Psychometric properties of the UK and Italian versions of the HSE Stress Indicator Tool: a crosscultural investigation

Journal Item

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.1027/1015-5759/a000122

ISSN-L 1015-5759 • ISSN-Print 1015-5759 • ISSN-Online 2151-2426

Official Organ of the European Association of Psychological Assessment

european journal of psychological assessment

www.hogrefe.com/journals/ejpa

Abstracted/Indexed in:
Current Contents/
Social & Behavioral Sciences
Social Sciences Citation Index (SSCI)
Social Scisearch
PsycINFO
Psychological Abstracts
PSYNDEX
ERIH
Scopus

Edited by
Karl Schweizer

HOGREFE

Original Article

Psychometric Properties of the UK and Italian Versions of the HSE Stress Indicator Tool

A Crosscultural Investigation

Stefano Toderi[1], Cristian Balducci[2], Julian A. Edwards[3], Guido Sarchielli[1], Marco Broccoli[4], and Gianpiero Mancini[4]

[1]Department of the Education Sciences, University of Bologna, Bologna, Italy, [2]Department of Political Science, University of Bologna, [3]Business School, The Open University, [4]Servizio di Prevenzione e Sicurezza negli Ambienti di Lavoro, AUSL di Ravenna, Ravenna, Italy

Abstract. The Health and Safety Executive Stress Indicator Tool is a short questionnaire developed in the UK for the screening of common psychosocial risk factors leading to work-related stress. Previous research showed that in both the UK and Italy this test fits the same seven-factor solution. Building upon previous findings and by using multiple-group confirmatory factor analysis we investigated the degree of measurement equivalence of the Italian version of the test with its original UK version. Furthermore, we advanced the validation of the tool in Italy by exploring the correlations of its subscales with stress-related outcomes. Two samples of workers were available for analyses: 1,298 Italian employees of private-sector organizations and a sample of 7,589 UK employees from organizations of the same type. Results replicated previous findings: A 7-factor solution of the Indicator Tool was found to be equivalent across the Italian and UK samples. Furthermore, the analyses supported factor variance and factor covariance equivalence in addition to metric equivalence. The Indicator Tool subscales showed significant correlations in the expected direction with stress-related outcomes. The Italian version of the Indicator Tool behaves very consistently with its UK version and may be used for the screening of psychosocial risks in Italian workplaces.

Keywords: HSE indicator tool, work-related stress, psychosocial risks, measurement equivalence, crosscultural validation

Introduction

In the area of work-related stress, recent research has seen an increasing interest in the practice (Cox, Karanika, Griffiths, & Houdmont, 2007) and quality of interventions (LaMontagne, Keegel, Louie, Ostry, & Landsbergis, 2007). In Europe this interest was driven mainly by the Framework Directive 89/391/EEC, which essentially states that employers have a legal obligation to protect the occupational safety and health of workers, a duty that also applies to problems of work-related stress. A second important step was the agreement signed between the European social partners in October 2004, known as the "Framework Agreement on Work-Related Stress," which gave preliminary guidelines for stress assessment and prevention (see, e.g., European Foundation, 2010). Following these initiatives, researchers from different European countries developed theoretically based comprehensive methods aimed at the management of work-related stress (e.g., Cousins et al., 2004; Malchaire, Pierre, D'Horre, & Stordeur, 2008; Satzer, 2009).

One of the most advanced of such methods is the Management Standards approach of the UK Health and Safety Executive (Cousins et al., 2004; Mackay, Cousins, Kelly, Lee, & McCaig, 2004). According to this approach, work-related stress may be assessed by considering the following seven psychosocial risks, which have the potential to lead to stress-related negative outcomes:

- Demands (mainly exploring issues such as workload, work patterns, and the work environment);
- Control (how much say the person has in the way they do their work);
- Supervisor support (which includes the encouragement and resources provided by the organization and line management);
- Peer support (the encouragement and resources provided by colleagues);
- Relationships (which includes promoting positive work-

DOI: 10.1027/1015-5759/a000122

ing practices to avoid conflict and dealing with unacceptable behavior);
- Role (whether people understand their role within the organization and whether the organization ensures that the person does not have conflicting roles);
- Change (how organizational change is managed and communicated).

