The Assessment of Psychological Strain in Work Contexts

Concerning the Structural Equivalency of Nine Language Adaptations of the Irritation Scale

Gisela Mohr[1], Andreas Müller[2], Thomas Rigotti[1], Zeynep Aycan[3], and Franziska Tschan[4]

[1]University of Leipzig, Germany, [2]TU Munich, Germany, [3]Koç University, Istanbul, Turkey, [4]Université de Neuchâtel, Switzerland

Abstract. Irritation as defined in this paper is the subjectively perceived emotional and cognitive strain in occupational contexts. The structural equivalence of the Arabic, Dutch, English, French, Italian, Russian, Spanish, Turkish, and Polish adaptations of the German Irritation Scale is examined. The Irritation Scale is recommended for application particularly in occupational contexts. In this field it can be used for evaluating interventions, for research on stress at work, and for individual counseling. Exploratory factor analyses, as well as single group and different multiple-group confirmatory factor analyses, were performed. Structural equivalency in terms of equal factor loadings as well as equal factor covariances across all adaptations can be claimed. However, the single group analyses suggest that the hypothesized factor model should be rejected in the English, Russian, Spanish, and Turkish versions, mainly because of borderline values in the RMSEA index.

Keywords: irritation, cognitive, emotional, measure, strain

Introduction

Technological innovations, as well as the growth of information and service sectors, have led to an overall reduction in physical strain at work. Statistics relating to work-accidents and sick-leaves mirror this physical work strain reduction. As a point of contrast, however, it appears that mental strain has increased over the last years (e.g., Burchell et al., 1999; Statistisches Bundesamt, 1998). The tense labor market and rapidly changing demands have forced organizations to be highly flexible, and have resulted in a type of intensified work characterized by time pressure and social interaction. This shift in work strain is not unique to a single country. Economic globalization has contributed to an amalgamation of nationalities, especially in international organizations (OECD, 2001). Clearly, instruments of measurement able to provide a comparable assessment of mental strain across languages need to be developed. This study will present data from nine language adaptations of the widely used German scale "Irritation." When we refer to different "cultures" in this paper, we are specifically referring to groups that "speak different languages and are governed by different political units" (Brislin, Lonner, & Thorndike, 1973, p. 5).

The Irritation Scale is recommended for application particularly in occupational contexts. In this field it can be used for evaluating interventions, for research on stress at work, and for individual counseling (Mohr, Müller, & Rigotti, 2005).

The construct "irritation" was first introduced in Germany by Mohr (1986). It deals with the experience of uncertainty and corresponding reactions. In psychological terms, uncertainty may arise when an individual experiences a discrepancy between a given situation and an important personal goal. The concept of goal orientation, in a general sense, can be seen as a fundamental psychological concept. Cropanzano, James, and Citeria (1993) describe personality as a system of interrelated, hierarchically organized goals that determine our daily experiences and actions. A perceived obstacle to an important goal can be regarded as a potential threat to personality and, in terms of the transactional concept (e.g., Lazarus & Folkman, 1984), as a source of mental stress: "[. . .] stress has to do with the – anticipated or experienced – thwarting of goals" (Semmer, 1996, p. 53).

In these terms, irritation can be seen as a state of mental impairment resulting from a perceived goal-discrepancy. The two aspects of irritation that we will describe in the following – namely, *rumination* (here called cognitive irri-

tation, CI) and *irritability* (here called emotional irritation, EI) – can be regarded as the two most typical patterns of behavior resulting from a perceived goal discrepancy (e.g., Martin & Tesser, 1996).

Rumination (CI) describes "a class of conscious thoughts that revolve around a common instrumental theme and that recur in the absence of immediate environmental demands requiring the thoughts" (Martin & Tesser, 1996, p. 7). In terms of Klinger's (1975) "incentive-disengagement cycle," which illustrates the successive processes of a lasting goal discrepancy and its consequences for the development of further mental impairments, rumination can be seen as a state of reinforced efforts toward goal-achievement. In spite of this positive intention, ruminative strategies often turn out to be ineffective and counter productive in dealing with perceived goal-discrepancy: Ruminations intensify negative emotions (e.g., Lyubomirsky & Nolen-Hoeksema, 1993, 1995; Nolen-Hoeksema, McBride, & Larson, 1997) and can cause impairments in cognitive abilities as well as in cognitive flexibility (e.g., Beckmann, 1994; Davis & Nolen-Hoeksema, 2000; Lyubomirsky, Kasri, & Zehm, 2003). As a result, daily demands may no longer be fulfilled adequately. This may cause general expectancy deficits, which in turn may lead to learned helplessness (Kuhl & Weiß, 1994). Not surprisingly, then, many studies report that rumination fosters the development of depression (e.g., Harrington & Blankenship, 2002; Nolen-Hoeksema, 1998, 2000; Nolen-Hoeksema & Harrell, 2002; Nolen-Hoeksema et al., 1997; Nolen-Hoeksema & Morrow, 1993; Nolen-Hoeksema, Parker, & Larson, 1994).

With respect to the aforementioned "incentive-disengagement cycle," the second aspect of irritation, *irritability* (EI), can be seen as a state of losing the incentive to achieve a certain goal (but not yet as a disengagement from these goals, cp. Kuhl, 1997) after intense efforts toward goal-achievement have not led to diminished goal-discrepancy. The "incentive-disengagement cycle" describes such emotionally irritable reactions to goal-discrepancy as a *direct* precursor of depression. Compatible empirical evidence was found in a longitudinal study by Dormann and Zapf (2002). It appears, then, that EI (irritability) indicates a more severe kind of mental strain than CI (rumination).

