Journal of Occupational and Organizational Psychology (2003), 76, 427–450

www.bps.org.uk

Specific determinants of intrinsic work motivation, emotional exhaustion and turnover intention: A multisample longitudinal study

Inge Houkes*, Peter P.M. Janssen[1], Jan de Jonge[2] and Arnold B. Bakker[2]

[1]Maastricht University, The Netherlands
[2]Utrecht University, The Netherlands

This longitudinal study tested a theoretically derived pattern of specific relationships between work characteristics and outcomes. The research model proposed four central domains of the work situation (i.e. task characteristics, workload, social support and unmet career expectations) and three important psychological outcomes (i.e. intrinsic work motivation, emotional exhaustion and turnover intention). More specifically, it was hypothesized that intrinsic work motivation is primarily predicted by challenging task characteristics; emotional exhaustion is primarily predicted by a high workload and lack of social support; and turnover intention is primarily predicted by unmet career expectations. Furthermore, we hypothesized that (i) the research model is generalizable over samples; (ii) work characteristics at Time 1 influence outcomes at Time 2; and (iii) the proposed causal pattern of relationships holds over different occupational groups. These hypotheses were tested by means of self-report questionnaires among two samples (bank employees and teachers) using a full-panel design with two waves (one-year interval). Results showed that Hypothesis 1 was confirmed in both samples. Hypothesis 2 was confirmed in sample 1, but not in sample 2. In the latter sample, we found evidence for reverse causation. Hence, Hypothesis 3 could not be confirmed.

In the area of work and organizational psychology, many studies exist that combine work stressors with mental, physical and behavioural stress reactions, such as burnout, depression and psychosomatic diseases (cf. Cooper, 1998; Cooper & Payne, 1988; Parker & Wall, 1998; Schaufeli & Enzmann, 1998). In order to describe and explain the psychological well-being of people at work, several heuristic stress models (e.g. the Michigan model and the demand–control–support model; Johnson & Hall, 1988; Kahn & Byosiere, 1992; Karasek & Theorell, 1990) have been developed in which work stressors are related to employees' health. These models have yielded valuable insights regarding work stress experiences and outcomes (cf. Cooper, 1998), but have also been the subject of criticism regarding, for instance, specificity. Most of these models are based upon global theoretical frameworks that hardly lead to specific

*Requests for reprints should be addressed to Inge Houkes, Department of Health Organisation, Policy and Economics, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands (e-mail: I.Houkes@beoz.unimaas.nl).

hypotheses (cf. Cooper, 1998). In meta-theoretical terms (e.g., Walker & Avant, 1995), the above-mentioned models can be characterized as *conceptual* models. A conceptual model is highly abstract, includes multidimensional concepts, and provides perspectives for research, but is difficult to test.

In this light, Janssen, De Jonge, and Bakker (1999) and Houkes, Janssen, De Jonge, and Nijhuis (2001) concluded that an interesting step forward in the refinement of the above-mentioned conceptual models might be the development of a *theoretical model* that incorporates both the general insights regarding the influence of work on the health and well-being of employees, and more refined *theories* regarding specific relationships between work characteristics and outcome variables (see also Cooper, 1998; Kasl, 1996). Such a model may provide clues for efficient interventions at the workplace (Cooper, 1998).

A theoretical model contains narrowly bounded specific variables, postulates relationships between these specific variables, and is predictive in nature. These models are testable, yet sufficiently general to be scientifically interesting (Walker & Avant, 1995). Using conceptual models as a framework, theoretical models are constructed from available theories (e.g. the Job Characteristics Theory; cf. Walker & Avant, 1995). Theories are generally aimed at describing one specific phenomenon (e.g. intrinsic work motivation). Current examples of theoretical models that cover a relevant range of the area of work and organizational psychology, and that are at the same time specific enough to be empirically tested, appear to be lacking (cf. Cooper, 1998). Janssen *et al.* (1999) and Houkes *et al.* (2001) based their theoretical model on previous findings and several existing refined theories in the domain of Work and Organizational Psychology.

In line with suggestions of Kompier and Marcelissen (1990), Janssen *et al.* (1999) distinguished four clear categories of work characteristics: work content, working conditions, labour relations and conditions of employment. These four categories were operationalized as task characteristics (such as autonomy and task variety), workload, social support and unmet career expectations, respectively. In addition, they selected three theoretically distinct psychological outcome variables that were considered organizationally and socially important issues: intrinsic work motivation, emotional exhaustion (as the core component of burnout, cf. Maslach, 1998) and turnover intention. This set of outcome variables seemed to be an adequate representation of the effects work may have on people (i.e. negative *and* positive; in many studies in this area, only negative outcomes are included; Diener, Suh, Lucas, & Smith, 1999).

In a next step, Janssen *et al.* (1999) and Houkes *et al.* (2001) reviewed the literature regarding these three outcome variables and their determinants in order to formulate a specific pattern of relationships. With regard to *intrinsic work motivation*, it appeared that theory (e.g. Job Characteristics Theory) as well as empirical studies (e.g., Fried & Ferris, 1987; Tiegs, Tetrick & Fried, 1992) reveal that intrinsic work motivation is primarily predicted by work content variables (task characteristics) such as skill variety and autonomy. When employees have high autonomy, receive feedback about their performance, and have an important, identifiable piece of work to do which requires skill variety (i.e. meaningful work), they may experience feelings of happiness, and hence intrinsic motivation to keep performing well (Hackman & Oldham, 1980).

With regard to emotional exhaustion, Schaufeli and Enzmann (1998) have conducted an extensive review of the burnout literature and have concluded that

emotional exhaustion is particularly influenced by workload, time pressure, lack of social support and role stress (i.e. variables from the categories of working conditions and social and labour relations). A recent meta-analytic study by Lee and Ashforth (1996) also provides evidence for these relationships. Based on these empirical findings, Janssen *et al.* (1999) hypothesized that emotional exhaustion is primarily predicted by workload and lack of social support. Demanding aspects of work (i.e. high workload) lead to constant overtaxing and in the long-term to exhaustion. Human responses to stress (such as an increase in heart rate and blood pressure, resulting from higher levels of adrenaline, cf. Selye, 1976) prepare a person for action in case of danger. However, when stress ensues, and people experience stress every day, this may eventually result in a draining of one's energy and a state of emotional exhaustion. In addition, when an employee receives no social support from co-workers or supervisor, or even has conflicts with them, exhaustion may occur also. The latter hypothesis is in line with, for instance, the Conservation of Resources (COR) theory (Hobfoll & Freedy, 1993), on which Leiter (1993) based his process model of burnout. According to the COR theory, people strive to obtain or maintain things (i.e. resources) they value. Social support is an example of such a resource. When resources are lost or threatened, stress may occur. When individuals cannot deal with this stress effectively by allocating or investing new resources, prolonged stress and eventually burnout may develop.

Furthermore, emotional exhaustion is conceptually related to depression, a health outcome for which the importance of social causes such as social support, has been stressed by several authors (e.g. Street, Sheeran, & Orbell, 2001).

Finally, literature regarding *turnover intention* suggests that, pertaining to work-related factors, particularly conditions of employment (e.g. salary, career opportunities) are important causes of turnover intention (cf. Iverson & Roy, 1994; Rosse & Miller, 1984; Van Breukelen, 1989). When employees consider their career opportunities within the organization as limited or absent (unmet career expectations), a withdrawal reaction may be evoked in order to cope with the frustrations. For the individual employee, turnover to an alternative job with better career opportunities may thus be an attractive solution.

Based on these findings, Janssen *et al.* (1999), formulated the pattern of relationships that is depicted in Figure 1.

Janssen *et al.* (1999) tested this model in a sample of Dutch nurses working in a general hospital. Their results revealed that the proposed pattern of relationships largely holds true. Houkes *et al.* (2001) performed a further, more extensive validation of the proposed pattern of relationships. A multigroup analysis showed that the proposed pattern of relationships was *significant* and *invariant* (i.e. relationships have the same strength and direction) across the two samples. In sum, the empirical results with regard to this model seem very promising. However, Janssen *et al.* (1999) and Houkes *et al.* (2001) tested the model only cross-sectionally. An important drawback of cross-sectional research is the impossibility to demonstrate causal relationships; only associations between variables can be shown. Longitudinal designs, on the contrary, can be used to reduce the problems associated with cross-sectional studies: they provide a better opportunity to validate theoretically hypothesized causal relationships between, for instance, work characteristics and stress outcomes (Frese & Zapf, 1988; Zapf, Dormann & Frese, 1996). In addition, longitudinal designs make it possible to investigate alternative causational patterns, and to eliminate the influence of third variables (even without actually measuring them) (Zapf *et al.*, 1996).

Aim of the study

The present study aims at testing the research model developed by Janssen *et al.* (1999) and Houkes *et al.* (2001) longitudinally. In the first place, we aim to find evidence for the *generalizability* of the pattern of relationships proposed by Janssen *et al.* (1999) and Houkes *et al.* (2001) across samples. Second, we aim to find evidence for the *causality* of the proposed pattern of relationships between work characteristics and outcome variables. With regard to the first aim of this study, we hypothesize that this pattern of relationships is generalizable across samples at both Time 1 and Time 2 (H1). With regard to the causality of relationships, we hypothesize that work characteristics at Time 1 influence the outcome variables at Time 2 (H2). This is in line with the expectations of Janssen *et al.* (1999) and Houkes *et al.* (2001), and with, for example, the propositions of the demand–control–support model (Johnson & Hall, 1988; Karasek & Theorell, 1990) and the Job Characteristics Model (Hackman & Oldham, 1980). Both these models propose that work characteristics are a direct cause of outcomes such as intrinsic work motivation, satisfaction and health.

