

More than half of prison employees work in shifts. Also nightwork is obligatory in certain occupational groups because guarding cannot be stopped, even during the nights. The average number of night shifts during the last month was five (range 0–18) and the number of evening shifts was four (range 0–15).

About one-fifth of the respondents had done overtime during the last 12 months. The average number of overtime hours was nine per month ($SD \pm 10.8$). These data are not, however, completely reliable. For instance some of the respondents had included the overtime they had done outside the prison in the number of hours of overtime reported.

Residence, family and travel to work

Almost one-third of the respondents lived on the prison grounds. About two-thirds had lived on the prison grounds at some time during their service.

The daily traveling from home to work and back took less than an hour for most of the prison employees (80 %). Only 3 % of all the respondents used more than 2 h for daily traveling. The time used for traveling cannot thus be considered as any major stress factor in the present study sample.

Because the sample of the study was formed mainly by the older age groups, the number of young children in the families was relatively low. Only less than half of the participants had children under 16 a of age still at home.

In light of the skewed age distribution, it can be assumed that the standard of living of the respondents was also probably fairly good. They had had, in general, a rather long professional career and therefore had been in a situation to acquire a satisfactory residence. On the other hand, in many families the children had already moved out on their own. Only 11 % of all the respondents lived in small apartments with one or two bedrooms, and one-third of the participants lived in big residences with more than three bedrooms. The number of rooms averaged 1.16 per person.

COLLECTION OF DATA

Methods of data collection

Two basically different types of methods were used for the data collection, questionnaires and laboratory measurements. One of the questionnaires was planned specifically for this study and included information mainly on the work conditions, job satisfaction, and stress, and the other questionnaire was a standard one used by the Institute of Occupational Health for many years to obtain information on the health history of the patients. The former is referred to in the text as the work and stress questionnaire and the latter as the health questionnaire.

Work and stress questionnaire

In the construction of the questionnaire used to study work and work conditions, one starting point was to cover all the theoretically important aspects of the work environment. The other important criteria were the specific features of the prison system and the prisons as a work environment. The latter aspect was particularly important because of the expectations in regard to the immediate practical applications of the results of the study.

The structure and contents of the work and stress questionnaire, which is presented in appendix IV¹ are the following: (i) individual and occupational background (items 1–19), (ii) family and living conditions (items 20–22, 55–57) (iii) characteristics of present work and organization (items 23–24), (iv) physical work environment (items 50 a and b), (v) shift work (items 51–54), (vi) income (items 58–61), (vii) social climate and contacts in work (items 62–69), (viii) special features of the care of prisoners and related attitudes (items 70–73, of which some could be included also in certain other more general dimensions), (ix) appreciation of work (items 74–77), (x) job satisfaction and some of its theoretically close determinants (items 78–88), (xi) affective state during

¹ If the appendix is not included in this issue, it is available from the author upon request.

the workday (items 89 a and b), (xii) symptoms of stress and resources (items 90—95), (xiii) alcohol consumption (items 96—99), (xiv) self concept (item 100), (xv) need for mental care (items 101—107).

The items concerning the characteristics of the employees' current work and work organization included the following dimensions that have been found theoretically to be important in regard to job satisfaction and stress: (i) specific demands on mental and physical functions and abilities (item 23), (ii) safety (items 24—25 and certain subclasses of items 70—71), (iii) variability (items 30, 34), (iv) demands on mental and physical activity (items 26—29, 49), (v) participation in decision making (items 28, 38, 46, 47), (vi) responsibility (item 31), (vii) information and feedback (items 36, 37, 39, 40, 48), (viii) role ambiguity and conflict (items 41—43), (ix) overall daily work load (item 32), and (x) work arrangements (33, 44, 45).

Most of the items on work and work conditions have been in use in various studies at the Institute of Occupational Health in Helsinki (158, 229, 243) and/or in questionnaires used in the clinical occupational health practices of the same institution. Some of the questions were also included in the Finnish part of the international study on automation (218) and a study on occupational health in the metal industry in Finland (188). The items had thus been previously applied quite widely in the Finnish language, and information was also available on their psychometric properties. Some of the items have originally been translated from other languages and reported in closely corresponding form (39, 268, various unpublished questionnaires of the Laboratory for Clinical Stress Research, Stockholm).

