may reflect all superior-related experiences of the respondents, not only what was modelled in the fourth floor of the house. Both added covariance-parameters were between items closely resembling each other in wordings, which could also produce covariation attributable to the measurement method but not to work ability dimensions.

After adding these parameters the fit of the modified house-model (Model 2) was tested. Chi-square test was still significant indicating unsatisfactory fit, but all AFIs were at least on acceptable level (*CFI > .94, TLI > .93, RMSEA and SRMR < .06)*. In addition, comparison of Models 1 and 2 favoured the latter, ΔS-Bχ^2 (2, N = 3754) = 628.04, $p < .001$. Therefore it was concluded that fit of the house-model was satisfactory after these modifications.

Next, one-factor model of general work ability (Model 3) was tested and compared against Model 2. Error covariances added for Model 2 were included also in Model 3. The overall fit of the Model 3 was unsatisfactory. Of AFIs, only RMSEA was on acceptable level. Additionally, comparison of Models 2 and 3 indicated that Model 2 had better fit (AICs for comparison of non-nested models; Model 2: 237725.55, Model 3: 246276.82) thus, the house-model was retained for subsequent analyses.

Factor loadings for the Model 2 are presented in the Table 3. All loadings were statistically significant (mean standardized loading: .75, range: .47 to .93). In general, the pattern of factor loadings was highly uniform, demonstrating that all the indicators were representative of their specified structures. To investigate the accuracy of the factor loadings, five thousand datasets (Ns ranging from 150 to 500) were simulated using population parameter values given by Model 2 with R package simsem (Pornprasertmanit, Miller, & Schoemann 2014). Confidence intervals of factor loadings in the simulated datasets frequently contained the population parameters (coverage for factor loadings ranged from .940 to .952), and also, the mean squared errors in the estimates were small (MSE for factor loadings ranged from .002 to .004) indicating that the factor loadings were accurately estimated.

Correlations between the factors of Model 2 representing interrelations between the structures of the work ability house are presented at the bottom of Table 3. All the factors correlated significantly with each other. The average correlation between two structures was .56 (range .38–.88), indicating that different structures of work ability were moderately to highly interrelated. Three upper floors of the house showed strongest correlations between them (range .80–.88). Overlap of this magnitude may indicate inability in discriminating these constructs. Because of this high structural covariation, an ad-hoc model (Model 4) with additional second order factor with factor loadings from first order factors CO, AM and WM was also tested. HF and FS and their indicators, and the added error covariances were similar to Model 2. HF, FS, and the ad-hoc second order factor were permitted to covary. While the overall fit of the Model 4 was acceptable (CFI > .94, TLI > .93, RMSEA and SRMR < .06), it didn't prove better fit to the data compared to Model 2 (AICs; Model 2: 237725.55; Model 4: 237741.08). Based on these results, the more theory-driven Model 2 was retained as the best fitting model.

3.3.Measurement invariance

The confirmed model (Model 2) was used for the next step of the WA-PR validity examination: assessment of measurement invariance across gender, age groups and employee groups. Separate models were constructed for different groupings (Model 5 for gender, Model 6 for age groups, and Model 7 for employee groups). Results are presented in Table 2.

The fits of the configural models (Models 5a, 6a, and 7a) were acceptable, indicating that the house-model with five factors was appropriate for each group. Metric invariance was tested by constraining the factor loadings to be equal across groups (Models 5b, 6b and 7b). According to the chi square difference tests, constraining factor loadings decremented the model fits compared to the configural models, but ΔAFIs indicated support for the invariant factor loadings across gender (ΔCFI = -0.001, ΔRMSEA = -0.001), age groups (ΔCFI = 0.000, ΔRMSEA = -0.002), and employee groups (ΔCFI = -0.006, ΔRMSEA = 0.002), demonstrating that the compositions of the WA-PR -structures are similar for men and women, and also for different age and employee groups.

In comparison to the metric models, scalar invariance models (Models 5c, 6c and 7c) with invariant factor loadings and invariant intercepts fitted the data worse for age groups (ΔCFI = -0.016) and employee groups (ΔCFI = -0.023). Instead, constraining item intercepts did not cause lack of fit in the gender model (ΔCFI = -0.004, ΔRMSEA = 0.000); thus, metric and scalar invariance of WA-PR between men and women was demonstrated. However, the lack of fit in models 6c and 7c illustrated differences in item intercepts between age groups, and between employee groups. To locate differentially functioning items, ill-fitting intercept parameters were examined from modification indexes. This approach for differential item functioning is also known as partial invariance, by which it is tested, if part of the measurement model is invariant in given parameters (Schmitt and Kuljanin 2008). In this exploratory procedure, item intercepts were allowed to vary freely across groups one intercept at a time (chosen by the magnitude of modification indexes), and each of these models were tested against the metric invariance model. This was repeated as many times as the pre-set criteria of invariance was achieved.

By conveying this procedure for the age group model, partial scalar invariance was achieved after two item intercepts were allowed to vary freely across age groups (Model 6d). The noninvariant items were (in order of unconstraining): support from colleagues (WM4) and support from supervisor (WM2). Examining the parameter estimates from Model 6c revealed that both these intercepts were at highest level for the youngest age group and declined in linear fashion by age (unstandardized estimates from youngest to oldest age group, WM4: 8.20, 8.01, 7.56, and 7.45; WM2: 6.76, 6.53, 6.34, and 6.13). The indicators for other structures of WA-PR were invariant across age groups.

