Psychometric analysis of the German version of the management standards indicator tool (MSIT-D)

Ekaterina Uglanova
Department of Work and Organisational Psychology, FernUniversität in Hagen, Hagen, Germany
Rosanna Cousins
Department of Psychology, Liverpool Hope University, Liverpool, UK, and
Jan Dettmers
Department of Work and Organisational Psychology, FernUniversität in Hagen, Hagen, Germany

Abstract

Purpose – This study aims to develop a reliable and valid German/Deutsch version of the management standards indicator tool (MSIT-D) to broaden the pool of instruments available to practitioners and to support international collaborations regarding this workplace management issue.

Design/methodology/approach – The MSIT-D was translated from English to German, then its psychometric properties examined using data from British employees (n = 321) and German employees (n = 358). Confirmatory factor analyses (CFAs) were used to evaluate the internal structure and measurement invariance, and Cronbach's alpha was used to assess internal consistency. Comparisons were made with the German language risk assessment tool Fragebogen zur Gefährdungsbeurteilung psychischer Belastungen (FGBU) to examine concurrent and incremental validity. Criterion validity was checked using established measures of work-related health.

Findings – The MSIT-D has an equivalent seven-factor structure (demands, control, managerial support, peer support, relationships, role and change) as the original; the analyses confirmed configural and metric measurement invariance with the original scale. The internal consistency of the scales ranged from 0.82 to 0.91. Regarding criterion validity, the MSIT-D was positively correlated with emotional exhaustion and psychosomatic complaints and negatively correlated with work engagement and workability. The analyses yielded meaningful correlations between the MSIT-D dimensions and the FGBU.

Originality/value – This is the first study to develop a German version of the MSIT and confirm metric measurement invariance. This will allow a comparison of MSIT scores with related constructs between German- and English-speaking samples. As a reliable and valid instrument for assessing work-related stressors, the outcome of this study presents opportunities for developing a unified surveillance system for work-related stress at the European level.

Keywords Work-related stress, Risk assessment, Validity, Reliability, Management standards, Indicator tool, Stress

Paper type Research paper

Introduction

According to the Council Directive 89/391/EEC of June 12, 1989, employers in the European Union are obliged to implement measures to improve the safety and health of employees at work. This requires workplace risk assessments to include psychosocial hazards. Numerous instruments – such as the Instrument zur stressbezogenen Arbeitsanalyse (ISTA) (Semmer *et al.*, 1999), the Kurzfragebogen zur Arbeitsanalyse (KFZA) (Prümper *et al.*, 1995), the Copenhagen Psychosocial Questionnaire (COPSOQ) (Nübling *et al.*, 2005), the Fragebogen zur Gefährdungsbeurteilung psychischer Belastungen (FGBU) (Dettmers and Krause, 2020) and the management standards indicator tool (MSIT) (Cousins *et al.*, 2004) – have been developed and established in psychological research and practice for psychosocial risk assessment. They are based on different stress models and target different stressors. An important practical criterion for the choice of the analysis instrument is conformity with national health regulations and recommendations (Dettmers and Krause, 2020). However, to compare psychosocial risk factors across countries, different language versions of the same analysis instrument are required, and similarly, it is important to use a tool with sound psychometric properties across the psychosocial work conditions in different countries. A succinct discussion of the various measures that have been used to measure work-related stress has been provided by Cassar *et al.* (2020). They point out that the MSIT has become the most cited of these measures, based on its practical ability to cover the key psychosocial hazards, and its excellent psychometric properties. The latter point has been recognised in many other studies (e.g. Bevan *et al.*, 2010; Edwards *et al.*, 2008).

The MSIT is a 7-factor, 35-item measure that was developed specifically to support organisations conduct a risk assessment process that informed them of whether their working conditions were in line with the UK Health and Safety Executive's six management standards: demands, control, support – managerial and colleague, relationship, role and change (Cousins *et al.*, 2004; Mackay *et al.*, 2004). The management standards include "states-to-be-achieved" that indicate working conditions that reflect a low risk of work-related stress for employees and associated harm to health and business costs (Edwards and Webster, 2012; Mackay *et al.*, 2004). Thus, findings from the MSIT provide managers with the information, procedures and tools needed to risk assess hazards in the workplace, as well as to demonstrate good practice in the management of employee health at work using evidence-based methods (Cousins *et al.*, 2004; Mackay *et al.*, 2004).

There is a plethora of research on the potential consequences of psychosocial hazards at work that has used the MSIT as a reliable and valid measure. This has repeatedly revealed that performance measures, health indicators, such as sickness absence, anxiety, depression and exhaustion, as well as indicators of well-being, such as job satisfaction, are associated with psychosocial stressors. Health outcomes have probably been studied most with regard to their association with psychosocial hazards at work. For example, the following MSIT scales significantly predict depression: change (Hackett *et al.*, 2009; Menghini *et al.*, 2022), demands, control, peer support (Menghini *et al.*, 2022) and managerial support (Kerr *et al.*, 2009). Anxiety is significantly predicted by demands (Kerr *et al.*, 2009; Menghini *et al.*, 2022), change (Kerr *et al.*, 2009; Menghini *et al.*, 2022), peer support (Menghini *et al.*, 2022) and relationships and role (Kerr *et al.*, 2009). The scores of psychological distress (General Health Questionnaire, GHQ) correlate with demands, change, relationships and peer support (Bridger *et al.*, 2016). Two Italian studies (Guidi *et al.*, 2012; Magnavita, 2012) presented evidence that all dimensions of the Italian version of the MSIT were associated with psychiatric caseness as determined by GHQ >2. Bevan *et al.* (2010) reported that exhaustion (an outcome indicative of cumulative work-related stress (Gaines and Jermier, 1983)) is strongly associated with role, whilst Ravalier *et al.* (2013) found that scores on the demands and controls scales predicted level of exhaustion. Similarly, Bridger *et al.* (2016) found need for recovery was correlated with the dimensions of demands and relationships and that

workability was positively associated with role in their sample of military personnel. For
Guidi's sample of Italian bank employees, workability was associated with all seven MSIT
scales (Guidi *et al.*, 2012).

Managing psychosocial hazards has also been associated with increased performance at
work. For example, based on previous research, the hypotheses of Toderi and Balducci (2015)
that demands, peer support and role would significantly predict job performance in a sample
of 326 employees were supported. Job satisfaction has been associated with all seven MSIT
dimensions, with managerial support being the strongest predictor (Kerr *et al.*, 2009).
Similarly, satisfaction with the quality of working life is predicted by the scores on demands,
peer support, managerial support and relationship scales (Bridger *et al.*, 2016).