As part of the Management Standards approach, the HSE developed a 35-item self-report questionnaire for the screening of these factors. The questionnaire, called the Indicator Tool, has provided evidence of adequate validity and reliability, and of having relationships with stress-related outcomes. Edwards, Webster, Van Laar, and Easton (2008), focusing on workers of mainly public sector organizations, found that a seven-factor solution indeed fit the data well, with all the Indicator Tool subscales achieving at least a sufficient internal consistency reliability. These authors, however, recognized the limit of considering mainly public organizations and pointed out the need for further evaluations using data collected from the private sector.

Furthermore, Kerr, McHugh, and McCrory (2009) found that the Indicator Tool subscales showed expected relationships with a number of stress-related outcomes such as anxiety, depression and job satisfaction.

Probably due to its simplicity and widely known theoretical underpinnings (Karasek & Theorell, 1990), the HSE approach crossed the UK boundary, and, for example, it was suggested by the Italian National Institute for Occupational Safety and Prevention (ISPESL) as a method for work-related stress assessment and prevention. Toderi, Broccoli, Sarchielli, and Mancini (in press), focusing on Italian workers of private sector organizations and using the same procedure of Edwards et al. (2008), found support for the same seven-factor solution that emerged with UK data. However, an investigation of stress-related correlates of the Indicator Tool in Italy is still lacking, meaning that the validation process is still not complete.

Thus, the main aim of the present study is to simultaneously contribute to the validation of the UK and Italian (IT) versions of the Indicator Tool through a crosscultural investigation, aimed at (1) replicating the validation of the UK Indicator Tool using private sector organizations data; (2) comparing the two (UK and IT) versions of the Indicator Tool by exploring their measurement equivalence (Cheung, 2008; Vanderberg & Lance, 2000); and (3) advancing the validation of the IT version of the Indicator Tool by exploring its correlations with criterion-related variables (e.g., job-related and general mental well-being), and differences according to main sociodemographic (e.g., sex) and occupational (e.g., job position) variables.

Specifically, we hypothesize (Hypothesis 1) that the UK version of the Indicator Tool will show the same 7-factor structure found in the previous UK and IT validations. Furthermore, we hypothesize (Hypothesis 2) that the IT version of the Indicator Tool should show at least configural equivalence (Cheung, 2008) with its original UK version, making it possible to test for even more stringent forms of equivalence (metric equivalence, uniqueness equivalence, scalar equivalence, etc.). Finally, we hypothesize (Hypothesis 3) that the IT version of the Indicator Tool subscales (for which higher scores indicate a better psychosocial work environment) should show a pattern of positive correlations with a measure of job-related well-being and a pattern of negative correlations with a measure of mental health problems.

Materials and Methods

Participants

The Italian sample consisted of 1,298 workers employed in 45 small, medium and large organizations operating in the private sector. Data were collected through a study carried out by a regional structure of the Italian National Health System, which randomly extracted 60 organizations and proposed their participation in a study aimed at promoting the quality of working life.

The overall workers response rate was 67.39%. Workers were mainly female (57.2%) and heterogeneous for age: 34% were 25–35 years old, 36% were 36–45, 22% were 46–55, 4% were younger than 25, and 4% were older than 56. In 83.4% of the cases the workers had a permanent job contract. Job position was distributed as follows: 61.2% of workers were blue-collar employees, 32.4% white-collar employees, and 6.4% had managerial responsibilities. In 47% of the cases workers had a shift work schedule, and in 27% of the cases they did night work. The UK sample, which is different from that used by Edwards et al. (2008), included 7,589 workers of 26 small, medium, and large private organizations. Organizations voluntarily participated in the HSE's piloting of the Management Standards. In return, the organizations were supported through the delivery of the Management Standards. Although no data collected was for the purpose of this study, organizations nevertheless consented to share their data on request. Information was provided on how to deliver the Management Standards via the Indicator Tool. In the overall UK dataset, only private firms were considered for this study, aimed to a better comparability with the Italian sample. Fifteen organizations had sample sizes of 50–249 and 11 organizations had 250+ responses. No further demographic and occupational variables were available.

Questionnaire

Both samples filled in the last version of the Indicator Tool developed by the HSE, in its UK (Edwards et al., 2008) or Italian (Toderi et al., in press) version depending on the sample.

The Indicator Tool consists of 35 items measuring seven work-related stressors, providing information about the performance of the organization on the Management Standards. Demands are measured by eight items (e.g., "I have to work very intensively"), Control by six items (e.g., "I have a say in my own work speed"), Supervisor support by five items (e.g., "I am given supportive feedback on the work I do"), Peer support by four items (e.g., "I get help and support I need from colleagues"), Relations by four items (i.e., "There is friction or anger between colleagues"), Role by five items (e.g., "I am clear what is expected of me at work") and Change by three items (e.g., "I have sufficient opportunities to question managers about change at work"). Response options varied from 1 (= *never or strongly disagree*, depending on the item) to 5 (= *often or strongly agree*, depending on the item).