Among the questions on social climate and contacts in work, appreciation of work, and job satisfaction most of the items had been used in the aforementioned studies. Some of the questions on social contacts and the appreciation of work have been used in a closely corresponding form by Porter et al (209) and the questions on income by Kornhauser (146).

The perceived symptoms of strain were approached with a number of questions covering all typical psychological and bodily stress reactions commonly studied

in stress research. To add reliability to the measurements, various items were included to measure each reaction dimension. In addition to various studies with the same type of items (127, 158, 188, 216) the Institute of Occupational Health has used the items in clinical practice for diagnosing patients occupationally exposed to industrial chemicals. Thus, the items have been continuously reformulated through long-term diagnostic practice.

In addition to measuring the stress reactions, an effort was made to gather data on the resources and positive aspects of health. Most of these items indicated social activity, participation, and the affects and emotions in regard to social situations.

The measurement of positive health indicators is rare in the literature (142, 271). No earlier reports on the application of such questions were available in Finnish at the time of the planning of the present questionnaire. Some of the items included in the questionnaire were in use in the clinical diagnostic methods of the Institute of Occupational Health. Other items were in use in the Students' Health Center in Helsinki and have been reported since then (230). Part of the items were planned specifically for this questionnaire.

The particular characteristics of prisons as a work environment and prison service as work required the development of certain items specifically for this purpose since earlier studies were lacking. Many of these items were based on consultations with the representatives of the prison administration, eg, health and other personnel or trade union representatives. An important source was also the information gathered in a preliminary study on the working climate in and employee opinions on Finnish prisons carried out by one member of the planning group of this study (257). The items developed in the research groups were accepted in the final form by the labor unions active in the prisons.

Those items of the questionnaire on work and the work environment that have not been described were not used in this report and have been published elsewhere (135, 136).

The questionnaire also included a self-concept test with 23 items. Technically,

the items were based on the semantic differential method (202). This same test had previously been used in a study on a representative sample of Finnish people (248). It is a modified, shortened version of the method used for many years by the Department of Vocational Counseling of the Finnish Ministry of Labor in studies on self concept and vocational interests (100). The test is based on the self theories of personality. The reliabilities of the original 11 scales vary between 0.66 and 0.85 (100, 101). The 23-item version used in the present study included items that have been found to measure the following dimensions of the conscious personality: ego strength, dominance/submissiveness, introversion/extroversion and self-esteem.

The problems of measuring self concept are considerable. A central problem is conceptual validity. No self theory has been systematically empirically tested. More studies have been done on the basis of Rogers' theory than on any other. These have primarily concentrated on the differences between the ideal self and the self concept as a determinant factor of prognosis in psychotherapy or general adaptation. Many other central concepts of the theory have remained unstudied. No generally accepted definition of the concept "self" has even been presented. It is not certain whether the items in the self-concept methods comprise a representative sample of the dimensions of the self. Also, the prognostic validity of the self-concept methods has remained largely unstudied. At least exact information on the validity and reliability of the methods used is often lacking.

One possible source of error with these methods is, eg, a person's unwillingness or inability to evaluate himself truthfully. The answers may be given on the basis of a motive to be socially acceptable, or they may be defensive. Other modifying factors are a person's situational motivation, state of mood, security in regard to anonymity, etc. As in all questionnaire studies a personal response tendency may also affect the ratings. Not uncommon is also a difficulty to look at oneself as an object, which reflects a weak knowledge of one's self. In addition to the sources of error related to the respondent the instrument itself and its administration can affect the results (2, 279, 280).

Health questionnaire

The health questionnaire is a standard method used by the Department of Epidemiology and Biostatistics of the Institute of Occupational Health for many years. The part of the questionnaire used in the present study is presented in appendix V.² It includes items on health history and present symptoms and illnesses.

Laboratory examinations

The laboratory examinations included several measurements, of which the following are used in this report: blood pressure, serum cholesterol, height, weight, and electrocardiography. All of these indicators were based on single measurements, which means that their reliability could not be checked. As is often indicated, the reliability of the measurement of the indicators listed, except height and weight, may be unsatisfactory.

Field work

Information on the study was distributed before the collection of the data mainly by the information channels of the prison administration. The registered nurses of the institutions acted as the contact persons. An information seminar was held in the Institute of Occupational Health for the contact persons. A member from the research group visited most of the prisons and presented information on the study and its purposes.