Thus, the two processes, rumination and irritability, which we are subsuming here under the label irritation, are not only indicators of mental strain but also precursors of more serious impairments. Mohr (1991) considered this idea and proposed a theoretical model concerning the development of psychological constraints. The model claims that stressors will only lead to further impairments, such as psychosomatic complaints, anxiety, depressive symptoms, and reduced self-esteem, when these stressors are mediated via irritation. Dormann and Zapf (2002) verified parts of this model (see above). Thus, an early assessment of such a process may help to predict severe mental impairments and aid in the planning and evaluating of organizational health programs.

Research Questions

The aim of this paper is to examine the cross-cultural equivalence of nine (Arabic, Dutch, English, French, Italian, Polish, Russian, Spanish, and Turkish) adaptations of the Irritation Scale (Mohr, 1986; Mohr, Rigotti, & Müller, 2005; Müller, Mohr, & Rigotti, 2004). More specifically, we tested the construct equivalence (= structural equivalence) between all versions to answer the question as to whether or not the Irritation Scale measures the same construct across all countries (Van de Vijver & Leung, 1997). We suppose that irritation is an etic construct by testing the same set of items across all countries.

Method

The Irritation Scale

Developed with the practical experiences of occupational health psychologists in mind, the instrument is economical and consists of only eight items. The items are almost literal quotations from qualitative interviews with German industrial workers (cp. Mohr, 1986). Because of the high frequency of spontaneous statements such as feeling "grumpy" or being unable to relax, the concept of irritation can be seen as highly socially relevant. Furthermore, the scale itself can be regarded as transparent and understandable across different occupational branches and hierarchies (Mohr, Rigotti, & Müller 2005). Because the items explicitly relate to the field of work, the scale should only be applied to persons with occupational experience. The confirmatory factor analysis of the German scale (Müller, Mohr, & Rigotti, 2004) supports a hierarchical model of the scale with the described two first-order factors, CI and EI, and one common second-order factor, "Irritation"

Items 1, 2, and 4 (e.g., Item 2: "Even at home I often think of my problems at work.") represent the first-order factor CI and refer to thoughts about problems at work (work specific ruminations). Items 3, 5, 7, and 8 represent the first-order factor EI – (e.g., Item 3: "I get grumpy when others approach me."). A 1 to 7 Likert-type response scale (high scores for high irritation) has been proven reliable and valid in various studies (for an overview see Mohr et al., 2005). Both subscales show good internal consistencies from .85 to .97, as well as high item-total correlations in over 15 studies with a total of 4,030 persons in the original German version. The second-order factor including all eight items correlates positively – as expected – with psychosomatic complaints, depression, social stressors, emotional exhaustion, and depersonalization. Negative correlations were found with constructs such as optimism, general self-efficacy, as well as occupational self-efficacy and life-satisfaction. A separate analysis of the two subscales allows more differentiated results, which can be interpreted in the theoretical context of goal-discrepancy (Müller et al.,

Table 1. Description of all samples with respect to gender and age

	N	Women	Men	Mean age in years *(SD)*
Arabic (Morocco)	111	45	66	35 (9)
Dutch (Netherlands)	123	77	44	40 (13)
English (USA)	151	111	35	33 (13)
French (Switzerland)	277	142	134	33 (13)
Italian	141	91	49	Categories
Polish	103	72	31	33 (10)
Russian (Russia, Ukraine)	204	127	77	35 (10)
Spanish (Spain)	152	78	74	41 (10)
Turkish	101	56	44	39 (11)
German	4.030	1.621	2.409	Categories

Note. Age categories exist only for the Italian sample. 18–25 years: 2.8%, 26–35 years: 58.5%, 36–45 years: 30.3%, 46–55 years: 8.5%

2004): CI exhibits positive correlations to different aspects of work-motivation but is at the same time connected with increased workload and diminished *occupational* self-efficacy, although not with diminished *general* self-efficacy. This can be interpreted in terms of the reinforcing processes described above. On the other hand, in the same study, EI showed negative or even no correlations to different aspects of work-motivation but higher positive correlations than CI to further mental impairments such as diminished general self-efficacy, depression, or psychosomatic complaints. Both subscales are correlated with negative work-home interference, which points toward a spill-over effect of irritation into leisure time and family interaction.

Scale Translations

The scale was translated into Arabic, Dutch, English, French (the French translation had already been made in 1991 by the Norbert Semmer team from the University of Bern, Switzerland), Italian, Polish, Russian, Spanish, and Turkish. Following the recommendations of Behling and Law (2000), as well as those of Werner and Campbell (1970), we put emphasis on an equivalent linguistic transfer. A good translation should lead to semantic, conceptual, and normative equivalence, and does not only include the literal translation process. A presumption of conceptual equivalence is that the constructs assessed are etic. We do not expect conceptual invariance for Irritation across different countries. Therefore, we used input harmonization (Hoffmeyer-Zlotnik & Wolf, 2003) by using the same set of items in each adaptation.