In addition to sociodemographic and occupational variables and to the Indicator Tool items, Italian participants were also asked how the job made them feel (1 = *very bad*, 10 = *very well*). A second question investigated how often the participant made use of medications for anxiety, depressed mood, insomnia, or pain (1 = *never*, 5 = *everyday/always*). Finally a third question asked whether the participant experienced accidents at work in the last year (*yes/no*).

Analyses

The core of the analyses consisted of a series of confirmatory factor analyses (CFA), which were run by using LISREL 8.71 (maximum likelihood estimation method). We first carried out a CFA on the IT and UK datasets separately, comparing the fit of a one-factor model to the fit of the seven-factor model in each country. The aim of these analyses was to ascertain whether the seven-factor model fit the data better than the one-factor model in each country, and to establish the baseline models over which to run multiple-group analyses. Next, a series of multiple-group CFA was run, through which different forms of equivalence were tested (Cheung, 2008; Vanderberg & Lance, 2000). In addition to configural equivalence (i.e., the number of constructs and the observed variables that are associated with each construct are the same across groups), the following forms of equivalence may be tested:

- Metric equivalence, or equivalence in factor loadings (λ_{ij} group 1 = λ_{ij} group 2)
- Uniqueness equivalence, or equivalence in item error variances (θ_{ii} group 1 = θ_{ii} group 2)
- Scalar equivalence, or equivalence in item intercepts (τ_i group 1 = τ_i group 2)
- Equivalence in factor means (κ_j group 1 = κ_j group 2)
- Equivalence in factor variances (φ_{jj} group 1 = φ_{jj} group 2)
- Equivalence in factor covariances ($\varphi_{jj'}$ group 1 = $\varphi_{jj'}$ group 2)

There is no clear consensus over the exact sequence in conducting the various forms of measurement equivalence tests (Marsh, Tracey, & Craven, 2006). Tests of configural equivalence and metric equivalence are the first two steps. Uniqueness equivalence may be the third step; however, this form of equivalence is rarely met and it is not required for testing the remaining forms of equivalence (see Cheung, 2008). Equivalence in factor means can be tested once scalar equivalence is supported. Equivalence in factor variances and covariances are independent each other and may be tested once metric equivalence has been established (Cheung, 2008).

CFA results were evaluated by using the χ^2 statistic and a variety of other more practical fit indices (see Schweizer, 2010): the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the comparative fit index (CFI). RMSEA values lower than .05 are usually considered good, while values lower than .08 are considered acceptable (Browne & Cudeck, 1993). Values at the SRMR below .10 are considered acceptable (Kline, 2005). Finally, values at the CFI equal or higher than .90 are considered acceptable, while values equal or higher than .95 are considered good (Hu & Bentler, 1999). Nested models were evaluated not only by using the $\Delta\chi^2$, which – like its absolute value – is sensitive to sample size (Meade, Johnson, & Braddy, 2008), but also by using the ΔCFI, with values in this statistic up to .002 indicating that the models are equivalent in terms of fit (Meade et al., 2008).

In addition to CFA, other well-known analytical tools such as correlations and multivariate analyses of variance (MANOVA) were also used, which were implemented by using SPSS 17.

Results

Multiple-Group CFA

Table 1 reports the results of a series of CFA carried out on the Indicator Tool data.

Preliminary analyses conducted separately on the Italian and UK datasets confirmed that the 7-factor solution of the Indicator Tool was a better fit than the 1-factor solution in both Italy (Table 1, Model 1a vs. Model 1b) and the UK (Table 1, Model 2a vs. Model 2b). Thus, Model 1b and Model 2b, which configured a 7-factor solution of the Indicator Tool, were taken as the baseline models for the following sequence of multiple group analyses through which measurement equivalence was tested (see Figure 1).