Altogether, this small sample of evidence from the literature shows the potential of the
MSIT to work in a wide variety of work settings with established performance and health
outcome measures to understand working conditions and to guide workplace interventions
where necessary. On the basis of the above discussion, the aim of the current project was to
develop and evaluate a German translation of the MSIT. The MSIT possesses suitable
psychometric properties, and importantly, findings from a staff survey that included the
MSIT provide organisations with comprehensive instructional support for benchmarking
health and safety standards. In Germany, since 2013, the Occupational Safety and Health Act
(*Arbeitsschutzgesetz*) has explicitly appealed to companies to conduct mental stress
assessments (GDA, 2022). The German Work Protection Strategy (GDA) designed a
special work programme aimed at developing recommendations for assessing the risk of
mental stress and providing informational and practical support for companies in
implementing psychosocial hazards' risk assessment. From a practical point of view,
companies are interested in valid and reliable measurement instruments that capture
relevant stress factors as economically as possible. In this context, the development of valid
and reliable instruments fosters the professionalisation of risk assessment (Dettmers and
Krause, 2020). Thus, making this reliable and valid instrument available to the German-
speaking labour market will support this pragmatic approach by broadening the pool of
instruments available to practitioners for assessing stress factors that are relevant to the
workplaces they are supporting. Furthermore, adapting the MSIT to a German sample will
help strengthen cross-cultural analyses of processes involved in the development and
management of work-related stress in different countries. The Management Standards
provide a solid background against which similar issues across Europe could be identified
(van Stolk *et al.*, 2014), and a German translation of the MSIT, an MSIT-D, could be an
important step towards developing a unified surveillance system for work-related stress at
the European level. Hence, the current study also includes analyses of the relationships
between MSIT scores and health and well-being indicators.

The study was aimed at (1) testing the internal structure and consistency of the MSIT-D
scales; (2) testing measurement invariance by comparing the MSIT-D with the original British
scale; (3) testing the criterion validity of the MSIT-D by evaluating specific correlations
between MSIT-D scores and scores for criterion-related constructs; (4) exploring concurrent
validity by evaluating correlations of the MSIT-D scales and scales of the Risk Assessment
for Mental Stress at Work Questionnaire (FGBU) and (5) exploring incremental validity by
analysing the additional amount of variance explained by the MSIT-D beyond FGBU in the
response variables.

Methods
Study design and participants
The study used a cross-sectional analytic survey design based on two sub-samples, namely a
German sample and a British one. The British data were used for testing measurement

invariance, whereas validity and internal consistency were investigated using only the German sample. The data for the German sample were collected in January and February 2022. Participants were recruited through social networks and the so-called "survey pool" of the FernUniversität in Hagen. The advertisement sent out contained information about the study (aim, benefits, inclusion criteria, confidentiality of data) as well as a link to the website of the questionnaire. To be eligible for the study, respondents had to: (1) be at least 18 years old, (2) work at least 20 h per week in paid employment and (3) not be self-employed. All job positions were eligible. Participants who were also psychology students could receive the credit hours necessary to complete their study requirements; there was no monetary compensation. 416 potential participants gave informed consent before entering the study. After excluding participants who did not meet the inclusion criteria, the final German sample size was $N = 358$. The average age of this sample was 37.8 ± 11.1 years and 34.34% were male. 58.1% had tertiary/higher education, 31.3% had vocational qualifications and 10.6% had secondary education. 18.7% were in a managerial position. The average service record was 5.8 ± 2.2 years; 78.2% had permanent employment status and 21.8% were on fixed-term or casual contracts.

The British data originated from an investigation of work-related stress in five categories of job roles working in the broad area of higher education in the Merseyside region of England. Similar to the German inclusion criteria for the study, recruitment was based on employment for at least 20 h per week and in post for at least 6 months. The participants filled in the English version of the questionnaire (Cousins *et al.*, 2004; see Table 2 for reliability coefficients). Information about the anonymous survey was provided in advance to allow informed consent to be taken. Of the 321 completed surveys that made up this sample, 104 were from academics (of various grades) who can participate in "hybrid" working, 18 were from senior leaders, 85 were from "at work" administrative staff, 68 were submitted by "at work" support staff and 46 were "homeworker" support staff. The average age was 45.29 ± 9.54 years, 43.5% were male, average service record was 12.39 ± 9.19 years and all had permanent employment status.

Management Standards Indicator Tool (MSIT)
The MSIT was developed by the UK Health and Safety Executive to support employers in undertaking a risk assessment for psychosocial hazards in their workplace that can lead to work-related stress in employees and costs from underperformance and sickness absence (Cousins *et al.*, 2004; Mackay *et al.*, 2004). The MSIT is a 35-item scale with 5-point Likert scale responses according to agreement or frequency. Scores are averaged for each of the seven scales (demands (8 items), control (6 items), managerial support (5 items), peer support (4 items), relationships (4 items), role (5 items) and change (3 items)) that make up the MSIT, so that each has the same score range of 1–5. For the scales control, managerial and peer support, role and change higher scores indicate better working conditions in that stressor area for the scales demands and relationships, vice versa [1]. Benchmarking data are available according to sector and regular risk assessment – as legally required across Europe – allows organisations to see improvement where intervention is necessary.

Translation of the MSIT
First, the English version of the MSIT was translated into German by a German psychologist. Then, a back translation from German into English was performed by a professional translator who had not read the original items. Finally, an expert group of occupational psychologists (employees of the Department of Work and Organisational Psychology of FernUniversität in Hagen) compared the English and the back-translated versions and created a preliminary German version after some corrections for words, meanings and content of each item (see Appendix, Table A1 for the German version).

To analyse the criterion validity of the MSIT, theoretically meaningful and empirically supported correlations with stress outcomes were assessed using three indicators of health and one indicator of well-being. Since the criterion validity was accessed only for the MSIT-D, these data were collected only on the German sample.

Psychosomatic complaints were measured using the 20-item psychosomatic complaints in a nonclinical context scale (Mohr and Müller, 2005). This self-report measure uses a five-point Likert scale (1 = never to 5 = almost daily) to determine the extent to which respondents suffer from various psychosomatic complaints, including headaches, nausea, or backache. In the present study, the internal consistency of the scale was $\alpha = 0.85$.