The translation teams were pointedly asked to uphold the colloquial character of the items in the new versions. The translation teams in most cases consisted of a minimum of four persons, including native speakers with excellent knowledge of the German language, as well as Germans with excellent knowledge of the particular foreign language. In a first step, the scales were translated into the target language. In a second step, these translations were translated back into German by independent teams. Differences between the back-translations and the original German items were discussed by the two independent teams of translators, and any necessary corrections were carried out. Special attention was also paid to the translation of the word-anchors of scales. In all cases, we found a compromise between using the most common verbal qualifiers for a national questionnaire in comparison to the translation of the German anchors (Harkness, 2003). A member of the research team supervised the whole process.

Samples and Data Preparation

The total sample consists of *N* = 1,363 persons from nine countries and *N* = 4,030 from Germany. Data was gathered in Morocco (Arabic), The Netherlands (Dutch), USA (English), Switzerland (French, data was partly provided by Tschan), Italy (Italian), Poland (Polish), Ukraine and Russia (Russian), Spain (Spanish data was provided by Enrique Merino Tejedor, Universidad de Valladolid), and Turkey (Turkish data was provided by Aycan).

The French data is from 1998, German data is from the years 1997 to 2004. All other data-sets are from the years 2004 and 2005. Questionnaires were distributed by foreign students in their home countries, or by contact persons in the countries. All were samples of convenience. The main characteristics of the sample are described in Table 1. In accordance with the comments concerning the Irritation Scale mentioned above, only employed persons were included in our study. For purposes of comparison, we will also present data from the standardized German sample, which consists of *N* = 4,030 persons and was compiled from various other studies (see Mohr et al., 2005).

As a result of previous multigroup analyses in German samples, no differences in the scale structure between the sexes as well as between different age groups is to be expected (Müller et al., 2004). Consequently, the data for both sexes and for different age groups were combined for further analyses for every sample. Missing values on single items were very scarce. The highest rate was found in the Russian sample with a rate of 3% missing values. When

Table 2. Item means, skewness, component loadings, total correlations, explained variance of the components as well as internal consistencies of the total scale and the two subscales of irritation

	German N = 4.030					Arabic N = 111					Dutch N = 123					English N = 151					French N = 277				
Item	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}
1	3.77	.05	.43	*.90*	.75	3.59	.01	.22	*.82*	.63	2.73	.77	.60	*.80*	.60	2.85	.85	.52	*.72*	.51	2.94	.61	.46	*.69*	.44
2	3.93	−.04	.46	*.92*	.82	3.83	.08	.28	*.84*	.61	3.47	.13	.43	*.88*	.73	3.36	.28	.27	*.88*	.68	4.20	−.35	.34	*.86*	.63
3	2.95	.56	*.79*	.46	.67	2.91	.14	*.66*	.45	.50	2.90	.47	*.71*	.46	.57	2.61	.79	*.77*	.22	.62	2.36	.76	*.72*	.32	.55
4	2.73	.76	.50	*.85*	.70	3.27	.26	.35	*.80*	.59	2.39	.84	.36	*.87*	.67	2.42	1.03	.22	*.87*	.65	3.10	.46	.33	*.82*	.54
5	2.49	.97	*.79*	.52	.69	3.56	.39	*.64*	.59	.53	3.28	.33	*.79*	.54	.68	3.57	.23	*.68*	.54	.57	3.10	.49	*.74*	.53	.63
6	2.89	.64	*.87*	.38	.75	3.54	.16	*.83*	.26	.66	3.02	.33	*.91*	.39	.80	2.87	.76	*.84*	.19	.71	2.91	.66	*.82*	.28	.67
7	2.95	.60	*.90*	.38	.80	3.34	.41	*.83*	.21	.64	2.93	.54	*.90*	.34	.79	3.11	.48	*.89*	.31	.79	2.94	.50	*.85*	.33	.71
8	3.08	.56	*.76*	.51	.66	3.44	.09	*.70*	.22	.52	2.24	1.57	*.69*	.59	.58	3.44	.20	*.77*	.42	.66	3.36	.35	*.71*	.51	.59
EV(%)	57	16				43	19				55	15				48	19				47	15			
Factor correlation	.55*					.41*					.59*					.44*					.51*				
α (Total)	.89					.80					.88					.84					.84				
α – (5, 8)	.86					.76					.84					.78					.77				
α CI	.87					.78					.81					.77					.71				
α EI	.88					.79					.86					.85					.83				

