Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Vocational Behavior 63 (2003) 532–559
Journal of Vocational Behavior
www.elsevier.com/locate/jvb

Does Work–Home Interference mediate the relationship between workload and well-being?

Sabine A.E. Geurts,[a,*] Michiel A.J. Kompier,[a] Susan Roxburgh,[b] and Irene L.D. Houtman[c]

[a] *Department of Work and Organizational Psychology, University of Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands*
[b] *Kent State University, Kent, OH, USA*
[c] *TNO Work and Employment, Hoofddorp, The Netherlands*

Received 3 January 2001

Abstract

Drawing on the Effort–Recovery (E–R) model, the current study investigated to what extent Work–Home Interference (WHI) mediated the relationship between workload and two indicators of well-being, that is, (a) affective well-being (i.e., work-related negative affect and depressive mood) and (b) subjective health (i.e., health complaints). In Part I of this study, Structural Equation Modeling (SEM) was used to test several competing models of mediation—full, partial, or no mediation—in three homogeneous samples (166 medical residents, 194 child care workers, and 224 bus drivers). In Part II of this study, we cross-validated the best fitting model in an independent heterogeneous sample (1421 Dutch workers). The results provided support for the E–R model in that WHI played a significant role in mediating the impact of workload on workers' well-being. WHI fully mediated the relationship of workload with depressive mood and health complaints, and partially mediated the relationship with work-related negative affect. This differential role of WHI indicates that WHI might play a more crucial mediating role with respect to general (context-free) indicators of well-being than with respect to work-related indicators of well-being. In general, the findings of the current study suggest that workload exerts its negative effects on well-being (at least partly) through a process of spillover of negative load-effects that impede recovery during the non-working hours.

Keywords: Work–Home Interference; Workload; Affective well-being; Subjective health

* Corresponding author. Fax: +31-24-361-5937.
E-mail address: S.Geurts@psych.kun.nl (S.A.E. Geurts).

0001-8791/$ - see front matter
doi:10.1016/S0001-8791(02)00025-8

1. Introduction

Striking changes have occurred during the past two decades in the composition of the work force, as well as in the nature of work itself. In both Europe and the United States, the majority of women are now employed in the paid labor force. Whereas in 1991 39% of the European working population was female, this has increased to 42% in 1996 (Paoli, 1997). Recent statistics in the Netherlands, the country for which this paper presents findings, indicate that 52% of the women have paid jobs (SCP, 2001). In the United States more than three quarters of women between the ages of 35–44 are in the labor force (Reskin & Padavic, 1994). Associated with this development is the growing number of dual-career couples. For instance, in the Netherlands approximately 60% of the (married) couples have two incomes (SCP, 2000). A recent survey of approximately 3000 American workers revealed that in three-quarters of the dual-earner couples, both partners held full-time jobs (Bond, Galinsky, & Swanberg, 1998). These developments have increased the likelihood that both female and male workers have considerable domestic responsibilities in addition to their work obligations (Allen, Herst, Bruck, & Sutton, 2000). In fact, the proportion of Dutch workers (i.e., 47%) that nowadays combine substantial work obligations (⩾12 h per week) with substantial domestic obligations (⩾8 h per week) is far above those proportions in 1995 and in 1975 (respectively, 38 and 20%) (SCP, 2001).

The nature of work is also changing, particularly in terms of how mentally and emotionally demanding workers perceive their jobs to be. A recent survey conducted among a representative sample of the European work force revealed that a growing number of workers (from 35% in 1991 to 42% in 1996), particularly in the Netherlands (from 47% in 1991 to 58% in 1996), reported working at high speed 'most of the time' (Paoli, 1997). A similar trend among US employees has been reported by Bond and his colleagues (1998). Thus, there is evidence that workers are facing greater pressures at home and at work because of the proliferation of dual-earner couples and because of changes in the nature of the workplace.

Research by Galinsky, Bond, and Friedman (1993) showed that a substantial proportion of employed parents (i.e., 40%) experienced problems or conflicts in combining work and family demands, often referred to as work-to-family conflict or Work-to-Home Interference. Empirical research has consistently shown that work demands are far more likely to interfere with domestic obligations than the other way around (Burke & Greenglass, 1999; Frone, Russell, & Cooper, 1992; Leiter & Durup, 1996). In fact, Frone et al. (1992) found that 'work interfering with home' was reported three times more frequently than 'home interfering with work' by both male and female employed adults. Bond et al. (1998) reported similar results for an American sample, with an even greater divergence in the prevalence of the two types of perceived interference. Whereas 32% of the US workers reported that during the last three months they have had not enough time for themselves because of their jobs, only 4% reported that during the same period family or personal life had kept them from getting paid labor done on time. In light of the higher prevalence of Work-to-Home Interference, and its potential relevance for understanding the impact of work and family experiences on health and psychological well-being, in this paper we focused exclusively on Work-to-Home Interference (WHI).

Social scientists from various disciplines have studied work and family issues, in particular work-to-family conflict. Within Occupational Health Psychology (OHP), the main focus has been on the relationship of WHI and stress-related outcomes for workers and their families (Barnett, 1998). Although work–family research has made important contributions to OHP, research within this field has at least two limitations. A methodological limitation is that many work–family studies have been conducted with very homogeneous groups and work contexts. According to Kossek and Ozeki (1998),

> samples range from female nurses and male engineers to male executives to Israeli prison guards and to high powered dual-career couples (p. 141).

We agree with Kossek and Ozeki in that restricting the range of a single study to specific gender or occupational groups limits the generalizability of the findings within this field.

A theoretical limitation is that work–family researchers have not based their definitions of and hypotheses about WHI on strong conceptual frameworks (cf. Grandey & Cropanzano, 1999). Where conceptual frameworks were found, researchers within OHP have relied on role stress theory (Kahn, Wolfe, Quinn, Snoek, & Rosenthal, 1964), which hypothesizes that participation in one role limits the allocation of resources (e.g., time and energy) to other roles (Greenhaus & Beutell, 1985). From this perspective, however, it remains unclear how WHI should be embedded in the stressor–strain relationship. WHI is often considered a potential source of stress that, in addition to other potential stressors, will have adverse effects on well-being, such as dissatisfaction (Burke, 1988; Leiter & Durup, 1996; Rice, Frone, & McFarlin, 1992), negative affect (e.g., Burke, 1993; Frone et al., 1992, 1994, 1996; MacEwen & Barling, 1994; Parasuraman, Greenhaus, & Granrose, 1992), anxiety (Kinnunen & Mauno, 1998), depressive complaints (Burke, 1988; Frone, Russell, & Cooper, 1997a; Kinnunen & Mauno, 1998), exhaustion, psychosomatic symptoms, poor physical health (e.g., Burke, 1988; Frone, Russell, & Cooper, 1991; Frone, Russell, & Barnes, 1996; Kinnunen & Mauno, 1998; Thomas & Ganster, 1985), and sleeping disorders (Burke, 1988; Geurts, Rutte, & Peeters, 1999).