The first multiple-group analysis thus tested a model of configural equivalence (Model 3) by simultaneously evaluating the fit of Model 1b and Model 2b. The practical fit indices (SRMR = .057; RMSEA = .063; CFI = .965) all indicated a quite good fit for this model, supporting an equivalent 7-factor solution of the Indicator Tool in the UK

Table 1. Goodness of fit statistics for tests of crosscultural equivalence of the HSE

	χ^2	*df*	SRMR	RMSEA	CFI
Model 1a (1-factor, IT)	8,260.994	560	.093	.129 (.127–.131)	.877
Model 1b (7-factor, IT)	2772.481	539	.054	.060 (.058–.062)	.964
Model 2a (1-factor, UK)	57000.316	560	.103	.148 (.147–.148)	.867
Model 2b (7-factor, UK)	15321.246	539	.057	.063 (.063–.064)	.965
Model 3 (m.g.[a] – configural equivalence)	18093.727	1078	.057	.063 (.062–.064)	.965
Model 4 (m.g. – metric equivalence)	18662.254	1106	.059	.063 (.062–.064)	.964
Model 5 (m.g. – uniqueness equivalence)	20817.274	1141	.059	.067 (.066–.067)	.960
Model 6 (m.g. – scalar equivalence[b])	22265.582	1134	.059	.071 (.071–.072)	.957
Model 7 (m.g. – factor variance equivalence)	18883.968	1113	.059	.063 (.063–.064)	.963
Model 8 (m.g. – factor covariance equivalence)	19256.989	1134	.059	.064 (.063–.064)	.963

Notes. [a]m.g = multiple-group. [b]To obtain an identified model, the intercept of the reference item for each factor was constrained to "0" across countries. UK: *N* = 7,589; IT: *N* = 1,298.

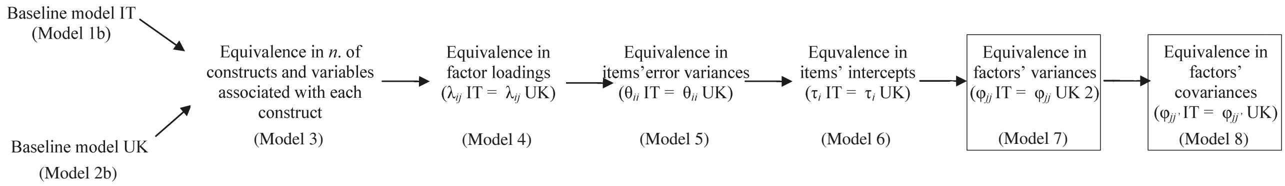

Figure 1. Flow diagram of the measurement equivalence tests conducted on the Italian and UK versions of the Indicator Tool.

and IT datasets. Model 4 tested for metric equivalence (see Figure 1). While the χ^2 of Model 4 deteriorated compared to that of Model 3 ($\Delta\chi^2_{M4\text{-}M3}(28) = 568.527$), all the fit indices of Model 4 were acceptable. More importantly, a $\Delta CFI_{M3\text{-}M4} = .001$ suggested that Model 4 could be considered equivalent to Model 3. Thus, metric equivalence was also supported. By applying the same logic we rejected uniqueness equivalence as tested by Model 5 ($\Delta\chi^2_{M5\text{-}M4}(35) = 2155.02$; $\Delta CFI_{M4\text{-}M5} = .004$). Thus, Model 6, testing for scalar equivalence (i.e., equivalence in item intercepts – see Figure 1) was nested within Model 4. However, constraining item intercepts to be equal across samples determined a drastic deterioration in the χ^{22} ($\Delta\chi^2_{M6\text{-}M4}(28) = 3603.328$); a $\Delta CFI_{M4\text{-}M6} = .007$ suggested that scalar equivalence was not supported. Since a lack of equivalence at this level may imply that the scores on the items in one country can be systematically biased upward or downward compared to the corresponding scores in the other country (Steenkamp & Baumgartner, 1998), we then investigated which items contributed the most to this potential bias. To this end, we followed a procedure described by Cheung and Rensvold (1999). Specifically, for each HSE factor, the intercept of the referent item was fixed at 0 in both countries and then the intercepts of the other items were tested one by one for equivalence. A second item was then used as a referent item in both countries, its intercept constrained at 0, and then the intercepts of the remaining items were tested one by one for equivalence. The process was repeated until all of the items had been used as a referent item. This analysis didn't reveal items with equivalent intercepts according to the method proposed by Cheung and Rensvold (1999). The items that contributed the most to the lack of scalar equivalence as indicated by the change in the χ^2 statistics were the following: Item 26 ("I have sufficient opportunities to question managers about change at work – Change"), determining an average drop in χ^2 of 490,5; item 32 ("When changes are made at work, I am clear how they will work out in practice – Change"), with an average drop in χ^2 of 359,1; and Item 22 ("I have unrealistic time pressures – Demand"), with an average drop in χ^2 of 350,0[1]. However, for none of the tests conducted did the CFI of a value drop higher than .002, suggesting that a cumulative deterioration in fit was responsible for the lack of scalar equivalence as tested by Model 6.