Similar procedure for employee groups resulted in non-reduced model fit after setting three parameters free for blue- and white-collar employees (Model 7d). The noninvariant items were physical work ability (HF3), received appreciation at work (AM1), and flexibility of working times (FS3). Parameter estimates for the groups indicated that intercepts of these indicators were lower for blue-collar employees (unstandardized estimates for blue-collar and white-collar employees, respectively; HF3: 7.76, 8.33; AM1: 6.21, 6.69; FS3: 7.92, 8.44). Noninvariance in these items implies that there are some systematic between-group differences – not attributable to work ability house and its structures – in how these items function when measuring different employee groups. Altogether, partial scalar invariance of WA-PR for age groups and employee groups was supported, as well as full scalar invariance across gender.

Demonstration of (partial) scalar invariance allowed mean comparisons in WA-PR structures between groups. In addition to latent means estimated in models 5c, 6d and 7d, observed means were calculated as well by aggregating items of each structure. Because there were four age groups, comparisons of latent means were conducted pairwise, by constraining single group's means to zero one at a time in the structural equation. The z-tests for deviation from zero for other groups were Bonferroni-adjusted to account for multiple comparisons. For gender model, men served as reference, likewise blue-collar employees for employee group model. Differences in observed means of WA-PR -structures were tested with ANOVAs using aggregates of all indicator variables, and additionally, using aggregates comprised of invariant items. Observed means for each group were adjusted for other group memberships, allowing controlled main effect comparisons between groups. Differences in observed aggregates were tested with contrasts. Means are presented in Table 4.

Women had higher latent means on all WA-PR -dimensions, but the differences in Health and Functional Capacity, Competence and Attitudes and Motivation were non-significant in employee and age group -adjusted observed means. Also, white-collar employees showed higher work ability levels on all dimensions than blue-collar employees, except on Work, Family and Spare-time Activities after noninvariant FS3 was excluded from the observed mean. In addition, there were age-group differences in all work ability dimensions. Younger age groups scored higher in Health and Functional Capacity, but lower in Competence, Attitudes and Motivation, Working conditions and Management and Work, Family and Spare-time Activities. In Working conditions and Management, the age group differences were illustrated also in observed means after the exclusion of noninvariant items (WM2 and WM4).

3.4. Convergent and discriminant validity

Convergent and discriminant validity of WA-PR was examined by studying the associations between its dimensions and three alternative work ability measures: WAS, RvD and RET. These measures were added to the Model 2 as dependent variables, and were regressed by all the WA-PR -factors simultaneously (Model 8)[‡]. Residuals of WAS, RvD and RET were permitted to covary. Fits of the structural equations are presented in the Table 2.

Based on AFIs, fit of the Model 8 was acceptable. Eleven of the 15 paths from house structures to WAS, RvD and RET were significant. After removing the non-significant paths, the more parsimonious model (Model 9) didn't prove worse fit and was selected for estimate presentation (Table 5). WA-PR -factors accounted for 67%, 48% and 43% of the total variances of WAS, RvD and RET, respectively. The association patterns between WA-PR -constructs and alternative measures varied by the dependent measure. In detail, Health and Functional Capacity was the strongest associate of WAS. Competence, Work Conditions and Management, and Work, Family

---

[‡] It should be noted that the chosen directionality of the paths was arbitrary. Although SEM is powerful statistical method for analyzing causality, the authors do not suggest that WA-PR – dimensions are the cause of WAS, RvD and RET. Actually, questions about causality cannot be answered with the current cross-sectional dataset. The purpose was solely to study overlap between constructs when controlled for other relevant constructs (unique associations and nomothetic overlap). To examine robustness of the results presented in the text, the models were tested also in reversed causal manner (alternative measures predicting WA-PR constructs). The model fits and interpretations of these models were identical to the models presented in the text. Results for models not presented in the article are available from authors.

and Spare-time Activities also explained unique variance in WAS, but these paths were notably weaker than the path from Health and Functional Capacity. Unexpectedly, the association between Competence and WAS was negative. Three WA-PR -constructs converged with RvD. Work Conditions and Management was the strongest associate, whilst Health and Functional Capacity and Work, Family and Spare-time Activities explained additional variance in RvD. In Model 9, all but Work Conditions and Management accounted for unique variance in RET, with Health and Functional Capacity and Work, Family and Spare-time Activities showing the strongest convergence.

Because of the high correlations between 2nd, 3rd and 4th floors of the house, and possibly resulting multicollinearity and inaccurate parameter estimates, similar structural equation model was constructed also for the ad-hoc second order factor model (Model 4). In the constructed Model 10, WAS, RvD and RET were regressed by HF, FS and the second order factor comprising of CO, AM and WM.

AIC-comparison with Model 10 (AIC: 289412.85) indicated better fit for Model 9 (AIC: 289370.24), but absolute fit of the Model 10 was nevertheless acceptable. Three WA-PR -factors in Model 10 accounted for 66%, 48% and 43% of the total variances of WAS, RvD and RET, respectively, illustrating that there were no differences in explanation power of Models 9 and 10. In addition, pattern of the path coefficients in Model 10 resembled that of the Model 9 (Table 5). WA-PR's convergence with WAS was almost entirely accounted by HF. The strongest correspondent with RvD was CO/AM/WM, along with substantial paths from HF and FS. Based on the path coefficients, all three factors explained similar proportions of variance in RET.

4. Discussion

The purpose of this study was testing the validity of WA-PR: a self-report measure of subjectively experienced work ability of the employee in his/her current occupation. The procedure considering three research questions was conducted with extensive data-set of employees participating in the Federation of Finnish Technology Industries 'Good Work – Longer Career' program.