In other studies, WHI has been treated as an *outcome* of stress (i.e., as an indicator of strain), caused by potential stressors, particularly those found in the work domain. Strong and consistent evidence has been provided for quantitative workload (i.e., having too many things to do and not enough time to do them) as the most, and often only relevant antecedent of work-to-family conflict (Burke, 1988; Frone et al., 1992, 1997b; Grzywacz & Marks, 2000; Parasuraman, Purohit, Godshalk, & Beutell, 1996; Wallace, 1997).

In addition to WHI as a source or outcome of stress, another line of research has provided evidence that WHI *mediates* the stressor–strain relationship. Parasuraman et al. (1996) found that male and female entrepreneurs who experienced work-role overload reported higher levels of WHI, which in turn was related to their general life stress. A similar intervening role was found by Grandey and Cropanzano

(1999) among university professors: job stressors (i.e., role conflict and ambiguity in the workplace) were not only directly related to job strain (i.e., work satisfaction and work tension), but also indirectly through WHI. Finally, Stephens, Franks, and Atziena (1997) showed that WHI was an important mediating mechanism in the relationship between job stressors and depression among female employees with care-giving responsibilities.

The present study was designed to overcome the methodological and theoretical limitations noted above and to further examine the mediating role of WHI on psychological well-being and subjective health among workers. First, we examined WHI in multiple homogeneous samples, and we validated our results using a large-sized sample that is heterogeneous, both in terms of individual and organizational diversity. This provides the opportunity to generalize results from specific groups of employees to the general population of employed workers. Second, our definition of WHI and our hypotheses about its position in the stressor–strain relationship were based on a clear theoretical model—the Effort–Recovery Model (Meijman & Mulder, 1998).

1.1. The theoretical perspective

The Effort–Recovery (E–R) Model (Meijman & Mulder, 1998) has its roots in the classical frame of thought of exercise physiology, and in particular in its applications to the study of workload in relation to a person's capacity. Although many prominent models in OHP focus on the relationship of workload (or, more general, of task characteristics or job stressors) with individuals' well-being (e.g., Karasek's Job-Demands-Job Control Model, Hackman & Oldhams' Job Characteristics Model, Warr's Vitamin Model), none of these models pay attention to what psychobiological processes are involved that may explain why well-being is likely to deteriorate in a stressful work setting. The E–R Model sheds light on the underlying mechanisms in the relationship of workload with well-being by assuming that recovery from workload effects during the non-working period plays a crucial role. A central idea of this work psychological model is that meeting work demands that require effort produces two kind of outcomes, the product itself (i.e., the tangible result of work activities) and the short-term physiological and psychological reactions (the 'costs' and 'benefits' for the individual). Under normal circumstances these reactions are reversible: after a short respite from work demands, the worker's psychobiological systems will stabilize again at a baseline level and recovery from the effects of work demands that have built up during the working period will be sufficient. As a result, fatigue and other effects of the demanding work situation will diminish and finally disappear.

But what if opportunities for recovery after being exposed to workload are insufficient? This might be the case when demands put on the individual do not cease after working time but continue to exist during the non-working period (e.g., when workers have extensive domestic obligations). Recovery may also be hampered when workers are slowly unwinding (also referred to as sustained activation, Ursin, 1980), because load-effects of a stressful working day do not unfold immediately but last during the evening hours at home (e.g., when workers have difficulty to relax after a demanding working period) (see Sonnentag, 2001).

If opportunities for recovery after being exposed to workload are insufficient, the psychobiological systems are activated again before they had a chance to stabilize at a baseline level. The person, still in a sub-optimal state, will have to make an additional compensatory effort to perform adequately during the next working period. This may result in an increased intensity of the load reactions during the working period, which in turn will make higher demands on the recovery process. As a result of such an insufficient recovery process, negative load-effects accumulate and result in longer term negative effects, such as prolonged fatigue, chronic stress, sleep deprivation and other psychosomatic complaints (Sluiter, 1999; Ursin, 1980). Under unchanged conditions these symptoms may develop into manifest health problems (cf. Kompier, 1988; Sluiter, 1999).

The pivotal role of recovery during the non-working period makes the E–R model a relevant perspective for studying WHI. Recently, Sonnentag (2001) used this theoretical perspective to study how individuals use their non-working hours to recover from work. The basic assumption in the current study was that load-effects that have built up during working hours, particularly when work demands have been excessive, spillover to the non-work situation and limit opportunities for recovery during non-working hours. As a consequence, workers may start the next working day in a sub-optimal state, requiring compensatory effort and making higher demands on recovery. The accumulative process just described may be started, resulting in impaired well-being that in the long run become increasingly serious and chronic by nature.

Drawing on this model, we defined WHI as a process whereby a worker's functioning and recovery in the home domain were hampered by load-effects that have built up in the work domain. Notice here that WHI was defined in terms of recovery: a high level of WHI implies that recovery in the home situation was impeded. It was hypothesized that WHI plays a crucial mediating role in the relationship between workload and well-being.

1.2. The conceptual model, samples, and design

The conceptual model that guided the present research is presented in Fig. 1. In the basic model, represented by the continuous arrows (paths 1 and 2), it was assumed that WHI plays a full mediating role in the relationship between workload

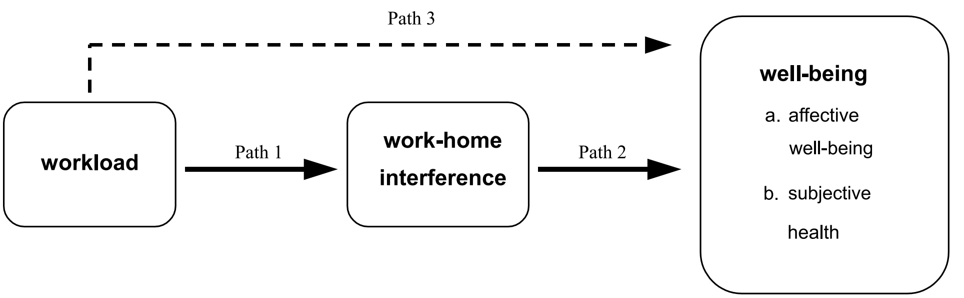

Fig. 1. The hypothetical model (for explanation, see text).

and two indicators of well-being, that is (a) affective well-being, merely reflecting feelings and moods (e.g., feeling depressed), and (b) subjective health, reflecting health complaints (e.g., dizziness, pain in the chest or heart area, fatigue, and backache). In the extended model, including the interrupted arrow (path 3) to each indicator of well-being, it was assumed that WHI partially mediates these relationships. In general, we expected that WHI is more likely to play a full mediating role in the relationship of workload with general (i.e., context-free) indicators of well-being than with work-related (i.e., not context-free) indicators of well-being (see also, Geurts et al., 1999).

The current study was conducted in two parts. In both parts the same and context-free indicator of subjective health (i.e., health complaints) was used. Therefore, in both Part I and Part II a full mediating role of WHI was expected with respect to health complaints. The indicator of affective well-being, however, was different in Part I and Part II. In Part I, we measured work-related negative affect, reflecting the three axes of affective well-being proposed by Warr (1987, 1990, see further the measure description). As we asked for negative affect experienced at work, this measure was not context-free (though 'work-related' does not necessarily mean 'caused by work'). In Part II, we measured depressive mood in a context-free way. In Part I it was, therefore, expected that an extended model (with an additional direct path between workload and work-related negative affect) would better fit the data than the basic model (assuming a full mediating role of WHI). In Part II, on the other hand, it was expected that the basic model (assuming a full mediating role of WHI in the relationship of workload with both indicators of well-being) would provide a sufficiently satisfactory fit.