Most longitudinal studies (including the present study) are designed to demonstrate causal relationships between stressors (x) at Time 1 and outcome variables (y) at Time 2 (i.e. 'regular causation'). This regular causal inference can be made plausible, by means of rejecting alternative explanations (cf. Cook & Campbell, 1979; Zapf *et al.*, 1996). Alternative explanations for regular causation are reverse causation of y on x, or reciprocal causation (x influences y, *and* y influences x). Reverse causation might be explained by the 'drift' hypothesis (i.e. unhealthy and unmotivated employees drifting to worse jobs) or the 'true strain–stressor' hypothesis (i.e. sometimes stressors are in fact influenced by outcomes) (see Zapf *et al.*, 1996). Empirical evidence shows that the reverse causation hypothesis receives support, albeit less than the regular causation hypothesis (cf. De Jonge *et al.*, 2001; Zapf *et al.*, 1996). Reciprocal relationships are also found in longitudinal studies (e.g. James & Jones, 1980; James & Tetrick, 1986). Bidirectional influences (which imply a sort of vicious circle) may, although they do not correspond entirely to the nature of most psychological and social systems, be present in job stress research. For instance, an increase of stressors caused by emotional exhaustion may in turn contribute to even more emotional exhaustion (cf. De Jonge *et al.*, 2001; Zapf *et al.*, 1996). In addition, reciprocal relationships may stabilize or equilibrate a causal system, if the effects are of opposite sign.

Both Lave and March (1980) and De Groot (1981) posit that, in order to test the generalizability of a theoretical model, its validation should be as broad and varied as possible. Kristensen (1995, 1996), however, argues that in order to test a causal pattern of relationships, the amount of variation in exposure to the model variables is more important. Kristensen, therefore, recommends including only a few, but well-defined occupations in a study sample. Considering the recommendations of both Lave and March, and Kristensen, we decided to test our hypotheses in two well-defined but different occupational groups: bank employees (profit sector) and teachers (not-for-profit sector). Related to this issue, a third hypothesis was formulated: the proposed causal pattern of relationships (Figure 1) holds over different occupational groups (H3).

Multiple samples and a longitudinal study in itself are not enough to demonstrate causal relationships. According to Zapf *et al.* (1996) and Finkel (1995), several methodological pitfalls and difficulties can be distinguished, which can impede the full realization of the power of longitudinal designs. Therefore, they recommend keeping the following methodological issues in mind. All variables should be measured at all time

points (i.e., a full panel design), the same measures should be used at all measurement points, the time lag should be adequately planned (ideally, measurement periods should match causal lags), a linear structural equation approach should be used to analyse the data (instead of, for instance, a cross-lagged correlation procedure; cf. Rogosa, 1980), multiple competing models should be tested, researchers should consider third variables as potential confounders, and finally, the stability of variables should be taken into account. We follow these recommendations as much as possible.

Method

Design and participants

We conducted a full panel design with two waves and used self-report questionnaires to measure the study variables. These questionnaires were administered at two time points with an interval of one year (April 1998 and April 1999). This period seems long enough to measure possible changes in individual scores, and not too long with regard to non-response. Furthermore, this one-year interval ensures that seasonal influence is stable (cf. Frese & Zapf, 1988; Zapf *et al.*, 1996). The questionnaires contained a code in order to identify participants in the second wave. As mentioned above, we selected two different study populations: (i) *bank employees* working at the local offices of a large Dutch bank (*profit sector*), and (ii) *teachers* working at a centre for technical and vocational training for 16–18 years old and adults (*not-for-profit* sector).

Sample 1 comprised 500 employees randomly selected from all people working at the bank on a permanent basis. At Time 1, 253 (51%) usable questionnaires were returned. At Time 2, the 470 employees that were still working at the bank received the second questionnaire. This time, 200 (43%) questionnaires were returned. Our final study sample (the 'panel group') consisted of 148 employees who filled out both questionnaires (30% of the initial group). The mean age in the panel group at Time 1 was 39.3 years (SD=8.7).[1] Sixty per cent of this group were male and 83% were employed full-time.

Unfortunately, a common feature of many panel studies is that a large part of the initial sample is lost (cf. Hagenaars, 1990; Kessler & Greenberg, 1981; Verbeek, 1991). This may have implications for the external validity of the findings. The non-response in the second wave is discussed in the Results and Discussion sections.

Sample 2 comprised teachers working at the training centre on a permanent basis (N=644). At Time 1 374 (58%) usable self-report questionnaires were returned by mail. At Time 2, the 612 teachers from the initial sample still working at the school, received the second questionnaire; 226 (37%) of them returned the questionnaire. The panel group in this sample consists of 190 teachers (30% of the initial sample). The mean age in the panel group at Time 1 was 47.4 years (SD=5.8). Seventy per cent of this group were male, 62% were employed full-time. Among the part-time employed teachers, 17% were contracted for fewer than 18 hours a week.

Measures

Work characteristics

Task characteristics were measured by means of a Dutch translation of the Job Diagnostic Survey (JDS; Hackman & Oldham, 1980). The JDS consists of 16 items and

[1] *Both study samples appeared not to differ significantly from the total working populations at the bank and the education centre, respectively, with regard to the demographic variables gender (sample 1:* $\chi^2(1)=.40$, *ns*; sample 2: $\chi^2(1)=.13$, *ns*) and age (sample 1: $t=.65$, *ns*; sample 2: $t=.93$, *ns*). Furthermore, a striking feature of employees in the Dutch education sector is that they are relatively old: 46% of the teachers are older than 45 years.

measures five task characteristics: autonomy, task variety, job feedback, task identity and task significance (range 1–7). An example item is: 'The job gives me considerable opportunity for independence and freedom in how I do the work'. In line with suggestions of Fried and Ferris (1987), we combined these five task characteristics into a single, unweighted, additive index that reflects the motivating potential of a job (Motivating Potential Score–additive index). According to Fried and Ferris (1987), this simple additive index is a better predictor of work outcomes than the multiplicative MPS index, which has been suggested by Hackman and Oldham (1980). In addition, the multiplicative MPS contains two cross-product terms, which may unnecessarily increase measurement error (Evans, 1991). Second-order factor analysis of the five task characteristics in this additive index (Principal Axis Factoring) showed that a one-factor solution was admissible in both samples and at both measurement points, indicating that it is permitted to combine the five task characteristics into one index (explained variance varied between 40 and 46%).

Workload was measured by means of an 8-item scale that was developed by De Jonge, Landeweerd, and Nijhuis (1993) (range 1–5). This scale consists of a range of quantitative and qualitative demanding aspects in the work situation, such as working under time pressure, working hard and strenuous work. An example item is: 'In the organization where I work, too much work has to be done'. As in earlier studies (e.g. De Jonge, 1995), this scale showed good validity and reliability properties.

Social support was measured by means of a 10-item scale, derived from a Dutch questionnaire on organizational stress (Vragenlijst Organisatie Stress-Doetinchem; Bergers, Marcelissen, & De Wolff, 1986; range 1–4). An example item is: 'If problems exist at your work, can you discuss them with your colleagues?'.

Unmet career expectations were measured by means of a 5-item scale (range 1–5), derived from an existing questionnaire called 'Desire for career progress' (cf. Buunk & Janssen, 1992; Janssen, 1992). We selected 5 of the 8 items of this scale: unmet expectations regarding salary, responsibility, opportunities to develop knowledge and skills, job security and position. For reasons of item overlap with other measures, the three remaining items (i.e. unmet expectations regarding support, self-determination and creativity) were not included.

Outcome variables

Intrinsic work motivation was measured by means of a 6-item scale derived from a questionnaire developed by Warr, Cook, and Wall (1979) (range 1–7). An example item is: 'I take pride in doing my job as well as I can'.

Emotional exhaustion was measured by means of a 5-item subscale of the Maslach Burnout Inventory–General Survey (Schaufeli, Leiter, Maslach, & Jackson, 1996) (range 1–7). This version of the Maslach Burnout Inventory is suitable for use in all professions and not merely the human services. An example item is: 'I feel tired when I get up in the morning and have to face another day on the job'.

Turnover intention was measured by means of a 4-item scale (range 1–2) derived from a Dutch questionnaire on the experience of work (Questionnaire on the Experience and Evaluation of Work; Van Veldhoven & Meijman, 1994). An example item is: 'I intend to leave this organization this year'.

