

Measuring psychological stress and strain at work: Evaluation of the COPSOQ Questionnaire in Germany

Methoden zur Erfassung psychischer Belastungen: Erprobung des COPSOQ in Deutschland

Abstract

The undisputed increase of the relevance of mental work load is confronted with a lack of qualified or at least well documented measuring instruments covering all important aspects.

The COPSOQ (Copenhagen Psychosocial Questionnaire), a comprehensive instrument for the assessment of psychosocial factors at work, was tested in a partly modified version in a large German sample (N = 2561 employees). The aims of the study were the detailed investigation of the psychometric measurement properties, and based on these results, the development of an abbreviated version of the instrument. The analysis of objectivity, acceptance, practicability, sensitivity and content validity of the questionnaire as a whole did not show any problematic results – with some limitations regarding the length of the questionnaire.

The assessment of the reliability, generalisability, construct validity, criterion validity and diagnostic power of the single scales showed medium to good measuring qualities for the majority of the scales (i.e. Cronbach's alpha mostly >0.7). In addition, the psychometric properties were very similar to those in the Danish COPSOQ-study.

Considering all aspects of the measurement quality, a shortened version of the instrument was created. It attempts to combine measuring qualities as high as possible with a number of questions as low as possible.

The German COPSOQ questionnaire is a free screening-instrument for the recording of psychosocial work load and strain for all enterprises and organisations interested. The next step is the construction of a "job exposure matrix" for psychosocial factors at work, that means a central database with work load profiles and reference values for as many occupational groups as possible.

Keywords: psychosocial work load, questionnaire, measurement qualities

Zusammenfassung

Der unbestrittenen Zunahme der Relevanz psychischer Belastungen bei der Arbeit steht bisher ein Mangel an inhaltlich umfassenden und gesichert qualifizierten oder zumindest transparent dokumentierten Messinstrumenten gegenüber.

Der inhaltlich sehr breit angelegte dänische COPSOQ-Fragebogen (Copenhagen Psychosocial Questionnaire) zur Erfassung psychosozialer Faktoren bei der Arbeit wurde in einer umfangreichen Stichprobe (N = 2561 Beschäftigte) in Deutschland in teilweise modifizierter Fassung erprobt. Ziele waren die detaillierte Prüfung der psychometrischen Messeigenschaften, sowie darauf aufbauend die Erstellung einer verkürzten Fassung des Instruments.

Bei der Prüfung der Messmethode und des Messinstrumentes insgesamt gab es bezüglich der Objektivität, Akzeptanz, Praktikabilität, Sensitivität

Matthias Nübling¹

Ulrich Stöbel²

Hans-Martin Hasselhorn³

Martina Michaelis¹

Friedrich Hofmann³

1 FFAS: Freiburg Research Centre of Occupational and Social Medicine, Freiburg, Germany

2 University of Freiburg, Department of Medical Sociology, Freiburg, Germany

3 University of Wuppertal, Applied Physiology, Occupational Medicine and Infectiology (FB D), Wuppertal, Germany

und Inhaltsvalidität - mit Abstrichen bezüglich der Länge des Fragebogens - keine problematischen Befunde.

Die Prüfung der einzelnen Skalen auf Reliabilität, Generalisierbarkeit, Konstruktvalidität, Kriteriumsvalidität und diagnostische Aussagekraft ergab für die Mehrzahl der Skalen brauchbare bis gute Messeigenschaften (z.B. Cronbach's alpha meist $>0,7$), die außerdem den dänischen Ergebnissen sehr ähnlich waren.

Unter Würdigung aller geprüften Aspekte der Messqualität wurde ein verkürztes Instrument erstellt, das versucht, möglichst hohe Messqualitäten mit einer möglichst geringen Fragenanzahl zu verbinden.

Die deutsche Version des COPSOQ steht nunmehr als Screening-Instrument zur Erfassung psychischer Belastungen und Beanspruchungen bei der Arbeit Betrieben und Organisationen kostenfrei zur Verfügung. Besonders lohnend erscheint nun der Aufbau einer „job-exposure matrix“ im Bereich der psychosozialen Faktoren bei der Arbeit, also einer zentralen Datenbank mit Belastungsprofilen und Referenzwerten für möglichst viele Berufsgruppen.

Schlüsselwörter: psychische Belastungen, Fragebogen, Messqualitäten

1. Introduction

Working conditions in the Western industrial countries have experienced numerous, partly fundamental changes over the last decades. Along side with comprehensive changes of production conditions and realities in industry, administration and service, demands on the employees are also changing. Time and local flexibility, high toughness or social competences are becoming more and more key qualifications. As a consequence of these developments, the psychological stress has markedly increased. As a consequence sick-leave days and treatment costs due to psychosocial factors have increased over the last years. The current DAK-health report 2005 (DAK = German employee insurance company) with its focus on psychological sickness states a continuously increasing rate for sick-leave days for this diagnosis group. Among DAK-members cases of illness as well as the total number of days off due to mental health problems increased by ~70% between 1997 and 2004 [1], while total sick leave increased only by 5% in this population in the same period. The direct and indirect costs of psychological stress are counted almost as high as those for physical work pressure (e.g. [2]). For the (workplace) health promotion it is mandatory, to minimize health risks caused by psychosocial work overloads of the employees.

For the evaluation of the stress and demand situation it is indispensable to collect basic data of the psychological factors on the workplace in the first place. According to German legislature (Arbeitsschutzgesetz, §5 ff) companies are obligated to assess potential risk factors for the different working places, to document the results and, if necessary, to apply appropriate safety measures. Although this also includes psychosocial risk factors, the lack of validated and easy to apply measuring instruments has been one barrier for companies to follow their obligations so far.

1.1 Models for psychological stress

In work sciences - and also in the norms for measuring psychological stress, for instance ISO 10075 - one distinguishes between work load and stress (the entirety of measurable external influences) and strain (effects of the stress on employee depending on his/her individual conditions) as well as consequences of strain (e.g. disease). All of the current models assume that the relationship between stress and strain is not deterministic but is mediated and moderated through intervening variables like personal resources [3], [4], [5].

Two models dealing with the relationship between stress factors and all kinds of stress consequences or discomforts are leading in European working sciences: the “demand-control model”, originally formulated by Karasek [6] which has later been expanded to the “demand-control-support model” [7] by adding the dimension of social support. This model assumes working situations to have negative psychological or physical consequences especially when high demands concur with limited decision latitude (and low social support at the workplace in the extended model).

The “effort-reward imbalance model” has been developed by Siegrist [8], [9]. It postulates that the concurrence of highly extrinsic and intrinsic efforts with low chances of reward has particularly negative effects such as poor health.

Intensified efforts have been made over the last years to develop and validate standardised instruments to assess psychological stress in the working environment based on these and other models. The goal from the perspective of the work place risk assessment obligatory in Germany is to provide instruments which are scientifically and methodologically sound but also applicable by the practitioners on-site in the companies (e.g. occupational physicians, safety engineers, psychologists, health promoting personnel etc.).