	Italian N = 141					Russian N = 204					Spanish N = 152					Turkish N = 101					Polish N = 103				
Item	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}	M	SKEW	F1	F2	r_{it}
1	3.00	.33	*.92*	.32	.83	3.68	.00	.54	*.85*	.68	3.70	.11	.35	*.93*	.80	3.52	.37	.42	*.76*	.40	2.97	.57	.40	*.84*	.53
2	3.45	.12	*.92*	.34	.84	4.39	−.28	.41	*.89*	.76	3.36	.35	.48	*.94*	.84	4.42	−.28	.59	*.64*	.67	3.90	.12	*.73*	.62	.64
3	2.80	.63	.32	*.79*	.63	3.36	.15	*.70*	.39	.53	2.03	1.52	*.72*	.29	.56	2.49	1.26	.39	*.69*	.50	4.10	−.21	*.74*	.06	.35
4	2.38	1.05	*.85*	.29	74	3.69	.13	.41	*.85*	.68	2.29	1.19	.60	*.81*	.67	3.37	.34	.59	*.59*	.53	2.89	.67	.54	*.70*	.63
5	2.38	1.06	.56	*.76*	.68	3.94	−.21	*.75*	.57	.63	3.64	.19	*.78*	.39	.67	4.22	−.29	*.70*	.61	.59	4.36	−.22	*.72*	.50	.54
6	2.60	.58	.28	*.87*	.68	3.50	.27	*.88*	.37	.73	2.84	.72	*.91*	.40	.82	4.01	−.21	*.87*	.39	.65	3.83	−.02	*.63*	.53	.59
7	2.50	.78	.29	*.89*	.71	3.37	.31	*.83*	.39	.67	3.08	.61	*.90*	.51	.81	3.85	.18	*.90*	.35	.61	4.36	−.44	*.78*	.46	.62
8	2.80	.61	*.80*	.46	.46	3.43	.24	*.66*	.63	.56	3.01	.74	*.79*	.52	.70	2.99	.78	.26	*.82*	.41	2.82	.90	.27	*.83*	.35
EV(%)	52	22				52	15				57	17				47	14				47	14			
Factor correlation	.52*					.59*					.53*					.65*					.67*				
α (Total)	.87					.87					.89					.84					.83				
α – (5, 8)	.81					.83					.86					.80					.79				
α CI	.90					.84					.88					.71					.77				
α EI	.83					.83					.88					.77					.72				

Notes. F1 und F2: Loadings of exploratory factor analysis (principal component analysis), rotation: promax. Cursive displayed are the items main component-loadings. Item-total correlations were analyzed with respect to the items hypothesized assignment to the two subscales (item 1, 2, 4 = cognitive irritation, item 3, 5, 6, 7, 8 = emotional irritation). EV = Explained Variance in percent. Factor-Correlation: *p < .01 (Pearson product-moment correlation, significance: two-tailed). α (total) = internal consistency (Cronbach's α) of the total scale including all eight items,. α – (5, 8) = internal consistency (Cronbach's α) of the total scale without item 5 and 8, α (CI) = internal consistency (Cronbach's α) of subscale cognitive irritation, α (EI) = internal consistency (Cronbach's α) of subscale emotional irritation.

dealing with missing value rates of such small proportion, they are not expected to bias the analyses in any serious sense (c.p. Tabachnick & Fidell, 2001, p. 59), so we performed a listwise deletion of missing values.

Results

Descriptive Properties

Table 2 shows the main descriptive item parameters. In most of the adaptations, the skewness indicates a partly nonnormal distribution of the items.

Structural Equivalence

To extract problematic items, we first employed separate exploratory principal component analyses (PCA) for all adaptations of the scale (see Table 2 for results). In all adaptations, two components with eigenvalues over 1.00 were extracted. In the promax-rotated factor solution, the hypothesized pattern of component loadings (see Appendix) could be replicated in six adaptations (Arabic, Dutch, English, French, Russian, and Spanish). In these adaptations, the first component (F1 in Table 2) can be interpreted in terms of EI, the second component in terms of CI (F2 in Table 2). Unexpected main loadings were found for Item 8 in the Italian, Turkish, and Polish adaptations, as well as for Item 3 in the Turkish and Item 2 in the Polish adaptation. For the Turkish, Polish, and Italian adaptations, the exploratory PCA allows for no clear interpretation of the components. On the whole, in addition to Item 8, Item 5, which had interpretable loadings of at least .42 (Comrey & Lee, 1992) on both factors in eight out of nine of the presented adaptations, showed the worst ability to represent the assumed factor.

After employing the PCA, we carried out confirmatory factor analyses with AMOS 5 (Arbuckle, 2003). With at least 5.9 observations per parameter to estimate ($N = 101/17$ parameters to estimate), we meet Bentler and Chou's (1987) suggestion of five observations per estimated model parameter to ensure the reliability of structural models. All fit-statistics are listed in Table 3. The classical χ^2 value is only presented for the sake of completeness. It is not regarded as reliable in evaluating model adequacy because it is highly affected by sample size and distributional assumptions (e.g., Hu & Bentler, 1998). Out of the variety of fit indices, we chose three of the most widely used and best-approved measures: the GFI (Jöreskog & Sörbom, 1984), the CFI (e.g., Bentler, 1990), and the RMSEA (Steiger, 1990). The GFI indicates the overall fit of the research model. The CFI can be characterized as an incremental fit index, because it indicates the improvement between the tested model and a hypothesized null-model. The RMSEA is a so-called parsimonious fit index because

it considers the fact that more complex models always have better fit. In other words, it "rewards" less complex models. CFI- and GFI- values of over .95 are thought to represent good fit, whereas values of over .90 represent acceptable fit. RMSEA values of .05 and below represent good fit; values of up to .08 represent acceptable fit. Models which exceed the RMSEA value of .10 have only poor fit (Browne & Cudeck, 1992).

All fit statistics of the assumed two-factorial model, including all eight items of the scale, are summarized in Table 3 under the label "A." The GFI indicates acceptable to good fit (.85–.95) for all adaptations, as does the CFI (.85–.98). On the other hand, all models have to be rejected on the basis of the RMSEA. Thus, an additional analysis (Model B in Table 3) was carried out excluding those two items (namely 5 and 8) that seemed to poorly represent their underlying factors. The results show an improvement of the model fit of five (Arabic, Dutch, French, Italian, Polish) out of nine adaptations. For these adaptations all three indices point to an (at least) acceptable fit of Model B to the empirical data. For the remaining four adaptations, the GFI as well as the CFI show good fit, but again, the RMSEA values indicate only poor fit.