Workability was measured using the German version of the Work Ability Index (WAI) (Hasselhorn and Freude, 2007). The WAI has 23 items and seven dimensions to capture the physical and mental demands of employees in relation to their work, diagnosed disease, sick leave during the past 12 months, workability prognosis in the next two years and psychological resources. A sample item is "How do you rate your current workability with respect to the physical demands of your work?" The index is derived as the sum score of ratings on each dimension. Scores range from 7–49 and are classified into poor (7–27), moderate (28–36), good (37–43) and excellent (44–49) workability.

Exhaustion was assessed through the exhaustion sub-scale of the German version of the Maslach Burnout Inventory (Büssing and Perrar, 1992). This sub-scale contains seven items (e.g. "I feel emotionally drained from my work") and uses a seven-point Likert scale response format, which ranges from 1 = never to 7 = daily. The internal consistency of the scale was $\alpha = 0.91$.

Work engagement was used as a measure to reflect work-related well-being. We used the German translation of the short version of the Utrecht Work Engagement Scale (UWES-9; Schaufeli *et al.*, 2006). The questionnaire has nine items (e.g. "At my work, I feel bursting with energy"), each rated on a seven-point Likert scale ranging from 1 = never to 7 = always. The internal consistency of the scale was $\alpha = 0.95$.

To access the concurrent validity of the MSIT-D, we used the German-language Risk Assessment for Mental Stress at Work Questionnaire (FGBU) (Dettmers and Krause, 2020). The questionnaire was developed against the background of recommendations of the Joint German Occupational Health and Safety Strategy to comprehensively record psychological job stressors that should be accounted for in occupational risk assessments. The FGBU has 19 scales and an index of 10 physical stressors. This results in a total of 67 items and one comment field in which the participants could indicate further stressors and add other comments. Each item is answered on a four-point Likert scale ranging from 1 = does not apply to 4 = applies. The FGBU was also used to evaluate the incremental validity of the MSIT-D. These data were collected only from the German sample.

Analytical strategy
Internal structure
Structural equation modelling with maximum likelihood (ML) estimation with robust standard errors (MLR) was applied to the German data to examine the internal structure of the MSIT-D in Mplus 6.0. The fit of the theoretically assumed seven-factor structure – which has been confirmed in numerous studies (e.g. Cousins *et al.*, 2004; Edwards *et al.*, 2008) – was compared with the fit of a one-factor model. To evaluate model fit, we used the chi-square goodness-of-fit statistic, the comparative fit index (CFI), the root mean square error of approximation (RMSEA) and standardised root mean residual (SRMR), using the cut-off criteria proposed by Hu and Bentler (1999) (RMSEA $\leq$0.06, CFI $\geq$0.95 and SRMR $\leq$0.08). In addition to applying the goodness of fit indices, and in case the specified model did not fit the

data well, we modelled and interpreted possible sources of misfit. Only theoretically feasible modifications were allowed.

The analysis comprised several steps. First, the multivariate normality requirements for the use of ML estimator were tested by means of Mardia's test. The results rejected the assumption of multivariate normality: multivariate skewness = 184.56, $p < 0.001$, multivariate kurtosis = 1421.20, $p < 0.001$. As these results do not allow assuming a distribution that is close to normal, we followed the suggestion of Kline (2015) and employed the Satorra–Bentler rescaled chi-square statistic (Satorra and Bentler, 1994), which compensates for thr non-normality of variables' distribution. Second, the first-order confirmatory factor analysis (CFA) (Model 1) was applied to test the seven-factor structure of the 35-item MSIT. Next, a first-order model, in which the load of all 35 items on a single factor (Model 2) was estimated and compared with the theoretically assumed seven-factor model. Finally, a second-order CFA (Model 3) was conducted to establish whether the instrument contains, besides seven first-order factors, a higher-order factor component – general work-related stress (Edwards *et al.*, 2008).

Measurement invariance

To test configural, metric and scalar invariances, we followed the stepwise approach of Rudnev *et al.* (2018). The approach is based on fitting multi-group CFA models to the data using different sets of specific constraints that correspond to the specific level of measurement invariance. To assess the differences in goodness-of-fit between the configural, factor loading and intercept models, we followed the recommendations of Chen (2007). According to this approach, to ensure measurement invariance, the values of ΔRMSEA should not exceed 0.015, the values of ΔCFI should not exceed 0.01 and the values of ΔSRMR should not exceed 0.03 (0.01 to compare the models with constrained factor loadings and with constrained intercepts). A chi-square difference test was also used.

Cronbach's alpha statistic was used to calculate the reliability of each sub-scale. To test the concurrent validity of the MSIT-D, product-moment correlations with the scales of the FGBU were assessed. To test the criterion validity, again, the product-moment correlations between the MSIT scales and outcomes of psychosocial hazards at work were calculated.

To test the incremental validity, stepwise hierarchical regression was used. The scales of the FGBU were entered as a first block of predictors, and the scales of the MSIT were entered as a second block of predictors. Subsequently, the explained variance in the response variables (stress outcomes) was analysed.

Results

Internal structure

The results of the model testing are shown in Table 1. The examination of fit indices of the theoretically assumed seven-factor model demonstrated a misfit. Therefore, following the suggestions of Bowen and Guo (2011), we analysed modification indices and the following parameter estimates: (1) inter-factor correlations, (2) R^2 for observed variables and (3) standardised residual covariances to determine the fit between sample covariances and expected covariances. None of the inter-factor correlations exceeded the recommended cut-off value of 0.8 (Bowen and Guo, 2011).

The R^2 for two variables – Control 6 ("My working time can be flexible") and Relationships 3 ("I am subject to bullying at work") – were under the recommended cut-off value of 0.40 (0.32 and 0.32, respectively). Several pairs of variable covariances were not reproduced well. The most affected variables in this sense were demands 3, demands 1, controls 2, relationship 4, role 2, role 4 and role 5. Although certain items contributed to the poor fit, deleting these

	χ^2/df/P	$\Delta\chi^2$/df/ p	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR	ΔSRMR
7-factor model	1358.76/ 539/<0.01		0.88		0.065		0.067	
7-factor model (re-specified)	1179.17/ 534/<0.01		0.907		0.058		0.067	
1-factor model	4543.42/ 560/<0.01		0.428		0.141		0.138	
Second-order factor model	1251.98/ 548/<0.01		0.899		0.060		0.076	
Measurement invariance								
Configural invariance	2187.54/ 1070/ <0.01		0.909		0.055		0.066	
Metric invariance	2276.76/ 1098/ <0.01	105.452	0.904	0.005	0.056	0.001	0.073	0.007
Scalar invariance	2552.05/ 1126/ <0.01		0.884	0.02	0.061	0.005	0.078	0.005
Scalar invariance – re-specified	2466.73/ 1124/ <0.01		0.890	0.014	0.059	0.003 (compared with metric)	0.076	0.003