Data analyses

We analysed the data in four steps. First, we determined whether there were differences between employees in the panel group and the so-called 'dropouts' (employees

who participated in the first wave, but not in the second) with regard to variable means as well as the pattern of relationships between variables. Second, prior to testing hypotheses 1 to 3, we performed several preliminary analyses (means, standard deviations, Cronbach's alphas, test–retest reliabilities, correlational analyses). Third, in order to test the first hypothesis (H1), we performed multi-sample analysis (MSA; by means of the LISREL 8 computer program). More specifically, we tested whether the proposed pattern of relationships (see Figure 1) was generalizable across the two study samples (bank employees and teachers) at both Time 1 and Time 2 (strictly, we should speak of the pattern of *associations*, because these analyses are cross-sectional). By means of MSA it is possible to investigate to what extent a proposed pattern of relationships is actually consistent with the observed data in two or more samples simultaneously (Jöreskog & Sörbom, 1993; Schumacker & Lomax, 1996). Furthermore, it is possible to investigate whether a proposed pattern of relationships is *invariant* (i.e. has the same strength and direction) across different groups, thereby providing a powerful validation of this pattern (cf. Byrne, 1994, 1998). In addition, MSA provides the possibility to test whether differences exist between groups. One can specify certain relationships as invariant, and specify additional relationships for each group separately, where appropriate (Schumacker & Lomax, 1996). In the fourth and final step, we analysed the panel data in order to test hypotheses 2 and 3. These analyses are longitudinal, while the analyses in step 3 are cross-sectional. We used a cross-lagged panel design (e.g. Finkel, 1995; Kessler & Greenberg, 1981) and we tested a sequence of competing, nested structural equation models in several steps (cf. De Jonge *et al.*, 2001; Hom & Griffith, 1991; Jöreskog & Sörbom, 1993).

In all LISREL-analyses, we simplified the covariance structure by assuming that the latent and observed variables were identical. In other words, we did not include measurement models in our analyses, but only structural equation models. Simultaneous consideration of all observed variables (i.e. including all measurement models and covariance structure models in one analysis) would result in unreliable parameter estimates and insufficient power (cf. De Jonge *et al.*, 2001; Schumacker & Lomax, 1996). This procedure seems justified because the measures used in this study have all proven to be valid and reliable in this study and in previous studies. To assess the overall model fit, several commonly used fit indices were used (cf. Bentler, 1990; Hu & Bentler, 1998; Jöreskog, 1993; Jöreskog & Sörbom, 1993; Schumacker & Lomax, 1996; Verschuren, 1991): the chi-square statistic (χ^2), the adjusted goodness-of-fit index (AGFI), the root mean square error of approximation (RMSEA), the Akaike information criterion (AIC), the non-normed fit index (NNFI) and the comparative fit index (CFI). With regard to specific relationships, LISREL provides *t*-values indicating the significance of the specified relationships, and the so-called 'modification indices'. These modification indices provide information as to what specific relationships should be added to the model, when theoretically plausible, in order to improve the fit between the hypothesized model and the data (Hayduk, 1987; Jöreskog & Sörbom, 1993). Finally, competing models were compared by means of the chi-square difference test (Bentler & Bonett, 1980; Jöreskog & Sörbom, 1993).

Results

Step 1: Analysis of differences between panel group and dropouts

In order to rule out selection problems due to panel loss, we determined whether there were differences between employees in the panel group and the dropouts with

regard to relevant demographic and job variables (the four work characteristics and three outcome variables). In both samples, *t*-tests revealed *no* significant differences between the panel group and the dropouts with regard to the mean values of work characteristics and outcome variables. Only two minor demographic differences were found between the panel group and the dropouts. It appeared that, at Time 1, in both samples (bank employees and teachers), the panel group was somewhat older than the dropouts (less than one quarter of the standard deviation in both samples). Furthermore, sample 2 included slightly more male employees in the panel group than in the group of dropouts ($\chi^2(1)=3.48$, $p>.05$).

This verification of mean differences between respondents and non-respondents is important, but not sufficient to exclude selection problems. Kessler and Greenberg (1981) additionally recommend investigating whether the causal relationships between all study variables are similar for the panel group and the dropouts. In order to examine this so-called causal homogeneity, they suggest using information from the waves in which responses have been obtained. However, in our study the dropouts only responded in one wave, hence, an absolute test of causal homogeneity is hardly possible (only cross-sectional data are available). Therefore, we examined at Time 1 whether the interrelationships between the study variables were homogeneous for the dropouts and the panel group. In order to test this, we first verified whether the intercorrelations between all study variables were similar for both the dropouts and the panel group (in both samples). We performed a cross-sectional, multisample structural equation analysis using two corresponding correlation matrices (one matrix for the panel cases and one matrix for the dropouts), for both sample 1 and sample 2 (Jöreskog & Sörbom, 1993). In both samples, no significant differences were found regarding the intercorrelations of the study variables between the panel group and the dropouts (*sample 1*: $\chi^2(28)=35.66$, $p>.05$; *sample 2*: $\chi^2(28)=36.12$, $p>.05$). Second, by means of the chi-square difference test, we tested whether the theoretically proposed pattern of relationships (see Figure 1) was invariant for both the panel group and the dropouts (cf. De Jonge *et al.*, 2001). For both samples, we tested whether a model in which the relationships were specified as non-invariant, fitted better than a model in which the relationships were specified as invariant. These analyses also revealed no significant differences between the panel group and the dropouts in both our study samples; the difference in chi-square was not significant (*sample 1*: $\Delta\chi^2(4)=6.75$, $p>.05$; *sample 2*: $\Delta\chi^2(4)=1.73$, $p>.05$). Thus, we may conclude that in both samples the panel group and the dropouts are quite comparable with regard to the interrelationships between the study variables and with regard to the pattern of relationships, and that no serious selection problems have occurred.

Step 2: Preliminary analyses

Next, we performed several preliminary analyses (i.e. means, standard deviations, Cronbach's alphas, test-retest reliabilities, zero-order Pearson correlations). The results of these analyses are depicted in Tables 1 and 2.

Table 1 shows that the internal consistencies are generally quite acceptable (according to the .70 criterion of Nunnally, 1978). In sample 1, two variables have reliabilities below .65 (i.e. Time 1 turnover intention and Time 2 emotional exhaustion). In sample 2, however, the reliabilities of these variables are quite acceptable. Furthermore, Table 2 shows that the test–retest reliabilities of most variables were moderate to high, with somewhat higher stabilities for sample 2. The cross-sectional correlations (i.e. Time 1

Table 1. Means, standard deviations, and Cronbach's alphas of sample 1 (bank employees) and sample 2 (teachers) at T1 and T2

	Sample 1 (N=148)			Sample 2 (N=188)		
	M	*SD*	α/KR20[a]	*M*	*SD*	α/KR20[a]
MPS (additive index) (1)[b]	5.30	.66	.78	5.00	.65	.70
Workload (1)	3.37	.64	.89	3.42	.63	.88
Social support (1)	3.13	.38	.83	2.98	.45	.83
Unmet career expectations (1)	3.41	.76	.79	3.37	.77	.76
Intrinsic work motivation (1)	5.96	.63	.67	6.05	.63	.68
Emotional exhaustion (1)	2.59	1.02	.90	3.02	1.21	.90
Turnover intention (1)	1.29	.28	.63[a]	1.29	.33	.79[a]
MPS (additive index) (2)[b]	5.32	.63	.79	5.11	.70	.80
Workload (2)	3.51	.58	.87	3.46	.62	.90
Social support (2)	3.15	.35	.80	3.07	.46	.85
Unmet career expectations (2)	3.31	.64	.70	3.23	.72	.68
Intrinsic work motivation (2)	5.96	.54	.63	5.99	.64	.76
Emotional exhaustion (2)	2.69	1.12	.91	2.97	1.08	.89
Turnover intention (2)	1.31	.31	.72[a]	1.29	.33	.78[a]

[a]KR20 is a measure of internal reliability that is particularly appropriate for dichotomous items. Its interpretation is equal to the interpretation of Cronbach's alpha.
[b](1)=Time 1, (2)=Time 2.

work characteristics to Time 1 outcome variables, and Time 2 work characteristics to Time 2 outcome variables) show that in both samples and at both measurement points the associations between work characteristics and outcome variables generally meet our expectations (see Figure 1). An exception is the relationship between the additive MPS index and intrinsic work motivation, which is not significant in sample 2 (teachers) at Time 1. At Time 2, however, this relationship is significant.

Step 3: Cross-sectional analyses: Testing the generalizability of the proposed pattern of relationships (H1)

In order test the generalizability of the pattern of relationships depicted in Figure 1 over samples (H1), we performed multi-sample analyses (MSA) for bank employees and teachers at both Time 1 and Time 2. We used covariance matrices as input and we controlled for gender and age, because these demographic variables may confound the results (cf. Karasek & Theorell, 1990; Zapf *et al.*, 1996). These variables were labelled as exogenous (cf. Bollen, 1989; De Jonge *et al.*, 2001), and all other variables (i.e. work characteristics and outcome variables) were labelled as endogenous. Because the endogenous variables might partly be predicted by variables that are not included in the present study, the error variances of the endogenous variables themselves and the error variances regarding the relationships among the endogenous variables (i.e. within the group of work characteristics, and within the group of outcome variables) were set free (cf. De Jonge *et al.*, 2001; MacCallum, Wegener, Uchino, & Fabrigar, 1993). Figure 1 shows the general structural equation model we used in these analyses.