In addition, in order to examine the fit of the model's overall structure to each single adaptation, we also wished to examine the equivalence across the different adaptations on a more sophisticated level. We tested this by means of

Table 3. Fit-indices of the two-factorial hierarchical model of irritation

	N	Model	χ^2	GFI	CFI	RMSEA
German	4.030	A	586.93	.95	.87	.09
		B	245.96	.98	.94	.09
Arabic	111	A	44.68	.92	.91	.11
		B	12.75	.96	.98	.07
Dutch	123	A	80.98	.88	.88	.16
		B	18.19	.95	.97	.10
English	151	A	85.94	.88	.88	.15
		B	25.33	.95	.95	.12
French	277	A	82.54	.93	.92	.11
		B	21.11	.98	.97	.08
Italian	141	A	108.46	.84	.86	.18
		B	6.14	.99	1.00	.00
Russian	204	A	107.40	.89	.89	.15
		B	35.94	.95	.95	.13
Spanish	152	A	69.43	.90	.94	.13
		B	35.56	.93	.95	.15
Turkish	101	A	64.12	.86	.84	.15
		B	17.05	.95	.95	.11
Polish	103	A	45.42	.90	.90	.12
		B	7.99	.97	1.00	.00

Note. Model A: includes all eight items ($df = 19$), Model B: without Item 5 and 8 ($df = 8$).

Table 4. Tests of structural invariance for multigroup model of the Irritation Scale (with and without Item 5 and 8)

Model	χ^2	df	$\Delta\chi^2$	Δdf	GFI	CFI	ΔCFI	RMSEA
	Including all 8 items							
Unconstrained model	1592.45	190	–	–	.93	.94	–	.04
Measurement weights	1705.82	244	113.37**	54	.93	.94	.00	.03
Structural covariances	1780.74	271	74.92**	27	.92	.94	.00	.03
Measurement residuals	2697.96	343	917.22**	72	.86	.90	.05	.03
	Without items 5, and 8							
Unconstrained model	485.71**	80	–	–	.97	.98	–	.03
Measurement weights	571.14**	116	85.43**	36	.96	.97	.01	.03
Structural covariances	641.50**	143	70.36**	27	.96	.97	.00	.03
Measurement residuals	1378.34**	197	736.84**	54	.90	.93	.04	.03

Note. **$p < .001$

several multigroup analyses with increasingly restricted specifications.

The *unconstrained model* simply tested for the same number of factors across all adaptations. The *measurement weight model* also assumed the same factor loadings across all adaptations. The *structural covariances model* additionally required equal factor covariances across all adaptations. The most restricted model was the *measurement residual model*, which also demanded equal residual variances. To indicate fit differences between models resulting from the added restrictions, we used the ΔCFI, which, unlike the traditional χ^2, is not sensitive to sample size (Cheung & Rensvold, 2002). ΔCFI values higher than .01 indicated a significant decrease of fit.

The results of the multigroup analyses across all adaptations including the German original scale are summarized in Table 4.

In both versions we did not observe a significant decrease of fit between the unconstrained model and the measurement weight model. Thus, we can assume that the differences in the factor loadings across the adaptations, which we had observed in the PCA, are within the range of statistical tolerance and allow us to accept equivalence across adaptations. Furthermore, the addition of the constraint of equal factor covariances did not lead to a significantly poorer fit, whether Items 5 and 8 were included in our analyses or not. We may conclude that the exclusion of these two items leads to an increased factorial validity of the Irritation Scale in all adaptations (as can be observed in Table 3). The structural equivalency of the poorer model (including Items 5 and 8) is also supported.

Only the residual variances were not equal across groups, which is to say that this restriction leads to a significant decrease of fit (ΔCFI > .01) in comparison to the structural covariances model.

Reliability

The results of the reliability analyses are summarized in Table 2. Nearly all items in all adaptations show at least acceptable item-total correlations ($r_{it} > .30$) with respect to the hypothesized assignment to one of the two subscales. All internal consistency values can be regarded as at least satisfying.

Validity

For the German version, a large number of studies report correlations between the Irritation Scale and stressors at work, further impairments, such as psychosomatic complaints or depression, and missing resources, such as social support (for an overview see, Mohr et al., 2005). When differentiating the subscales, it could be shown in a German context that CI is – not unexpectedly – positively related to work motivation, which could indicate a kind of dysfunctional overcommitment. In contrast to this finding, EI was shown to be negatively related to motivation (Müller et al., 2004).

So far, further validity analyses in our data-sets are only available for the Arabic and Turkish adaptations with respect to work-family overload and depression. In the Arabic sample, a translation of the Depression scale by Mohr (1986; Mohr & Müller, 2004) measuring preclinical depressive symptoms (eight items, Cronbach's $\alpha = .83$) was assessed. Additionally, in the Turkish sample, a scale assessing work-family overload (Duxbury & Higgins, 1994) with 10 items (Cronbach's $\alpha = .90$) and the CES-D by Santor and Coyne (1997) measuring depression (7 items, Cronbach's $\alpha = .78$) were employed. In the Arabic and Turkish samples, both subscales show positive correlations to depression. As hypothesized, in the Arabic sample, EI ($r = .37$, $p < .001$) is more closely related to depression than CI ($r = .29$, $p < .01$). In the Turkish sample, however, CI ($r = .33$, $p < .01$) correlates more strongly to depression than EI ($r = .28$, $p < .05$). Further findings in the Turkish sample nevertheless show the hypothesized results: both aspects of irritation are positively correlated to work-family overload (CI: $r = .61$, $p < .01$, EI: $r = .40$, $p < .01$, total: $r = .62$, $p < .01$). Authors reporting on the Spanish version of the Irritation Scale have shown substantial correlations between