Note(s): Compared with metric invariance
Source(s): Authors' own work

Sub-scale	M	SD	α
Demands	3.34/2.94	0.87/0.88	0.9/0.88
Control	3.57/3.58	0.89/0.87	0.86/0.85
Managerial support	3.26/3.66	0.98/1.01	0.91/0.91
Peer support	3.66/3.93	0.86/0.83	0.89/0.86
Relationships	3.90/3.94	0.78/0.86	0.88/0.77
Role	4.20/4.22	0.68/0.68	0.88/0.81
Change	3.00/3.06	0.99/1.14	0.82/0.88

Note(s): The number before the slash refers to the German sample, the number after the slash – to the British sample
Source(s): Authors' own work

problematic variables resulted in an unacceptably poor fit. The interpretation of modification indices suggested that the highest degree of misfit lay in the error covariance matrix and represented correlated errors of measurement between role items 4 and 5, 1 and 2, control items 1 and 2, demands items 7 and 8 and peer support items 1 and 4. Therefore, the residuals of these pairs of items were allowed to correlate. It was considered that these five pairs of items, each representing the same factor, were similar enough to be correlated in the current analysis. Based on both theoretical and empirical grounds, the 35-item model was re-specified with the five additional error covariances.

After re-specification by allowing correlated residuals, the chi-square test still produced a statistically significant value of 1179.17 ($df = 534$, $p < 0.001$). The CFI (0.91), however, did not reach the recommended cut-off value of 0.95 and could still represent a reasonably good fit

(Kline, 2015; Schumacker and Lomax, 2016). The other two fit statistics – RMSEA and SRMR were also acceptable. Figure 1 shows the 35-item factor loadings for Model 1.

Model 2, in which all 35 items were loaded on a single factor, demonstrated a poor fit. This confirmed that the MSIT is better represented by a seven-factor structure. This result is consistent with Edwards *et al.* (2008).

Model 3, in which a higher-order factor component was included, demonstrated marginally acceptable fit (CFA and RMSEA barely reached the cut-off values and SRMR was acceptable). In order for the model to converge, the factor loadings of the factor demands
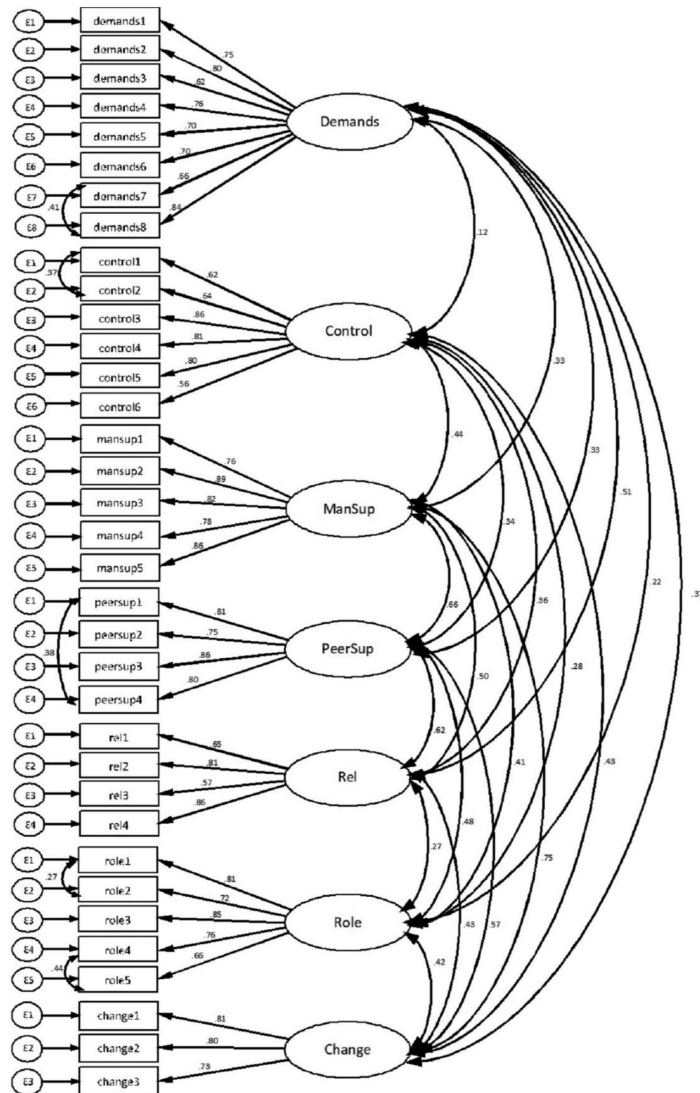

Figure 1.
CFA for the 35 items of
the MSIT-D

Source(s): Authors' own work

were freed following the suggestion of Muthen and Muthen (2007). These findings are consistent with previous work that explored the factor structure of the scale (e.g. Edwards *et al.*, 2008); they suggest that the MSIT has a hierarchical factor structure. Whereas the seven sub-scales test distinct concepts, at the same time, they touch upon aspects of the same underlying concept of work stress. At that point, it is not the 35 items, but the seven sub-scales that would measure the overall stress. The instrument could therefore be used to assess an overall score for psychosocial hazards in their organisation.

Measurement invariance

The results of testing for measurement invariance are presented in Table 1. Identification of a variance-covariance structure of the seven factors was achieved by constraining one factor loading per factor to 1 and fixing one indicator intercept per factor to 0 (always the first item) (Rudnev *et al.*, 2018). The models reached their best fit with these parameters. Whilst selecting the marker indicator, whose loading is fixed to 1, one should choose the most reliable and invariant item, which is conceptually closest to the latent variable underlying the factor. After examining the results of the CFA for the joint sample, the following items were selected as marker indicators: Demands 8 for the factor "Demands", control 3 for the factor "Control", managerial support 5 for the factor "Managerial Support", peer support 1 for the factor "Peer Support", relationships 2 for the factor "Relationships", role 1 for the factor "Role" and change 2 For the factor "Change".