1.2 Measuring psychological work load and strain

There is a wide consent that measuring psychological work load is highly complicated due to the indefiniteness of the theoretical construct and therefore various ways to define and operationalise load and strain (e.g. [10], [11], [12]). Besides expert based assessment of risks (examples: REBA [13]; SIGMA [14]; and others) and experimental techniques (example: [15]), directly asking the employees for psychological work load and strain experienced is the most commonly used measuring method. Such employee surveys are usually performed as anonymous cross-sectional surveys using standardised questionnaires so that large samples can be realized with relatively small financial and temporal efforts.

Most commonly used methods are postal inquiries, telephone interviews or online-surveys. Examples for such questionnaire instruments in Germany are the BASA-questionnaire (psychological assessment of working conditions) Richter [16], the ERI (effort-reward imbalance) questionnaire [8], the “short-questionnaire for work analysis” (KFZA) [17] or the “salutogenetic subjective work analysis” (SALSA) [18] among others. The COPSOQ also belongs to this group of self-administered questionnaires.

1.3 COPSOQ – the Danish original study

The COPSOQ (Copenhagen Psychosocial Questionnaire) has been developed and validated by Kristensen and Borg of the Danish National Institute for Occupational Health in Copenhagen [19]. The questionnaire was aimed to be “theory-based without being based on one specific theory.” Therefore, the COPSOQ is covering a broad range of aspects of currently leading concepts and theories. The following are mentioned [20]: “1. the job characteristics model. 2. the Michigan organizational stress model. 3. the demand-control-(support) model. 4. the sociotechnical approach. 5. the action-theoretical approach. 6. the effort-reward-imbalance model. 7. the vitamin model.” The COPSOQ tries to deal with the broadness respectively indefiniteness of the construct “psychosocial factors” by applying a multidimensional approach with a very wide spectrum of ascertained aspects [20]. Most COPSOQ questions were taken from already existing and well approved and validated instruments, for instance from the “Setterlind Stress Profile” [21], the “Whitehall II Study” [22] or the “Job Content Questionnaire” [23]. Only a small portion of items has newly been developed, in the long version 13 out of 141 items regarding stress, demands and discomforts. Most items are ordinal with five answer categories. In the Danish study the psychometrical qualities of the instrument have been tested on the basis of a representative sample of 1858 Danish employees (49% female, response rate 62%) between 20 and 60 years. The Danish authority for occupational safety has acknow-

ledged the COPSOQ (short version) as an instrument to evaluate psychosocial work load.

In numerous countries (Spain, Belgium, Norway, Brazil, Sweden and others) translations respectively adaptations of COPSOQ are currently tested. Therefore, first results concerning the methodological suitability as well as international comparisons of psychosocial stress in the workplace can be expected soon.

1.4 The German COPSOQ-validation study – goals

The German COPSOQ validation study was conducted by a project group headed by the Freiburg Research Centre of Occupational and Social Medicine (FFAS) and funded by the Federal Institute for Occupational Safety and Health (BAuA/FIOSH).

The scientific goal was the establishment of a German version of the COPSOQ questionnaire and a detailed examination and assessment of its’ measuring qualities on basis of a broad sample (N>2000). The practical goal in a next step was the provision of a shortened instrument to evaluate psychosocial factors at work.

2. Methods

The validation study comprised the following working steps:

- Translation and adaptation of the questionnaire: semantic adaptation, validation of additional items and scales.
- Performance of a pilot study (N=300); adaptation of the instrument.
- Performance of the main study: reference sample with a heterogeneous spectrum of jobs (N=2500) in Germany.
- Reanalysis of the measurement criteria of the questionnaire: objectivity, sensitivity, validity, reliability, diagnostic power, generalisability. (Assessment of the suitability in terms of the ISO 10075-3).
- Proposal of a shortened measuring instrument.

2.1 Content of the questionnaire

Considering the great extent of the original questionnaire and the primary objective of the study – to test the existing COPSOQ in Germany – supplements and changes were restricted to a minimum. Some scales with copyrights and not free for use had to be substituted, new questions were integrated to cover additional topics. (For details see the project report [24]; the German questionnaires (long and shortened version) are available as downloads in PDF from <http://www.copsoq.de/>).

Newly added in the German version (partly as a replacement) were: the WAI (Work Ability Index) of Tuomi and Ilmarinen [25] in the German translation of the BAuA/FIOSH [26], the question on subjective general

health from the EQ-5D [27], the Satisfaction with life scale (SWLS) [28], one single item concerning bullying/mobbing taken from the BIBB/IAB survey [29], one single item concerning intention to leave the job from the NEXT study [30], the work-family conflict scale (five categories instead of seven; term “family” extended to “privacy”) according to Netemeyer [31], the scales procedural justice and relational justice developed in Finland [32], [33], [34] and the scale personal burnout from the Copenhagen Burnout Inventory (CBI) [35].

The wording of the sociodemographic questions was taken from the German demographic standards [36]. Some questions were also taken from the Eurobarometer 44.2 [37] in its German translation.

In Figure 1 the scales of the German COPSQ are arranged according to the thematic fields.

2.2 Translation/adaptation of the questionnaire

For the development of a German COPSQ questionnaire a two-stage procedure was applied: First, the English and the Danish original versions were translated into German and after that back into the respective source language („translation-retranslation process“ or „forward-backward procedure“). In a second step all available versions were examined by a project group and the final question wording was fixed (so-called “Committee Assessment”). For details of the procedure see “Questionnaires in Translation” [38].

2.3 Data collection

The survey was advertised through the FFAS and open for all organisations and enterprises interested – they were included according to their date of application and the establishment of a broad mix of professions and branches. Companies received all materials from the FFAS and distributed the survey-kits consisting of the questionnaire, a covering letter of the FFAS, a recommendation letter of the company and a free return envelope to all employees. The return period was set to three weeks; shortly before the end of that period one reminder was sent out. The pilot study was conducted in September 2003, the main study from February till October 2004. Participating companies were asked to answer a short questionnaire concerning the accomplishment and acceptance of the survey in their organisation. All companies and organisations received “their” report in a written form as well as on CD until the end of 2004.

2.4 Methods of data analysis

All categorical items on work load and strain were transformed on a value range from 0 (minimum value, for instance: “do not agree at all”) to 100 points (maximum value, for instance: “fully agree”). Answer refusals and the category “does not apply” possible for some questions (e.g. B.8: “I have no superior/no colleagues”) were treated

as missing values. Scale scores were computed as the average of the values of the single aspects, if at least half of the single items had valid answers. Like all of the items all scales therefore have a theoretical range from 0 to 100.

Data analysis included

- descriptive statistics,
- parametric and non-parametric correlation analyses,
- explorative and confirmatory factor analyses,
- simple and multiple regression analyses and
- reliability analyses.

Additionally to Cronbach’s alpha and intra-class-correlation (ICC) for the total population, G-coefficients in a single-facet design [39] were determined during the analysis of the scale reliability to verify the generalisability of the item information [40]. The applied statistical methods were in particular geared to the recommendations of Bortz [41], Devellis [42], and Wirtz and Caspar [43]. Except for the structural equation models which were computed with Amos 5, all analyses were performed using SPSS 11[®] and 12.

3. Results

After carrying out the pilot study (N=352) and an extensive assessment of missing responses and bottom and ceiling effects, the instrument for the main study was optimised in some few points. From February 2004 until October 2004, the COPSQ main study was realized. Since differences between pilot and final questionnaire were small, the 352 participants of the pilot study were integrated into the total database for all questions that remained unchanged, raising the total number of participants to N= 2561 persons.