In Part I, the assumed mediating role of WHI was investigated in three homogeneous samples (i.e., medical residents, child-care workers, and bus drivers) by employing Structural Equation Modeling. First, the basic model was tested against the extended models in each of the three homogeneous samples separately, in order to find out whether WHI plays a full, a partial or no mediating role in the relationship between workload and the two indicators of well-being. Second, a multi-sample analysis was conducted across the three homogeneous samples simultaneously to test whether the basic model (or one of the extended models) holds true across the three samples, assuming equal parameter estimates. In Part II of the current study, the model that best fitted the three homogeneous (Part I) samples was cross-validated in an independent and large-sized heterogeneous sample.

2. Part I

2.1. Data and method

2.1.1. Samples and procedure

Data for the analysis among the homogeneous samples was drawn from three cross-sectional studies. *Sample* 1 consisted of 166 medical residents working at an academic hospital in the Netherlands (see also Geurts et al., 1999). The final response rate was 60%. The sample consisted of 105 men and 59 women, of which 75% was married (or co-habiting) and 36% had one or more home-living children. Sixty-one percent of the sample was employed on a full-time basis—that is, they were contracted to work 46 h a week. The age of the respondents ranged from 25 to 49 years, with an average age of 32 years. *Sample* 2 consisted of 194 employees working in a daily child care center in the Netherlands. The response rate was 77%. All the respondents were female, of which 73% was married (or co-habiting) and 38% had one or more home-living children. Fifty-three percent was employed on a full-time basis. The age of the respondents ranged from 20 to 50 years, with an average age of 30 years. *Sample* 3 consisted of 224 employees working for a single bus driving company in the Netherlands. Because three-quarters of the employees were bus drivers, we refer to this sample as 'bus drivers.' The response rate was 75%. The sample consisted of 185 men and 37 women. Information regarding the marital and parental status of the respondents was unavailable. Eighty-four percent of the respondents worked full-time. Half of the respondents were over 45 years, 39% between 35 and 44 years and 11% under 35 years of age. In all three samples the respondents were informed about the purpose and the interests of the study both in writing (by a publication in the company magazine) and face-to-face by a joint presentation from the company's personnel department and the researchers.

2.1.2. Survey measures

Basic descriptive statistics (means, *SD*s, Cronbach's α) of the variables included in the current study are presented in Table 1 for each of the three homogeneous samples (as well as of the heterogeneous sample, which are discussed in Section 3). Table 2 shows the intercorrelations of the variables under study for each sample separately.

Workload was assessed by using seven items from the (Dutch) Questionnaire on Experience and Evaluation of Work (Van Veldhoven, Meijman, Broersen, & Fortuin, 1997). This questionnaire has been extensively validated in Dutch samples (Van Veldhoven, Broersen, & Fortuin, 1999). The original scale of 11 items measuring 'work pace and work load' appears to be a unidimensional and reliable scale. The items used in the current study were: (a) "Do you have too much work to do?"; (b) "Do you work under time pressure?"; (c) "Do you have to hurry?"; (d) Can you do your work at ease?" (reversed); (e) Do you have problems with the pace of work?"; (f) Do you have problems with the pressure of work?"; and (g) "Would you prefer a calmer work pace?." Respondents answered on a four-point scale, ranging from 'never' (1) to 'always' (4), with a higher mean score indicating a higher perceived work load. The internal consistency in all three samples was satisfactory (see Table 1). A One-way Analysis of Variance (ANOVA) showed that the average workload score among medical residents ($M = 2.34$) and child care workers ($M = 2.33$) was significantly higher than the average score among bus drivers ($M = 2.17$) (($F_{(2,575)}$) = 5.1, $p < .01$). Compared to the average score of a large-sized reference group ($n = 13{,}491$; $M = 2.31$; Van Veldhoven et al., 1999), only the bus drivers reported significantly lower workload ($z = -4.57$, $p < .001$).

Work–Home Interference (*WHI*) was measured with five items that were adopted from Kopelman, Greenhaus, and Connolly (1983). The participants had to indicate

Table 1
Descriptive statistics (Means, *SD*s, and Cronbach's α) for the three homogeneous samples (Part I) and the large heterogeneous sample (Part II)

Statistics	Part I									Part II		
	Sample 1 Medical residents ($n=166$)			Sample 2 Child care workers ($n=194$)			Sample 3 Bus drivers ($n=224$)			Sample 4 Heterogeneous ($n=1421$)		
	M	*SD*	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*	α
1. Workload[a]	2.34	.54	.87	2.33	.64	.91	2.17	.54	.85	2.59[a]	.47	.71
2. Work-to-Home interference	2.47	.54	.81	1.64	.47	.75	1.80	.52	.76	1.47	.44	.72
3a. Work-related negative affect	1.83	.38	.78	1.74	.44	.83	1.79	.44	.87	–	–	–
3b. Depressive mood	–	–	–	–	–	–	–	–	–	1.34	.35	.85
4. Health complaints	1.25	1.69	.69	2.97	2.70	.78	2.67	2.58	.79	2.62	2.67	.79

[a] The workload scale used in Sample 4 is not identical but highly similar to the workload measure used in the other three samples.

Table 2
Zero-order correlations for the three homogeneous samples and the large heterogeneous sample

Statistics	Sample 2 Child care workers ($n = 194$)					Sample 4 Heterogeneous ($n = 1421$)				
	1	2	3a	3b	4	1	2	3a	3b	4
1. Workload	–	.59	.70	–	.30	–	.33	–	.18	.15
2. Work-to-Home interference	.55	–	.64	–	.43	.54	–	–	.42	.32
3a. Work-related negative affect	.36	.51	–	–	.47	.50	.46	–	–	–
3b. Depressive mood	–	–	–	–	–	–	–	–	–	.53
4. Health complaints	.25	.38	.26	–	–	.36	.37	.41	–	–
	Sample 1 Medical residents ($n = 166$)					Sample 3 Bus drivers ($n = 224$)				

Notes. All correlations are significant at $p < .01$.
Below the diagonal interrcorrelations are presented from, respectively, Samples 1 and 3. Above the diagonal, intercorrelations are presented from, respectively, Samples 2 and 4.

how often their work interfered with their private life. The items were: "How often... (a) "do you work so hard that you cannot do some of the things you would like to do?"; (b) "are you irritable at home because your work is so demanding?"; (c) "do the demands of your job make it difficult for you to feel relaxed at home?"; (d) "does you work schedule interfere with your private life?"; and (e) "does your work take up time that you would like to spend on your private life?." Participants responded on a four-point scale ranging from 'never' (1) to 'always' (4), with a higher mean score indicating a higher level of WHI. A closer inspection of these items revealed that two components may be distinguished, that is a time-based component (one's functioning at home being impeded by work that takes up too much time; see items a, d, and e) and a strain-based component (one's functioning at home being impeded by strain-effects that have built up at work; see the items b and c). Drawing on the E–R Model, we assumed that the time-based component would limit opportunities for quantitative recovery during non-working hours, and that the strain-based component would limit opportunities for qualitative recovery in the home domain. Confirmatory Factor Analysis (CFA) revealed that all five items loaded on the same factor (see also the results of CFA in the following section). The internal consistency of the full scale in all three samples was satisfactory (see Table 1). A One-way ANOVA revealed that medical residents experienced a higher level of WHI ($M = 2.47$) than child care workers ($M = 1.64$) and bus drivers ($M = 1.80$) (($F_{(2,573)}$) = 133.4, $p < .001$). Among the medical residents almost 90% reported to experience WHI at least occasionally, whereas among the child care workers and the bus drivers these percentages were, respectively, about 50 and 60%.