The differences in chi-square between all three models remained significant; however, given the large (>300) sample size, it might be the case that the chi-square test rejects models even when violations are minor. The differences in RMSEA, CFI and SRMR between the configural and metric models were below the cut-off criteria proposed by Chen (2007). The fit of the scalar model was somewhat worse than the fit of the metric model; although the differences in RMSEA (0.005) and SRMR (0.005) were below the cut-off criteria proposed by Chen (2007), the difference in the CFI was above the cut-off criterion (0.02). That is, the CFI was below 0.9, thus demonstrating a poor fit. Examination of modification indices suggested that intercepts of two items (namely, Relationships 1 and Role 4) should be estimated freely in order to improve the model's fit. Indeed, after releasing these intercepts, the fit improved, although the CFI was still somewhat above the desirable cut-off criteria: ΔRMSEA = 0.003, ΔCFI = 0.014, ΔSRMR = 0.003. The overall goodness-of-fit indices RMSEA and SRMR reached an acceptable level, whilst the CFI did not. Thus, at best, only partial scalar invariance could be established in this study.

Reliability estimates. To assess the reliability of the seven dimensions of the instrument, Cronbach's alphas were calculated. These were good for all dimensions (see Table 2 for reliability coefficients). The values of the MSIT-D were similar (for the sub-scales demands, control, managerial support and peer support), somewhat higher (for the sub-scales relationships and role) and somewhat lower (for the scale change) than the values for the MSIT from the British sample.

Relationships with potential consequences of psychosocial hazards at work. Table 3 shows the product-moment correlations between the seven MSIT scales and theoretically meaningful consequences of psychosocial hazards at work – exhaustion, psychosomatic complaints, work ability and engagement. All seven scales were negatively related to exhaustion and psychosomatic complaints. All scales were positively related to work ability, and six, with the exception of demands, were positively related to engagement.

Concurrent validity. Table 4 shows correlations between the scales of the MSIT-D and the FGBU. The correlation analyses yielded plausible and meaningful results. The highest correlations were found between scales with a similar meaning. The scale Demands was strongly negatively correlated with the following scales of the FGBU (in order of descending

Table 3.
Product-moment correlations between seven dimensions of the MSIT and indicators of job-related well-being

	Demands	Control	Managerial support	Peer support	Relationships	Role	Change	Engagement	Exhaustion	Work ability	Psychosomatic complaints
Demands	1.00	0.15**	0.29**	0.28**	0.46**	0.16**	0.30**	0.01	−0.58**	0.25**	−0.29**
Control	0.05	1.00	0.40**	0.31**	0.35**	0.23**	0.38**	0.35**	−0.21**	0.27**	−0.15**
Managerial support	0.22**	0.63**	1.00	0.60**	0.41**	0.39**	0.66**	0.48*	−0.38**	0.46**	−0.12*
Peer support	0.25**	0.60**	0.71**	1.00	0.51**	0.43**	0.49**	0.33**	−0.34**	0.39**	−0.15**
Relationships	0.43**	0.48**	0.60**	0.62**	1.00	0.25**	0.34**	0.13*	−0.49**	0.37**	−0.34**
Role	0.29**	0.47**	0.62**	0.52**	0.47**	1.00	0.38**	0.30**	−0.28**	0.43**	−0.16**
Change	0.22**	0.64**	0.76**	0.61**	0.51**	0.68**	1.00	0.40**	−0.33**	0.36**	−0.17**
Engagement								1.00	−0.25**	0.37**	−0.18**
Exhaustion									1.00	−0.53**	0.57**
Work ability										1.00	−0.48**
Psychosomatic complaints											1.00

Note(s): Correlations for the German sample are presented in the above the diagonal, correlations for the British sample – below

*$p < 0.05$, **$p < 0.01$

Source(s): Authors' own work

	Demands	Control	Managerial support	Peer support	Relation-ships	Role	Change
Completeness_ work task	−0.226**	0.339**	0.207**	0.193**	0.086	0.268**	0.192**
Autonomy	0.065	0.711**	0.316**	0.236**	0.269**	0.151**	0.214**
Variability	−0.295**	0.372**	0.307**	0.188**	0.089	0.194**	0.200**
Lack of information	−0.444**	0.197**	−0.161**	−0.163**	−0.183**	−0.173**	−0.164**
Information overload	−0.546**	0.126*	−0.133**	−0.108*	−0.182**	−0.228**	−0.153**
Role ambiguity	−0.441**	−0.133*	−0.393**	−0.349**	−0.384**	−0.446**	−0.381**
Qualification problems	−0.469**	0.034	−0.143**	−0.138**	−0.266**	−0.274**	−0.175**
Underutilisation of skills	0.194**	−0.188**	−0.246**	−0.120*	−0.105*	−0.085	−0.176**
Social and emotional stress	−0.423**	−0.248**	−0.229**	−0.177**	−0.541**	−0.061	−0.204**
Emotional dissonance	−0.375**	−0.305**	−0.272**	−0.217**	−0.446**	−0.070	−0.276**
Work time design	−0.320**	−0.282**	−0.210**	−0.072	−0.300**	−0.030	−0.184**
Overtime	−0.689**	−0.093	−0.195**	−0.193**	−0.329**	−0.110*	−0.176**
Work intensity	−0.810**	−0.114*	−0.209**	−0.201**	−0.359**	−0.110	−0.212**
Work interruptions	−0.527**	0.090	−0.143**	−0.087	−0.229**	−0.017	−0.183**
Lack of communication	−0.222**	−0.103*	−0.222**	−0.337**	−0.265**	−0.263**	−0.230**
Social support_ colleague	0.261**	0.312**	0.589**	0.780**	0.458**	0.363**	0.437**
Social stressors_ colleague	−0.391**	−0.284**	−0.430**	−0.506**	−0.772**	−0.261**	−0.323**
Social support_ supervisor	0.282**	0.380**	0.843**	0.507**	0.425**	0.394**	0.619**
Feedback and recognition	0.175**	0.381**	0.814**	0.507**	0.368**	0.311**	0.590**
Physical stressors	−0.371**	−0.406**	−0.200**	−0.182**	−0.411**	−0.032	−0.183**

Note(s): *$p < 0.05$ and **$p < 0.01$
Source(s): Authors' own work

correlations, $r > 0.5$): work intensity ($r = −0.81$), overtime ($r = −0.69$), work interruptions ($r = −0.53$), information overload ($r = −0.55$). The scale Control was strongly positively correlated with the FGBU scale Autonomy ($r = 0.71$). Managerial support was strongly positively correlated with support from supervisor ($r = 0.84$), feedback and recognition ($r = 0.81$) and support from colleagues ($r = 0.59$). Peer support was strongly positively correlated with support from colleagues ($r = 0.78$), feedback and recognition ($r = 0.51$) and support from supervisor ($r = 0.51$) and negatively correlated with social stressors from colleagues ($r = - 0.51$). The scale Relationships was strongly negatively correlated with the FGBU scales Social Stressors from Colleagues ($r = −0.77$) and Social and Emotional Stress ($r = −0.54$). The scale Role was negatively correlated with Role Ambiguity ($r = −0.45$). Finally, change was strongly positively correlated with support from supervisor ($r = 0.62$) and feedback and recognition ($r = 0.59$).