4.1.Factorial structure

Confirmatory factor analysis revealed that the house of work ability fitted the data after inclusion of two covariance-parameters. Thus, the work ability house structures are obtained with WA-PR -questionnaire and the assumption of WA-PR's subscale-formation based on work ability house is acceptable. This claim was additionally supported by the result that the fit of the competing one-dimensional General work ability -model was poor.

In the confirmed model, factor loadings were at least moderate in size and uniformly patterned, indicating that the items are measuring the WA-PR -constructs reliably, and that the factors show good internal convergence. Based on the simulation experiment conducted with the population parameters given by the confirmed house-model, accuracy of the factor loadings was also supported.

It is depicted in the house model of work ability that its structures are all interrelated (Gould et al. 2008; Ilmarinen, Tuomi, and Seitsamo 2005). This was also confirmed as all the structures of the house correlated with each other at least moderately strong in the accepted model. In fact, the floors of Competence, Attitudes and Motivation and Work were found to be very highly correlated, indicating extensive overlap between these dimensions. Correlations of modest magnitude have also been found in other multidimensional approaches to work ability (Martus et al. 2010). However, based on the highly correlated upper floors, an alternative model with second order factor combining Competence, Attitudes and Motivation and Work was also evaluated. The model had acceptable, although somewhat worse fit than house-model. In the higher order factor model, correlations between factors were more modest, suggesting less overlap and more distinction between work ability dimensions.

Future studies should address the dimensionality issues of work ability more in detail. Possible explanations for high correlations between WA-PR -factors should be examined. It would be interesting to disentangle the causality between the attitudinal domain of the employee, and experiences of the work. For instance, understanding and differentiating between autonomous motivation (experiences of volition and self-endorsement at work) and controlled motivation (motivation controlled by rewards and punishment at work) would be of high interest in the light of work ability (Deci and Ryan 2008). The third floor of the house should be able to capture both of these types of motivations, but based on the high overlap with the work-dimensions, it is possible that the motivational domain controlled by experiences at work overplays the floor of attitudes and motivations. As for work ability promotions, potential sources of lack of motivation should be the main scope of investigations.

It should also be noted that an item measuring the second floor (Competence) was excluded before testing the house model. CO2 was nonlinearly associated with all the other WA-PR items, demonstrated by sharp reversal of linearity towards the higher tails of joint distributions. Participants who responded to have highest professional competence reported lower work ability in other items than those who reported second highest professional competence. Thus, CO2 was negatively correlated with other measures of work ability at the higher tail, but at other parts of joint distributions the correlation was positive. As the item considered self-reporting of professional competence, it is possible that the observed nonlinearity results from inability of those who are not competent to assess their own competence accurately (Kruger and Dunning

1999). In other words, those who are less skilled are also unaware and less accurate in assessing their proficiency, and tend to see themselves notably more competent than they actually are. This effect, also known as “Dunning-Kruger -effect”, has been presented for numerous self-judgments (Dunning et al. 2003). Based on the possibility for this effect, direct self-report measures of professional competence should be avoided. In future development of WA-PR, detailed attention should be paid on developing alternatives for competence measurement.

4.2.Measurement invariance

As the aim of WA-PR is to efficiently locate dimensions and groups for work ability promotion, the test for measurement invariance is a prerequisite for justified group-comparison (Schmitt and Kuljanin 2008). In the invariance test procedure conducted for WA-PR, equivalence of factor loadings (metric invariance) and item intercepts (scalar invariance) were assessed.

Tests of metric invariance of WA-PR demonstrated equal factor loadings between men and women, between age groups, and between blue-collar and white-collar employees. This result illustrates that the factors are comprised in similar fashion and have same meanings across these groupings. For example, when measuring under 35-year-olds and over 54-year-olds with WA-PR, the work ability dimensions depicted in the house-model have same interpretation for both age groups.

The scalar invariance assumption of WA-PR across groups was partially supported. Only males and females had equivalent item intercepts in all constrained WA-PR -items. In model comparing blue-collar and white-collar workers, partial scalar invariance was demonstrated after readjusting the model by unconstraining three item intercepts:physical ability to work (HF3), received appreciation at work (AM1), and flexibility of working times (FS3). In similar manner, model comparing four age groups demonstrated invariance after setting two intercepts free: support from supervisor (WM2) and support from colleagues (WM4). The age group -variant items are indicators of the work-dimension, and both deal with received support in difficult and challenging work situations (from supervisor and colleagues, respectively). Additionally for these items, the intercepts were higher for younger age groups and showed linear decline by age. This indicates that when employees of different age with similar standing in the ‘Work’ -dimension give responses to these items, younger employees rate themselves higher. Whether this reflects the general experience of lack of support (independent of the fourth floor of the house of work ability) for older people, cannot be concluded from analyses conducted here. However, the fact that the noninvariant items were different for age groups and employee groups, and that all items were invariant across gender, is likely indicating that the group differences in the item responses are not product of the general inequality of the items when addressed to people with different group membership. It is more likely that these items have meaningful and possibly unequal properties for certain groups producing variant intercepts for the groups studied. Nevertheless, one noninvariant item was concluded to be somewhat problematic also in the confirmatory phase of the house structure: WM2 had correlated errors with WM3. In further development of WA-PR, revising WM2 should be considered.