Work-related negative affect was measured by using nine items of the scale 'psychological complaints' of the Dutch Questionnaire Organizational Stress-Doetinchem (VOS-D; Bergers, Marcelissen, Wolff, & de , 1986). Participants were asked how frequently they experienced the following feelings at work: (a) "angry," (b) "relaxed" (reversed), (c) "confused," (d) "cheerful" (reversed), (e) "worried," (f) "depressed," (g) "calm" (reversed), (h) "frustrated," and (i) irritated." The responses ranged from never (1) to always (4), with higher scores indicating a higher level of work-related negative affect.The nine items reflect the three principal dimensions of affective well-being proposed by Warr (1987, 1990): displeasure–pleasure (items a, h, and f), anxiety–comfort (items b, c, e, and g), and depression–enthusiasm (items d and f).The internal consistency in all three samples was satisfactory (see Table 1). A one-way ANOVA revealed that with respect to their level of negative affect the three samples did not differ significantly (medical residents: $M = 1.83$; child care workers: $M = 1.74$; bus drivers: $M = 1.79$; $((F_{(2,591)}) = 1.77$, *ns*). Compared to the average score of a large-sized reference group of 2800 workers in various occupations and professions ($M = 1.75$, $SD = .39$; Bergers et al., 1986), the three samples experienced a similar level of work-related negative affect.

Health complaints were measured with a Dutch questionnaire on subjective health developed by Dirken (1969), the so-called VOEG. In this study the 13-item version (VOEG13) was used (Joosten & Drop, 1987), which is extensively validated in Dutch samples (Cronbach's α's are around .80). The Central Bureau of Statistics (CBS) in the Netherlands uses this instrument to assess the general health of the Dutch population. The VOEG13 consists of dichotomous items ('no' (0) and 'yes' (1)), each asking whether or not one occasionally suffered from a range of health complaints, for instance, pain in the chest or heart area, short of breath, headache, an upset stomach, dizziness, fatigue, tingling limbs, being listless, pain in muscles or bones, and backache. The total number of 'yes' responses constituted the measure of health complaints (i.e., 'sum score'). In order to approach a normal distribution of this measure, a $\log_{10}$ transformation was performed (because a logarithm of '0' does not exist, a '1' was added to the raw score, prior to the $\log_{10}$ transformation), resulting in an acceptable skewness in all three samples. The internal consistency in all samples was satisfactory. A One-way Variance Analysis showed that child care workers ($M = 2.97$) and bus drivers ($M = 2.67$) experienced a significantly higher number of health complaints than medical residents ($M = 1.25$) $((F_{(2,558)}) = 25.2$, $p < .001$). Compared to a large-sized reference group of 7717 workers from various branches ($M = 2.78$, Houtman et al., 1995), medical residents experienced relatively few health complaints.

2.1.3. Data analysis

A comprehensive test of the basic model (and extended models) in Fig. 1 was performed using Structural Equation Modeling (SEM) (Jöreskog & Sörbom, 1993). In general, with help of SEM it is possible to investigate to what extent a postulated pattern of relationships is consistent with the observed data in one or more samples. SEM can be conducted in one sample at a time as well as in multiple samples simultaneously (Multi-Sample Analysis, MSA). We tested our basic model (and extended models) on the scale level, rather than on the item level, because it is generally undesirable to have too many manifest indicators for a latent variable. Due to the high number of parameters to be estimated, it is likely to find an unsatisfactory fit (cf. Bagozzi & Heatherton, 1994), as well as less reliable and unstable results. Therefore, Bentler (1985) has recommended the ratio of at least five subjects per parameter in order to derive reliable parameter estimates.

To make sure, however, that the variables included in our model were indeed different constructs, we first conducted two preliminary Confirmatory Factor Analyses (CFA) (Jöreskog & Sörbom, 1993). Given the relatively high intercorrelations between workload and WHI (see Table 2), we performed a CFA on the seven workload items and the five WHI items. We also wanted to ensure that the two indicators of well-being (i.e., work-related negative affect and health complaints) were not both indicators of the same underlying construct but indeed represented distinct constructs. Therefore, a CFA was performed on the nine items measuring work-related negative affect and the 13 items measuring health complaints. By using CFA, we compared the fit of a correlated two-factor model (where each factor represented one construct) to the fit of an alternative (one-factor) model that hypothesized that one single factor underlay the relationships among the items. When the two-factor model fitted the data significantly better than the one-factor model, it is plausible that the two constructs under study were indeed distinct constructs.

Next, we tested the basic model (paths 1 and 2) against the extended models (including additional path 3) in each of the three samples separately. The demographic variables age and gender were included as control variables, because they may confound the postulated relationship between work demands and workers' well-being. With regard to each indicator of well-being, three alternative models were tested:

- In Model 1 we assumed that WHI fully mediated this relationship (only paths 1 and 2 were specified);
- In Model 2, we assumed that WHI only partially mediated this relationship (Model 1 was extended by an additional path 3); When path 2 was still significant, WHI played at least a partial mediating role.
- Model 3 was similar to Model 2, except that there was a constraint on path 2: This parameter estimate had to be equal to the one in Model 1, assuming full mediation.

The difference in χ^2 between Model 2 and Model 3 indicated to what extent WHI played a mediating role in the relationship between workload and each of the two indicators of well-being:

1. When the difference in χ^2 between Model 2 and Model 3 was *not* significant, then the model assuming that WHI played a *full* mediating role was the best fitting model.
2. When the difference in χ^2 between Model 2 and Model 3 was significant, but path 2 was significant in Model 2, then the model assuming a *partial* mediating role of WHI was the best fitting model.
3. When the difference in χ^2 between Model 2 and Model 3 was significant, and path 2 was *not* significant in Model 2, then the model assuming *no* mediating role of WHI was most consistent with the observed data.

In the final step, an MSA was carried out to find out whether the path estimates were equal or different across the three samples. By using MSA, it is possible (1) to estimate separately the parameters for each of the independent samples, and (2) to test whether specified parameters are equal across different samples. In Step 1 a fit function was calculated, hypothesizing that the parameters of the common paths were the same in the three samples. In Step 2 a fit function was calculated, hypothesizing that all parameters were different across the three samples. If the fit in Step 2 did not significantly improve compared to the fit is Step 1, the assumption that the parameters of the common paths were equal in the three samples was confirmed. This result, then, implies that the model was robust not only in terms of type of relationships (i.e., model structure) but also in terms of strengths of relationships.