The response rate calculated from the distributed questionnaires as documented by the companies is 40.4% (minimum: 24% in a hotel business company, maximum: 68% in a sample of priests). One possible explanation for the relatively low response rate derives from the free text statements on working conditions: a lot of employees expressed demotivation due to worsened working conditions in the companies (increased working hours per week, loss of extra pay etc.). Another reason is the questionnaire itself, especially its length - with a shortened COPSQ a response rate of at least 50% should be possible in future.

3.1 Description of the sample (and external data for comparisons)

57% of the 2561 respondents were female, 43% male, wherewith women were over-represented compared to the working population in Germany. The average age was 42.6 years (standard deviation 10.6 years) with a range from 18 to 80 years. Persons older than 65 years (retirement age in Germany) were all priests. One percent of the respondents stated to have no secondary school

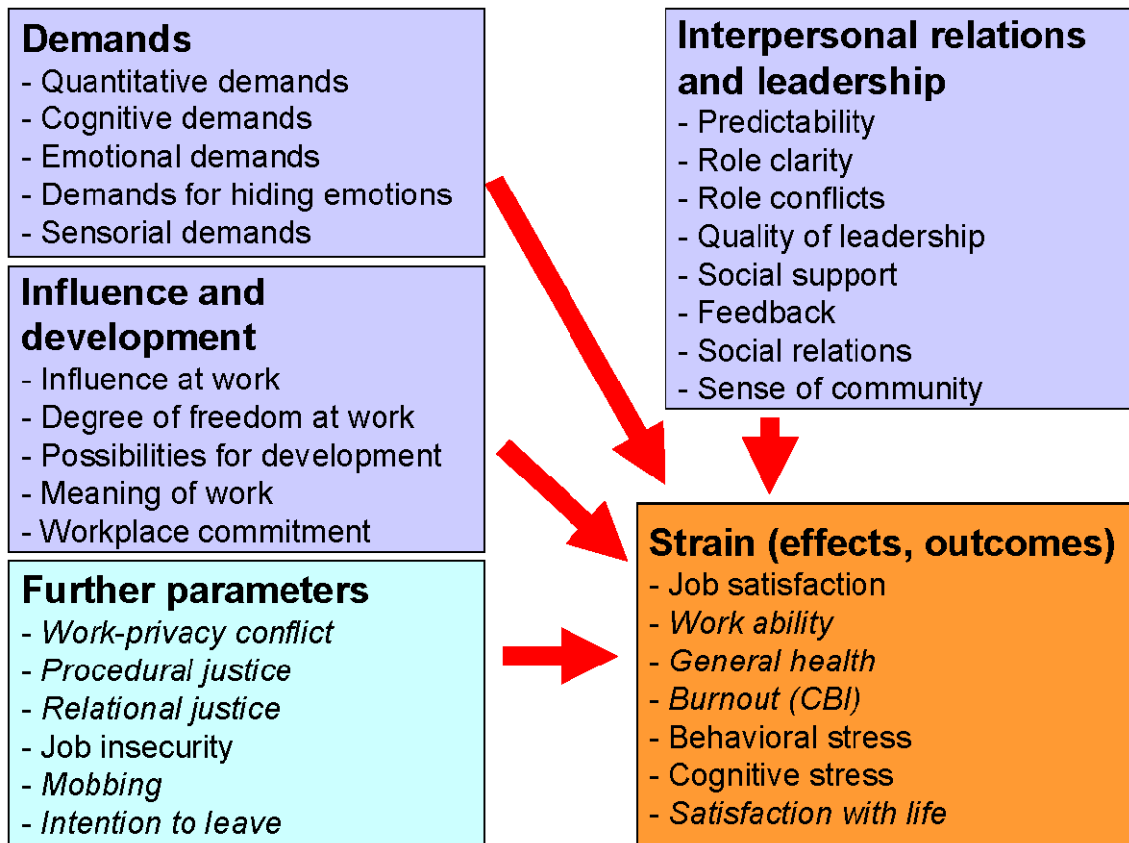


Figure 1: Structure of the German COPSQ questionnaire (changes compared to the Danish original are italicised)

Table 1: Distribution of professions according to the KdB classification

Job areas, job sectors and professional groups according to the KdB classification	N
I. Plant breeders, animal breeders, fishing occupation	4
II. Miners, mineral extracting occupations	0
III. Manufacturing occupations	30
IV. Technical occupations	39
V. Service occupations: Va: Goods salesmen, Vb: Service salesmen, Vc: Transport occupations, Vi. Other Service Occupations	64
Vd. Service: higher organisational, administrative, office occupations (groups 75, 76)	244
Vd. Service: lower organisational, administrative, office occupations (groups 77,78)	319
Ve. Service: protective services	146
Vf. Service: authors, artists	30
Vg. Service: Health care professions, physicians (codes 8410-8419)	77
Vg. Service: Health care professions, nurses (codes 8530-8539, 8541)	602
Vg. Service: Health care professions, rescue service (code 8542)	202
Vg. Service: Health care professions, others	73
Vh. Service: Social and educational professions, social work (codes 8610-8619)	112
Vh. Service: Social and educational professions, teachers (group 87)	360
Vh. Service: Social and educational professions, catholic priests (code 8911)	52
Vh. Service: Social and educational professions, protestant priests (code 8912)	68
Vh. Service: Social and educational professions, others	103
VI. Other professions	36
Total	2561

qualification, 10% had completed secondary school, 29% had a secondary school level I certificate, 4% a polytechnical school certificate, 9% had an advanced technical college entrance qualification and 46% have a general qualification for university entrance. With these results higher graduated people are clearly over-represented in the COPSQ sample as compared to the normal German population: the microcensus 2000 lists 43% to have completed secondary school, 19% secondary school level I certificate and 16% general qualification for university entrance [44]. The job titles stated were coded according to the classification of occupations of the Federal Statistical Office Germany (KdB 92, Klassifikation der Berufe) [45]. Table 1 gives case numbers for all six job areas (first hierarchy of the classification; I.–VI.) as well as subgroup specifications and KdB codes for those job sections, professional groups and job classes with larger case numbers.

Compared to the official statistics or other surveys the results of the COPSQ sample reveal a clearly increased social status of the respondents participating. However, other work correlated parameters like weekly hours of work or the type of employment are not noticeable different in the COPSQ sample.

3.2 Reanalysis of the validity criteria of the instrument

A comprehensive assessment of the psychometrical properties of the questionnaire was carried out using the data of the 2561 respondents of the German COPSQ study. Topics were:

- Objectivity (of the measuring process, analysis, and interpretation),
- Answer refusals, missing values,
- Bottom and ceiling effects (sensitivity),
- Content validity,
- Reliability, internal consistency of the scales,
- Generalisability of the measuring qualities,
- Construct validity, factorial validity and
- Diagnostic validity.

3.2.1 Answer refusals

Goal of the missing value analysis was to identify problematic accumulations of missing values in certain questions. The results for the main study reconfirmed findings of the original Danish study and the pilot study in which all presented questions were well answerable by the respondents; the rate of missing values was normally below 2%.