Estimated group-means revealed gender differences in Work Conditions and Management and in Work, Family and Spare-time Activities (higher mean-levels for women) and employee group differences in all but Work, Family and Spare-time Activities (white-collar employees showing higher levels). Additionally, comparisons of age groups revealed that there are somewhat linear age trends in all WA-PR -structures. Older age groups scored higher in Competence, Attitudes and Motivation, Work conditions and Management and Work, Family and Spare-time activities; but in Health and Functional capacities the trend was reverse as younger age groups had higher mean-levels. The fact that health-related dimensions of work ability decline with age is not surprising as it has been frequently reported in work ability literature (van den Berg et al. 2009). Of interest is that all other work ability dimensions showed no decline but improvement with age, even when effects of gender and employee status were taken into account. Based on this, the WA-PR behaves in part similarly and in part differently, than traditional work ability measures that show one-dimensional work ability decline with age (Ilmarinen, Tuomi, and Klockars 1997).

4.3. Convergent and discriminant validity

The dimensions of WA-PR were strongly associated with alternative work ability measures. Besides showing convergence, the present results suggest discrimination between WA-PR -dimensions illustrated by unique explanatory patterns.

Health and Functional Capacity (HF) was the strongest associate of traditional work ability measure: work ability score (WAS). This is not unexpected, given that the work ability measurement based on WAI emphasizes health and functional capacity. WAS has been recently proposed as one-item measure of work ability for screening purposes, based on its high convergence with the WAI, and similar associations with diseases and physical activity (El Fassi et al. 2013; Ahlstrom et al. 2010). In the present study, the overlap between WAS and HF illustrated in a similar fashion that the traditional work ability is captured to a high degree by the first floor of the house (health and functional capacity). There was also an unexpected – although weak – negative link between Competence and WAS. This likely resulted from the multicollinearity between Competence, Attitudes and Motivation, and Work conditions and Management, backed by the evidence that the raw correlations of CO-items and

WAS were all positive, indicating no evidence for negative associations. Results from the modelling with second order factor further affirmed this as there was a weak positive association from the CO/AM/WM -factor to WAS.

'Resources versus demands' -operationalization (RvD) of work ability converged with three house structures: Health and Functional Capacity, Work Conditions and Management and Work, Family and Spare-time Activities. The strongest path was from Work conditions and Management, probably reflecting the demands of work in the fourth floor of the house. In addition, the personal resources to cope with the demands (HF), as well as the interface of work and other life (FS) accounted for unique variance in work ability conceptualized as a balance between resources and demands.

The 'ability to perform in the current job until retirement' (RET) had unique associations with four work ability dimensions. Health and Functional Capacity and Work, Family and Spare-time activities had somewhat stronger convergence with RET than Competence and Attitudes and Motivation. It is noteworthy that although Competence, and Attitudes and Motivation had extensive overlap, both had unique links to RET, which seemed incremental, as the combined second order factor in Model 10 was notably stronger associate, than either of single paths in Model 9. This result gives also some support for the discrimination between these highly correlated work ability dimensions. Additionally, unique links from various house structures to ability to work until retirement implies potential for extending working careers by multidimensional approach to work ability.

Altogether, Health and Functional Capacity showed unique convergence with all alternative work ability measures, emphasizing its importance as a basis of work ability, and supporting its location as a base-element in house model. But Health and Functional

Capacity was the strongest associate only with the most health-focused measure, WAS. For balance of resources and work demands (RvD) and ability to work until retirement (RET), other house structures showed at least comparable associations, supporting the multidimensional approach to work ability measurement that is the underlying assumption in WA-PR.

4.4. Conclusions

Altogether, self-reports with WA-PR based on the dimensions of the work ability house showed satisfying psychometric properties. WA-PR reflected five interrelated dimensions of work ability as depicted in the house-model. It should be mentioned, however, that alternative modelling with higher order factor combining competence, attitudes and motivation, and work characteristics was also appropriate, although somewhat worse fitting. In addition, sufficient invariance of the measure was demonstrated, allowing for group-comparison that may serve as a baseline for solid and accurate work ability promotions at group level. Finally, the convergent and discriminant associations of WA-PR structures and alternative measures indicated that WA-PR obtains relevant and multidimensional information about work ability.

In sum, conceptualizing work ability by multiple dimensions is of high relevance. All five structures presented in the house of work ability, and measured with WA-PR showed associations with work ability conceptualized as either, the balance of personal resources and work demands, or as ability to work until retirement. The conclusion is that although health and functional capacity is enormously important factor, and remains as the base-element of work ability, the one-dimensional approach to work ability is too restrictive. Based on evidence gathered in this study, there are various highly important factors that underlie work ability. Besides, a notable difference between health and functional capacity and other factors is that only the former declines while other factors show increase by age, implicating, that senior employees should not be the only target group for work ability promotions. Attempts for work ability promotions should utilize the multidimensional approach to work ability by detecting and discriminating between different factors, and by locating areas, as well as groups, that are most in demand for actions. Because WA-PR is easily administered self-report method, it provides an effective opportunity to measure work ability in workplaces with low implementation costs.

Acknowledgements

The authors would like to acknowledge the Good Work - Longer Career project of the Federation of Finnish Technology Industries (2010-2015; http://www.tyohyvinvointi.info/in-english), the Metalworkers' Union, the Federation of Professional and Managerial Staff YTN, the Trade Union Pro and the Professional Engineers in Finland, for providing the data for this article.

Conflicts of interest

The authors declare that they have no conflict of interest.