Several commonly used fit indices were used to assess the overall fit of the model (Bollen & Scott Long, 1993; Jöreskog, 1993). These were as follows:

1. Chi-square statistic (χ^2), which is a measure of overall fit of the model to the data. A non-significant small χ^2 corresponds to a good fit.
2. Adjusted Goodness-of-Fit Index (AGFI; Jöreskog, 1993), which is also an overall measure of fit, but with taking into account the number of degrees of freedom. Values greater than .90 indicate a reasonable fit of the model, whereas values equal to or greater than .95 indicate a close fit (Browne & Cudeck, 1993; Verschuren, 1991).
3. Root-Mean-Square Error of Approximation (RMSEA; Browne & Cudeck, 1993), which gives useful information for the assessment of approximation of the population, and has the potential to reward more parsimonious models; Values of .08 or less indicate a reasonable fit of the model, whereas values lower than .05 indicate a close fit (Browne & Cudeck, 1993; Verschuren, 1991).
4. Non-Normed Fit Index (NNFI; Bentler & Bonett, 1980), indicating the *incremental* fit of a particular model as compared to a null model that assumes zero relationships among the variables. Values less than .90 usually mean that the model can be improved substantially.
5. Comparative Fit Index (CFI; Bentler, 1990), with a 0–1 interval in which the value 1 stands for a perfect model fit. This index seems to be one of the best indices in the context of model misspecifications and variations in sample size (Bentler, 1990).
6. Expected Cross-validation Index (ECVI); Browne & Cudeck, 1993). This is a measure of "discrepancy between the fitted covariance matrix in the analyzed sample and the expected covariance matrix that would be obtained in another sample of the same size" (Jöreskog, 1993, p. 307). The model with the smallest ECVI can be considered as the most robust model across samples.

2.2. Results

2.2.1. Preliminary Confirmatory Factor Analyses

Confirmatory Factor Analysis was used to examine the discriminant validity of the seven workload items and the five WHI items. In all three samples, the two-factor model (Sample 1: $\chi^2_{(53)} = 149.2$, $p = .00$, GFI = .85, RMSEA = .12, NNFI = .86, and CFI = .89; Sample 2: $\chi^2_{(53)} = 124.6$, $p = .00$, GFI = .90, RMSEA = .09, NNFI = .92, and CFI = .93; Sample 3: $\chi^2_{(53)} = 141.1$, $p = .00$, GFI = .90, RMSEA = .09, NNFI = .89, and CFI = .91) provided a significantly better fit than the alternative one-factor model (Sample 1: $\Delta\chi^2_{(1)} = 108.78$, $p < .001$; Sample 2: $\Delta\chi^2_{(1)} = 64.04$, $p < .001$; Sample 3: $\Delta\chi^2_{(1)} = 1108.58$, $p < .001$). All workload items had a factor loading of .61 or higher in each sample (with median loadings in the three samples of .70, .77, and .69, respectively). All WHI items had a factor loading of. 45 or higher in each sample (with median loadings in the three samples of .72, .63, and .65, respectively). These results indicate that, although workload and WHI were substantially interrelated ($r = .55$, .59, and .54 in Samples 1, 2, and 3, respectively), they could be empirically distinguished.

Confirmatory Factor Analysis was also used to examine the factorial validity of the nine items measuring work-related negative affect and the 13 items measuring health complaints. In all three samples, the two-factor model (Sample 1: $\chi^2_{(208)} = 359.3$, $p = .00$, GFI = .82, RMSEA = .07, NNFI = .73, and CFI = .76; Sample 2: $\chi^2_{(208)} = 412.8$, $p = .00$, GFI = .84, RMSEA = .07, NNFI = .79, and CFI = .81; Sample 3: $\chi^2_{(208)} = 860.9$, $p = .00$, GFI = .75, RMSEA = .11, NNFI = .63, and CFI = .66) provided a significantly better fit than the alternative one-factor model (Sample 1: $\Delta\chi^2_{(1)} = 110.04$, $p < .001$; Sample 2: $\Delta\chi^2_{(1)} = 99.84$, $p < .001$; Sample 3: $\Delta\chi^2_{(1)} = 308.58$, $p < .001$). In spite of the fact that the two-factor model fitted the data significantly better than the one-factor model, the fit of the first still fell below commonly accepted cut-off values for good fit. Follow-up analysis revealed that permitting correlated error terms between a limited number of item pairs belonging to the same theoretical construct improved the fit substantially. However, as our purpose was not to obtain the best possible fit, but only to show that a two-factor model better accounted for the observed data than the one-factor model, we refrained from doing so. All items measuring work-related negative affect had factor loadings varying from .23 to .71 in Sample 1 (with median loading of .53), varying from .38 to .77 in Sample 2 (with a median loading of .61), and varying from .55 to .73 in Sample 3 (with a median loading of .68). All items measuring health complaints had factor loadings varying from .18 to .67 in Sample 1 (with a median factor loading of .32), varying from. 26 to .64 in Sample 2 (with a median factor loading of .51), and varying from .34 to .76 in Sample 3 (with a median factor loading of .47). These results indicate that, although work-related negative affect and health complaints were not independent from one another ($r = .26$, .47, and .41 in Samples 1, 2, and 3, respectively), the two variables were empirically distinct indicators of well-being.

2.2.2. Testing the basic model against the extended models in each of the three samples

Table 3 shows for each sample separately the fit measures of the basic model (M1, including only paths 1 and 2 (a + b) of Fig. 1) and various extended models, assuming a partial mediating role of WHI in the relationship between workload and each of the two indicators of well-being. In Model 2a (M2a) the basic Model 1 was extended by including path 3a, assuming that WHI played a partial mediating role in the relationship between workload and work-related negative affect. In Model 2b (M2b) the basic Model 1 was extended by including path 3b, assuming that WHI played a partial mediating role in the relationship between workload and health complaints. Model 3a was similar to Model 2a, but with a constraint on the parameter estimate of path 2a to be equal to the estimate in Model 1 (assuming a full mediating role of WHI). Model 3b was similar to Model 2b, but with a constraint on the parameter estimate of path 2b to be equal to the estimate in Model 1 (assuming a full mediating role of WHI).