3.2.2 Bottom and ceiling effects (sensitivity)

The answer distributions of all questions were controlled for bottom and ceiling effects. Strong bottom or ceiling effects may be an indicator for methodological shortcomings in the selection and wording of the questions and/or

answer categories. This could lead to insensitivity towards real existing differentiations. With the exception of the scale sensorial demands (mean: 82 from 100 points) all scale means are within the range of 20 to 80 points. Only few single items, among which are three aspects of cognitive and three items addressing sensorial demands, show distinct ceiling effects (more than 80 points). As expected, bottom effects (less than 20 points) are found in the single questions concerning bullying/mobbing at work and intention to leave the job.

3.2.3 Content validity

Proving content validity is a task that can not be achieved solely by statistical methods; more important is to ensure a complete coverage of the aspects of interest in the planning of the study - as it was done in the original Danish study. In the interview situation content validity can additionally be reviewed by open questions like C.3 in COPSQ: "Are there any important aspects missing concerning psychological demands? If yes, which?". Some of the additional topics mentioned in the German COPSQ study are useful hints (for instance: "mobbing to be assessed in more detail"). But due to the length of the questionnaire we decided to add no more aspects.

3.2.4 Objectivity

The COPSQ is a standardised self administered survey instrument. Therefore, interviewer bias is not relevant. Moreover, the survey method was identical for all participants in the German study as well as the reminding procedure. The analysis procedure (scale construction, calculation of scale values) is given by the test author and cannot be altered. Scales and indices that were not taken from the original Danish study were computed and treated according to the methods proposed by their respective authors. Therefore, data collection as well as the analysis procedure present a good objectivity in terms of independence from influences by interviewers or data analysts.

3.2.5 Scale Reliability

Most efforts have been put into the assessment of the reliability of the measurement. There are several different aspects of reliability, especially test-retest reliability (stability over time), inter-rater reliability (independence from the assessing person), procedure reliability (consistency of the procedure) and scale reliability (internal consistency of the scales). Scale reliability assesses how well single aspects theoretically defined as a thematic unit are really seen as such a unit by the respondents in their answering patterns. This means, that a high inter-correlation among the single items as well as a strong correlation of these variables with their common virtual core, the so-called "latent variable" is expected (see [42]). Most commonly coefficient alpha according to CRONBACH (first: 1951 [46]) is used for assessing internal consistency of scales.

Table 2: Reliability of the COPSOQ scales on psychosocial factors at work, overall means and SD

Scale	COPSOQ DK		COPSOQ D				
	Long version: N items (alpha)	Shortened version: N items (alpha)	Long version: N items (alpha)	Shortened version: N items (alpha)	Long version: ICC	Shortened version: ICC	Long version overall means (SD)
Demands							
Quantitative demands	7 (0.80)	4 (0.65)	7 (0.82)	4 (0.69)	0.39	0.36	54 (17)
Cognitive demands	8 (0.86)	4 (0.78)	8 (0.83)	4 (0.71)	0.38	0.39	72 (15)
Emotional demands (*)	3 (0.87)	3 (0.87)	3 (0.82)	3 (0.82)	0.60	0.60	63 (21)
Demands for hiding emotions (*)	2 (0.59)	2 (0.59)	2 (0.65)	2 (0.65)	0.48	0.48	52 (21)
Sensorial demands	5 (0.70)	4 (0.66)	5 (0.74)	4 (0.72)	0.37	0.37	82 (17)
Influence and development							
Influence at work	10 (0.83)	4 (0.73)	10 (0.76)	4 (0.64)	0.24	0.30	53 (15)
Degree of freedom at work (*)	4 (0.68)	4 (0.68)	4 (0.78)	4 (0.78)	0.47	0.47	46 (25)
Possibilities for development	7 (0.82)	4 (0.75)	7 (0.80)	4 (0.73)	0.36	0.40	67 (15)
Meaning of work (*)	3 (0.77)	3 (0.77)	3 (0.82)	3 (0.82)	0.60	0.60	77 (19)
Workplace commitment (*)	4 (0.74)	4 (0.74)	4 (0.72)	4 (0.72)	0.39	0.39	54 (19)
Interpersonal relations and leadership							
Predictability (*)	2 (0.78)	2 (0.78)	2 (0.75)	2 (0.75)	0.60	0.60	52 (21)
Role clarity (*)	4 (0.77)	4 (0.77)	4 (0.83)	4 (0.83)	0.59	0.59	77 (16)
Role conflicts (*)	4 (0.72)	4 (0.72)	4 (0.79)	4 (0.79)	0.49	0.49	47 (20)
Quality of leadership	8 (0.93)	4 (0.87)	8 (0.94)	4 (0.89)	0.66	0.68	53 (23)
Social support (*)	4 (0.74)	4 (0.74)	4 (0.80)	4 (0.80)	0.50	0.50	65 (20)
Feedback (*)	2 (0.64)	2 (0.64)	2 (0.58)	2 (0.58)	0.40	0.40	40 (21)
Social relations (*)	2 (0.65)	2 (0.65)	2 (0.68)	2 (0.68)	0.51	0.51	49 (30)
Sense of community (*)	3 (0.80)	3 (0.80)	3 (0.79)	3 (0.79)	0.56	0.56	76 (17)
Additional scales							
Job insecurity (*)	4 (0.61)	4 (0.61)	4 (0.67)	4 (0.67)	0.33	0.33	29 (21)
Strain (effects, outcomes)							
Job satisfaction	7 (0.84)	4 (0.75)	7 (0.79)	4 (0.69)	0.35	0.36	62 (14)
Behavioral stress symptoms	8 (0.79)	4 (0.65)	8 (0.90)	4 (0.85)	0.53	0.58	28 (20)
Cognitive stress symptoms (*)	4 (0.85)	4 (0.85)	4 (0.87)	4 (0.87)	0.63	0.63	29 (19)

(*) scale unchanged in shortened version

But, due to the fact that Cronbach's alpha is influenced by the number of items in the scales - long scales usually have a higher alpha [47] -, the intra-class correlation (ICC) as a pure measurement of homogeneity was additionally computed. For the purpose of group comparisons (like in this COPSOQ study) scale reliabilities with alpha >0.7 (besides other criteria) are seen as an indication for the

suitability of the scale (e.g. [43]); however the (not non controversial) ISO 10075-3 fixes >0.8 as threshold value. Table 2 gives the reliability values of the scales on psychosocial work load and strain in the original Danish COPSOQ study (N=1858) and in the German main study (N=2561). The reliability coefficients on base of the shortened Danish version are given, too. (Please note:

the final shortened German version (see below) differs in some points from the shortened Danish version used here.)

Furthermore, overall means (range 0-100, and standard deviations) for all scales of the long version are documented in Table 2.

Obviously, the results of the German study were very similar to those in the Danish study. Taking Cronbach's alpha >0.7 as a threshold value, in the long version five out of 22 scales in the Danish and four in the German study missed this criteria. The same applied for the shortened version of the questionnaire.

Independently from country and version of the questionnaire especially short scales (2-4 items) were affected by low reliabilities in terms of Cronbach's alpha. This reflects the known dependency of this coefficient from the number of items - which is not the case for the ICC.