The results of the MSA at Time 1 indicate that the basic pattern of relationships (Figure 1) holds in both samples (i.e. the proposed relationships are significant in both

Table 2. Pearson correlations of the study variables of sample 1 (bank employees, *N* = 148, left-lower corner) and sample 2 (teachers, *N* = 188, right-upper corner)

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Gender	—	−.18*	−.02	−.20**	.05	−.08	−.04	−.09	−.04	.05	−.08	.06	−.08	.02	.02	−.06
2. Age	−.41**	—	−.00	−.07	.03	−.27**	.08	.06	−.31**	−.06	.15*	.00	−.27**	.19*	−.02	−.31**
3. MPS (additive index) (1)[a]	−.20*	.05	—	−.19**	.49**	−.11	.12	−.42**	−.13	.62**	−.16*	.32**	−.07	.13	−.31**	−.16*
4. Workload (1)	−.18*	.15	.09	—	−.21**	.15*	.16*	.43**	.15*	−.16*	.71**	−.18*	.04	.01	.37**	.13
5. Social support (1)	.10	−.23**	.19*	−.11	—	−.20**	.04	−.42**	−.27**	.37**	−.10	.56**	−.14	.06	−.29**	−.19**
6. Unmet career expectations (1)	−.21*	−.05	.10	.23**	−.14	—	.07	−.02	.29**	.02	.09	−.14	.74**	.01	.00	.21**
7. Intrinsic work motivation (1)	.06	.12	.23**	.16	−.01	−.02	—	.12	−.09	.15*	.18*	−.04	−.04	.47**	.05	−.01
8. Emotional exhaustion (1)	−.09	.18*	.02	.45**	−.21*	.10	.09	—	−.18*	−.27**	.34**	−.31**	−.06	.07	.69**	.12
9. Turnover intention (1)	−.02	−.24*	−.09	.06	−.19**	.20*	−.10	.11	—	−.07	.20**	−.17*	.34**	−.06	.12	.62**
10. MPS (additive index) (2)[a]	−.10	.07	.49**	.14	.08	.08	.25**	−.14	−.14	—	−.25**	.38**	.00	.22**	−.38**	−.30**
11. Workload (2)	−.11	.12	.06	.57**	−.15	.17*	.08	.31**	.05	−.00	—	−.28**	.01	.03	.50**	.25**
12. Social support (2)	.10	−.19*	.29**	−.01	.53**	−.08	.00	−.20*	−.07	.17*	−.26**	—	−.12	.06	−.36**	−.28**
13. Unmet career expectations (2)	−.13	−.14	.01	.14	−.11	.56**	−.19*	.04	.17*	−.04	.18*	−.22**	—	−.04	−.04	.29**
14. Intrinsic work motivation (2)	.05	.06	.24**	.15	.06	−.02	.49**	.14	−.15	.29**	.13	.06	−.09	—	−.01	−.08
15. Emotional exhaustion (2)	−.01	.09	−.02	.18**	−.17*	.04	.06	.66**	.06	−.26**	.40**	−.28**	.07	.07	—	.22**
16. Turnover intention (2)	.05	−.23**	−.15	−.05	−.06	.17*	−.08	−.09	.45**	−.27**	.05	−.16	.33**	−.17*	.03	—

[a](1) = Time 1, (2) = Time 2.
*$p \leq .05$, two-tailed; **$p \leq .01$, two-tailed.

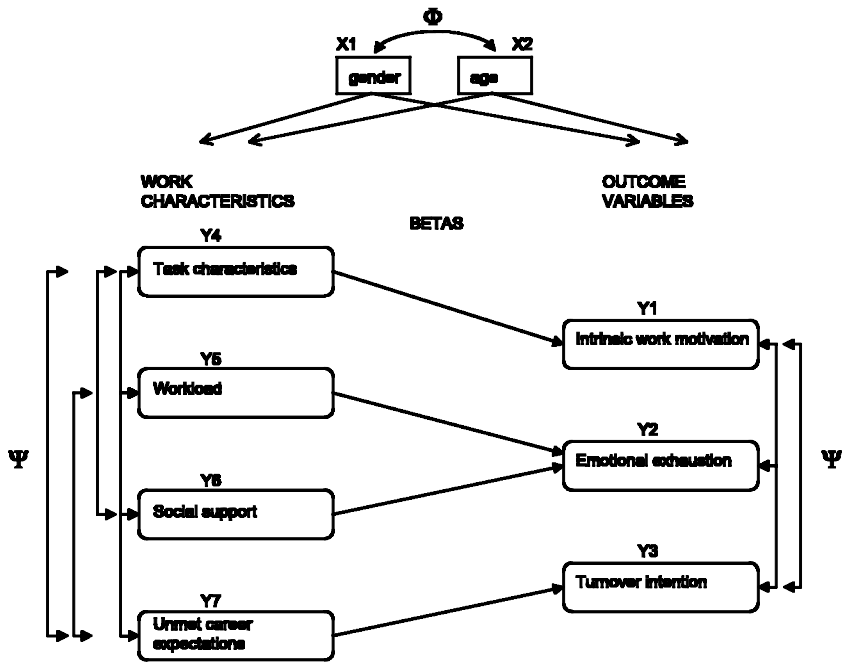

Figure 1. The proposed pattern of relationships between work characteristics and outcome variables: A cross-sectional structural equation model.

samples as indicated by the *t*-values). The overall model fit (the fit of all models over all samples) was, however, not optimal. In line with the suggestions provided by LISREL (modification indices), we added the relationship between social support and turnover intention in both samples, and the relationship between the additive MPS index and emotional exhaustion in sample 2. This improved the model fit considerably (see Discussion section).

In order to find more evidence for the generalizability of the basic pattern of relationships over samples (H1), we compared a model, M1, in which the basic pattern of relationships (Figure 1) was specified as invariant across the two samples, with a less restrictive model, M2, in which the basic pattern of relationships was specified as non-invariant, by means of the chi-square difference test. The additional relationships were specified for each sample separately, in both M1 and M2. It appeared that the difference in chi-square between M1 and M2 was not significant ($\Delta\chi^2$ (4)=3.54, p>.05), indicating that M2 has no better fit than M1. The proposed pattern of relationships is *invariant* across both samples. The practical fit indices of M1 appeared to be acceptable (i.e. RMSEA=.06, AIC=105.61, NNFI=.89, CFI=.95, AGFI=.95).

The MSA results for Time 2 also showed that the basic pattern of relationships holds in both samples (i.e. significant *t*-values). This time we had to add the relationship between social support and turnover intention, and the relationship between the additive MPS index and emotional exhaustion only in sample 2 (see Discussion section). The Time 2 results also indicated that the basic pattern of relationships was invariant across the two samples (the chi-square difference between M1 and M2 was not significant: $\Delta\chi^2$ (4)=1.85, p>.05). The practical fit indices of M1 were reasonable (i.e. RMSEA=.08, AIC=114.95, NNFI=.81, CFI=.95, AGFI=.95). In conclusion, the

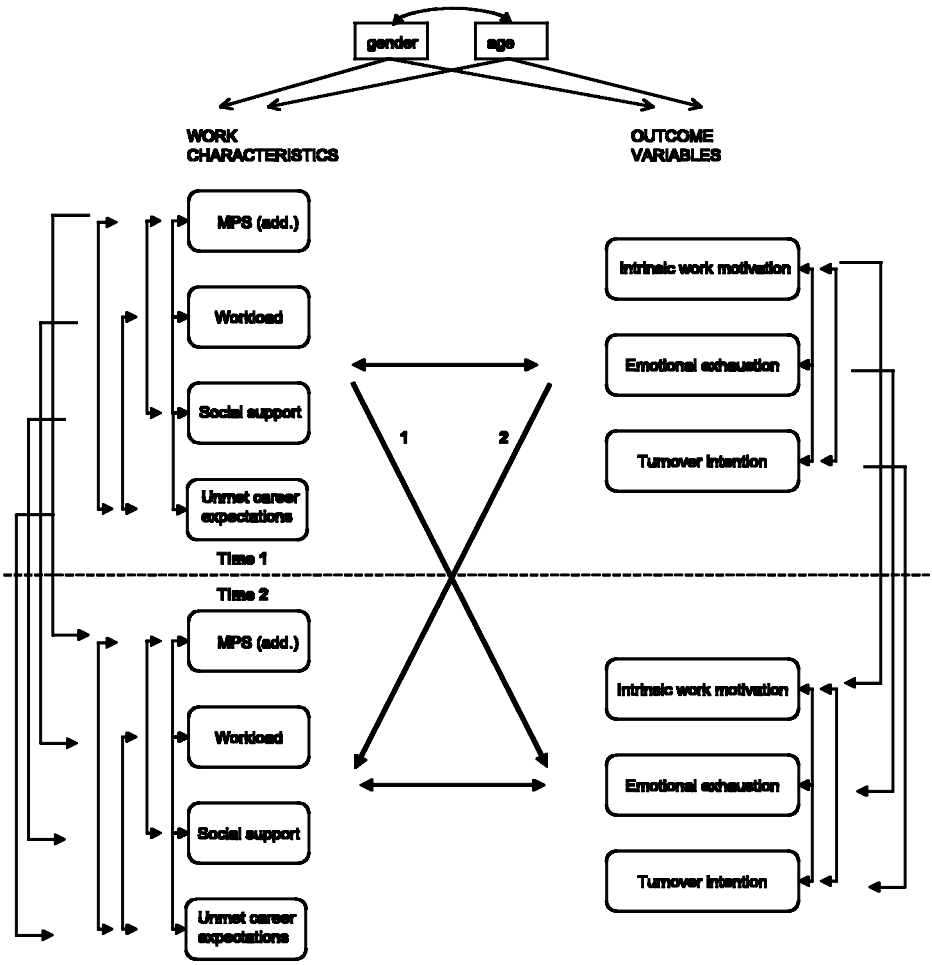

Figure 2. A cross-lagged structural equation model.

proposed basic pattern of relationships (Figure 1) seems to be invariant across the study samples at both Time 1 and Time 2: H1 seems to be confirmed.

Step 4: Panel analyses: Testing H2 and H3

In order to test H2 (regular causation between Time 1 work characteristics and Time 2 outcome variables), we used structural equation modelling for longitudinal analyses. Covariance matrices were used as input for these analyses. The general panel model we used is depicted in Figure 2.