Table 3
Results of covariance structure analyses: Fit indices of the hypothesized model and alternative models assuming different mediating roles of WHI

Model	χ^2	*df*	*p*	AGFI	RMSEA	NNFI	CFI	Model comparison	$\Delta\chi^2$	*df*	*p*
Sample 1: medical residents (n = 166)											
M1	2.52	3	.47	.96	.00	1.02	1.00				
M2a	.50	2	.78	.99	.00	1.08	1.00	M2a ↔ M1	2.02	1	*ns*
M3a	.50	3	.92	.99	.00	1.09	1.00	M3a ↔ M2a	0	1	*ns*
M2b	2.19	2	.33	.95	.02	.99	1.00	M2b ↔ M1	.33	1	*ns*
M3b	2.26	3	.52	.97	.00	1.03	1.00	M3b ↔ M2b	.07	1	*ns*
Sample 2: child care workers (n = 194)											
M1	68.2	3	.00	.37	.31	.08	.78				
M2a	.79	2	.67	.99	.00	1.03	1.00	M2a ↔ M1	67.5	1	<.001
M3a	25.9	3	.00	.72	.19	.62	.92	M3a ↔ M2a	25.1	1	<.001
M2b	65.6	2	.00	.91	.37	.57	.79	M2b ↔ M1	2.6	1	*ns*
M3b	66.6	3	.00	.38	.31	−.05	.79	M3b ↔ M2b	1.0	1	*ns*
Sample 3: bus drivers (n = 224)											
M1	30.1	3	.00	.72	.20	.39	.88				
M2a	8.7	2	.01	.87	.12	.77	.97	M2a ↔ M1	21.4	1	<.001
M3a	14.9	3	.00	.85	.13	.73	.95	M3a ↔ M2a	6.2	1	<.05
M2b	27.8	2	.00	.60	.23	.13	.88	M2b ↔ M1	2.3	1	*ns*
M3b	28.5	3	.00	.73	.19	.42	.88	M3b ↔ M2b	0.7	1	*ns*

Notes. χ^2, chi-square; *df*, degrees of freedom; AGFI, Adjusted Goodness of Fit Index; RMSEA, Root Mean Square Error of Approximation; NNFI, Non-Normed Fit Index; CFI, Comparative Fit Index; $\Delta\chi^2$, chi-square difference. M1, basic model (paths 1, 2a, and 2b); M2a, extended model including path 3a from workload to work-related negative affect; M3a, M2a but with the parameter estimate of path 2a fixed as in M1; M2b, extended model including path 3b from workload to health complaints; and M3b, M2a but with the parameter estimate of path 2a fixed as in M1.

This basic Model 1 provided a very good fit in Sample 1. However, the fit of M1 in both Samples 2 and 3 could be improved. The extended Model 2a, in which a direct relationship between workload and work-related negative affect (path 3a) was unconstrained, provided a significantly improved fit in both Sample 2 ($\Delta\chi^2_{(1)} = 67.5$, $p < .001$) and Sample 3 ($\Delta\chi^2_{(1)} = 21.4$, $p < .001$). A significant parameter estimate of path 2a from WHI to work-related negative affect ($\beta = .35$, $p < .001$ and $\beta = .28$, $p < .001$ in Samples 2 and 3, respectively) indicates that WHI played at least a partial mediating role in these samples. The fit of the alternative Model 3a in which the parameter estimate of path 2a was fixed to be equal to the one in Model 1 (assuming a full mediating role of WHI) deteriorated significantly in both samples ($\Delta\chi^2_{(1)} = 25.1$, $p < .001$ in Sample 2 and $\Delta\chi^2_{(1)} = 6.2$, $p < .05$ in Sample 3). In Sample 1, none of the competing models 2a and 3a fitted the data significantly better than the basic Model 1. These results indicate that with regard to the relationship between workload and work-related negative affect, WHI played a *full* mediating role in Sample 1 and a *partial* mediating role in Samples 2 and 3.

The alternative Model 2b, in which a direct relationship between workload and health complaints (path 3b) was unconstrained, did not provide a significantly improved fit in both Sample 2 ($\Delta\chi^2_{(1)} = 2.6$, $p = ns$) and Sample 3 ($\Delta\chi^2_{(1)} = 2.3$, $p = ns$) compared to the basic Model 1. In both samples, path 2b from WHI to health complaints was significant ($\beta = .44$, $p < .001$ and $\beta = .37$, $p < .001$ in Samples 2 and 3, respectively), indicating that WHI played a mediating role in these samples. The fit of alternative Model 3b (M3b) in which the parameter estimate of path 2b was fixed to be equal to the one in Model 1 (assuming a full mediating role of WHI), did not deteriorate significantly in both samples ($\Delta\chi^2_{(1)} = 1.0$, $p = ns$ in Sample 2 and $\Delta\chi^2_{(1)} = 0.7$, $p = ns$ in Sample 3). In Sample 1, none of the competing models 2b and 3b fitted the data significantly better than the basic Model 1. These results indicate that with regard to the relationship between workload and health complaints, WHI played a *full* mediating role in all three samples.

2.2.3. Testing for invariance of parameter estimates across the three samples

Our next step was to test for invariance of the basic pattern of relationships across the three samples. Thus we performed an MSA in which we compared four different models by means of a so-called χ^2 difference test (cf. Bentler & Bonett, 1980; Jöreskog & Sörbom, 1993). According to Byrne (1991), testing for invariance involves specifying a model in which certain parameters are free to take on any value across groups (the variant model), and then comparing that model with the more restrictive model in which these parameters are constrained to be equal across groups (the invariant model). If the difference in fit ($\Delta\chi^2$) is not significant, the hypothesis of equal parameter estimates across multiple samples is considered tenable. Table 4 shows the fit measures of the variant and invariant models that were tested using MSA.

In Step 1 a fit function was calculated of the basic Model 1 (paths 1, 2a, and 2b) hypothesizing that the parameter estimates were *variant* (i.e., not equal) across the three samples. In Step 2 a fit function was calculated of the same basic Model 1, hypothesizing that the parameter estimates of paths 1, 2a, and 2b were *invariant*

Table 4
Fit measures of the variant and the invariant structural model in a multi-sample analysis with the three homogeneous samples 1, 2, and 3

Model	χ^2	*df*	*p*	Model comparison	$\Delta\chi^2$	*df*	*p*
Basic Model 1 (paths 1, 2a, and 2b) variant parameter estimates	100.9	9	.00				
Basic Model 1 (paths 1, 2a, and 2b) *in*variant parameter estimates	107.1	15	.00	M1 variant ↔ M1 invariant	6.19	6	*ns*
Extended Model 2a (paths 1, 2a, 2b, and 3a) variant parameter estimates	10.01	6	.12	M2a variant ↔ M1 variant	90.0	3	<.001
Extended Model 2a (paths 1, 2a, 2b, and 3a) *in*variant parameter estimates	22.06	13	.05	M2a variant ↔ M2a invariant	12.05	7	*ns*

Notes. χ^2, chi-square; *df*, degrees of freedom; $\Delta\chi^2$, chi-square difference.

(i.e., equal) across the three samples. In Step 3 a fit function was calculated of the extended Model 2a (including an additional path 3a), hypothesizing that the parameter estimates of paths 1, 2a, and 2b were *variant* (i.e., not equal) across the three samples. In Step 4, a fit function was calculated of the more restrictive Model 2a, hypothesizing that the parameter estimates of paths 1, 2a, and 2b were *invariant* (i.e., equal) across the three samples, and that path 3a was equal in Samples 2 and 3.