Scales taken from other sources than the Danish COPSQ offer limited possibilities for such comparisons with external reliability coefficients, since there are major differences concerning structure of the study, research methods and composition of the samples. Nevertheless, contrasting the reliability coefficients of the German COPSQ study to the respective original studies in the following table may give an impression, if the scales' properties are similar. Furthermore, overall means (range 0-100, and standard deviations) for all additional scales are documented in Table 3.

All additional scales showed very high reliabilities, some of them even better than in the original studies. This is true for Cronbach's alpha (all around 0.9) as well as for the ICC (all values between 0.6 and 0.7).

3.2.6 Generalisability

In addition to Cronbach's alpha and ICC, G-Coefficients in the single-facet design [39] were computed for the whole sample to test the generalisability of the item information [40]. In our case it was tested, to what extent the overall results remain stable when tested separately for two age groups (median-split), both sexes and four selected professional groups. Only little differences in age or sex were found in this sub-group-analysis. There were differences concerning the professional groups but with no consistent pattern that would raise doubts on the general appropriateness of the instrument for all professional groups (for further details see study report [24]).

3.2.7 Construct validity and factorial validity

Main aspects of validity besides content validity (see above) are construct validity and criterion validity respectively factorial validity.

To test the construct validity, it was analysed if the scales were associated with each other as postulated by theory: Scales thought to measure similar constructs should be positively correlated with each other, and conversely, scales that are believed to measure different issues, should show no or only weak correlations.

The four distinctive main constructs of the COPSQ (groups of scales, see Figure 1) under the terms "demands", "influence and possibilities for development", "interpersonal relations and leadership" and "strain (effects/outcomes)" were tested in this manner.

The correlation analysis was followed by an exploratory factor analysis, to test how well the four textual topics could be empirically replicated with the data of the German COPSQ study. Further analysis performed (confirmatory factor analysis, structural equation modelling) are documented in the project report [24].

The internal criterion validity (factorial validity, respectively) was analysed by examining the coherences theoretically postulated or assumed of the psychosocial work factors with the outcome parameters: stress parameters are supposed to have higher values when the demands are high respectively when the influence at work and the social support are low.

Bivariate correlations among related constructs

Concerning the bivariate correlations among presumably related scales (each analysis done separately within the four subareas), the resulting values were mostly low to moderate (ca. 0.2-0.5) as postulated. Thus, these scales are measuring similar issues without being redundant. In addition, the results in the German study were very close to those of the original Danish study. Exceptions were found for two scales: the scale degree of freedom at work (concerning breaks and vacations) was only weakly affiliated to the presumed sector influence and development at work; the same was the case for the scale social relations at the workplace which was only weakly correlated with some of the other scales in the field interpersonal relations and leadership.

Among the scales measuring strain (effects/outcomes) high correlations between work ability (WAI) and general health were found ($r=0.73$) as well as between personal burnout (CBI) and behavioural stress symptoms ($r=0.74$). This indicates a redundancy of these two pairs of scales used.

Factor analysis of the COPSQ scales

In order to test if the four theoretically defined areas of psychosocial factors at work can be replicated with the empirical data, an exploratory factor analysis was performed with the COPSQ scales (principal component analysis (PCA), varimax-rotation, listwise deletion of missing cases, number of factors to be extracted set to four). Scales included for the four theoretical areas were: demands (5 scales), influence and development (5 scales), interpersonal relations and leadership (8 scales), strain (effects/outcomes) (7 scales). The 4-factor solution explained 52.7% of the total variance which is a good result considering the reduction from 25 parameters to four factors. The communalities of the single parameters were between 0.30 and 0.77, with one exception (scale: social relations: 0.25). Altogether the exploratory factor

Table 3: Reliability of the additional scales on psychosocial factors at work, overall means and SD

Scale	Original studies: N items (alpha)	COPSOQ D: N items (alpha)	COPSOQ D: ICC	COPSOQ D: overall means (SD)
Additional scales				
Work-family (privacy) conflict [31]	5 (0.88)	5 (0.92)	0.70	45 (28)
Procedural justice [32]	4 (0.80)	4 (0.90)	0.69	46 (26)
Relational justice [32]	4 (0.90)	4 (0.88)	0.66	65 (23)
Complaints, outcomes				
Copenhagen Burnout Inventory (CBI); personal burnout [35]	6 (0.80)	6 (0.91)	0.61	42 (19)
Satisfaction with life scale [28]	5 (0.87-0.88)	5 (0.90)	0.65	64 (19)

analysis fitted the underlying theoretical construct very well: only four of the 25 scales had their highest factor loading on another factor than the one theoretically postulated and only three parameters showed second loadings above 0.3.

Including the additional aspects work-privacy conflict, procedural justice, relational justice, mobbing and intention to leave in the analysis resulted in a similar model (30 parameter, four factors, variance explained: 50.5%). Work-privacy conflict was mainly associated with the demand factor, relational justice and mobbing belonged to the factor interpersonal relations and leadership, procedural justice belonged to the factor influence and possibilities for development and intention to leave was mainly associated to the outcome factor.

Criterion validity, regression models on the outcomes

Regression analyses of the work load scales on the eight outcome parameters on health, well being and work satisfaction were conducted for two reasons:

- methodically: to find out the most important genuine predictors for the particular outcome parameter (criterion validity) in order to build parsimonious models and
- practically: identification of those scales which do not add additional explanatory power and therefore are abdicable.

Because of the numerous scales and models involved it was necessary to present the results in a compressed manner. Table 4 sums up the central findings: column two: explained variance (R^2 , determination coefficient) of the regression models of all 18 COPSOQ work load scales (demands: DEM, influence and development: INF, and interpersonal relations and leadership: REL) on the particular outcome parameter. Column three: explained variance R^2 when the five supplementary aspects (SUP): work-privacy conflict, procedural justice, relational justice, job insecurity and mobbing were included in the models. Columns four and five: explained variance R^2 and names of the first five predictors included into the model (in order of the integration; method: forward stepwise).

First of all it becomes obvious that the outcome factor job satisfaction could be explained better (up to 59% of explained variance) through the stress factors than any of the other outcomes: the six health and life quality related parameters and the new outcome parameter intention to leave. The - statistically speaking - worst prediction using psychosocial work load was found for the factor general health. The comparisons of the R^2 -values for every criterion in the columns two to four showed, that the integration of supplementary aspects besides the 18 COPSOQ scales improved the models significantly (with the exception of job satisfaction). Thus, these additional five parameters explained an additional portion of variance in the outcomes and are valuable supplies of the COPSOQ questionnaire.

In each of the eight models one or more representatives of this group (SUP) was involved as a predictor in prominent place. The work-privacy conflict scale acted as a predictor for all eight outcomes, the COPSOQ scale job insecurity was comprised in six models, the predictor mobbing was in one case among the five most important predictors.

Out of the 18 original COPSOQ scales the scales meaning of work and sense of community were most relevant with being an important predictor in six models out of eight. The scales emotional demands and possibilities for development were incorporated in three models each as one of the most important factors. Role conflicts was included in two models, quantitative demands, workplace commitment, predictability, role clarity and quality of leadership in one model each - the latter as the most important predictor for job satisfaction. On the other hand, eight (from 18) original COPSOQ scales and two out of five additional scales never made it in the "charts" of the most important predictor parameters.