We then proceeded by testing equivalence in factor variances (see Figure 1 and Table 2) and found that this form of equivalence was tenable ($\Delta\chi^2_{M7\text{-}M4}(7) = 221.714$; $\delta CFI_{M4\text{-}M7} = .001$). We finally tested for equivalence in factor covariances by nesting the respective model (Model 8) within Model 7 and found that this latter form of equivalence was also tenable ($\Delta\chi^2_{M8\text{-}M7}(21) = 373.021$; $\Delta CFI_{M7\text{-}M8} = .000$). Common metric completely standardized factor loadings for the final model (Model 8) were all statistically significant and varied between .516 and .815, with a mean

1 The complete results from this analysis are available upon request from the first author.

Table 2. Descriptive statistics of HSE scales as obtained in the UK and IT datasets

Dimension	No. items	α	Mean	*SD*	Percentiles		
					25	50	75
Demands							
Item 3, 6, 9, 12, 16, 18, 20, 22	8	.85 (.80)	3.14 (3.36)	0.71 (0.63)	2.63 (3.00)	3.13 (3.38)	3.63 (3.88)
Control							
Item 2, 10, 15, 19, 25, 30	6	.86 (.84)	3.32 (3.36)	0.87 (0.92)	2.83 (2.83)	3.50 (3.50)	4.00 (4.00)
Managerial support							
Item 8, 23, 29, 33, 35	5	.87 (.80)	3.48 (3.19)	0.85 (0.93)	3.00 (2.60)	3.60 (3.20)	4.00 (3.80)
Peer support							
Item 7, 24, 27, 31	4	.79 (.85)	3.80 (3.75)	0.68 (0.86)	3.50 (3.25)	4.00 (3.75)	4.25 (4.50)
Relations							
Item 5, 14, 21, 34	4	.77 (.75)	3.95 (4.01)	0.73 (0.68)	3.50 (3.50)	4.00 (4.00)	4.50 (4.50)
Role							
Item 1, 4, 11, 13, 17	5	.85 (.78)	4.14 (4.28)	0.66 (0.65)	3.80 (4.00)	4.20 (4.40)	4.60 (4.80)
Change							
Item 26, 28, 32	3	.78 (.72)	2.93 (3.16)	0.90 (0.97)	2.33 (2.33)	3.00 (3.33)	3.67 (4.00)

Notes. ( ) = results obtained in the IT dataset. Lower scores indicate higher risk of work-related stress.

Table 3. Pearson correlations (IT dataset: n = 1,298) of HSE dimensions with age, occupational, and well-being variables

	1	2	3	4	5	6	7	8	9	10	11	12
1. Demands	–											
2. Control	.26**	–										
3. Supervisor support	.31**	.56**	–									
4. Peer support	.32**	.46**	.54**	–								
5. Relationships	.48**	.40**	.50**	.57**	–							
6. Role	.31**	.40**	.52**	.36**	.37**	–						
7. Change	.34**	.50**	.74**	.49**	.45**	.56**	–					
8. Age	–.01	–.01	.06*	–.04	.00	.12**	.06*	–				
9. Shift work (0 = *no*, 1 = *yes*)	.03	–.41**	–.19**	–.14**	–.16**	–.03	–.15**	–.02	–			
10. Night work (0 = *no*, 1 = *yes*)	–.01	–.29**	–.14**	–.11**	–.11**	–.04	–.12**	–.02	.63**	–		
11. Job-related well-being (1 = *very bad*, 10 = *very well*)	.40**	.45**	.64**	.51**	.57**	.48**	.54**	.06*	–.11**	–.09**	–	
12. Use of medications for anxiety, depressed mood, insomnia, or pain (1 = *never*, 5 = *always*)	–.20**	–.25**	–.23**	–.23**	–.27**	–.17**	–.20**	.13**	.05	.08**	–.27**	–
13. Accidents at work this year (0 = *no*, 1 = *yes*)	–.01	–11**	–.04	–.06*	–.07*	–.03	–.06*	.02	.06*	.04	–.03	.05

Notes. $*p < .05$, $**p < .01$.

loading of .70, while correlations between latent factors varied between .39 (Demands/Control and Peer support/Control) and .76 (Managerial support/Change), with a mean correlation of .53. On the whole these results supported our first and second hypotheses, that a 7-factor solution would fit the UK dataset and that the same solution would be equivalent in the IT Italian dataset.