The fit in Step 2 (Model 1, with invariant parameter estimates) ($\chi^2_{(15)} = 107.1$, $p = .00$, GFI = .96, RMSEA = .16, NNFI = .58, CFI = .86, and ECVI = .33) did not deteriorate significantly compared to the fit in Step 1 ($\Delta\chi^2_{(6)} = 6.19$, $p = ns$), indicating that the parameters of the common paths (1, 2a, and 2b) were equal across the three samples. However, when we compare fit of the unconstrained (i.e., variant) basic Model 1 (Step 1) and the unconstrained (i.e., variant) but more extended Model 2a (with the additional path 3a, Step 3), we notice that the latter provided a significantly better fit ($\Delta\chi^2_{(3)} = 90.0$, $p < .001$), indicating that the extended Model 2a (with the additional path 3a) ($\chi^2_{(6)} = 10.01$, $p = .12$, GFI = .99, RMSEA = .058, NNFI = .95, CFI = .99, and ECVI = .21) did better account for the data than the basic Model 1 ($\chi^2_{(9)} = 100.9$, $p = .00$, GFI = .96, RMSEA = .21, NNFI = .30, CFI = .86, and ECVI = .34).

Comparing the fit in Step 3 (Model 2a, with variant parameter estimates) with the fit in Step 4 (Model 2a, with invariant parameter estimates), we notice that the fit in Step 3 was not significantly better than the fit in Step 4 ($\Delta\chi^2_{(7)} = 12.05$, $p = ns$). This result confirmed the assumption that the parameters of the common paths (1, 2a, and 2b) were equal in the three samples, and that the parameter of path 3a was equal in Samples 2 and 3. These findings showed that not only the basic pattern of relationships was robust across the three samples, but that also the strength of relationships was stable. Fig. 2 shows the invariant standardized regression coefficients that resulted from the MSA.

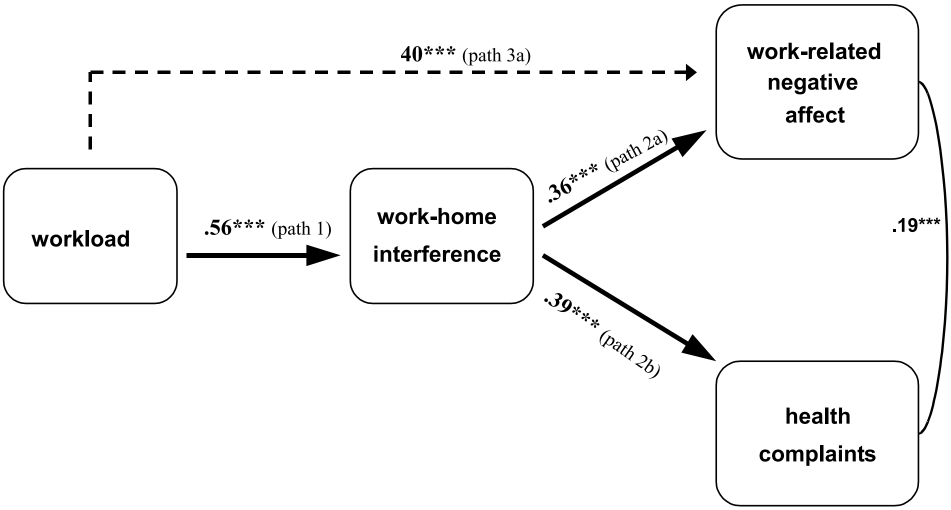

Fig. 2. The invariant standardized regression coefficients of the three homogeneous samples 1, 2, and 3 (the dotted arrow is sample-specific, explanation see text).

3. Part II

3.1. Data and method

3.1.1. Sample and procedure

Part II of this study was part of a large 3-year prospective cohort study, called the Study on Musculoskeletal disorders, Absenteeism, Stress and Health (SMASH). For a full description of survey design, we refer to Hoogendoorn et al. (2000). The measures used were part of a comprehensive self-administered questionnaire. The initial population consisted of 2064 workers from 34 Dutch companies from various sectors (i.e., industry and services). The response to the questionnaire in the baseline study was 87% (i.e., 1786 respondents). Workers with less than one year experience in their current job, who had temporary contracts, and those who received a benefit because of (partial) disability for work were excluded. Consequently, the number of respondents in the baseline study consisted of 1738 workers. In total 317 workers (18%) were lost to the second follow-up. Consequently, the data used in Part II of the current study were provided by 1421 workers (i.e., 69% of the initial population of 2064 workers). Of this sample 70% was male and 30% was female. Information regarding the marital and parental status of the respondents was unavailable. Seventy-nine percent of the respondents was full-time employed. Their age ranged from 18 to 59 years, with an average age of 36 years. Of all respondents 11% did not complete or only completed primary education, 40% completed lower vocational education, and 29% completed lower secondary education or middle vocational education. Twenty percent of the workers in the sample reported higher vocational and academic education.

3.1.2. Survey measures

Basic descriptive statistics (means, *SD*s, and Cronbach's α) as well as the zero-order correlations among the variables included in the current study were presented in Tables 1 and 2, respectively.

Workload was assessed using five items that were adopted from the well-known Job Content Questionnaire (Karasek, 1985), and that were highly similar to the items used in Part I. To improve its theoretical and psychometrical quality, the item about conflicting demands in the JCQ was replaced by the statement "My job is very hectic" on the basis of a PAF analysis (De Jonge, Reuvers, Houtman, & Kompier, 2000). Respondents responded on a four-point scale, ranging from 'strongly disagree' (1) to 'strongly agree' (4), with higher scores indicating a higher perceived workload. The internal consistency was sufficient (see Table 1).

Work–Home Interference (WHI) was measured with the same five items that were used in Part I, with higher scores indicating a higher level of WHI. The internal consistency is satisfactory (see Table 1). Compared to the three homogeneous samples that were used in Part I of this study, only Sample 1 consisting of medical residents experienced a significantly higher level of WHI than the respondents in this heterogeneous sample ($t = 27.01$, $p < .001$). In general, about 40% of the respondents in the heterogeneous sample reported to experience WHI at least occasionally.

Depressive mood, the affective well-being measure used in this second part of the study, was somewhat different from the measure used in Part I. A Dutch translation of a short version (Iowa form) of the Center for Epidemiologic Studies Depression (CES-D) scale (Kohout, Berkman, Evans, & Cornoni-Huntley, 1993; Radloff, 1977) was used to measure the current level of depressive symptoms, with emphasis on the affective component, that is, depressive mood. Each participant was offered 11 brief statements describing feelings or behaviors and was asked to indicate how often he or she felt that way during the past two weeks. Some illustrative statements are: "I felt everything I did was an effort"; "I felt depressed"; "I enjoyed life" (reversed); "I felt lonely"; and "I was happy" (reversed). Items were answered on a three-point rating scale ranging from hardly ever or never (1) to much or most of the time (3), with higher scores indicating a higher level of depressive symptoms. The internal consistency was satisfactory.

Health complaints were measured with the same 13 items that were used in Part I, with the individual scores indicating the number of complaints reported by the respondent. The average number of health complaints in the heterogeneous sample under study ($M = 2.62$) was similar to the level of health complaints in a large-sized reference group of 7717 workers from various branches ($M = 2.78$, Houtman et al., 1995), and significantly higher than the level of complaints among medical residents (Sample 1; $t = -6.46$, $p < .001$).

3.1.3. Data analysis

Similar to Part I, Confirmatory Factor Analyses (CFA) was conducted to ensure that workload and WHI were two different constructs, and that the same was true for the two indicators of well-being (i.e., depressive mood and health complaints). Next, the model that best fitted the three homogeneous samples in Part I was cross-validated in the heterogeneous sample under study by means of Structural Equation Modeling (SEM). For more details about SEM and MSA, we refer to the data analysis section in Part I.