We compared the following competing structural equation models (see also De Jonge *et al.*, 2001).

M1: a model without cross-lagged paths, but with temporal stabilities and synchronous correlations (i.e. off-diagonal psi values) between the Time 1 work characteristics and Time 1 outcome variables, and between the Time 2 work characteristics and the time 2

outcome variables. These synchronous correlations were specified according to the step 3 results.

M2: a model which is identical to M1, but also includes cross-lagged paths from Time 1 work characteristics to Time 2 outcome variables, according to the pattern of relationships depicted in Figure 1 (i.e. regular causation, reflected by arrow 1 in Figure 2).

M3: a model which is identical to M1, but also includes cross-lagged paths from Time 1 outcome variables to Time 2 work characteristics, according to the pattern of relationships depicted in Figure 1 (i.e., reverse causation, reflected by arrow 2 in Figure 2).

M4: a model which is identical to M1, but also includes both cross-lagged patterns (i.e. reciprocal causation, reflected by arrows 1 *and* 2 in Figure 2).

Models M3 and M4 were included in order to test the alternative explanations for a regular causal pattern of relationships between Time 1 work characteristics and Time 2 outcome variables (H2). If model M2 appears to be the best model, H2 will be accepted. If M3 (reverse causation) appears to be the best model, H2 will be rejected. When M4 is found to be the best model, reciprocal causation exists, and H2 will be confirmed only partially.

Maximum likelihood was used to estimate the fit of the structural equation models (Jöreskog & Sörbom, 1993). Again, we controlled for Time 1 gender and age (cf. Zapf *et al.*, 1996): All work characteristics and outcome variables were labelled as endogenous. The structural equation models consist of the following parameters: regression coefficients representing the differential cross-lagged structural paths (arrows 1 and 2 in Figure 2); temporal stability coefficients between the measurement scales of the work characteristics and the outcome variables; covariances between the background variables; and residual covariances among the work characteristics, among the outcome variables, and between the work characteristics, on the one hand, and the outcome variables, on the other hand (i.e. the synchronous off-diagonal psi values) (cf. De Jonge *et al.*, 2001).

Testing causal predominance in sample 1: Model comparisons

The results of the model comparisons are summarized in Table 3.

First, we compared M1 (no cross-lagged paths) and M2 (cross-lagged paths from Time 1 work characteristics to Time 2 outcomes). The chi-square difference between M1 and M2 was significant. This means that a model with cross-lagged relationships between work characteristics at Time 1 and outcome variables at Time 2 (M2) fits better to the data than a model without cross-lagged relationships (M1). Next, we compared M1 and M3 (cross-lagged paths from Time 1 outcomes to Time 2 work characteristics). The chi-square difference test between M1 and M3 was not significant, indicating that M3 has no better fit than M1. A model with cross-lagged relationships between Time 1 outcome variables and Time 2 work characteristics (M3) had no better fit than a model without these cross-lagged relationships (M1): Time 1 outcome variables generally do not have unique causal effects upon Time 2 work characteristics. Furthermore, the chi-square difference test between M1 and M4 was not significant either. The model with reciprocal cross-lagged relationships (M4) did not have a better statistical fit than a model without cross-lagged relationships (M1). In addition, the chi-square difference test between the models M2 and M4 was not significant. This means that addition of the reverse cross-lagged paths between Time 1 outcome

Table 3. Fit measures and likelihood ratio tests of nested structural equation models in the panel analyses (sample 1, bank employees and sample 2, teachers)

Model	χ^2	*df*	Comparison	$\Delta\chi^2$	Δdf	RMSEA	AIC	AGFI	NNFI	CFI
Sample 1										
M1: no cross-lagged	100.12*	59				.062	246.42	.82	.84	.92
M2: cross WC_{T1}–OV_{T2}[a]	90.34*	55	M1 vs. M2	9.78*	4	.058	242.56	.83	.86	.93
M3: cross OV_{T2}–WC_{T1}	95.13*	55	M1 vs. M3	4.99	4	.067	250.73	.82	.84	.93
M4: both cross	85.55*	51	M1 vs. M4	14.57	8	.061	246.92	.83.	.85	.94
			M2 vs. M4	4.79	4					
Sample 2										
M1: no cross-lagged	107.35*	56				.067	261.04	.84	.89	.95
M2: cross WC_{T1}–OV_{T2}	103.81*	52	M1 vs. M2	3.54	4	.070	264.86	.84	.88	.95
M3: cross OV_{T1}–WC_{T2}	97.08*	52	M1 vs. M3	10.27*	4	.065	259.62	.84	.90	.96
M4: both cross	93.96*	48	M1 vs. M4	13.39	8	.068	263.90	.84	.89	.95
			M3 vs. M4	3.12	4					

[a]WC=work characteristics, OV=outcome variables.
*$p \leq .05$.

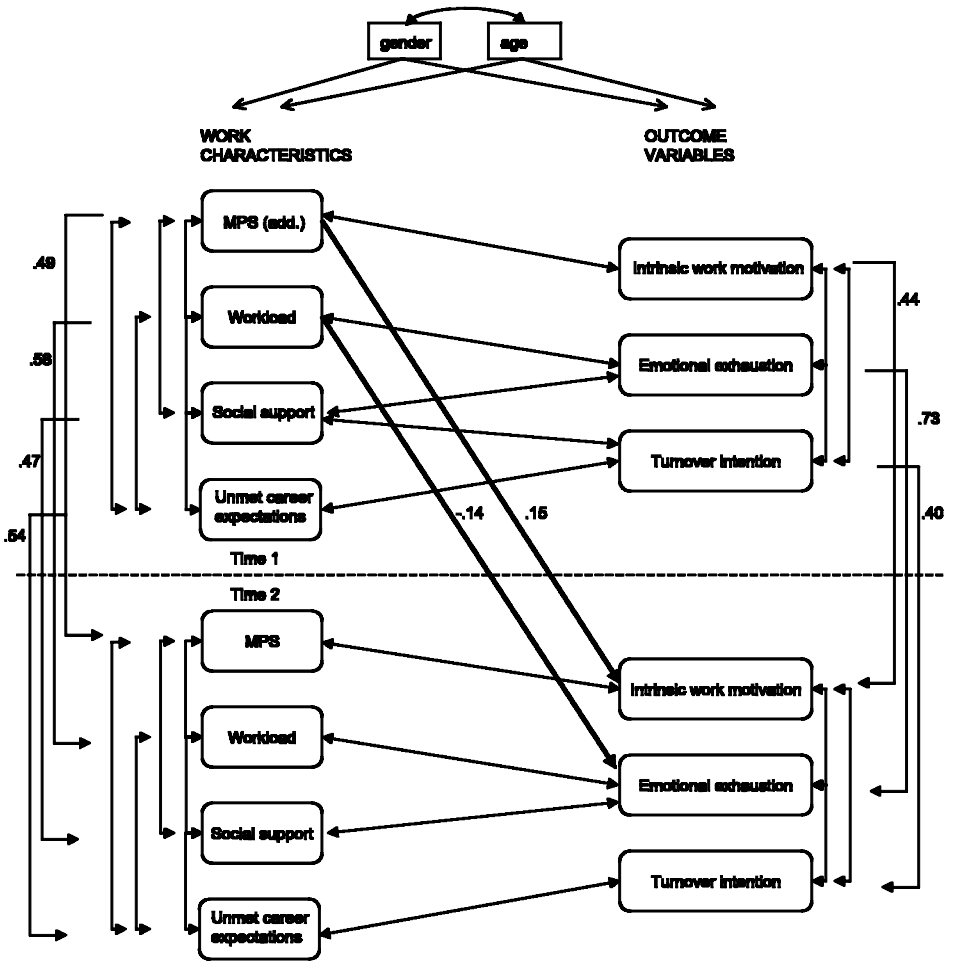

Figure 3. The final longitudinal structural equation model with standardized coefficients and significant cross-lagged relationships for sample 1 (bank employees, *N*=148).

variables and Time 2 work characteristics does not improve the fit of a model with cross-lagged paths between Time 1 work characteristics and Time 2 outcome variables (M2). We may conclude from these results that in terms of chi-square relative to the degrees of freedom (parsimony), model M2 has the best fit of the four competing models. When we consider the practical fit indices, model M2 also appears to have the best model fit. For instance, NNFI (fit criterion for model comparison) is highest, and AIC and RMSEA (fit criteria for parsimony) are lowest (cf. De Jonge *et al.*, 2001; Schumacker & Lomax, 1996). Thus, for sample 1 (bank employees), hypothesis 2 seems to be confirmed. Time 1 work characteristics influence time 2 outcome variables.

Specific causal effects in sample 1: Cross-lagged relationships

In addition to the determination of causal predominance, it is very informative to look at specific causal effects and measures of strength of individual relationships (cf. De Jonge *et al.*, 2001; Rogosa, 1980). Figure 3 represents the estimated structural

coefficients of the best fitting model in sample 1 (i.e. M2; the parameters have been standardized and only significant cross-lagged relationships and temporal stabilities are shown).

It appears from Figure 3 that the cross-lagged relationships between the Time 1 additive MPS index and Time 2 intrinsic work motivation, and between Time 1 workload and Time 2 emotional exhaustion were significant, even after controlling for stability effects. The relationship between the Time 1 additive MPS index and Time 2 intrinsic work motivation was positive (as expected). Higher levels of MPS at Time 1, lead to higher levels of intrinsic work motivation at Time 2. However, the cross-lagged relationship between Time 1 workload and Time 2 emotional exhaustion appeared to be negative, which is contrary to expectations and not in line with the bivariate Pearson correlation (see Table 2). We address this issue in the Discussion section.