References

Ahlstrom, L., A. Grimby-Ekman, M. Hagberg, and L. Dellve. 2010. "The Work Ability Index and Single-Item Question: Associations with Sick Leave, Symptoms, and Health–a Prospective Study of Women on Long-Term Sick Leave." *Scandinavian Journal of Work, Environment & Health*, 36 (5): 404–412.

Akaike, H. 1987. "Factor Analysis and AIC." *Psychometrika* 52 (3): 317–332. doi: 10.1007/BF02294359

Bentler, P. M. 1995. "*EQS Structural Equations Program Manual.*" Encino, CA: Multivariate Software.

Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88 (3): 588–606. doi: 10.1037/0033-2909.88.3.588

Bentler, P. M. 1990. "Comparative Fit Indexes in Structural Models." *Psychological Bulletin* 107 (2): 238–246.

Browne, M. W., and R. Cudeck. 1993. "Alternative Ways of Assessing Model Fit." In *Testing Structural Equations Models.*, edited by. K. A. Bollen, J. S. Long. 136–162. Newbury Park, CA: Sage Publications.

Brynin, M. 2002. "Overqualification in Employment." *Work, Employment and Society* 16 (4): 637–654. doi: 10.1177/095001702321587406

Chen, F. F. 2007. "Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance." *Structural Equation Modeling* 14 (3): 464–504.

Cheung, G. W., and R. B. Rensvold. 2002. "Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance." *Structural Equation Modeling* 9 (2): 233–255. doi: 10.1207/S15328007SEM0902_5

Curran, P. J., S. G. West, and J. F. Finch. 1996. "The Robustness of Test Statistics to Nonnormality and Specification Error in Confirmatory Factor Analysis." *Psychological Methods* 1 (1): 16–29.

Deci, E. L., and R. M. Ryan. 2008. "Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health." *Canadian Psychology* 49 (3): 182–185. doi: 10.1037/a0012801

Donaldson, S. I., and E. J. Grant-Vallone. 2002. "Understanding Self-Report Bias in Organizational Behavior Research." *Journal of Business and Psychology* 17 (2): 245–260. doi: 10.1023/A:1019637632584

Dunning, D., K. Johnson, J. Ehrlinger, and J. Kruger. 2003. "Why People Fail to Recognize Their Own Incompetence." *Current Directions in Psychological Science* 12 (3): 83–87. doi: 10.1111/1467-8721.01235

El Fassi, M., V. Bocquet, N. Majery, M. L. Lair, S. Couffignal, and P. Mairiaux. 2013. "Work Ability Assessment in a Worker Population: Comparison and Determinants of Work Ability Index and Work Ability Score." *BMC Public Health* 13 (1): 305–315. doi: 10.1186/1471-2458-13-305

Demerouti, E., F. Nachreiner, A.B. Bakker, and W.B. Schaufeli. 2001. "The Job Demands-Resources Model of Burnout." *Journal of Applied Psychology*, 86 (3): 499–512. doi: 10.1037//0021-9010.86.3.499

Goedhard, R. G., and W. J. A. Goedhard. 2005. "Work Ability and Perceived Work Stress." *International Congress Series* 1280: 79–83. doi: 10.1016/j.ics.2005.02.051

Gould, R., J. Ilmarinen, J. Järvisalo, and S. Koskinen, eds. 2008. *Dimensions of Work Ability - Results of the Health 2000 Survey*. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

Gregorich, S. E. 2006. "Do Self-Report Instruments Allow Meaningful Comparisons Across Diverse Population Groups? Testing Measurement Invariance Using the Confirmatory Factor Analysis Framework." *Medical Care* 44: 78–94. doi: 10.1097/01.mlr.0000245454.12228.8f

Grzywacz, J. G., and N. F. Marks. 2000. "Reconceptualizing the Work–Family Interface: An Ecological Perspective on the Correlates of Positive and Negative Spillover Between Work and Family." *Journal of Occupational Health Psychology* 5 (1): 111–126.

Ilmarinen, J., and K. Tuomi. 1993. "Work Ability Index for Aging Workers." *Aging and Work.* Helsinki, Finland: Finnish Institute of Occupational Health: 142–151.

Ilmarinen, J., K. Tuomi, and M. Klockars. 1997. "Changes in the Work Ability of Active Employees Over an 11-Year Period." *Scandinavian Journal of Work, Environment and Health* 23: 49–57.

Ilmarinen, J., K. Tuomi, and J. Seitsamo. 2005. "New Dimensions of Work Ability." *International Congress Series* 1280: 3–7. doi: 10.1016/j.ics.2005.02.060

Jöreskog, K. G., and D. Sörbom. 1984. *LISREL VI User's Guide.* Mooresville, IN: Scientific Software.

Kaplan, D. 2009. "Structural Equation Modeling: Foundations and Extensions." Thousand Oaks, CA: SAGE Publications. doi: 10.4135/9781452226576

Karasek, R. A. 1979. "Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign." *Administrative Science Quarterly*: 285–308.