Incremental validity. Table 5 presents the results of the stepwise hierarchical regression analysis with the scales of the FGBU (the 1st block of predictors) and the MSIT-D (the 2nd block of predictors). The MSIT-D scales explained significantly more variance (ΔR^2) in work engagement (3%) and in exhaustion (4%), and the additional variance explained in work ability (3%) and in psychosomatic complaints (1%) was not significant.

Discussion

The aim of this study was to validate the German version of the MSIT on a sample of German employees. In doing so, we conducted a series of confirmatory factor analyses to evaluate the internal structure of the questionnaire and investigated theoretically meaningful relationships with potential consequences of psychosocial hazards at work to evaluate criterion validity. Through the cross-cultural comparison, we established configural and metric measurement invariance. Furthermore, we compared the MSIT-D with another measure of psychosocial hazards at work, the FGBU, to examine concurrent and incremental validity. In addition, we examined internal consistency to evaluate the reliability of the MSIT-D. In sum, our results indicate acceptable psychometric properties of the German MSIT. A series of CFAs revealed that the hypothesised seven-factor model fit the data better than the one-factor model, suggesting that psychosocial hazards should be described in terms of distinct dimensions. The second-order model did not demonstrate a very good fit but was still acceptable.

Tests of measurement invariance using the German sample and the British sample indicated that the MSIT exhibits a certain degree of invariance in measurement of psychosocial hazards at work across German- and English-speaking samples. In this study, configural and metric invariance could be established; however, only partial scalar (intercept) invariance could be established. This means that the number of factors and the factor loading are equal across groups; however, intercepts are probably not equal across groups. Factor loading equivalence is needed to compare correlations of the MSIT with constructs from its nomological network, whereas intercept equivalence is needed to compare latent means of the scales (Sass, 2011). Our results suggest that correlations with possible predictors and consequences of psychosocial hazards at work can be compared across German- and English-speaking samples. However, latent means and, consequently, the levels of psychosocial hazards at work should be compared with caution. All in all, the MSIT allows for partial comparisons between German- and English-speaking samples and can be applied to some research questions concerning the intercultural measurement of psychosocial hazards at work. Toderi *et al.* (2013) arrive at similar results regarding the scalar invariance in measurement across English and Italian samples and suggest that the result may occur due to the heterogeneity of the sample characteristics. The explanation could also be relevant for the present study. The issue should be addressed further by re-analysing the measurement invariance on highly comparable samples. Researchers and

	Engagement	Exhaustion	Work ability	Psychosomatic complaints
FGBU	0.42	0.46	0.35	0.24
FGBU + MSIT-D	0.46	0.49	0.38	0.25
p	0.010	0.049	0.077	0.348

Table 5.
Hierarchical regression analyses

Note(s): The numbers show the amount of variance in an outcome explained by the scale, *p* refers to the significance level of the additional variance explained by the MSIT-D questionnaire
Source(s): Authors' own work

practitioners should pay special attention to the comparability of the samples when attempting to compare scale means across countries.

The internal consistency of each of the seven sub-scales was sufficient ($0.82 < \alpha < 0.91$), meeting the stringent criterion of 0.7, and most of the sub-scales were comparable with the values obtained from the British sample. As far as the relationship with potential consequences of psychosocial hazards at work is concerned, all scales of the MSIT-D were positively associated with work ability and work engagement (except for demands) and negatively associated with emotional exhaustion and psychosomatic complaints. These results are in line with earlier research (Guidi *et al.*, 2012; Kerr *et al.*, 2009; Menghini *et al.*, 2022; Toderi and Balducci, 2015) and confirm that individuals with high levels of psychosocial hazards at work have worse indicators of health and well-being. The study also established the concurrent validity of the MSIT-D by showing high correlations with similar scales of another instrument, the FGBU. Incremental validity was partly confirmed by showing that the MSIT-D explains a significant amount of additional variance in outcomes exhaustion and engagement beyond the FGBU.

Practical implications

The study made the MSIT-D available to the German-speaking labour market, thus making an important step towards developing a unified surveillance system for work-related stress at the European level. The tool can be applied in comparative studies that aim to identify similar issues (as well as differences) in occupational health across Europe and, thus, contribute to the discussion about standards for assessing psychosocial risks at work.

Limitations

There are some limitations to the study that need to be acknowledged. First, compared to similar research on English-speaking samples (e.g. Edwards and Webster, 2012), our study had a relatively small sample size ($N = 358$). Second, the cross-sectional study design implies that all inferences about causal relationships between psychosocial hazards at work and health and well-being should be made with caution. Finally, common method variance might have affected the results, suggesting that the true associations between variables might be weaker than those observed in this study.

Conclusion

The current study reached its five aims: (1) it found that a seven-factor solution of the indicator tool is equivalent across the German and the British samples; (2) it established (metric) measurement invariance with the original British sample; (3) it confirmed the concurrent validity of the instrument by showing that the seven scales of the MSIT-D are correlated with theoretically close scales of the FGBU questionnaire; (4) it demonstrated the criterion validity of the tool by showing negative associations of psychosocial hazards at work with health and well-being indicators and (5) it confirmed the incremental validity of the MSIT-D by showing that the tool explains additional variance in some outcome variables. The results once again confirm that psychosocial hazards at work are detrimental to health and well-being of employees.

Note

1. For the simplicity of interpretation, individual values for the scales demands and relationships were recoded, so higher scores also indicate better conditions.

References

Bevan, A., Houdmont, J. and Menear, N. (2010), "The management standards indicator tool and the estimation of risk", *Occupational Medicine*, Vol. 60 No. 7, pp. 525-531, doi: 10.1093/occmed/kqq109.

Bowen, N. and Guo, S. (2011), *Structural Equation Modeling*, Oxford University Press, New York, NY.

Bridger, R.S., Dobson, K. and Davison, H. (2016), "Using the HSE stress indicator tool in a military context", *Ergonomics*, Vol. 59 No. 2, pp. 195-206, doi: 10.1080/00140139.2015.1057544.