3.2.8 Diagnostic power

Concerning the diagnostic power of the instrument it was tested if the COPSOQ scales were able to replicate known or plausible differences in work load and stress profiles of different professional groups. The results showed, that some of the work factors are obviously dependent on the

Table 4: Regression models on the outcomes (multiple regression)

Criterion/ Predictors	R ² : 18 COPSOQ scales	R ² : 18 COPSOQ scales + 5 additional aspects	R ² : 5 most important aspects	5 most important predictors
Job satisfaction	0.57	0.59	0.52	REL: Quality of leadership (+) INF: Meaning of work (+) REL: Sense of community (+) REL: Predictability (+) SUP: Work-privacy conflict (-)
Work ability (WAI)	0.26	0.30	0.26	SUP: Work-privacy conflict (-) INF: Meaning of work (+) SUP: Job insecurity (-) REL: Sense of community (+) DEM: Emotional demands (-)
General health	0.13	0.17	0.15	SUP: Work-privacy conflict (-) INF: Poss. for development (+) REL: Sense of community (+) SUP: Job insecurity (-) SUP: Mobbing (-)
CBI: personal burnout	0.31	0.37	0.32	SUP: Work-privacy conflict (+) INF: Poss. for development (-) DEM: Emotional demands (+) DEM: Quantitative demands (+) INF: Meaning of work (-)
Behavioral stress	0.30	0.39	0.35	SUP: Work-privacy conflict (+) INF: Meaning of work (-) SUP: Job insecurity (+) REL: Sense of community (+) REL: Role conflicts (+)
Cognitive stress	0.21	0.24	0.19	SUP: Work-privacy conflict (+) REL: Role clarity (-) REL: Sense of community (-) SUP: Job insecurity (+) REL: Role conflicts (+)
Satisfaction with life (SWLS)	0.18	0.24	0.22	INF: Meaning of work (+) SUP: Work-privacy conflict (-) SUP: Job insecurity (-) INF: Poss. for development (-) REL: Sense of community (+)
Intention to leave	0.26	0.30	0.27	INF: Meaning of work (-) SUP: Work-privacy conflict (+) INF: Workplace commitment (-) DEM: Emotional demands (+) SUP: Job insecurity (+)

profession, whereas this was not the case in other spheres. Expectedly the mean value for emotional demands for instance was higher in the medical profession and for nursing staff compared to the average of all employees or to office clerks for instance. On the other hand, the scale sense of community showed almost no differences between the professional groups, which was expected too, because this aspect is less determined by the profession but more importantly by the concrete situation at the workplace.

3.3 Proposal of a shortened questionnaire

One important practical goal of this study was the construction of a shortened measuring instrument to be used in future. The main criterion was to achieve a maximum of shortening with a minimum loss of quality. Another goal was to stay as close as possible at the shortened Danish original version in order to facilitate international comparisons further on. For this reason we decided to

Table 5: Question catalogue of the shortened German COPSOQ

Scale (resp. single item)	Origin	N Items	Question Number
Demands			
Quantitative demands	COPSOQ	4	B1: 1-4
Emotional demands	COPSOQ	3	B1: 5-7
Demands for hiding emotions	COPSOQ	2	B1: 8, 9
Work-privacy conflict	Netemeyer	5	B2: 1-5
Influence and development			
Influence at work	COPSOQ	4	B3: 1-4
Degree of freedom at work	COPSOQ	4	B3: 5-8
Possibilities for development	COPSOQ	4	B4: 1, B5: 1-3
Meaning of work	COPSOQ	3	B5: 4-6
Workplace commitment	COPSOQ	4	B5: 7-10
Interpersonal relations and leadership			
Predictability	COPSOQ	2	B6: 1-2
Role clarity	COPSOQ	4	B6: 3-6
Role conflicts	COPSOQ	4	B6: 7-10
Quality of leadership	COPSOQ	4	B7: 1-4
Social support	COPSOQ	4	B8: 1-4
Feedback	COPSOQ	2	B8: 5-6
Social relations	COPSOQ	2	B8: 7-8
Sense of community	COPSOQ	3	B8: 9-11
Mobbing (single item)	BIBB/IAB	1	B8:12
Additional scales			
Job insecurity	COPSOQ	4	B9: 1-4
Strain (effects, outcomes)			
Intention to leave (single item)	NEXT	1	B10
Job satisfaction (long)	COPSOQ	7	B11: 1-7
General health	EQ-5D	1	B12
Copenhagen Burnout Inventory (CBI), scale: personal burnout	Borritz/ Kristensen	6	B13: 1-6
Cognitive stress symptoms	COPSOQ	4	B14: 1-4
Satisfaction with life scale (SWLS)	Diener	5	B15: 1-5
Sum of items		87	

keep some scales in the shortened German instrument although they were seen critically from a psychometric point of view. As a result of the analysis, we propose the following question catalogue for the shortened German COPSOQ (Table 5). The questionnaire itself (in German) is available from the project report [24] or at <http://www.copsoq.de/>.

The shortened instrument assesses the broad spectrum of psychosocial factors at work with 87 items in 25 scales; the long version of the German COPSOQ study used 157 items to measure 31 constructs. Multiple regression

models to predict the outcomes with the scales of the shortened version (as shown in Table 4 with the long scales) showed a very limited loss of measuring quality at least under the aspect of criterion validity. The explained variance decreased only slightly for all eight outcome factors, in addition the models remained almost unchanged (lists of most important predictor parameters). The standard-questionnaire of the shortened COPSOQ version is supplemented with some few sociodemographic variables. For general assessments we propose to limit these to age, sex, and profession. In addition some sup-

plementary aspects may be included for use in specific professional groups. Furthermore, an open question for suggestions of the employees concerning the psychosocial work situation at the end of the questionnaire should be considered.

Since 2005 the shortened questionnaire finds a widespread use as a paper & pencil questionnaire and as an online tool. In a cooperation model between science (FFAS) and companies or organisations using the , data of new surveys are added to a dynamically growing database with profession-specific profiles of psychosocial factors at work. In return, companies get a comparison (“benchmarking”) of their results compared to the external data in the database which facilitates the interpretation of their results and the implementation of improvement measures.

4. Discussion and summary

In a broad survey (N=2561), the Danish Copenhagen Psychosocial Questionnaire (COPSOQ) on psychosocial factors at work was tested in Germany. Main goal besides the detailed evaluation of the measuring quality was the construction of a shortened instrument with still good measurement properties. According to ISO 10075-3 the COPSOQ is a screening instrument on level 2, placed between instruments for orientation (level 3) and those for exact measures (level 1). The purpose of an employee survey using the COPSOQ should therefore be the detection of possibilities for improvements of the psychosocial work situation in the company (for instance by the internal comparison of sub-units or by external comparison with similar companies and professional groups).

The extensive psychometric testing of the COPSOQ showed the following results:

Objectivity: The survey procedure as well as the analysis are standardised and therefore objective in a statistical sense. With variation in the survey techniques in the future (telephone interview, online interview, internal company survey) possible method-effects must be examined.

Practicability/acceptance: Both, companies (implementation of the survey) and employees (filling out the questionnaires) stated that they had no difficulties with the instrument. This is backed with statistical test criteria (missing values etc.). The questions seem to be presented in an understandable way and are sensitive for differentiation. The main problem with the (long) questionnaire is its length, why a shortening considering the criteria of measuring quality was performed.