Kruger, J., and D. Dunning. 1999. "Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments." *Journal of Personality and Social Psychology* 77 (6): 1121–1134. doi: 10.1037/0022-3514.77.6.1121

Martus, P., O. Jakob, U. Rose, R. Seibt, and G. Freude. 2010. "A Comparative Analysis of the Work Ability Index." *Occupational Medicine* 60 (7): 517–524. doi: 10.1093/occmed/kqq093

McKee-Ryan, F. M., and J. Harvey. 2011. '"I Have a Job, But. . .": A Review of Underemployment.' *Journal of Management* 37 (4): 962–996. doi: 10.1177/0149206311398134

Nabe-Nielsen, K., K. Thielen, E. Nygaard, S. V. Thorsen, and F. Diderichsen. 2014. "Demand-Specific Work Ability, Poor Health and Working Conditions in Middle-Aged Full-Time Employees." *Applied Ergonomics* 45 (4), 1174–1180. doi: 10.1016/j.apergo.2014.02.007

Pornprasertmanit, S., P. Miller, and A. Schoemann. 2014. "simsem: SIMulated Structural Equation Modeling. R package version 0.5-8." http://CRAN.R-project.org/package=simsem

R Core Team. 2014. "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria.

Radkiewicz, P., and M. Widerszal-Bazyl. 2005. "Psychometric Properties of Work Ability Index in the Light of Comparative Survey Study." *International Congress Series* 1280: 304–309. doi: 10.1016/j.ics.2005.02.089

Rosseel, Y. 2012. "lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software* 48 (2): 1–36.

Satorra, A., and P. M. Bentler. 1994. "Corrections to Test Statistics and Standard Errors in Covariance Structure Analysis." In *Latent Variables Analysis: Applications for Developmental Research.*, edited by A. von Eye, and C. C. Clogg, 399–419. Thousands Oaks, CA: Sage Publications.

Schmitt, N., and G. Kuljanin. 2008. "Measurement Invariance: Review of Practice and Implications." *Human Resource Management Review* 18 (4): 210–222. doi: 10.1016/j.hrmr.2008.03.003

Siegrist, J. 1996. "Adverse Health Effects of High-Effort/Low-Reward Conditions." *Journal of Occupational Health Psychology*, 1 (1): 27–41.

Sturesson, M., C. Edlund, A. Fjellman-Wiklund, A. H. Falkdal, and B. Bernspång. 2013. "Work Ability as Obscure, Complex and Unique: Views of Swedish Occupational Therapists and Physicians." *Work: A Journal of Prevention, Assessment and Rehabilitation* 45 (1): 117–128. doi: 10.3233/WOR-2012-1416

Tengland, P.-A. 2011. “The Concept of Work Ability.” *Journal of Occupational Rehabilitation* 21 (2): 275–285. doi: 10.1007/s10926-010-9269-x

Tucker, L. R., and C. Lewis. 1973. “A Reliability Coefficient for Maximum Likelihood Factor Analysis.” *Psychometrika* 38 (1): 1–10. doi: 10.1007/BF02291170

Tuomi, K., J. Ilmarinen, A. Jahkola, L. Katajarinne, and A. Tulkki. 1998. *Work Ability Index*, Finnish Institute of Occupational Health. Helsinki, Finland.

Tuomi, K., S. Vanhala, E. Nykyri, and M. Janhonen. 2004. “Organizational Practices, Work Demands and the Well-Being of Employees: A Follow-Up Study in the Metal Industry and Retail Trade.” *Occupational Medicine* 54 (2): 115–121. doi: 10.1093/occmed/kqh005

Van Den Berg, T. I. J., L. A. M. Elders, B. C. H. De Zwart, and A. Burdorf. 2009. “The Effects of Work-Related and Individual Factors on the Work Ability Index: A Systematic Review.” *Occupational and Environmental Medicine*: 211–220. doi: 10.1136/oem.2008.039883

Van Den Berg, T. I. J., L. A. M. Elders, and A. Burdorf. 2010. “Influence of Health and Work on Early Retirement.” *Journal of Occupational and Environmental Medicine* 52 (6): 576–83. doi: 10.1097/JOM.0b013e3181de8133

Table 1. Means, standard deviations and correlations between variables

	1	2	3	4	5	6	7	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Gender	-																						
2.Age group	.06	-																					
3. Employee	.23	.06	-																				
4. HF1	.08	-.14	.16	-																			
5. HF2	.06	-.13	.10	.81	-																		
6. HF3	.05	-.14	.28	.64	.59	-																	
7. CO1	.09	.05	.16	.28	.29	.27	-																
8. CO3	.06	.01	.27	.29	.31	.32	.50	-															
9. AM1	.04	.04	.21	.32	.34	.33	.47	.50	-														
10. AM2	.05	.04	.09	.33	.36	.30	.45	.42	.63	-													
11. AM3	.08	.12	.06	.31	.39	.29	.34	.43	.43	.55	-												
12. AM4	.07	.10	.08	.36	.45	.33	.38	.49	.50	.57	.79	-											
13. AM5	.01	-.02	.11	.33	.36	.32	.43	.45	.72	.66	.45	.55	-										
14. WM1	.12	.10	.12	.33	.36	.30	.48	.39	.52	.54	.45	.51	.53	-									
15. WM2	.06	.01	.16	.30	.33	.30	.50	.41	.56	.55	.40	.46	.61	.61	-								
16. WM3	.06	.08	.21	.26	.28	.26	.45	.44	.54	.47	.35	.41	.53	.52	.75	-							
17. WM4	.01	-.11	-.03	.25	.27	.24	.27	.27	.36	.27	.26	.29	.38	.35	.37	.31	-						
18. FS1	.11	.10	.05	.32	.33	.29	.24	.19	.25	.30	.29	.32	.27	.34	.28	.24	.22	-					
19. FS2	.06	.11	.04	.37	.38	.32	.29	.24	.31	.33	.27	.34	.30	.38	.30	.29	.22	.71	-				
20. FS3	.12	.08	.17	.31	.30	.32	.23	.26	.27	.30	.34	.35	.28	.34	.29	.28	.21	.70	.51	-			
21. WAS	.05	-.14	.09	.68	.72	.64	.29	.29	.37	.38	.40	.46	.38	.40	.35	.28	.30	.37	.44	.31	-		
22. RET	.09	.11	.21	.46	.49	.43	.33	.38	.40	.42	.44	.50	.39	.42	.38	.36	.20	.39	.46	.38	.52	-	
23. RvD	.06	.02	.09	.46	.51	.45	.37	.38	.44	.44	.44	.50	.43	.50	.42	.35	.34	.41	.45	.38	.57	.50	-
Mean	0.28	2.23	0.54	8.22	8.20	8.24	6.35	7.35	6.70	6.62	8.13	7.67	7.05	6.44	6.73	5.99	7.97	7.88	6.84	8.29	7.64	6.76	7.63
Std. Deviation	0.45	1.02	0.50	1.50	1.51	1.59	2.52	2.24	2.18	2.40	1.77	2.02	2.39	2.08	2.50	2.60	1.75	1.97	2.40	1.96	1.74	2.84	1.70