Additional Psychometric Analyses

Table 2 reports the additional descriptive statistics of the seven subscales of the Indicator Tool, divided up into UK and IT datasets. The internal consistency of all the scales was adequate, varying from .77 (Relation) to .87 (Managerial support) in the UK sample, and from .72 (Change) to .85 (Peer support) in the IT sample. The mean values obtained on the scales fell in all the cases in the positive range, and in both the datasets Change had the lowest mean value while Role had the highest. The three quartiles are also reported in Table 2 for the UK data (IT in brackets). Note, however, that comparisons between UK and IT observed scores may be misleading, since these comparisons would require strict factorial equivalence (metric, scalar, and uniqueness equivalence), which was not supported in the present study.

Additional analyses focused on the Italian dataset (n = 1298) in order to evaluate Hypothesis 3. Table 3 reports Pearson's correlations between the seven Indicator Tool scales and some "criterion" variables included in the Italian study. As expected, the Indicator Tool scales correlated positively with a measure of job-related well-being: Supervisor support showed the strongest correlation with this measure ($r = .64$), while Demands had the weakest ($r = .40$).

Furthermore, the better the psychosocial working conditions as measured by the seven Indicator Tool scales, the lower the frequency of use of medications for anxiety, depressed mood, insomnia, or pain. These results were consistent with our Hypothesis 3. Of interest were also the correlations between the Indicator Tool scales and the variables accounting for shiftwork and nightwork. In all circumstances these correlations were negative, suggesting that shiftworkers and nightworkers tended to experience worse psychosocial conditions at work. This was particularly the case for Control, which showed a moderate and strongly negative correlation with shift work and night work, respectively. Also, there was a negative correlation between the occurrence of accidents at work in the last year and a number of Indicator Tool scales, particularly Control ($r = -.11$).

Using MANOVA, we then examined differences in psychosocial conditions as measured by the Indicator Tool scales and a number of sociodemographic and occupational variables. MANOVA (Wilk's criterion) revealed an overall significant difference in psychosocial conditions at work according to sex and age: $F(7, 1257) = 9.52$, $p < .001$ for sex; $F(28, 4616{,}53) = 3{,}26$, $p < .001$ for age. As far as sex is concerned, follow-up F-tests revealed significant differences in all the cases but Role, with females consistently reporting worse psychosocial conditions than males. As for Age, significant differences emerged in all the cases but Control. Mean values for the different age classes were consistent with a U-shaped relationship (i.e., better conditions for the extreme classes) in the case of Demands, Supervisor support, Peer support, Relationship, and Change, while a linear relationship emerged between age and Role, with older workers perceiving a more clear role.

Differences also emerged in psychosocial conditions according to type of job contract and position, $F(7, 1279) = 12.21$, $p < .001$, and $F(7, 1210) = 18.29$, $p < .001$, respectively. Workers with a permanent job contract reported better conditions for Control and Role, but worse for Demands. As far as job position is concerned, employees at the managerial level reported better conditions than other employees in all the cases but Demands, where their mean values suggested that they had to deal with a higher level of job demands.

Discussion

The present study had the main aim of advancing the UK and Italian validation of the HSE Stress Indicator Tool, a self-report questionnaire that can be used in the control cycle of work-related stress (Cox, Karanika, Griffiths, & Houdmont, 2000), specifically for the characterization of the psychosocial environment of work organizations. In line with what we hypothesized (Hypothesis 1), we replicated the results of Edwards et al. (2008), namely, that a 7-factor solution would best fit the UK data of the Indicator Tool, focusing on private sector organizations.

Second, we found that the same 7-factor solution was invariant on Italian data. This implies that Italian workers conceptualize the psychosocial work environment as measured by the Indicator Tool in the same way as UK workers do (Cheung & Rensvold, 2002). Furthermore, we found evidence for metric equivalence, factor variance equivalence, and factor covariance equivalence, which means that, overall across the two samples, the strength of the relationship between items and their underlying constructs (i.e., the way in which the constructs are manifested) is the same: Participants use equivalent ranges of the seven constructs continua, and the relationships between constructs are equivalent, respectively. However, we failed to find evidence for scalar equivalence, which would have permitted meaningful comparisons (at least) at the level of construct means between countries. This may be a consequence of the heterogeneity of the sample characteristics. We further discuss this issue below when we address limitations of the present study.