The study questionnaires (health questionnaire and work and stress questionnaire) were sent to the contact persons in the prisons; they were distributed to the staff members by the health personnel. A week thereafter, the mobile clinic of the Institute of Occupational Health visited the outlying prisons to carry out the health examinations and laboratory tests. The questionnaires were returned to the personnel of the mobile clinic at the time of the clinical examination. The work and stress questionnaire was returned in a sealed envelope and forwarded to the research group.

² If the appendix is not included in this issue, it is available from the author upon request.

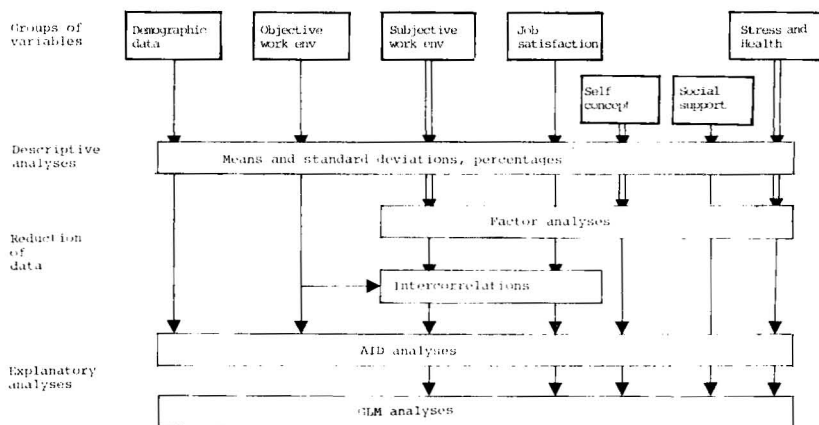


Fig 10. Schematic flow of the analysis of the data. (env = environment, AID = automatic interaction detector, GLM = general linear model)

The mobile clinic unit carried out the health examinations in most of the institutions. The staff members of the prisons in and around Helsinki came to be examined at the Institute of Occupational Health. This was the most practical arrangement because of the location of the research institute in Helsinki. Each participant personally received an epicrisis of his or her health status and the results of the laboratory examinations. At the same time possible suggestions for a follow-up health examination were given. The place of the follow-up examination was to be chosen by the participants themselves. The data were collected during the late fall of 1975 and the spring of 1976.

STATISTICAL TECHNIQUES

The statistical analysis of the data was divided into three parts: descriptive analyses, reduction of data, and explanatory analyses of selected indicators and variables (see fig 10).

The number of subjects in the analyses vary somewhat because of the omission of missing information, the proportion of which was relatively small. In the items with little missing data, these data were included in the middle category.

Descriptive analyses

Descriptive analyses were carried out as the first step of the analysis of the data.

They included the calculation of percentage distribution and arithmetic means with standard deviations by, eg, occupational groups and other subgroups of the sample.

The results of the descriptive analyses have been previously reported (135, 136, 242), and only summaries of these results are given in this report.

Data reduction

The reduction of the data was conducted first through an examination of the marginal distributions of the variables. Variables with extremely low variation were deleted from further analyses. (The marginal distributions of the variables used in the further analyses are given in appendices IV & V³). Because of the large number of variables, the interrelationships among variables in the same conceptual domain were then examined by a correlational multivariate technique, factor analysis, in order to examine the dimensionality of the variables for selecting the appropriate variables for explanatory analyses.

A factor analysis was conducted within each of the following groups of variables: (i) characteristics of work, (ii) characteristics of self concept, (iii) characteristics of stress reactions and health.

The first two factor analyses were carried out on groups of explanatory vari-

³ If the appendices are not included in this issue, they are available from the author upon request.

ables, ie, predictors, and the third one on the dependent variables of the study.

The factor analyses were based on the product-moment correlations among the variables. The missing data were omitted when the correlation coefficients were calculated. The factors were extracted by the principal axis technique, the highest correlation coefficient being used irrespective of its sign as the estimate of the communality of the variable. The maximum number of common factors was decided on the basis of a mathematical criterion recommended by Kaiser (131), according to which the sum of the squared factor loadings for each factor has to be at least one. The rotation of the extracted factors was conducted by the varimax method (103), which produces orthogonal, mutually uncorrelated rotated factors. The rotation was done for the mathematically defined maximum number of factors and, in addition, for the factors less than that amount. Of these results the solution with the most easily interpretable factors was used. In the factor analyses of the characteristics of work and of the characteristics of stress and health status, the rotated factor solutions with one factor less than the mathematically defined maximum number were used. These factor analyses are described in this study. The factor analysis of the characteristics of self concept has been described separately by Senne (242).