N = 3754, All correlations, p < .05; Men, blue-collar employees, and youngest age group serve as reference group (0); HF = Health and Functional Capacity; CO = Competence; AM = Attitudes and Motivation; WM = Work conditions and Management; FS = Work, Family and Spare-time Acitivites; WAS = Work Ability Score; RET = Ability to Perform in the Current Job until Retirement; RvD = Resources versus Demands

Table 2 Model parameters, chi square tests and fit indexes for estimated models

Model and description	Comparison model	S-B χ²	df	Scaling factor	ΔS-B χ² / Δdf	p (Δ)	CFI	ΔCFI	RMSEA	ΔRMSEA	TLI	SRMR
1. House-model (5 factors)		2516.71	109	1.642			.879		.077		.849	.055
2. House-model (5 factors; 2 residual covariances)	1	1196.32	107	1.601	314.02	***	.945	.067	.052	-.025	.930	.046
3. General Work Ability (1 factor; 2 residual covariances)	2	6387.15	117	1.642	411.99	***	.684	-.261	.119	.067	.633	.090
4. House-model + 2nd order factor for CO, AM and WM (2 residual covariances)	2	1216.30	111	1.594	4.18	**	.944	-.001	.052	.000	.932	.047
5a. Gender: Configural model		1331.90	214	1.600			.943		.053		.928	.048
5b. Gender: metric invariance	5a	1360.59	226	1.600	2.38	**	.942	-.001	.052	-.001	.930	.050
5c. Gender: scalar invariance	5b	1446.47	238	1.563	8.09	***	.938	-.004	.052	.000	.930	.050
6a. Age: Configural model		1560.45	428	1.599			.940		.053		.924	.051
6b. Age: metric invariance	6a	1597.61	464	1.614	1.30	.109	.940	.000	.051	-.002	.930	.054
6c. Age: scalar invariance	6b	1903.69	500	1.546	15.06	***	.926	-.016	.055	.004	.920	.057
6d. Age: partial scalar invariance	6b	1776.86	494	1.564	8.43	***	.932	-.008	.053	.002	.926	.056
7a. Employee groups: Configural model		1358.26	214	1.581			.942		.053		.927	.049
7b. Employee groups: metric invariance	7a	1496.42	226	1.590	11.04	***	.936	-.006	.055	.002	.923	.057
7c. Employee groups: scalar invariance	7b	1966.15	238	1.530	134.82	***	.913	-.023	.062	.007	.901	.062
7d. Employee groups: partial scalar invariance	7b	1662.15	235	1.557	31.85	***	.928	-.008	.057	.002	.917	.062
8. SEM: Five house structures → WAS, RvD, RET		1614.30	143	1.637			.939		.052		.919	.045
9. SEM: Model 8 + ns. paths removed	8	1621.13	147	1.635	1.30	.266	.939	.000	.052	.000	.921	.045
10. SEM: Three (2nd order) house structures → WAS, RvD, RET		1662.20	153	1.628			.937		.051		.922	.046

Note: For all models, absolute chi square based fit was significant p < .001; S-B χ² = Satorra-Bentler scaled chi-square value (All AFIs except SRMR were calculated using this); df = Degrees of freedom; Scaling factor = Coefficient for calculating the original chi-square value from S-B -value; Δ = Difference between model and comparison model; CFI = Comparative fit index; RMSEA = Root mean square error of approximation; TLI = Tucker-Lewis Index (non-normed fit index); SRMR = Standardized root mean square of residuals; *** p < .001 ** p < .01 * p < .05

Table 3. Standardized factor loadings of indicator variables and correlations between latent factors (Model 2)

Indicator variables	Health and Functional Capacity (HF)	Competence (CO)	Attitudes and Motivation (AM)	Work conditions and Management (WM)	Work, Family and Spare-time acitivites (FS)
HF1	.92				
HF2	.89				
HF3	.69				
CO1		.73			
CO2		.25			
CO3		.68			
AM1			.81		
AM2			.79		
AM3			.60		
AM4			.68		
AM5			.84		
WM1				.76	
WM2				.79	
WM3				.70	
WM4				.48	
FS1					.80
FS2					.89
FS3					.60
Latent factors					
HF	-				
CO	.49	-			
AM	.49	.79	-		
WM	.48	.83	.88	-	
FS	.47	.43	.45	.50	-

Table 3. NO CO2 Standardized factor loadings of indicator variables and correlations between latent factors (Model 2)