Büssing, A. and Perrar, K.M. (1992), "Die Messung von Burnout. Untersuchung einer deutschen Fassung des Maslach Burnout Inventory (MBI-D)", *Diagnostica*, Vol. 38 No. 4, pp. 328-353.

Cassar, V., Bezzina, F. and Buttigieg, S. (2020), "Investigating the psychometric properties and assessment capabilities of the short version of the health and safety Executive's management standards indicator tool", *The International Journal of Human Resource Management*, Vol. 31 No. 16, pp. 2115-2140, doi: 10.1080/09585192.2018.1431955.

Chen, F. (2007), "Sensitivity of goodness of fit indexes to lack of measurement invariance", *Structural Equation Modeling*, Vol. 14 No. 3, pp. 464-504, doi: 10.1080/10705510701301834.

Cousins, R., Mackay, C.J., Clarke, S.D., Kelly, C., Kelly, P.J. and McCaig, R.H. (2004), "Management standards' work-related stress in the UK: practical development", *Work and Stress*, Vol. 18 No. 2, pp. 113-136, doi: 10.1080/02678370410001734322.

Dettmers, J. and Krause, A. (2020), "Der Fragebogen zur Gefährdungsbeurteilung psychischer Belastungen (FGBU)", *Zeitschrift für Arbeits-und Organisationspsychologie A&O*, Vol. 64 No. 2, pp. 99-119, doi: 10.1026/0932-4089/a000318.

Edwards, J.A. and Webster, S. (2012), "Psychosocial risk assessment: measurement invariance of the UK health and safety executive's management standards indicator tool across public and private sector organizations", *Work and Stress*, Vol. 26 No. 2, pp. 130-142, doi: 10.1080/02678373.2012.688554.

Edwards, J., Webster, S., Van Laar, D. and Easton, S. (2008), "Psychometric analysis of the UK health and safety executive's management standards work-related stress indicator tool", *Work and Stress*, Vol. 22 No. 2, pp. 96-107, doi: 10.1080/02678370802166599.

Gaines, J. and Jermier, J. (1983), "Emotional exhaustion in a high stress organisation", *Academy of Management Journal*, Vol. 26 No. 4, pp. 567-586.

GDA (2022), "Berücksichtigung psychischer Belastungen in der Gefährdungsbeurteilung: empfehlungen zur Umsetzung in der betrieblichen Praxis", available at: https://www.gda-portal.de/DE/Downloads/pdf/Psychische-Belastung-Gefaehrdungsbeurteilung-4-Auflage.html (accessed 23 October 2023).

Guidi, S., Bagnara, S. and Fichera, G.P. (2012), "The HSE indicator tool, psychological distress and work ability", *Occupational Medicine*, Vol. 62 No. 3, pp. 203-209, doi: 10.1093/occmed/kqs021.

Hackett, A., Palmer, S. and Farrants, J. (2009), "Phase 1 of an investigation into the levels of stress in United Kingdom hospice services", *International Journal of Palliative Nursing*, Vol. 15 No. 2, pp. 66-72, doi: 10.12968/ijpn.2009.15.2.39802.

Hasselhorn, H.M. and Freude, G. (2007), *Der work-ability-Index: ein Leitfaden*, Wirtschaftsverlag NW, Verlag für Neue Wissenschaft GmbH, Bremerhaven.

Hu, L.T. and Bentler, P.M. (1999), "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 6 No. 1, pp. 1-55, doi: 10.1080/10705519909540118.

Kerr, R., McHugh, M. and McCrory, M. (2009), "HSE Management Standards and stress-related work outcomes", *Occupational Medicine*, Vol. 59 No. 8, pp. 574-579, doi: 10.1093/occmed/kqp146.

Kline, R. (2015), *Principles and Practice of Structural Equation Modeling*, 4th ed., Guilford Press, New York, NY.

Mackay, C.J., Cousins, R., Kelly, P.J., Lee, S. and McCaig, R.H. (2004), "'Management standards' and work-related stress in the UK: policy background and science", *Work and Stress*, Vol. 18 No. 2, pp. 91-112, doi: 10.1080/02678370410001727474.

Magnavita, N. (2012), "Validation of the Italian version of the HSE indicator tool", *Occupational Medicine*, Vol. 62 No. 4, pp. 288-294, doi: 10.1093/occmed/kqs025.

Menghini, L., Balducci, C. and Toderi, S. (2022), "Italian adaptation of warr's job-related affective well-being scale: factorial structure and relationships with the HSE management standards indicator tool", *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, Vol. 29 No. 3, pp. 309-325, doi: 10.4473/TPM29.3.3.

Mohr, G. and Müller, A. (2005), "Psychosomatic complaints in a non-clinical context", in Glöckner-Rist, A. (Ed.), *ZUMA-Information System. Digital Handbook of Measures in the Social Sciences*, Zentrum fur Umfragen, Methoden und Analysen, Mannheim, Vol. ZIS 8.00.

Muthén, L.K. and Muthén, B.O. (2007), *Mplus User's Guide*, Muthén & Muthén, Los Angeles, CA.

Nübling, M., Stößel, U., Hasselhorn, H.M., Michaelis, M. and Hofmann, F. (2005), *Methoden zur Erfassung psychischer Belastungen. Erprobung eines Messinstrumentes (COPSOQ)*, Wirtschaftsverlag NW, Verlag für Neue Wissenshaft GmbH, Bremerhaven.

Prümper, J., Hartmannsgruber, K. and Frese, M. (1995), "KFZA. Kurz-Fragebogen zur Arbeitsanalyse", *Zeitschrift für Arbeits-und Organisationspsychologie*, Vol. 39 No. 3, pp. 125-131.

Ravalier, J.M., McVicar, A. and Munn-Giddings, C. (2013), "The management standards indicator tool and evaluation of burnout", *Occupational Medicine*, Vol. 63 No. 2, pp. 145-147, doi: 10.1093/occmed/kqs217.

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P. and Zick, A. (2018), "Testing measurement invariance for a second-order factor: a cross-national test of the alienation scale", *Methods, Data, Analyses: A Journal for Quantitative Methods and Survey Methodology (MDA)*, Vol. 12 No. 1, pp. 47-76, doi: 10.12758/mda.2017.11.

Sass, D. (2011), "Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework", *Journal of Psychoeducational Assessment*, Vol. 29 No. 4, pp. 347-363, doi: 10.1177/0734282911406661.

Satorra, A. and Bentler, P.M. (1994), "Corrections to test statistics and standard errors in covariance structure analysis", in von Eye, A. and Clogg, C.C. (Eds), *Latent Variables Analysis: Applications for Developmental Research*, Sage Publications, Thousand Oaks, CA, pp. 399-419.