Finally, in order to explore the data somewhat more, we examined the modification indices LISREL provided with regard to the cross-lagged paths and the synchronous paths within Time 2 (betas). It appeared that the following paths had modification indices >5.00: Time 2 additive MPS index to Time 2 turnover intention, Time 2 workload to Time 2 emotional exhaustion, and Time 2 unmet career expectations to Time 2 turnover intention. We added these relationships to the best-fitting model one by one. It appeared that all these relationships were significant, except for the relationship between the Time 2 additive MPS index and Time 2 turnover intention. The betas of the two significant relationships were .81 and .85, respectively. Note that these additional paths are in line with our model (see Figure 1). In the analyses described above, however, only the synchronous psi values were specified, and the synchronous betas were not. The fit indices of the model that includes the two additional paths are: χ^2 (53)=72.22, p>.05, RMSEA=.04, AIC=231.10, NNFI=.92, CFI=.96, AGFI=.86. There were no cross-lagged paths with modification indices >5.00.

Testing causal predominance in sample 2: Model comparisons

The results of the model comparisons for sample 2 (teachers) can also be found in Table 3. First, we compared M1 and M2. The chi-square difference test between M1 and M2 was not significant. This finding indicates that a model with cross-lagged relationships between work characteristics at Time 1 and outcome variables at Time 2 (M2) does not fit better than a model without cross-lagged relationships (M1). Next, we compared M1 and M3 (cross-lagged paths from Time 1 outcomes to Time 2 work characteristics). The chi-square difference test between M1 and M3 was significant. This means that there is evidence for *reverse* causation: a model with cross-lagged relationships between Time 1 outcome variables and Time 2 work characteristics (M3) fits the data better than a model without cross-lagged relationships (M1). Furthermore, the chi-square differences between M1 and M4 (model with both cross-lagged patterns), and between M3 and M4 were not significant. There exists no evidence for reciprocal causation. Thus, in terms of chi-square relative to the degrees of freedom (parsimony), model M3 has the best fit of the four competing models. When we consider the practical fit measures, it appears that M3 shows the best combination of NNFI and CFI and the lowest RMSEA. Thus, we may conclude that M3 is the best model. Time 1 outcome variables influence Time 2 work characteristics. H2 (regular causation) could not be confirmed in sample 2.

Specific causal effects in sample 2: Cross-lagged relationships

With regard to specific relationships, it appears that one reverse cross-lagged relationship was significant, even after controlling for stability effects. The standardized

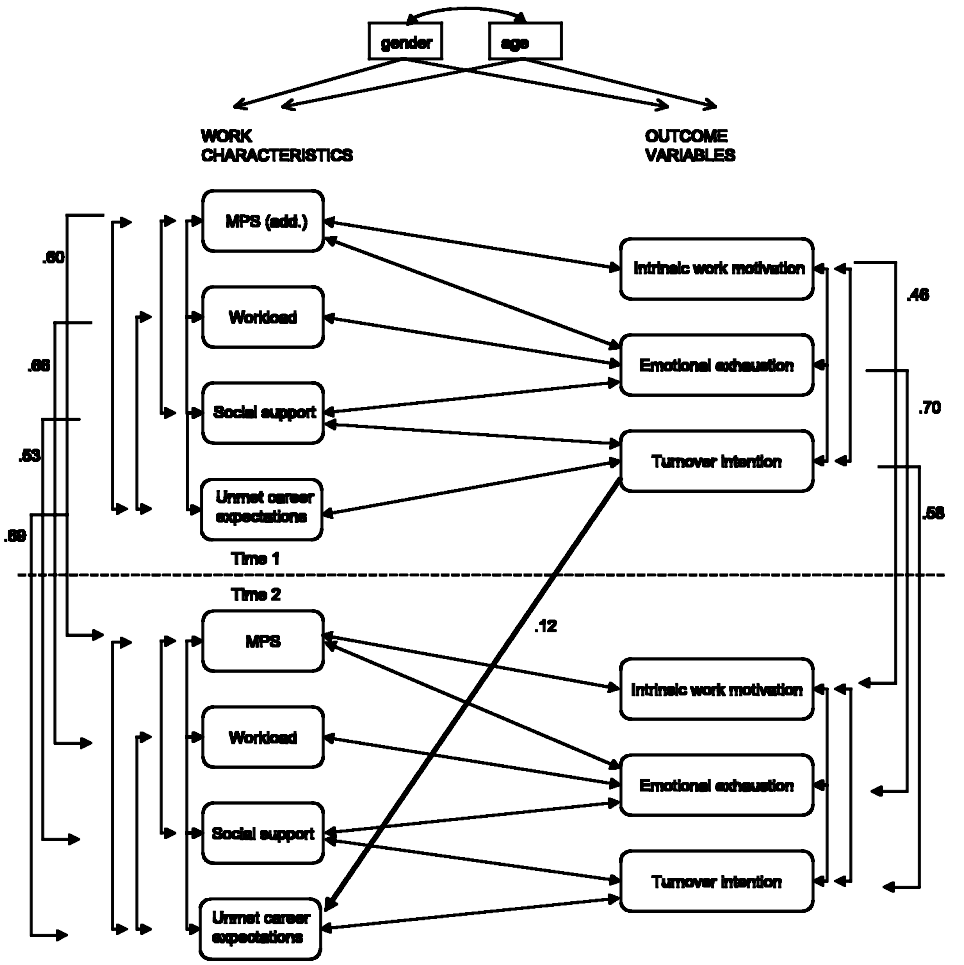

Figure 4. The final longitudinal structural equation model with standardized coefficients and significant cross-lagged relationships for sample 2 (teachers, *N*=188).

estimated structural coefficients of model M3 are depicted in Figure 4. Only significant cross-lagged relationships and temporal stabilities are shown.

It appears from Figure 4 that the cross-lagged relationship between Time 1 turnover intention and Time 2 unmet career expectations was significant. A higher level of turnover intention at Time 1 leads to a higher level of unmet career expectations at Time 2.

Finally, just as in sample 1, we examined the modification indices LISREL provided with regard to the cross-lagged paths and the synchronous paths at Time 2 (betas). It appeared that the following paths had modification indices which were >5.00: Time 2 additive MPS index to Time 2 emotional exhaustion, Time 2 additive MPS index to Time 2 turnover intention, and Time 2 social support to Time 2 turnover intention. We added these relationships to the best-fitting model one by one. It appeared that all relationships were significant. The betas of the three significant relationships were −.20, −.24 and −.15, respectively. These additional relationships are partly in line with the step 3 results. The fit indices of the model that includes the three additional paths

are: χ^2 (49)=61.11, p>.05, RMSEA=.04, AIC=233.92, NNFI=.97, CFI=.99, AGFI=.89. There were no cross-lagged relationships with modification indices >5.00.

With regard to hypothesis 3 (H3), it may be concluded that the causality of the proposed pattern of relationships was not stable over the two groups. In sample 1, the cross-lagged pattern of relationships between time 1 work characteristics and time 2 outcome variables is causally predominant, and in sample 2, a reverse pattern of relationships is predominant. Thus, we conclude that H3 has not been confirmed in this study.

Discussion

The main purpose of this study was to find evidence for the specific pattern of relationships between four work characteristics (i.e. task characteristics, workload, social support and unmet career expectations) and three outcome variables (i.e. intrinsic work motivation, emotional exhaustion and turnover intention). By means of a longitudinal design, we aimed to test the causality of the pattern of relationships depicted in Figure 1. We used a two-wave panel design and followed the suggestions of Zapf *et al.* (1996) and Finkel (1995) with regard to longitudinal research as much as possible.

Our first hypothesis was that the pattern of relationships between work characteristics and outcome variables, as it is shown in Figure 1, is generalizable across samples (H1). Our results indicated that the proposed pattern of relationships (Figure 1) was invariant (i.e. had the same strength and direction) across the two samples at both Time 1 and Time 2. However, in both samples, the relationship between social support and turnover intention had to be added in order to reach optimal model fit. In the present study, this additional relationship appears to be generalizable over samples. With regard to this relationship, indeed some theoretical and empirical evidence exists (cf. McFadden & Demetriou, 1993; Schaufeli & Enzmann, 1998; Stremmel, 1991). According to, for instance, Schaufeli and Enzmann, conflict with colleagues and/or supervisors might be a plausible reason to leave an organization, especially when employees find no alternative solution to resolve the conflict. Empirical results regarding this relationship are mixed, however. Janssen *et al.* (1999), for instance, did not report this relationship. This non-finding might be explained by the fact that they only measured co-worker support, whereas in the present study a combined scale was used to measure social support (i.e. measuring social support from co-workers and from supervisors). It might be that conflicts with supervisors are more salient to individual employees and their careers than conflicts with colleagues, and may as such have more impact. In sample 2 (teachers) we also added the relationship between the additive MPS index and emotional exhaustion (i.e. the lower the additive MPS index, the higher the level of emotional exhaustion; lack of, for instance, autonomy and variety may lead to stress and emotional exhaustion). This relationship appeared not be generalizable over the samples in our study, but it has been reported before by, for instance, Xie and Johns (1995). They found a curvilinear relationship between job scope and emotional exhaustion. That is, both when job scope is too low, and when job scope is too high, emotional exhaustion occurred.