Indicator variables	Health and Functional Capacity (HF)	Competence (CO)	Attitudes and Motivation (AM)	Work conditions and Management (WM)	Work, Family and Spare-time acitivites (FS)	Factor loading within simulated CI
HF1	.92					.946
HF2	.88					.948
HF3	.69					.950
CO1		.72				.946
CO3		.70				.947
AM1			.81			.955
AM2			.79			.948
AM3			.60			.951
AM4			.68			.951
AM5			.84			.948
WM1				.76		.940
WM2				.79		.946
WM3				.70		.946
WM4				.47		.952
FS1					.93	.947
FS2					.76	.946
FS3					.74	.949
Latent factors						
HF	-					
CO	.47	-				
AM	.49	.80	-			
WM	.48	.82	.88	-		
FS	.43	.38	.42	.47	-	

Table 4 Latent and observed group-means for WA-PR -dimensions

	Health and Functional Capacity (HF)			**Competence (CO)**		**Attitudes and Motivation (AM)**			**Work conditions and Management (WM)**			**Work, Family and Spare-time activities (FS)**		
	Latent	Observed		Latent	Obs.	Latent	Observed		Latent	Observed		Latent	Observed	
		All items	HF3 excl.		All items		All items	AM1 excl.		All items	WM2 and WM4 excl.		All items	FS3 excl.
Gender														
Male (N = 2706)	0.00	8.12[a]	8.11	0.00	6.79[a]	0.00	7.23[a]	7.36	0.00	6.73	6.18	0.00	7.63	7.35
Female (N = 1048)	0.19	8.21[a]	8.26	0.25	6.93[a]	0.11	7.33[a]	7.50	0.26	6.92	6.42	0.29	7.98	7.70
Employee Group														
Blue-collar (N = 1714)	0.00	7.88	7.98	0.00	6.38	0.00	7.05	7.26	0.00	6.56	5.93	0.00	7.67	7.46[a]
White-collar (N = 2040)	0.34	8.45	8.39	0.70	7.34	0.26	7.51	7.60	0.39	7.09	6.67	0.12	7.95	7.58[a]
Age group														
[–35] (N = 1128)	0.00	8.55	8.58	0.00[a]	6.84[a]	0.00[a]	7.17[a]	7.32[a]	0.00[a]	6.80[a]	6.09[a]	0.00[a]	7.66[a]	7.36[a]
[35–44] (N = 1137)	-0.26[a]	8.18	8.15[a]	0.01[a]	6.76[a]	0.01[a]	7.18[a]	7.30[a]	0.06[a]	6.74[a]	6.12[ab]	-0.09[a]	7.48[a]	7.10
[45–54] (N = 991)	-0.36[ab]	8.02[a]	8.05[a]	0.05[a]	6.83[a]	0.00[a]	7.24[a]	7.40[a]	0.19	6.77[a]	6.33[b]	0.16	7.87	7.59[a]
[54–] (N = 498)	-0.40[b]	7.91[a]	7.96[a]	0.13[a]	7.00[b]	0.17	7.52	7.68	0.40	6.98[a]	6.65	0.42	8.22	8.03

Note. Latent means are standardized. Observed means (obs.) are unstandardized and adjusted for memberships of other groups. Observed means presented for original scales including all items and for scales including items demonstrated to be invariant across all grouping structures are presented separately. Observed mean-differences were tested using ANOVA contrasts. For all group comparisons, lower-case letter indicates non-significant (p ≥ .05) pairwise group-difference.

Table 5. Path coefficients from WA-PR -factors to alternative work ability measures

Model 9

Dependent variable (R^2)	Independent variable	Unstd. coefficient	Std. Error	Z	p	Std. Coefficient
WAS (66%)	HF	0.954	0.029	33.052	***	.730
	CO	-0.141	0.039	-3.642	***	-.146
	WM	0.250	0.044	5.676	***	.229
	FS	0.059	0.017	3.418	**	.061
RvD (47%)	HF	0.370	0.027	13.584	***	.290
	WM	0.403	0.027	14.705	***	.378
	FS	0.162	0.020	8.242	***	.172
RET (42%)	HF	0.633	0.046	13.738	***	.297
	CO	0.213	0.066	3.208	**	.135
	AM	0.283	0.066	4.315	***	.176
	FS	0.331	0.032	10.388	***	.210

Model 10

Dependent variable (R^2)	Independent variable	Unstd. coefficient	Std. Error	Z	p	Std. Coefficient
WAS (66%)	HF	0.933	0.029	32.229	***	.714
	CO/AM/WM	0.108	0.020	5.552	***	.097
	FS	0.073	0.017	4.421	***	.076
RvD (47%)	HF	0.339	0.028	12.020	***	.266
	CO/AM/WM	0.420	0.030	14.129	***	.385
	FS	0.172	0.019	8.961	***	.182
RET (42%)	HF	0.622	0.046	13.429	***	.291
	CO/AM/WM	0.566	0.045	12.538	***	.310
	FS	0.303	0.033	9.225	***	.192

** $p < .01$, *** $p < .001$, R^2 = variance explained in the dependent variable; HF = Health and Functional Capacity; CO = Competence; AM = Attitudes and Motivation; WM = Work conditions and Management; FS = Work, Family and Spare-time Acitivites; WAS = Work Ability Score; RET = Ability to Perform in the Current Job until Retirement; RvD = Resources versus Demands. If the path coefficients would have been drawn the other way around (from Alternative Work Ability measures to WA-PR-factors) the associations and their interpretations would have been almost identical.