the Irritation Scale and facets of burnout (Merino, Carbonero, Moreno, & Morante, 2006). Results from the English version applied in Scotland and Ireland (Rotheiler & Metz, 2003) show correlations with an inability to relax ($r = .73$, $p < .001$) or with impatience ($r = .70$, $p < .001$), both of which are subscales of the Fragebogen zur Analyse belastungsrelevanter Anforderungsbewältigung (FABA; Richter, Rudolf, & Schmidt, 1996).

Thus, all correlations to date are as expected, except for the stronger correlation found in the Turkish sample between CI and depression.

Discussion

It was the aim of this paper to examine the equivalence of nine adaptations of the Irritation Scale, a measure to assess mental strain in occupational contexts. For these purposes, exploratory factor analyses, as well as single group and multiple confirmatory factor analyses, were carried out. With the exclusion of two culturally diverse items (Item 5 and 8), measurement equivalency, that is to say, structural invariance, can be claimed between the approved original German version and the nine tested adaptations. Hence, these adaptations can be adequately described by the same overall structure as well as by group-invariant factor loadings. This still holds true, even though the English, Russian, Turkish, and Spanish adaptations did not fully fit to the hypothesized model. Taking into account that the unsatisfying fit values, which should lead to the rejection of the model, show borderline values to an acceptable fit, and also taking into account the good internal consistencies and item-total correlations, a further examination of the measurement equivalence even of these four adaptations seems worthwhile.

For all adaptations, good to acceptable internal consistencies were found throughout. Both excluded items focus on perceived physiological arousal in connection with emotional irritation (Item 5: "From time to time I feel like a bundle of nerves."; Item 8:" When I come home tired after work, I feel rather irritable."). The indicator function of these aspects varies between the considered cultures. This may also be caused by the colloquial term used in Item 5, which makes it hard to transfer to other languages, and the conditional phrase in Item 8. The exclusion doesn't affect the assumed characteristic of the construct.

For the examination of the cross-cultural equivalence of psychological constructs and measures, it is also important to ensure that the constructs are organized in a system of psychological constructs that is also cross-culturally equivalent (Paunonen, 2000). The validity analyses performed so far provide only preliminary evidence for the integration of irritation and its two aspects into such a system. Regarding our present results, we can say that the construct is *generally* related to mental strain within all of the contexts we have analyzed so far (Arabic, English, Spanish, Turkish).

We still do not know if the above described *specific* "mechanisms" of goal disengagement (Klinger, 1975) that connect the two aspects of irritation to further mental impairments are also valid across different cultural contexts. A first special cultural feature in the processing of occupational stress was found in the Turkish population. Here, in contrast to the described assumptions of goal disengagement, and in contrast to previous findings in Germany, the aspect CI seems to indicate more serious mental impairments than EI. The outlined theoretical background of irritation enables the derivation of checkable hypotheses and provides a valuable framework to conduct necessary future research.

From our point of view, it is important to interpret all results while keeping the discussion concerning the adequate assessment of etic, or universal constructs (cp. Harkness, 1999, p. 128), in mind. Etic assessment strategies include: using different items in different cultures to measure the same construct (Triandis & Marin, 1983), aiming for structural equivalency as sufficient to ensure cross-cultural equivalency (Paunonen & Ashton, 1998), or taking metric invariance as the only satisfactory criterion (Byrne & Watkins, 2003). It is beyond the scope of this paper to cover all the pros and cons of these standpoints. Instead, the aim of the present study was to provide researchers from *all* of the above-mentioned "schools" with as broad a range of information as possible about the analyzed adaptations of the Irritation Scale.

A limitation of our study can be seen in the convenient samples, which may hamper comparisons between the country samples, as they cannot be seen as representative. However, the present results encourage the application of all nine tested adaptations of the Irritation Scale to assess work-related mental strain in cross-cultural (cross-national) contexts. Further research is necessary to test whether findings can be transferred to other national entities speaking the same languages, such as, for example, samples from other Arab-speaking countries as compared to the Moroccan sample, or samples from Latin-American countries as compared to the Spanish sample.

Another essential step should be the assessment of the external construct validity of irritation in order to further support the information presently available concerning the measurement equivalency of the Irritation Scale.

Acknowledgments

References

Arbuckle, J.L. (2003). *AMOS (Version 5).* Chicago: Small Waters.
Beckmann, J. (1994). Ruminative thought and the deactivation of an intention. *Motivation and Emotion, 18*, 317–334.

Behling, O., & Law, K.S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. New York: Wiley.

Bentler, P.M. (1990). Comparative fix indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Bentler, P.M., & Chou, C. (1987). Practical issues in structural modeling. *Sociological Methods and Research, 16,* 78–117.

Brislin, R.W., Lonner, W.J., & Thorndike, R.M. (1973). *Cross-cultural research methods*. New York: Wiley.

Browne, M.W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research, 21*, 230–258.

Byrne, B.R., & Watkins, D. (2003). The issue of measurement invariance revised. *Journal of Cross-Cultural Psychology, 34*, 155–175.