Explanatory analyses

Selection of variables

The selection of the variables for the explanatory analyses was based primarily on the factor analyses. Indicator variables of the factors were chosen on the basis of the rotated factor solutions.

The selection of the final variables on the basis of factor analysis helped to pinpoint the variables with a distinct and logical pattern of relationships with other variables in the same conceptual domain and therefore probably contributed to the selection of variables with higher reliability and validity in their domain. Since the communality of the variable is the lower limit of its reliability, variables

with the lowest communalities were avoided in their selection for the final analysis.

Automatic interaction detector analyses

The interrelationships between the selected explanatory and dependent variables were examined mainly by the automatic interaction detector (AID) technique (251). It was calculated separately for four dependent indicator variables of stress and health status. The purpose of these multivariate analyses was to determine which predictors had the closest relationships with each dependent variable, what the interactions among the predictors were like, and how great the explanatory power of the predictors was.

AID is applicable to a data set with one dependent variable and usually more than one predictor. The dependent variable has to be a quantitative variable or dichotomous. The predictors can be quantitative or qualitative.

AID proceeds stepwise by dividing the observations at each step into two groups according to that predictor and a cutting point in it which yields the maximum amount of variance explained. This predictor and its cutting point are determined empirically by the computer after the variance explained by a division into two groups at each possible cutting point is calculated. After the first step the observations are divided into subsets, and the number of later steps depends on the criteria specified by the user for the minimum amount of variance to be explained at each step and for the minimum number of observations in any subgroup. The analysis leads to groups of observations which are internally as homogeneous as possible and mutually as heterogeneous as possible. The arithmetic means of the final groups in the dependent variable usually vary a great deal, but there are also final groups with similar means in the dependent variable, although their composition is determined by a different mix of the predictors.

As its name suggests, an important feature of AID is its ability to reflect interactions among the predictors. It is to be emphasized, however, that AID treats, as applied in this study, all predictors in an identical fashion. The result of an AID analysis is thus a description of the inter-

relationships between the dependent variable and the predictors according to a specific statistical model, which may not reveal all relevant interrelationships among the variables. AID is more readily applicable to studies like the present one in which no explicit hypotheses are to be tested about the relations of the dependent variable with the explanatory ones. AID can be supplemented by additional multivariate analyses, eg, the multiple regression analysis (261).

Statistical tests of significance cannot be applied to the results of AID in the conventional way, since the analysis maximizes the variance in a specified direction in forming the subgroups of observations and, hence, some of the basic assumptions of significance tests do not hold. Significance tests can however be applied to assess the size of the differences or the explanatory ability of the model, but the conclusions drawn should relate to these questions only. Therefore, no results of statistical tests of significance are presented for the AID data.

The criteria used for the overall guidance of the AID analyses were as follows. Each dichotomous split had to explain at least 0.6% of the total variance of the dependent variable; the minimum size of a subgroup of observations was 20; and the maximum number of subgroups was 30.

The results of the AID analyses are depicted as AID trees, which describe visually the stepwise formation of the subgroups, the base predictor and its cutting point for splitting the group into two, the arithmetic mean of the subgroup in the dependent variable, and the number of observations (persons) in the subgroup.

Applications of the general linear model

After the application of the AID technique, the explanatory power of the predictors was analyzed further by the general linear model, which is called GLM in this report. This technique applies the general linear model with one dependent variable and one or more predictors to the analysis of their relationships (53). The predictors may be at any level of measurement, including nominal. The random error of the measurement of the

dependent variable must be normally distributed.

If all predictors are categorical, the GLM is an analysis of variance for unbalanced data without interactions, ie, the fitting and analysis of the "main-effects-only model" (238). The related computer program can be considered a further development of the multiple classification analysis (MCA), a computer program for multiple regression using categorical predictors (8). The GLM computer program (Nenonen, unpublished), developed at the Computing Center of the Social Insurance Institution, Finland, also performs a multiple regression analysis when all the predictors are measured at at least the level of an interval scale, ie, they are continuous. When there are both categorical and continuous predictors, the GLM performs an analysis of covariance. The GLM computer program was used in the present study in a stepwise fitting of predictors (238), which is not possible with the MCA computer program. The selection of the predictors was based on the reductions in sums of squares as indicators of the extent to which different predictors accounted for variation in the dependent variable.

