Construction of scores

Extrinsic effort is measured by five or six items respectively that refer to demanding aspects of the work environment: ERI1-ERI6. The 5-item version excluding physical load (item ERI5) has been found to be psychometrically appropriate in samples characterized predominantly by white collar jobs whereas the 6-item version was appropriate in blue collar samples and occupational groups with manual workers.

All questions refer to the present resp. last occupation and subjects are asked to indicate in how far the items reflect their typical work situation. The rating procedure is defined as follows with higher ratings pointing to higher efforts: (1) does not apply; (2) does apply, but subject does not consider herself or himself distressed; (3) does apply and subject considers herself or himself somewhat distressed; (4) does apply and subject considers herself or himself distressed; (5) does apply and subject considers herself or himself very distressed (see figure below).

A sum score of these ratings is computed as the unidimensionality of the scale has been documented. A total score based on the five items measuring extrinsic effort varies between 5 and 25 (or 6 and 30 with 6 items). The higher the score, the more extrinsic effort at work is assumed to be experienced by the subject.

Reward is measured by eleven items (items ERI7-ERI17). We postulate a three-factorial structure of the construct of occupational reward with a first factor defined by financial and status-related aspects (ERI11, ERI14, ERI16-ERI17), a second factor defined by esteem rewards (ERI7-ERI10, ERI15), and a third factor (ERI12-ERI13) defined by the gratification of job security. The three dimensions are assumed to load
on one latent factor (reward). Therefore, a second-order factor analysis is expected to define a one-dimensional scale.

The rating procedure is performed in analogy to extrinsic effort (see figure below).

After variable recoding procedures (see below), lower ratings point to lower rewards. A sum score of these ratings is computed which varies between 11 and 55. The lower the score, the fewer occupational rewards are supposed to be received by the person.

To compute the ER-ratio, the effort score is put in the enumerator and the reward score in the denominator whereas the latter score has been multiplied by a correction factor before to adjust for the unequal number of items. The correction factor is 0.4545 if the enumerator contains five effort items (5/11), and it is 0.5454 if the enumerator contains six effort items (6/11).

Overcommitment is measured by a short version (items OC1-OC6) based on the original psychometric test containing 29 items. Items range from 1 (low) to 4 (high overcommitment) (see figure). Note that item OC3 has to be reversed. The scale score is computed by adding item values.

If you need additional information on coding please mail to: siegrist@uni-duesseldorf.de.

2. Statistical Analysis

It is the main purpose of statistical analyses to test associations of ERI at work with health. Different approaches have been developed to this end. Examples are: testing main effects and interaction terms of the two variables effort and reward (e.g. van Veghel et al. 2005, in press) or testing a composite variable with three or four categories (e.g. Bosma et al. 1998, Jonge et al. 2000).

However, based on theoretical considerations and empirical evidence, the following procedure is currently recommended:

Use the ER-ratio in addition to single components. While this ratio indicates a clear cut point (values > 1.0 indicating an imbalance between high effort and low reward), a reduction of information to this dichotomous variable seems inappropriate because of observed dose-response-effects of the ratio on health and because of large variation of scores across different samples. Therefore, we propose to compute a continuous ratio score. Based on this, you can either estimate the effects of the ratio per 1 SD increase or define tertiles or quartiles based on the score distribution. See for example

three countries of Central and Eastern Europe. Social Science & Medicine, 58, 8, 1475-1482.


For a sophisticated test of different interaction terms between the scales effort and reward see


Use the sum score of Overcommitment to analyse associations with health. A standard procedure has been so far to define the upper tertile as risk condition and to compare health effects between the risk and non-risk group.

According to the theoretical assumption, the combined effect of the ER-ratio and Overcommitment on health maybe estimated, using appropriate statistical techniques (interaction term to test effect modification).

A further refinement may consist in computation of effort-reward ratios based on the three sub scales of reward (see above) with respective correction factors. This may be useful e.g. in the context of intervention studies. Examples can be taken from:


Psychometric information

1. Scale reliability
Published data document satisfactory internal consistency in terms of Cronbach’s a (usually >.70) of the three scales of effort, reward and overcommitment.

Test-retest-reliability has been analysed in one study so far with satisfactory results. Correlation coefficients for the three scales varied between r=.60 and r=.44 in a one-year interval and between r=.41 to r=.47 in a two-year interval (van der Linden; personal communication 2005). More recently, multiple assessment of scales has been conducted, using 'Ecological Momentary Assessment' technique documenting a strong correlation between the summary estimate based on the self-administered questionnaire and the momentary estimate based on EMA technique (r = .46; Johnston, Beedie, Montgomery, Trott & Young; personal communication 2004).

2. Factorial structure
Exploratory and confirmatory factor analyses were conducted with satisfactory results. In particular, confirmatory factor analyses based on data from five international samples resulted in a good model fit for the unidimensional "effort" and "overcommitment" scales and the three factorial structure of the "reward" scale.
Goodness of fit was assessed by the GFI- and the AGFI-index, in addition to Chi-square and root-mean square residual. For details see


These results were replicated and further validated in a third order confirmatory factor analysis in Germany. Model fit indices were: GFI = .91; AGFI = .89; CFI = .90; RMSEA = .06. (see figure). In addition, concurrent model testing revealed the best model fit for the theoretically postulated structure of scales.


3. Convergent validity

Several studies have documented the independent explanatory power of the ERI scales compared to the scales of the demand-control-model (Karasek et al. 1998) despite the fact that the scales 'demand' and 'effort' show modest to strong correlations (ranging between r=.30 and r=.60; Calnan et al. 2001, Tsutsumi et al., 2001, among others).


4. Discriminant validity

Several investigations found statistically significant differences in mean scores of effort, reward and overcommitment according to gender and gender role orientation (Starke and Niedhammer 2001), age, socioeconomic status and occupational group (Siegrist et al 2001). For instance, effort decreases with age, while reward increases. In several, but not in all studies reward decreases with lower socioeconomic status. Overcommitment is more pronounced in women compared to men, in some studies, and also has more explanatory power in employed women. These and other findings point to discriminant validity of the scores.

5. Criterion validity

see 'Selected publications on research evidence'
6. Sensitivity to change over time

Importantly, studies are now underway to test change of ERI over time and to explore health effects of cumulative work stress. One published study documented significant changes of ERI scales during a one year observation period in association with downsizing in a Japanese company.