Our second hypothesis (H2) was that Time 1 work characteristics influence Time 2 outcome variables. Comparison of four competing models (by means of the chi-square difference test) indicated that in *sample 1* (bank employees) the regular causation pattern (M2) was predominant. Time 1 work characteristics influence Time 2 outcome

variables. More specifically, we found two significant cross-lagged relationships. The first significant cross-lagged relationship we found in sample 1 was the relationship between the additive MPS index and intrinsic work motivation (a higher MPS at Time 1 leads to higher intrinsic work motivation at Time 2). Challenging work leads to intrinsic motivation, also in the long-term (cf. Hackman & Oldham, 1980). This implies that the positive effects of job redesign will not fade away in a short period, contrary to the positive effects of, for instance, a salary raise (Thierry, 1998).

Second, we found that Time 1 workload influences Time 2 emotional exhaustion. Opposite to our expectations, however, this relationship was negative, while the bivariate relationship between workload and emotional exhaustion was positive. This negative direct effect of Time 1 workload on Time 2 emotional exhaustion might be due to the so-called *suppressor* phenomenon, more precisely, negative suppression (Maassen & Bakker, 2000; Tzelgov & Henik, 1991). Negative suppression occurs when two or more independent variables (predictors, e.g. Time 1 workload and Time 1 emotional exhaustion) have positive predictive validity and intercorrelate positively, and one of the predictors receives a negative path coefficient. According to Maassen and Bakker (2000), there is a fair chance on a suppressor situation in a structural model that includes variables that are measured at two time points, and in which the stability coefficients and synchronous correlations are substantially larger than the cross-lagged coefficients. Maassen and Bakker (2000) state that we should not immediately conclude that there is a negative direct effect of Time 1 workload on Time 2 emotional exhaustion. This finding might be due to a statistical artifact.

The so-called 'adjustment' or 'habituation' processes may be another, more substantial explanation for this finding (Frese & Zapf, 1988; Gaillard, 1996). These processes imply that even although the stressor (e.g. workload) is still present, the dysfunctioning (e.g. emotional exhaustion) decreases. People get used to the presence of the stressor and learn to cope with it more effectively over time. We believe, however, that the first explanation (suppression) is more likely, because of the positive bivariate correlation between Time 1 workload and Time 2 emotional exhaustion.

In sample 2 (teachers), we found evidence for reverse causation: M3 was predominant. Time 1 outcome variables influenced Time 2 work characteristics. This finding is partly in line with earlier findings and theoretical considerations of, for instance, De Jonge *et al.* (2001) and Zapf *et al.* (1996), but contradicts with our hypothesis. Zapf *et al.* (1996) state that it can be useful to test synchronous effects (betas) too. In a study with a time interval of one year a true three months causal lag may in fact be better represented by the synchronous paths (at Time 2) than by the cross-lagged paths. Perhaps among teachers the causal lags are shorter than among bank employees and is this the reason we did not find regular causation among teachers. Hence, the results may depend on the specific method of analysis that is used.

More specifically, we found the reverse relationship between Time 1 turnover intention and Time 2 unmet career expectations to be significant. In the present study, the most likely explanation for this finding is probably the true strain–stressor hypothesis. We also found a significant cross-sectional correlation between Time 1 unmet career expectations and Time 1 turnover intention. In the short run, people may have unmet career expectations and therefore think about leaving the organization. Once they have reached this stage, these employees and their supervisors, might put even less effort into career development. As a consequence, the level of unmet career expectations rises even further (self-fulfilling prophecy). However, for Dutch teachers, job alternatives are scarce. Therefore, it is rather difficult for them to actually leave the

organization (bank employees have more job alternatives). Thus, among teachers, turnover intention cannot easily be transferred in turnover. This might be frustrating and the frustration about the career possibilities within the organization might increase in turn. Or, in other words, employees decide they want to leave the organization, and then look for reasons to justify their decision (e.g. their level of unmet career expectations rises further and further). In social psychology this phenomenon is known as the urge for confirmation: people consider information that confirms their decisions and convictions more important than information that refutes their beliefs (e.g. Ditto & Lopez, 1992). Hence, they absorb confirming information more easily.

In the panel analyses (step 4) we also found several additional relationships, which were largely in line with our hypotheses in Figure 1 and the above-mentioned step 3 results.

Our final, third hypothesis (H3) was that the proposed causal pattern of relationships holds over different occupational groups. We did not explicitly test this hypothesis. However, we may conclude from the results discussed above, that hypothesis 3 was not confirmed in this study. In sample 1, we found evidence for regular causation, but in sample 2 (teachers), we found evidence for reverse causation. As we have concluded above, there is still an ongoing discussion about the best way to analyse panel models (Zapf *et al.*, 1996). It is, for instance, not clear whether synchronous effects should be modelled in panel analysis. Some researchers do (e.g., Zapf *et al.*, 1996), whereas others do not (e.g., De Jonge *et al.*, 2001).

In sum, the present results indicate that our theoretical expectations with regard to the causality of relationships between work characteristics and outcome variables seem to hold to some extent. That is, the proposed pattern of relationships holds over samples, and work characteristics influence outcome variables in one sample. We were, however, not able to exclude the alternative explanation of reverse causation totally. These longitudinal results generally strengthen earlier cross-sectional results of Janssen *et al.* (1999) and Houkes *et al.* (2001) and the existing theories in the area of work and organizational psychology. Practically, our results provide a basis for interventions aimed at improving the worker's health and motivation. The results show that task redesign may influence the intrinsic work motivation of employees in the short-term as well as in the long-term. In addition, reducing levels of workload may prevent emotional exhaustion. Finally, by means of paying attention to career expectations of employees (e.g. career development programmes, education programmes), retention of valuable employees might be accomplished.

Limitations of the study

Some limitations with regard to the present study should be mentioned. First, we measured the study variables at two fixed time points, while the processes we observed are continuous. Thus, if the time lag between the two measurements does not match with the actual causal lag, it is possible that our results are not completely valid (cf. Kessler & Greenberg, 1981). This time-lag problem is very difficult to solve and is intrinsic to longitudinal research in general. As mentioned earlier, in a study that uses a time lag of one year, a true three-month time-lag, might be better represented by synchronous effects than by the lagged effects (Zapf *et al.*, 1996). Moreover, in our study, the time-lag problem is even more pronounced, because we studied several different relationships simultaneously. It is, for instance, theoretically plausible that the true time lag between the additive MPS index and intrinsic work motivation is shorter

than the true time lag between workload and emotional exhaustion. Moreover, the true causal lags may differ across samples.

Second, there is the problem of attrition (i.e. panel mortality). The panel groups in our study consisted of 30% of the initial samples. Thus, we have lost a considerable part of the respondents during our study. If this non-response has occurred completely at random, there seems to be no threat to validity. If, however, non-response did not occur completely at random, the conclusions drawn from the study, may not be valid for the total population (i.e. no external validity) (Hagenaars, 1990). Our non-response analyses showed that in both samples, the panel group did not differ from the dropouts regarding the variables under study. Therefore, we concluded that no serious selection problems had occurred.

In addition, considering our results, we may conclude that even structural equation models, full panel designs, correcting for background variables, etc. are not sufficient to exclude problems such as negative suppression. When the stability of, for instance, an outcome variable is large (i.e. a high test–retest reliability), the chance on negative suppression increases. Furthermore, the results of the longitudinal analyses may depend on the strategy of analysis that is used (cf. Zapf *et al.*, 1996). Hence, challenges for future research may be found in developing better techniques which are particularly appropriate to analyse longitudinal data and to determine adequate time lags.

In sum, we believe that the results of our study are noteworthy, as we (partly) validated a model that describes specific relationships between work characteristics and psychological outcomes by means of a longitudinal design which largely meets the present criteria for longitudinal research formulated by Zapf *et al.* (1996).

References

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

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Bergers, G. P. A., Marcelissen, F. H. G., & De Wolff, C. J. (1986). *VOS-D. Vragenlijst Organisatie-stress-D: Handleiding* [VOS-D. Work Stress Questionnaire Doetinchem: Manual]. Nijmegen, The Netherlands: University of Nijmegen.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Buunk, A. P., & Janssen, P. P. M. (1992). Relative deprivation, career issues and mental health among men in midlife. *Journal of Vocational Behavior*, *40*, 338–350.

Byrne, B. M. (1994). Burnout: Testing for the validity, replication, and invariance of causal structure across elementary, intermediate, and secondary teachers. *American Education Research Journal*, *31*, 645–673.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS*. Mahwah, NJ: Erlbaum.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis for field settings.* Boston: Houghton Mifflin.

Cooper, C. L. (1998). Introduction. In C. L. Cooper (Ed.), *Theories of organizational stress* (pp. 1–5). Oxford: Oxford University Press.

Cooper, C. L., & Payne, R. (1988). *Causes, coping and consequences of stress at work.* Chichester: Wiley.

De Groot, A. D. (1981). *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen* [Methodology: Foundations of research in behavioral sciences]. 's Gravenhage, The Netherlands: Mouton.

De Jonge, J. (1995). *Job autonomy, well-being, and health.* PhD thesis. Datawyse, Maastricht, The Netherlands.

De Jonge, J., Dormann, C., Janssen, P. P. M., Dollard, M. F., Landeweerd, J. A., & Nijhuis, F. J. N. (2001). Testing reciprocal relationships between job characteristics and psychological outcomes: A cross-lagged structural equation model. *Journal of Occupational and Organizational Psychology*, *74*, 29–46.