The internal consistencies of the Italian version of the seven Indicator Tool's subscales were adequate, in other words, they were in all cases above the suggested minimum threshold of .70 (Nunnaly & Bernstein, 1994). Correlations between the Indicator Tool subscales and individual job-related and context-free well-being were significant and in the expected direction, indicating that the worse the psychosocial environment of work, the worse the reported well-being. The strongest relationship with job-related well-being emerged for Supervisor support, which parallels what has been recently found with UK data by Kerr et al. (2009). Note also that Kerr et al. found significant relationships between the Indicator Tool subscales and errors/near misses at work, which is consistent with the negative pattern of relationships that emerged in the present study between the Indicator Tool subscales (particularly Control) and accidents at work.

The differences in the Indicator Tool subscales according to sociodemographic and occupational variables generally tended to be consistent with previous findings in this area. Women report worse psychosocial working conditions compared to men in all the cases, bar Change. Although previous research does not show such a clear general disadvantage for women, women do consistently report a lower level of job control (see for a review Warr, 2007). Furthermore, more often than men women tend to report harassment at work, particularly sexual harassment (European Agency for Safety and Health at Work, 2009) – though there may be important third variables that could explain this relationship (Einarsen, Hoel, Zapf, & Cooper, 2010). Thus, the difference between men and women on the Relationship subscale may also be expected.

Concerning age, Folkman, Lazarus, Pimley, and Novacek (1987) report greater family, financial, and occupational pressures for workers in their middle years (40–50), which agrees with findings of a J or U shaped relationship frequently reported between general and job-related well-being, on the one hand, and age on the other (e.g., Clark, 1996; Teachman, 2006). We found a U-shaped relationship between a number of psychosocial conditions as measured by the Indicator Tool subscales (i.e., Demands, Supervisor support, Peer support, Relationship and Change) and age may be at least in part justified.

Finally, the Italian version of the HSE Indicator Tool detected widely known differences in the psychosocial work environment between managers and other workers, for example, the more active job profile of managers (Karasek & Theorell, 1990), with greater demand but also greater control. Furthermore, there is also some evidence that workers with a permanent job contract report higher control than their temporary counterparts, albeit also higher demands (Aronsson, Gustafsson, & Dallner, 2002; Gallie, White, Cheng, & Tomlinson, 1998; Parker, Lajunen, & Summala, 2002), a finding that also emerged in the present study.

An important limitation of the present study is that we did not obtain two large, well-matched crosscultural samples (e.g., same occupation, job position, sex, and age distributions) to test for measurement equivalence between the Italian and UK versions of the Indicator Tool, which would have been the best condition for this kind of test. We cannot exclude that the lack of scalar equivalence (and perhaps the uniqueness equivalence) we found is a consequence of the imperfect comparability between the samples. Since it could be important to compare data of the Indicator Tool across countries, we believe that future research should try to obtain well-matched samples first – and then re-examine the measurement equivalence of the two versions of the tool under these conditions. As a second limitation, we had single item measures of well-being available, by which we tested the criterion-related validity of the Italian version of the Indicator Tool. Single-item measures, however, may have a low reliability (Nunnally & Bernstein, 1994). Thus, to obtain a more reliable assessment of the magnitude of relationships between the Indicator Tool subscales and criterion-related constructs, it is necessary to use multi-item measures and perhaps also objective measures of stress-related phenomena in future research.

To conclude, we provided further evidence for the robustness of the 7-factor solution of the HSE Stress Indicator Tool on UK data. Furthermore, we provided evidence that the Italian version of the Tool has psychometric properties comparable to those of the original UK version, with adequate factorial, construct, and criterion-related validity and reliability. We hope these encouraging results will motivate other researchers to validate their country-specific versions of the Indicator Tool and to share their data. The availability of common tools with characteristics of validity, reliability, comprehensiveness, and brevity, which are all revealed by the Indicator Tool, would be an important step toward developing a unified surveillance system (Dollard, Skinner, Tuckey, & Bailey, 2007) on work-related stress at the European level.

References

Aronsson, G., Gustafsson, K., & Dallner, M. (2002). Work environment and health in different types of temporary jobs. *European Journal of Work and Organizational Psychology, 11*, 151–175.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Beverly Hills, CA: Sage.

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.

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25*(1), 1–27.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Clark, A. E. (1996). Job satisfaction in Britain. *British Journal of Industrial Relations, 34,* 189–217.

Cousins, R., Mackay, C. J., Clarke, S. D., Kelly, C., Kelly, P. J., & McCaig, R. H. (2004). "Management Standards" and work-related stress in the UK: Practical development. *Work and Stress, 18*, 113–136.

