Appendix S1

SUPPLEMENTARY MATERIAL

Orthogonal Partial Least Square Regressions (OPLS)

Common methods such as logistic regression (LR) and multiple linear regression (MLR) can quantify the level of relationships of individual factors but disregard interrelationships among different factors and thereby ignore system-wide aspects. Moreover, such methods assume variable independence when interpreting the results, and there are several risks when considering one variable at a time. To account for our aims, the problems related to handling missing data (see below), and the risks associated with multicollinearity problems, we used advanced multivariate analyses (MVDA) for the regression analysis. Hence, using SIMCA-P+ (version 15; Umetrics, Sartorius Stedim Biotech, Umeå, Sweden), we applied advanced principal component analysis (PCA) for the multivariate correlation analyses of all investigated variables to detect multivariate outliers and orthogonal partial least square regressions (OPLS) for the multivariate regressions. These techniques do not require normal distributions of the variables. Note that the PCA of SIMCA-P+ differs considerably from the simpler version implemented, for example, in SPSS. PCA extracts and displays systematic variation in the data matrix. A cross-validation technique was used to identify nontrivial components (p). Variables loading on the same component p were correlated, and variables with high loadings but with opposing signs were negatively correlated. Variables with high absolute loadings were considered significant. The loading plot reports the multivariate relationships between variables. A corresponding plot reporting the relationships between subjects (i.e., t scores) was used to check for multivariate outliers. Outliers were identified using two methods: score plots in combination with Hotelling's T^2 and distance to model in the X space. No outliers were detected.

OPLS was used for the multivariate regression analyses of satisfaction with occupational performance at 1-year using the other variables presented (except occupational performance at 1-year) as regressors (X variables). OPLS is based on the assumption that the independent variables may be intercorrelated (in unknown ways) and takes advantage of this multicollinearity pattern. The variable influence on projection (VIP) indicates the relevance of each X variable pooled over all dimensions and Y variables, the group of variables that best explain Y. VIP(VIPpred) ≥1.0 was considered significant if VIP had a 95% jack-knife uncertainty confidence interval not equal to zero. *P*(corr) was used to note the direction of the relationship (positive or negative). This is the loading of each variable scaled as a correlation coefficient, and thus standardized in the range from −1 to +1. *P*(corr) is stable during iterative variable selection and comparable between models. Thus, a variable/regressor was considered significant when VIP >1.0. For each regression, we report R^2, Q^2, and the *P* value of a cross-validated analysis of variance (CV-ANOVA). R^2 describes the goodness of fit, the fraction of the sum of squares of all the variables explained by a principal component. Q^2 describes the goodness of prediction, the fraction of the total variation of the variables that can be predicted using principal component cross-validation methods. SIMCA-P+ uses the Non-linear Iterative Partial Least Squares (NIPALS) algorithm to handle missing data: maximum 50% missing data for variables/scales and maximum 50% missing data for subjects.