An important feature of the GLM, as that of the MCA, is its ability to analyze all categories of a predictor as dichotomous variables without having to leave out one, as is necessary in ordinary regression analyses.

The output of the GLM computer program varies according to the level of measurement of the predictors, and it can be specified by the user. Those predictors with a multiple correlation differing from zero with a lower probability of chance (p) than any other combination of predictors were not included in the explanatory model.

In the present study, the GLM was mainly applied as an analysis of variance of categorical predictors, although these were actually ordinal or interval scales. They were categorized and treated like MCA predictors so as to make it possible to calculate adjusted means of the dependent variable in the categories of the predictors also.

The multiple correlation (R) of the linear combination of the selected predictors and the statistical significance of its difference

from zero were estimated. The explanatory power of each selected predictor in the model is described by its simple and partial correlation ratios (η) with the dependent variable. The statistical significance of the difference of both η s from zero was tested by the F-test, and the resulting levels of significance are reported. The squared partial η describes the proportion of variance estimated by this predictor and not estimated by the other predictors. In addition, the squared semipartial η is given. It reflects the proportion of variation explained when the effects of other predictors have been removed from the predictor in question but not from the dependent variable (53). The MCA does not calculate the partial correlations, semipartial correlations, and corresponding tests of statistical significance.

The arithmetic means and adjusted means of the dependent variable are given for groups of subjects by the category of each predictor. The adjusted means calculated by the GLM are statistically equivalent to the adjusted means of the MCA, and they "are estimates of what the mean would have been if the groups had been exactly like the total population in its distribution over all the other predictor classifications [p 101]" (8). The mathematical estimation of the adjusted means is different and more exact in the GLM computer program when compared with the MCA. The biserial correlation of each category of the predictor with the dependent variable, as well as the level of significance of the difference of the r_{bis} from zero, are also given.

The GLM was also applied to continuous predictors, ie, as a stepwise multiple regression analysis. In this case the output of the computer program includes the multiple correlation coefficient R , its square, and its level of statistical significance from zero. In addition, for each predictor selected to the model, the following statistics are given: standardized regression coefficient (β) and its standard error, partial and semipartial r^2 , and the t-ratio, which gives the level of the statistical significance of the β coefficient. The correlation coefficient with the dependent variable and the level of the statistical significance of the latter coefficient

calculated by the t-test is also given. The covariance ratio of the predictor is given to describe the multicollinearity of this predictor with the other predictors, and it is expressed as the proportion of the variation of the predictor which it shares with one or more other predictors.

The statistical tests were applied two-directionally. The following levels of statistical significance were used: $p \geq 0.05$, not significant (NS); $p < 0.05$; $p < 0.01$; and $p < 0.001$.

The output of the GLM computer program also included graphic plots of the residuals of the dependent variable, ie, the difference between the actual and predicted values of the subjects in the dependent variable, plotted against the actual values, on one hand, and against the predicted values, on the other. These charts have not been reproduced since they did not reflect any major hidden effects of predictors not included in the model.

A linear additive model, and correspondingly also the GLM, accounts for interactions among the predictors only if the interactions are included in the analysis as separate variables. The interactions among the predictors were examined before the GLMs were calculated by the AID technique. The AID results were used as one basis for selecting predictors for the analysis with the GLM, and it was not considered necessary to include the interactions between any original predictors as additional predictors in the GLM analyses.

Assessment of reliability

In the present study the assessment of the reliability of the variables was not straightforward because of the cross-sectional nature of the data and the lack of parallel indicators. An exception was formed by the sum scales that were composed of two or more original items of the questionnaire.

The assessment of the reliability of the sum scales was based on the assumption that all the items of the scale were parallel indicators of the trait in question. Hence, the coefficient of reliability was to reflect the internal consistency of the items of the sum scale. The coefficient α ,

which in the case of dichotomous items is known as the Kuder-Richardson Formula 20, was chosen as the coefficient of the reliability of the sum scales (199). Its calculation is based either on the inter-correlations of the items or on the variance of the items and the total scale and on the number of items in the total scale.

In view of the foregoing reasoning, a lower limit of the reliability of most of

the other variables was obtained on the basis of the results of the factor analyses. Since single variables were selected to represent factors in subsequent statistical analyses, a lower limit of the reliability of the selected variables was estimated as their highest intercorrelation with other variables loading highly (> 0.40) on the same factor and having low loadings (< 0.40) on other factors.