Content validity: It is not possible to test the content validity with a standardised statistical method. However, the broad spectrum of psychosocial aspects in the COPSOQ as well as the takeover of most questions and scales from already validated instruments indicates a high content validity. Furthermore we have attempted to fill possible gaps regarding content in the German version by integrating additional scales. While shortening the questionnaire, it was emphasised to retain a broad

spectrum of topics. The national and international experiences in the next years will show if new aspects have to be integrated or existing aspects need to be inquired in a more detailed, advanced psychometrical way (e.g. mobbing as a scale which is so far only assessed with a single item).

Scale reliability: Almost all scales used reached a Cronbach’s alpha of >0.7 , about half of them exceeded the limit of 0.8 which is stated in ISO 10075-3. Shortening scales in the new questionnaire of course is accompanied by a loss of scale reliability in Cronbach’s alpha. But in our point of view this is no big issue as long as the ICC as another reliability criteria remains stable and - more importantly - as long as construct and criterion validity of the instrument are not decreased considerably. After this examinations we propose the use of the shortened scales in almost all cases (exception: job satisfaction).

But, since measuring properties of the scales are stated in a comprehensive and transparent way, future users are free to decide themselves, which scales respectively which version (long or shortened) they want to use.

Generalisability: The generalisability of the measured scale reliabilities was tested in different subgroups. There were only few differences in terms of age and sex but some concerning selected professional groups. Since no systematic weaknesses in measurement qualities for specific professions were found, the suitability of the COPSOQ as an instrument for all professions seems not to be threatened. However, experiences for the professions in the production sector are quite limited yet.

Construct validity: The analysis of the construct validity with the sales of the long and shortened version showed that the four theoretically defined areas demands, influence and possibilities for development, interpersonal relations and leadership and strain (effects, outcomes) can be confirmed by correlation analysis as well as by factor analysis.

Criterion validity: The analysis using the workplace factors as causes and the outcome parameters as consequences in regression models showed a correspondence of the theoretically postulated relations and the empirical reality. As expected, job satisfaction as the criterion closest to work related factors was predicted better through these factors than health and quality of life related outcome factors. In the prediction models some scales were found to be dispensable in terms of the prediction of the outcomes (analysis of risk factors at work).

Shortened version: The measurement qualities of the scales were one criterion to decide if a scale was included into the shortened questionnaire. Another criterion for the decision was sort of a cost-benefit consideration between the number of items (costs) and the increase of the measuring quality (benefit). For this reason the single question concerning general health is included instead of the rather long item catalogue of the WAI, although the latter demonstrated a slightly better measurement quality. The third criterion was the preservation of the broad content of psychosocial factors in the questionnaire; that is why some short scales were kept which do not have a

direct impact on the outcomes and why some scales were kept with critical measurement characteristics. Another advantage of this procedure is the preservation of the comparability with other international COPSOQ surveys. The performed analysis could show that the proposed shortened version of the questionnaire almost reaches the same level of criterion validity as the long version.

Outlook

The main advantage of the COPSOQ is its generic usability in all sorts of professions and industries. Profiles of psychosocial factors at work can be established and compared for different professional groups. Since 2005 the shortened version is used to build up a broad database with professional specific reference data on psychological factors at work. For organisations using the COPSOQ, the supplementary external comparison of their data (benchmarking) is valuable in terms of a better evaluation of their own results. For the work sciences this job exposure matrix offers manifold insight possibilities in the distribution of stress and demands in distinct professional groups or at distinct tasks and regarding the effect mechanisms of stress on strain.

Another task for the future lies in the international cooperation for the development of contents of the COPSOQ and measuring techniques (for instance integration of new aspects, expansion/improvement of short scales), but also in the international comparison of psychosocial stress values.

Notes

Conflicts of interest: none declared

Funding: The study was funded by the Federal Institute for Occupational Safety and Health (BAuA/FIOSH).

Acknowledgements

The authors would like to thank:

- Federal Institute for Occupational Safety and Health (FIOSH; project supervisor: Beate Beermann) for the financial support of the project (F1885, Fb1058),
- the “father” of COPSOQ, Tage S. Kristensen, Copenhagen for his sedulous support with so many questions;
- Markus Wirtz, Freiburg for his competent methodological support with the complex computing regarding generalisability and the structural equation models, and
- all participating companies and especially all employees for taking part in the survey.

References

1. presse.dak.de Hamburg. DAK Gesundheitsreport. 2005 [cited 2006 February 13]. Available from: <http://www.presse.dak.de/ps.nsf/sbl/38A5A5A6BBF15309C1256FE0005578E2>

2. Kuhn K. Kosten arbeitsbedingter Erkrankungen. In: Gesundheitsschutz in Zahlen 2000. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin; 2002. (Amtliche Mitteilungen 17, Sonderausgabe). p.12-21.
3. Richter G. Psychische Belastung und Beanspruchung. Stress, psychische Ermüdung, Monotonie, psychische Sättigung. Bremerhaven: Wirtschaftsverlag NW; 2000. (Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Fa 36, 3. überarbeitete Auflage).
4. Cox T, Griffiths A, Rial-González E. Research on work-related stress. Luxembourg: Office for Official Publications of the European Communities; 2000.
5. Harrach A. Arbeitswissenschaftliche Psychosomatik - arbeitsbedingte psychische und psychosomatische Störungen. In: Teske U, Witte B, editors. Prävention arbeitsbedingter Erkrankungen. Band 2: Gesundheitliche Auswirkungen und Erkrankungsschwerpunkte. Hamburg: VSA; 2000. p. 51-103.
6. Karasek RA. Job demands, job decision latitude and mental strain: Implications for job redesign. *Admin Sci Quart.* 1979;24:285-308.
7. Johnson JV. Control, collectivity and the psychosocial work environment. In: Sauter SL, Hurrell Jr JJ, Cooper CL, eds. Job control and worker health. Chichester: Jon Wiley & Sons; 1989.
8. Siegrist J. Adverse health effects of high effort - low reward conditions at work. *J Occupat Health Psychol.* 1996;1:27-43.
9. Siegrist J. A theory of occupational stress. In: Dunham J, editors. Stress in the workplace. London/Philadelphia: Whurr Publishers; 2001. p. 52-66.
10. Ertel M. Möglichkeiten und Grenzen bei der Erfassung psychischer Belastungen in der Arbeitswelt. In: Flake C, Freigang-Bauer I, Gröben F, Wenchel K-T. Psychischer Stress in der Arbeitswelt. Erkennen - mindern - bewältigen. Eschborn: RKW; 2001. p. 32-3.
11. Schmidtke H. Vom Sinn und Unsinn der Messung psychischer Belastung und Beanspruchung. *Z Arbeitswiss.* 2002;56:4-9.
12. Nachreiner F. Über einige aktuelle Probleme bei der Erfassung, Messung und Beurteilung psychischer Belastung und Beanspruchung. *Z Arbeitswiss.* 2002;56:10-21.
13. Pohlandt A, Jordan P, Rehnisch G, Richter P. REBA - ein rechnergestütztes Verfahren für die psychologische Arbeitsbewertung und -gestaltung. *Z Arb Organisationspsychol.* 1996;40:63-74.
14. Windel A. Entwicklung und Aufbau des Screening-Instruments zur Bewertung und Gestaltung von menschengerechten Arbeitstätigkeiten - SIGMA. In: Benda HV, Bratge D. Psychologie der Arbeitssicherheit. 9. Workshop 1997. Heidelberg: Asanger; 1998. p. 285-9.
15. Nickel P, Eilers K, Seehase L, Nachreiner F. Zur Reliabilität, Validität, Sensitivität und Diagnostizität von Herzfrequenz- und Herzfrequenzvariabilitätsmaßen als Indikatoren psychischer Belastungen. *Z Arb wiss.* 2002;56:22-36.
16. Richter G. Psychologische Bewertungen von Arbeitsbedingungen. Screening für Arbeitsplatzinhaber - BASA. Bremerhaven: Wirtschaftsverlag NW; 2001. (Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: F 1493)
17. Prümper J, Hartmannsgruber K, Frese M. KFZA. Kurz-Fragebogen zur Arbeitsanalyse. *Z Arb Organisationspsychol.* 1995;39:125-32.
18. Rimann M, Udris I. Subjektive Analyse: Der Fragebogen SALSA. In: Strohm U, Ulich E, editors. Unternehmen arbeitspsychologisch bewerten. Ein Mehrebenen-Ansatz unter besonderer Berücksichtigung von Mensch, Technik und Organisation. Zürich: vdf Hochschulverlag; 1997.