De Jonge, J., Landeweerd, J. A., & Nijhuis, F. J. N. (1993). Constructie en validering van de vragenlijst ten behoeve van het project "autonomie in het werk" [Construction and validation of the questionnaire for the "job autonomy project"]. *Studies bedrijfsgezondheidszorg nummer 9.* Maastricht, The Netherlands: University of Limburg.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. I. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*, 267–302.

Ditto, P. H., & Lopez, D. F. (1992). Motivated scepticism: Use of differential decision criteria for preferred and non-preferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584.

Evans, M. G. (1991). The problem of analysing multiplicative composites. Interactions revisited. *American Psychologist*, *46*, 6–15.

Finkel, S. E. (1995). *Causal analysis with panel data.* Thousand Oaks, CA: Sage.

Frese, M., & Zapf, D. (1988). Methodological issues in the study of work stress: Objective versus subjective measurement of work stress and the question of longitudinal studies. In C. L. Cooper & R. Payne (Eds.), *Causes, coping, and consequences of stress at work* (pp. 375–411). Chichester: Wiley.

Fried, Y., & Ferris, G. R. (1987). The validity of the job characteristics model: A review and a meta-analysis. *Personnel Psychology*, *40*, 287–322.

Gaillard, A. W. K. (1996). *Stress. Produktiviteit en gezondheid* [Stress. Productivity and health]. Amsterdam, The Netherlands: Nieuwezijds.

Hackman, J. R., & Oldham, G. R. (1980). *Work redesign.* Reading, MA: Addison-Wesley.

Hagenaars, J. A. (1990). *Categorical longitudinal data: Log-linear panel, trend and cohort analysis.* Newbury Park, CA: Sage.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances.* Baltimore: John Hopkins University Press.

Hobfoll, S. E., & Freedy, J. (1993). Conservation of resources. A general stress theory applied to burnout. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout. Recent developments in theory and research* (pp. 115–131). Washington, DC: Taylor & Francis.

Hom, P. W., & Griffith, R. W. (1991). Structural equations modeling test of a turnover theory: cross-sectional and longitudinal analysis. *Journal of Applied Psychology*, *76*, 350–366.

Houkes, I., Janssen, P. P. M., De Jonge, J., & Nijhuis, F. J. N. (2001). Specific relationships between work characteristics and intrinsic work motivation, burnout and turnover intention: A multi-sample analysis. *European Journal of Work and Organizational Psychology*, *10*, 1–23.

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

Iverson, R. D., & Roy, P. (1994). A causal model of behavioral commitment: Evidence from a study of Australian blue-collar employees. *Journal of Management*, *20*, 15–41.

James, L. R., & Jones, A. P. (1980). Perceived job characteristics and job satisfaction: An examination of reciprocal causation. *Personnel Psychology*, *33*, 97–135.

James, L. R., & Tetrick, L. E. (1986). Confirmatory analytic tests of three causal models relating job perceptions to job satisfaction. *Journal of Applied Psychology*, *71*, 77–82.

Janssen, P. P. M. (1992). *Relatieve deprivatie in de middenloopbaanfase bij hoger opgeleide mannen. Een vergelijking tussen drie leeftijdsgroepen* [Relative deprivation in midcareer among highly educated men. A comparison between three age groups]. PhD thesis. University of Nijmegen, Nijmegen, The Netherlands.

Janssen, P. P. M., De Jonge, J., & Bakker, A. B. (1999). Specific determinants of intrinsic work motivation, burnout and turnover intentions: A study among nurses. *Journal of Advanced Nursing*, *29*, 1360–1369.

Johnson, J. V., & Hall, E. M. (1988). Job strain, work place support and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*, *78*, 1336–1342.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. Scott Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's reference guide.* Chicago: Scientific Software International.

Kahn, R. L., & Byosiere, P. (1992). Stress in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, 2nd ed., pp. 571–650). Palo Alto, CA: Consulting Psychologists Press.

Karasek, R. A., & Theorell, T. (1990). *Healthy work: Stress, productivity and the reconstruction of working life.* New York: Basic Books.

Kasl, S. V. (1996). The influence of the work environment on cardiovascular health: A historical, conceptual and methodological perspective. *Journal of Occupational Health Psychology*, *1*, 42–56.

Kessler, R. C., & Greenberg, D. F. (1981). *Linear panel analysis. Models of quantitative change*. New York: Academic Press.

Kompier, M. A. J., & Marcelissen, F. H. G. (1990). *Handboek werkstress* [Handbook work stress]. Amsterdam, The Netherlands: NIA.

Kristensen, T. S. (1995). The demand–control–support model: Methodological challenges for future research. *Stress Medicine*, *11*, 17–26.

Kristensen, T. S. (1996). Job stress and cardiovascular disease: A theoretical critical review. *Journal of Occupational Health Psychology*, *1*, 246–260.

Lave, C. A., & March, J. G. (1980). *Modellen in de sociale wetenschappen. Een inleiding* [Models in social science. An introduction]. Alphen aan den Rijn, The Netherlands: Samsom.

Lee, R. T., & Ashforth, B. E. (1996). A meta-analytic examination of the correlates of the three dimensions of job burnout. *Journal of Applied Psychology*, *81*, 123–133.

Leiter, M. P. (1993). Burnout as a developmental process: Consideration of models. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout. Recent developments in theory and research* (pp. 237–250). Washington, DC: Taylor & Francis.

Maassen, G. H., & Bakker, A. B. (2000). Suppressor variabelen in padmodellen: Definities en interpretaties [Suppressor variables in path models: Definitions and interpretations]. *Kwantitatieve Methoden*, *21*, 43–69.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, *114*, 185–199.

Maslach, C. (1998). A multidimensional theory of burnout. In C. L. Cooper (Ed.), *Theories of organizational stress* (pp. 68–85). Oxford: Oxford University Press.

McFadden, M., & Demetriou, E. (1993). The role of immediate work environment factors in the turnover process: A systematic intervention. *Applied Psychology: An International Review*, *42*, 97–115.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Parker, S., & Wall, T. (1998). *Job and work design: Organizing work to promote well-being and effectiveness.* Thousand Oaks, CA: Sage.

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258.

Rosse, J. G., & Miller, H. E. (1984). Relationship between absenteeism and other employee behaviors. In P. S. Goodman & R. S. Atkin (Eds.), *Absenteeism* (pp. 194–227). San Francisco, CA: Jossey-Bass.

Schaufeli, W. B., & Enzmann, D. (1998). *The burnout companion to study and practice. A critical analysis.* London: Taylor & Francis.

Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). The MBI-General Survey. In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *Maslach Burnout Inventory* (3rd ed., pp. 19–26). Palo Alto, CA: Consulting Psychologists Press.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling.* Mahwah, NJ: Erlbaum.

Selye, H. (1976). *The stress of life* (2nd ed.). New York: McGraw-Hill.

Street, H., Sheeran, P., & Orbell, S. (2001). Exploring the relationships between different psychosocial determinants of depression: A multidimensional scaling analysis. *Journal of Affective Disorders*, *64*, 53–67.

Stremmel, A. J. (1991). Predictors of intention to leave child care work. *Early Childhood Research Quarterly*, *6*, 285–298.

Thierry, H. (1998). Motivation and satisfaction. In P. J. D. Drenth, H. Thierry, & C. J. de Wolff (Eds.), *Handbook of work and organizational psychology* (Vol. 4, 2nd ed., pp. 253–289). Hove, UK: Psychology Press/Erlbaum.

Tiegs, R. B., Tetrick, L. E., & Fried, Y. (1992). Growth need strength and context satisfaction as moderators of the Job Characteristics Model. *Journal of Management*, *18*, 575–593.

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin*, *109*, 524–536.

Van Breukelen, J. W. M. (1989). Personeelsverloop in organisaties. Een literatuuroverzicht en een model [Turnover in organisations. A literature review and a model]. *Gedrag & Organisatie*, *6*, 37–65.

Van Veldhoven, M., & Meijman, T. (1994). *Het meten van psychosociale arbeidsbelasting met een vragenlijst. De Vragenlijst Beleving en Beoordeling van de Arbeid (VBBA)* [Measuring psychosocial workload by means of a questionnaire. Questionnaire on the Experience and Evaluation of Work]. Amsterdam, The Netherlands: NIA.

Verbeek, M. J. C. M. (1991). *The design of panel surveys and the treatment of missing observations.* PhD thesis. NWO, Nijmegen, The Netherlands.

Verschuren, P. J. M. (1991). *Structurele modellen tussen theorie en praktijk* [Structural models between theory and practice]. Utrecht, The Netherlands: Spectrum.

Walker, L. O., & Avant, K. C. (1995). *Strategies for theory construction in nursing* (3rd ed.). Norwalk, CT: Appleton & Lange.

Warr, P. B., Cook, J. C., & Wall, P. B. (1979). Scales for the measurement of some attitudes and aspects of psychological well-being. *Journal of Occupational Psychology*, *52*, 129–148.

Xie, J. L., & Johns, G. (1995). Job scope and stress: Can job scope be too high? *Academy of Management Journal*, *38*, 1288–1309.

Zapf, D., Dormann, C., & Frese, M. (1996). Longitudinal studies in organisational stress research: A review of the literature with reference to methodological issues. *Journal of Occupational Health Psychology*, *1*, 145–169.

Received 19 April 2001; revised version received 30 December 2002