Schaufeli, W.B., Bakker, A.B. and Salanova, M. (2006), "The measurement of work engagement with a short questionnaire: a cross-national study", *Educational and Psychological Measurement*, Vol. 66 No. 4, pp. 701-716, doi: 10.1177/0013164405282471.

Schumacker, R.E. and Lomax, R.G. (2016), *A Beginner's Guide to Structural Equation Modeling*, 4 ed., Routledge, New York, NY.

Semmer, N., Zapf, D. and Dunckel, H. (1999), "Instrument zur stressbezogenen Tätigkeitsanalyse (ISTA)", in Dunchel, H. (Ed.), *Handbuch Psychologischer Arbeitsanalyseverfahren*, VdF Hochschulverlag, Zürich, Vol. 14, pp. 179-204.

Toderi, S. and Balducci, C. (2015), "HSE management standards indicator tool and positive work-related outcomes", *International Journal of Workplace Health Management*, Vol. 8 No. 2, pp. 92-108, doi: 10.1108/IJWHM-11-2013-0044.

Toderi, S., Balducci, C., Edwards, J.A., Sarchielli, G., Broccoli, M. and Mancini, G. (2013), "Psychometric properties of the UK and Italian versions of the HSE stress indicator tool", *European Journal of Psychological Assessment*, Vol. 29 No. 1, pp. 72-79, doi: 10.1027/1015-5759/a000122.

van Stolk, C., Staetsky, L., Hassan, E. and Kim, C.W. (2014), *Management of Psychosocial Risks at Work: an Analysis of the Findings of the European Survey of Enterprises on New and Emerging Risks (ESENER): A European Risk Observatory Report*, EU-OSHA.

(The Appendix follows overleaf)

Appendix

Item	Scale	English	German
1	Demands	Different groups at work demand things from me that are hard to combine	Unterschiedliche Stellen bei der Arbeit fordern Dinge von mir, die schwer miteinander zu vereinbaren sind
2		I have unachievable deadlines	Ich habe Terminvorgaben, die unmöglich einzuhalten sind
3		I have to work very intensively	Ich muss sehr intensiv arbeiten
4		I have to neglect some tasks because I have too much to do	Ich muss einige Aufgaben vernachlässigen, weil ich zu viel zu tun habe
5		I am unable to take sufficient breaks	Es ist mir nicht möglich, genügend Pausen einzulegen
6		I am pressured to work long hours	Ich werde unter Druck gesetzt, lange zu arbeiten
7		I have to work very fast	Ich muss sehr schnell arbeiten
8		I have unrealistic time pressures	Ich arbeite unter unrealistischem Zeitdruck
1	Control	I can decide when to take a break	Ich kann selbst entscheiden, wann ich eine Pause mache
2		I have a say in my own work speed	Ich habe Einfluss auf meine Arbeitsgeschwindigkeit
3		I have a choice in deciding how I do my work	Ich kann selbst entscheiden, wie ich meine eigene Arbeit erledige
4		I have a choice in deciding what I do at work	Ich kann selbst entscheiden, was ich bei meiner Arbeit tue
5		I have some say over the way I work	Ich habe ein gewisses Mitspracherecht bei meiner Arbeitsweise
6		My working time can be flexible	Meine Arbeitszeit kann flexibel sein
1	Managerial support	I am given supportive feedback on the work I do	Ich erhalte positives Feedback zu meiner Arbeit
2		I can rely on my line manager to help me out with a work problem	Bei Arbeitsproblemen kann ich mich auf Unterstützung durch meinen Vorgesetzten/meine Vorgesetzte verlassen
3		I can talk to my line manager about something that has upset or annoyed me about work	Mit meinem/meiner Vorgesetzten kann ich über Dinge sprechen, die mich bei der Arbeit geärgert haben
4		I am supported through emotionally demanding work	Bei emotional anspruchsvoller Arbeit werde ich unterstützt
5		My line manager encourages me at work	Mein Vorgesetzter/meine Vorgesetzte ermutigt mich bei der Arbeit
1	Peer support	I get help and support I need from colleagues	Ich bekomme von Kolleginnen und Kollegen die Hilfe und Unterstützung, die ich brauche
2		I receive the respect at work I deserve from my colleagues	Von meinen Kolleginnen und Kollegen bekomme ich bei der Arbeit den Respekt, den ich verdiene
3		My colleagues are willing to listen to my work-related problems	Meine Kolleginnen und Kollegen sind bereit, sich meine arbeitsbezogenen Probleme anzuhören
4		If work gets difficult, my colleagues will help me	Wenn die Arbeit schwierig wird, helfen mir meine Kolleginnen und Kollegen

(continued)

Table A1.
The management standards indicator tool, 35-item version

Item	Scale	English	German
1	Relationships	I am subject to personal harassment in the form of unkind words or behaviour	Ich bin persönlichen Angriffen in Form von unfreundlichen Worten oder Verhaltensweisen ausgesetzt
2		There is friction or anger between colleagues	Es gibt Spannungen oder Streit zwischen Kollegen/Kolleginnen
3		I am subject to bullying at work	Ich werde bei der Arbeit gemobbt
4		Relationships at work are strained	Die Beziehungen bei der Arbeit sind angespannt
1	Role	I am clear what is expected of me at work	Mir ist klar, was bei der Arbeit von mir erwartet wird
2		I know how to go about getting my job done	Ich weiß, was zu tun ist, um meine Arbeit zu erledigen
3		I am clear what my duties and responsibilities are	Meine Pflichten und Verantwortlichkeiten sind mir klar
4		I am clear about the goals and objectives for my department	Ich kenne die Ziele meiner Abteilung
5		I understand how my work fits into the overall aim of the organisation	Ich verstehe, wie meine Arbeit in das Gesamtziel der Organisation passt
1	Change	I have sufficient opportunities to question managers about change at work	Ich habe genügend Möglichkeiten, Führungskräfte zu Veränderungen bei der Arbeit zu befragen
2		Staff are always consulted about change at work	Veränderungen bei der Arbeit werden immer mit den Beschäftigten besprochen
3		When changes are made at work, I am clear how they will work out in practice	Wenn bei der Arbeit Änderungen vorgenommen werden, ist mir klar, wie sie in der Praxis funktionieren werden

Source(s): Authors' own work

Table A1.

Corresponding author
Ekaterina Uglanova can be contacted at: ekaterina.uglanova@fernuni-hagen.de