Burchell, B.J., Day, D., Hudson, M., Lapido, D., Mankelew, R., Nolan, J.P., Reed, H., Wichert, C.I., & Wilkinson, F. (1999). *Job insecurity and work intensification: Flexibility and the changing boundaries of work*. York: York Publishing Services.

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

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Cropanzano, R., James, K., & Citeria, M. (1993). A goal hierarchy model of personality, motivation, and leadership. *Research in Organizational Behavior, 15*, 267–322.

Davis, R.N., & Nolen-Hoeksema, S. (2000). Cognitive inflexibility among ruminators and nonruminators. *Cognitive Therapy & Research, 24*, 699–711.

Dormann, C., & Zapf, D. (2002). Social stressors at work, irritation, and depressive symptoms. Accounting for unmeasured third variables in a multi-wave study. *Journal of Occupational and Organizational Psychology, 75*, 33–58.

Duxbury, L., & Higgins, C. (1994). Work-family conflict: A comparison by gender family-type and perceived control. *Journal of Family Issues, 15*, 449–466.

Harkness, J. (1999). In pursuit of quality: Issues for cross-national survey research. *Research Methodology, 2*, 125–140.

Harkness, J. (2003). Questionnaire translation. In J. Harkness, P.P. Mohler, & F. Van de Vijver (Eds.), *Cross-cultural survey methods* (pp. 35–56). Hoboken, NJ: Wiley.

Harrington, J.A., & Blankenship, V. (2002). Ruminative thoughts and their relation to depression and anxiety. *Journal of Applied Social Psychology, 32*(3), 465–485.

Hoffmeyer-Zlotnik, J.H.P., & Wolf, C. (2003). Comparing demographic and socioeconomic variables across nations: Synthesis and recommendations. In J.H.P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison. A European working book for demographic and socioeconomic variables* (pp. 389–406). New York: Kluwer Academic/Plenum.

Horn, J.L., & McArdle, J.J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research Special Issue: Quantitative topics in research on aging, 18*, 117–144.

Hu, L., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453.

Jöreskog, K.G., & Sörbom, D. (1984). *LISREL-IV User's reference guide*. Mooresville, IN: Scientific Software.

Klinger, E. (1975). Consequences of commitment to and disengagement from incentives. *Psychological Bulletin, 82*, 1–25.

Kuhl, J. (1997). Wille und Freiheitserleben: Formen der Selbststeuerung [Volition and sense of freedom: Forms of self-control]. In J. Kuhl & H. Heckhausen (Eds.), *Motivation, Volition und Handlung* [Motivation, volition, and action] (Vol. 4, S. 665–765). Göttingen: Hogrefe.

Kuhl, J., & Weiß, M. (1994). Performance deficits following uncontrollable failure: Impaired action control or global attributions and generalized expectancy deficits? In J. Kuhl & J. Beckmann (Eds.), *Volition and personality: Action versus state orientation* (pp. 317–328). Göttingen: Hogrefe.

Lazarus, R.S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lyubomirsky, S., Kasri, F., & Zehm, K. (2003). Dysphoric rumination impairs concentration on academic tasks. *Cognitive Therapy & Research, 27*, 309–330.

Lyubomirsky, S., & Nolen-Hoeksema, S. (1993). Self-perpetuating properties of dysphoric rumination. *Journal of Personality & Social Psychology, 65*, 339–349.

Lyubomirsky, S., & Nolen-Hoeksema, S. (1995). Effects of self-focused rumination on negative thinking and interpersonal problem solving. *Journal of Personality & Social Psychology, 69*(1), 176–190.

Martin, L.L., & Tesser, A. (1996). Some ruminative thoughts. In R.S. Wyer (Ed.), *Ruminative thoughts* (pp. 1–47). Hillsdale, NJ: Erlbaum.

Merino, E., Carbonero, M.A., Moreno, B., & Morante, M.E. (2006, in press). La Escala de Irritación como instrumento de evaluación del estrés laboral [The Irritation Scale as an instrument to assess stress at work]. *Psicothema*.

Mohr, G. (1986). *Die Erfassung psychischer Befindensbeeinträchtigungen bei Industriearbeitern* [The assessment of mental strain of industrial workers]. Frankfurt: Lang.

Mohr, G. (1991). Fünf Subkonstrukte psychischer Befindensbeeinträchtigungen bei Industriearbeitern: Auswahl und Entwicklung [Five subconstructs of mental strain of industrial workers: Selection and development]. In S. Greif, E. Bamberg, & N. Semmer (Eds.), *Psychischer Streß am Arbeitsplatz* [Psychological strain in the workplace]. Göttingen: Hogrefe.

Mohr, G., & Müller, A. (2004). Depressivität im nichtklinischen Kontext [Depressivity in a nonclinical context]. In A. Glöckner-Rist (Ed.), *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente. Version 8.00* [ZUMA information system. Electronic handbook of assessment tools for the social sciences. Version 8.00]. Mannheim: Zentrum für Umfragen, Methoden, und Analysen.

Mohr, G., Müller, A, & Rigotti, T. (2005). Normwerte der Skala Irritation: Zwei Dimensionen psychischer Beanspruchung [Standardization data of the Irritation Scale. Two dimensions of mental strain]. *Diagnostica, 51*, 12–20.