Cox, T., Griffiths, A., & Rial González, E. (2000). *Research on work-related stress*. Luxembourg: Office for Official Publications of the European Communities.

Cox, T., Karanika, M., Griffiths, A., & Houdmont, J. (2007). Evaluating organizational-level work stress interventions: Beyond traditional methods. *Work and Stress, 21*, 348–362.

Dollard, M. F. Skinner, N., Tuckey, M. R., & Bailey, T. (2007). National surveillance of psychosocial risk factors in the workplace: An international overview. *Work and Stress, 21*, 1–29.

Edwards, J. A., Webster, S., Van Laar, D., & Easton, S. (2008). Psychometric analysis of the UK Health and Safety Executive's Management Standards work-related stress Indicator Tool. *Work and Stress, 22*, 96–107.

Einarsen, S., Hoel, H., Zapf, D., & Cooper, C. L. (2010). *Bullying and harassment in the workplace. Developments in theory, research, and practice*. New York: CRC Press/Taylor & Francis.

European Agency for Safety and Health at Work. (2009). *Workforce diversity and risk assessment: Ensuring everyone is covered*. Luxembourg: Office for Official Publications of the European Communities.

European Foundation for the Improvement of Living and Working Conditions. (2010). *Work-related stress*. Retrieved from http://www.eurofound.europa.eu/docs/ewco/tn1004059s/tn1004059s.pdf

Folkman, S., Lazarus, R. S., Pimley, S., & Novacek, J. (1987).

Age differences in stress and coping processes. *Psychology and Aging, 2*, 171–184.

Gallie, D., White, M., Cheng, Y., & Tomlinson, M. (1998) *Restructuring the employment relationship*. Oxford, UK: Clarendon.

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), 1–55.

Karasek, R., & Theorell, T. (1990). *Healthy work: Stress, productivity, and the reconstruction of working life*. New York: Basic Books.

Kerr, R., McHugh, M., & McCrory, M. (2009). HSE Management Standards and stress-related work outcomes. *Occupational Medicine, 59*, 574–579.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

LaMontagne, A. D., Keegel, T., Louie, A. M., Ostry, A., & Landsbergis, P. A. (2007). A systematic review of the job stress intervention evaluation literature: 1990–2005. *International Journal of Occupational & Environmental Health, 13*, 268–280

Mackay, C. J., Cousins, R., Kelly, P. J., Lee, S., & McCaig, R. H. (2004). Management standards and work-related stress in the UK: Policy background and science. *Work and Stress, 18*, 91–112.

Malchaire, J., Pierre, A., D'Horre, W., & Stordeur, S. (2008). *The Sobane Strategy applied to the management of psychosocial aspects.* Retrieved from http://www.deparisnet.be/PSY/Eng/Sobane_guide_psychosocial_aspects.pdf

Marsh, H. W., Tracey, D. K., & Craven, R. G. (2006). Multidimensional self-concept structure for preadolescents with mild intellectual disabilities: A hybrid multigroup- mimic approach to factorial invariance and latent mean differences. *Educational and Psychological Measurement, 66*, 795–818.

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.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Parker, D., Lajunen, T., & Summala, H. (2002). Anger and aggression among drivers in three European countries. *Accident Analysis and Prevention, 34*, 229–235.

Satzer, R. (2009). *Stress – Mind – Health: The START procedure for the risk assessment and risk management of work-related stress*. Düsseldorf: Hans-Böckler-Stiftung. Retrieved from http://www.boeckler.de/pdf/p_arbp_174.pdf

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

Steenkamp, J. B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in crossnational consumer research. *Journal of Consumer Research*, *25*, 78–90.

Teachman, B. A. (2006). Aging and negative affect: The rise and fall and rise of anxiety and depression symptoms. *Psychology and Aging*, *21*(1), 201–207.

Toderi, S., Broccoli, M., Sarchielli, G., & Mancini, G. (in press). I management standards per la gestione dei rischi psicosociali e prima validazione italiana dell'Indicator Tool [The management standards for psychosocial risk management and a first validation of the Italian Indicator Tool]. *Risorsa Uomo*.

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

Warr, P. (2007). *Work, happiness, and unhappiness*. Mahwah, NJ: Erlbaum.

Published online: March 15, 2012

Stefano Toderi

Department of the Education Sciences
University of Bologna
Via Filippo Re, 6
40126 Bologna (BO)
Italy
Tel. +39 051 209-1466
Fax +39 0547 338532
E-mail stefano.toderi@unibo.it