19. Kristensen TS, Borg V. AMI's spørgeskema om psykisk arbejdsmiljø. Copenhagen: National Institute of Occupational Health; 2000.
20. Kristensen TS, Hannerz H, Høgh A, Borg V. The Copenhagen Psychosocial Questionnaire (COPSOQ) - a tool for the assessment and improvement of the psychosocial work environment. *Scand J Work Environ Health*. 2005;31:438-449.
21. Setterlind S, Larsson G. The stress profile: A psychosocial approach to measuring stress. *Stress Medicine*. 1995;11:85-92.
22. Marmot MG, Smith GD, Stansfeld S, Patel C, North F, Head J, White I, Brunner E, Feeney A. Health inequalities among British civil servants. The Whitehall II Study. *Lancet*. 1991;337(8):1387-93.
23. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. The Job Content Questionnaire (JCQ): An Instrument for internationally comparative assessments of psychosocial job characteristics. *J Occupat Health Psychol*. 1998;3(4):322-55.
24. Nübling M, Stöbel U, Hasselhorn H-M, Michaelis M, Hofmann F. Methoden zur Erfassung psychischer Belastungen - Erprobung eines Messinstrumentes (COPSOQ). Bremerhaven: Wirtschaftsverlag NW; 2005. (Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Fb 1058). Available from: www.copsoq.de
25. Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A. Work Ability Index. Helsinki: Institute of Occupational Health, Occupational Health Care; 1994. p. 19.
26. BAuA, editors. Europäische Erfahrungen mit dem Arbeitsbewältigungsindex (Work Ability Index). Bremerhaven: Wirtschaftsverlag NW; 2002. (Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Tb 126).
27. EuroQol Group, editors. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199-208.
28. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction with Life Scale. *J Pers Assess*. 1985;49:71-5.
29. Zentralarchiv für Empirische Sozialforschung Köln, editors. Erwerb und Verwertung beruflicher Qualifikationen. BIBB/IAB-Erhebung 1998/99. Köln; Zentralarchiv für Empirische Sozialforschung; 1998/99. Machine-readable code book ZA 3379. Köln o.J. (no year given)
30. Hasselhorn H-M, Müller BH. Arbeitsbelastung und -beanspruchung bei Pflegepersonal in Europa - Ergebnisse der NEXT-Studie. In: Badura B, Schellschmidt H, Vetter C, editors. Fehlzeiten-Report 2004 - Gesundheitsmanagement in Krankenhäusern und Pflegeeinrichtungen. Berlin: Springer; 2004. p. 21-47.
31. Netemeyer RG, Boles JS, McMurrin R. Development and validation of Work-Family Conflict and Family-Work Conflict Scales. *J Appl Psychol*. 1996;81(4).
32. Kivimäki M, Elovainio M, Vahtera J, Ferrie JE. Organisational justice and health of employees: prospective cohort study. *Occup Environ Med*. 2003;60:27-34.
33. Elovainio M, Kivimäki M, Helkama K. Procedural justice, job control and occupational strain. *J Appl Psychol*. 2001;81:418-24.
34. Elovainio M, Kivimäki M, Vahtera J. Organizational justice: Evidence of a new psychosocial predictor of health. *Am J Pub Health*. 2002;92:105-8.
35. Borritz M, Kristensen TS. Copenhagen Burnout Inventory. Copenhagen Denmark: National Institute of Occupational Health; 1999.
36. Statistisches Bundesamt, editors. Demografische Standards, Ausgabe 1999. Wiesbaden 1999;(3) [cited 2006 Feb 13]. Available from: http://www.gesis.org/Methodenberatung/Untersuchungsplanung/Standarddemografie/dem_standards/demsta99.pdf
37. Paoli P. Second European survey on working conditions. Dublin: European Foundation for the Improvement of Living and Working Conditions; 1997.
38. Harkness JA, Schoua-Glusberg A. Questionnaires in translation. In: ZUMA, editors. ZUMA-Nachrichten Spezial Nr. 3: Cross-cultural survey equivalence. Mannheim: ZUMA; 1998.
39. Brennan RL. Generalizability theory. New York: Springer; 2001.
40. Cronbach LJ, Gleser GC, Nanda H, Rajaratnam N. The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: John Wiley; 1972.
41. Bortz J. Statistik für Sozialwissenschaftler. 5. Auflage. Berlin: Springer; 1999.
42. DeVellis RF. Scale development. Theory and applications, 1st ed. London: Sage Publications; 1991.
43. Wirtz M, Caspar F. Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen: Hogrefe; 2002.
44. Statistisches Bundesamt, editors. Demografische Standards, Ausgabe 2004. Wiesbaden; 2004 [cited 2006 Feb 13] Available from: http://www.gesis.org/Methodenberatung/Untersuchungsplanung/Standarddemografie/dem_standards/demsta2004.pdf
45. Statistisches Bundesamt, editors. Personensystematik. Klassifizierung der Berufe - Systematisches und alphabetisches Verzeichnis der Berufsbenennungen, Ausgabe 1992. Stuttgart: Metzler-Poeschel; 1992.
46. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-334.
47. Cortina JM. Coefficient alpha? An examination of theory and applications. *J Appl Psychol*. 1993;78:98-104.

Corresponding author:

Matthias Nübling

FFAS: Freiburg Research Centre of Occupational and Social Medicine, Bertoldstraße 27, 79098 Freiburg, Tel.: +49 (0)761-82526, Fax.: +49 (0)761-83432
ffas.freiburg@t-online.de

Please cite as

Nübling M, Stöbel U, Hasselhorn HM, Michaelis M, Hofmann F. Measuring psychological stress and strain at work: Evaluation of the COPSOQ Questionnaire in Germany. *GMS Psychosoc Med*. 2006;3:Doc05.

This article is freely available from

<http://www.egms.de/en/journals/psm/2006-3/psm000025.shtml>

Copyright

©2006 Nübling et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>). You are free: to Share — to copy, distribute and transmit the work, provided the original author and source are credited.