Mohr, G., Rigotti, T., & Müller, A. (2005). Irritation – ein Instrument zur Erfassung psychischer Befindensbeeinträchtigungen im Arbeitskontext. Skalen- und Itemparameter aus 15 Studien [Irritation – an instrument assessing mental strain in working contexts. Scale- and item-parameters from 15 studies]. *Zeitschrift für Arbeits- und Organisationspsychologie, 49*, 44–48.

Müller, A., Mohr, G. & Rigotti, T. (2004). Differentielle Aspekte psychischer Beanspruchung aus Sicht der Zielorientierung. Die Faktorstruktur der Irritations-Skala [Differential aspects of psychological strain with respect to goal orientation: The

factor-structure of the Irritation Scale]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 25*, 213–225.

Nolen-Hoeksema, S. (1998). Ruminative coping with depression. In J. Heckhausen & C.S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 237–256). New York: Cambridge University Press.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109*, 504–511.

Nolen-Hoeksema, S., & Harrell, Z.A. (2002). Rumination, depression, and alcohol use: Tests of gender differences. *Journal of Cognitive Psychotherapy: Depressive rumination in cognitive theory and therapy, 16*, 391–403.

Nolen-Hoeksema, S., McBride, A., & Larson, J. (1997). Rumination and psychological distress among bereaved partners. *Journal of Personality & Social Psychology, 72*, 855–862.

Nolen-Hoeksema, S., & Morrow, J. (1993). Effects of rumination and distraction on naturally occurring depressed mood. *Cognition & Emotion, 7*, 561–570.

Nolen-Hoeksema, S., Parker, L.E., & Larson, J. (1994). Ruminative coping with depressed mood following loss. *Journal of Personality & Social Psychology, 67*, 92–104.

OECD. (2001). The employment of foreigners. Outlook and issues in OECD countries. In OECD (Ed.), *Employment outlook* (pp. 167–206). Paris: Author.

Paunonen, S.V. (2000). Construct validity and the search for cross-cultural consistencies in personality. In R.D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 123–140). New York: Kluwer Academic/Plenum.

Paunonen, S.V., & Ashton, M.C. (1998). The structured assessment of personality across cultures. *Journal of Cross-Cultural Psychology: Personality and its measurement in cross-cultural perspective, 29*, 150–170.

Richter, P., Rudolf, M., & Schmidt, C.F. (1996). *Fragebogen zur Analyse belastungsrelevanter Anforderungsbewältigung (FABA). Handanweisung* [Questionnaire to analyze stress-related coping. Testmanual]. Frankfurt: Swets & Zeitlinger B.V.

Santor, D.A., & Coyne, J.C. (1997). Shortening the CES-D to improve its ability to detect cases of depression. *Psychological Assessment, 9*, 233–243.

Rotheiler, E., & Metz, U. (2003). Beziehungen zwischen Faktoren der Arbeits-Anforderungsbewältigung (FABA, Richter et al.) und den Konzepten "Overcommitment" (Siegrist) und "Irritation" (Mohr) [Relations between factors of coping with stress at work and the concepts of overcommitment, and Irritation]. *Wirtschaftspsychologie, 5*, 150–152.

Schwarzer, R. (1994). Optimistische Kompetenzerwartung: Zur Erfassung einer personellen Bewältigungsressource [Generalized self-efficacy: Assessment of a personal coping resource]. *Diagnostica, 40*, 105–123.

Semmer, N. (1996). Individual differences, work stress, and health. In M.J. Schabracq, J.A.M. Winnubst, & C.L. Cooper (Eds.), *Handbook of work and health psychology*. Chichester: Wiley.

Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173–180.

Statistisches Bundesamt (1998). *Gesundheitsbericht für Deutschland 1998* [Health report for Germany 1998]. Stuttgart: Metzler-Poeschel.

Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics*. Boston: Allyn & Bacon.

Triandis, H.C., & Marin, G. (1983). Etic plus emic versus pseudoetic: A test of a basic assumption of contemporary cross-cultural psychology. *Journal of Cross-Cultural Psychology, 14*, 489–500.

Van de Vijver, F., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.

Werner, O., & Campell, D. (1970). Translating, working through interpreters, and the problem of decentering. In R. Naroll & R. Cohen (Eds.), *A handbook of method in cultural anthropology* (pp. 398–420). New York: American Museum of Natural History.

Gisela Mohr

Universität Leipzig
Institut für Angewandte Psychologie
Arbeits- und Organisationspsychologie
Seeburgstr. 14–20
D-04103 Leipzig
Germany
E-mail rsplitt@uni-leipzig.de

Appendix: The Irritation Scale (English adaptation)

1. I have difficulty relaxing after work. (cognitive irritation)
2. Even at home I often think of my problems at work. (cognitive irritation)
3. I get grumpy when others approach me. (emotional irritation)
4. Even on my vacations I think about my problems at work. (cognitive irritation)
5. From time to time I feel like a bundle of nerves. (emotional irritation)
6. I anger quickly. (emotional irritation)
7. I get irritated easily, although I don't want this to happen. (emotional irritation)
8. When I come home tired after work, I feel rather irritable. (emotional irritation)

Response Format
1 = strongly disagree to 7 = strongly agree

Note. All further adaptations can be ordered from the first author. In addition to the adaptations presented in this paper, Swedish and Portuguese versions also exist, but have not yet been tested. Preliminary versions exist in Czech and Vietnamese. An application for research purposes only is exempt from charges. Data feedback is welcome.