3.2. Results

3.2.1. Preliminary Confirmatory Factor Analyses

Confirmatory Factor Analysis was used to examine the discriminant validity of the five workload items and the five WHI items. As in Part I, the two-factor model ($\chi^2_{(34)} = 525.3$, $p = .00$, GFI = .92, RMSEA = .11, NNFI = .80, and CFI = .85) provided a significantly better fit than the alternative one-factor model ($\Delta\chi^2_{(1)} = 804.01$, $p < .001$). All workload items had a factor loading of .35 or higher (with a median factor loading of .47), and all WHI items had a factor loading of .37 or higher (with a median factor loading of .56). Similar to Part I, these results indicate that workload and WHI, though interrelated ($r = .33$), could empirically distinguished.

Confirmatory Factor Analysis was also used to examine the factorial validity of the nine depressive mood items and the 13 items measuring health complaints. Again, a correlated two-factor model ($\chi^2_{(251)} = 4777.3$, $p = .00$, GFI = .82, RMSEA = .09, NNFI = .60, and CFI = .64) provided a significantly better fit than the alternative one-factor model ($\Delta\chi^2_{(1)} = 955.3$, $p < .001$). All items measuring depressive mood had factor loadings varying from .33 to .77 (with a median factor loading of .55). All items measuring health complaints had factor loadings varying from .36 to .70 (with a median factor loading of .49). As in Part I, the fit of the two-factor model fell short of acceptable values, but could be substantially improved by permitting error terms between a limited number of item pairs belonging to the same theoretical construct. Our only purpose was, however, to show that a two-factor model better accounted for the data than the one-factor model. The results support our assumption that depressive mood and health complaints were, although substantially interrelated ($r = .53$), empirically distinct indicators of well-being.

3.2.2. Cross-validation of the basic and an extended model

The model that best fitted the data of the three homogeneous samples in Part I was the extended Model 2a (see Table 3). Model 2a assumed that WHI played a partial mediating role in the relationship between workload and affective well-being and a full mediating role between workload and health complaints. The goodness-of-fit measures indicate that the extended model (with depressive mood as indicator of affective well-being, rather than work-related negative affect) also fitted the data of the heterogeneous sample quite good ($\chi^2_{(2)} = 7.44$, $p = .024$, AGFI = .98, RMSEA = .044, NNFI = .96, and CFI = .99). However, path 3a from workload to depressive mood was not significant ($\beta = .01$, $t = .24$, and $p = ns$) and could therefore be constrained to be zero, which did not result in a significant deterioration of the model ($\Delta\chi^2 = .05$, $p = ns$). The best (and also good) fitting model in the heterogeneous sample was the basic Model 1, assuming a full mediating role of WHI with regard to each indicator of well-being ($\chi^2_{(3)} = 7.49$, $p = .058$, AGFI = .99, RMSEA = .033, NNFI = .98, and

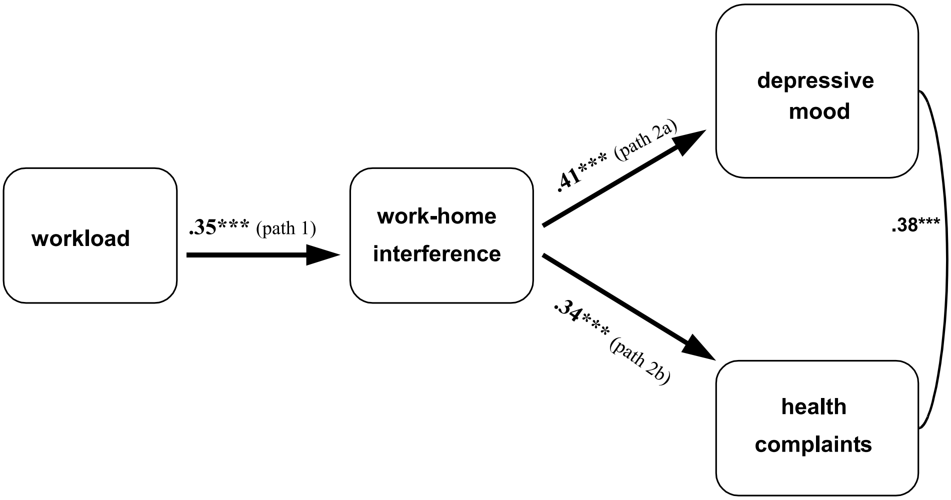

Fig. 3. The standardized regression coefficients in the large-sized heterogeneous Sample 4.

CFI = 1.00). The standardized regression coefficients in the heterogeneous Sample 4 are presented in Fig. 3.

4. Discussion

The purpose of the current study was to examine to what extent WHI mediated the impact of workload on affective well-being and subjective health. We also sought to overcome the methodological problems of previous studies, by examining this relationship in three homogeneous Dutch samples, and cross-validating the results in an independent large-sized heterogeneous sample of Dutch workers. In addition, we addressed the theoretical nature of much of the previous literature on Work–Home Interference by drawing on the insights of the Effort–Recovery (E–R) Model. In line with this work-psychological model, we argued that load-effects that have built up during working hours may impede one's functioning and recovery during the non-working hours (WHI). Due to a daily recurrence of insufficient quantitative and qualitative recovery, an unfavorable accumulative process may be triggered, resulting in impaired well-being that in the long run may take more serious and chronic forms (Sluiter, 1999; Ursin, 1980).

In Part I of this study, we systematically investigated in three homogeneous samples whether WHI played a full, a partial or no mediating role in the relationship between workload and two indicators of well-being, that is work-related negative affect and health complaints. The results of a Confirmatory Factor Analysis (CFA) revealed that in all three samples these two indicators of well-being could be regarded as both theoretically and empirically distinct (though interrelated) constructs. In general, the results of Part I showed that in all three homogeneous samples WHI played a *full* mediating role in the relationship between workload and health complaints, with the direction and the strength of the relationships being stable across the three different occupational groups. According to the results of a CFA, the strong association of workload with WHI was not based on the existence of one common underlying construct. The findings of Part I suggest that the direct relationship that is often found between workload and health complaints may be better understood as an indirect relationship in which WHI is a crucial intervening pathway. Workload, as such, may not cause health complaints, but rather workload exerts its negative effects on health through a process of 'spillover' of load-effects accompanied by insufficient quantitative and qualitative recovery during the non-working hours.

With regard to work-related negative affect (i.e., as indicator of affective well-being used in Part I), the results showed that in two of the three samples (i.e., among child care workers and bus drivers) WHI played not a full but a *partial* mediating role. This indicates that workload was both indirectly (through WHI) and directly associated with work-related negative affect. Among medical residents, WHI did play a full mediating role. This difference among the three occupational groups may be a result of the unique work circumstances of medical residents. WHI was most prevalent among medical residents (compared to the other two occupational groups), which is consistent with previous research in which it was argued that medical residents are particularly vulnerable to WHI due to long working hours (61% of the residents in the sample used in our analysis had formal contracts of at least 46 h), working under time pressure, and taking work home are characteristic of residents' 'day-to-day job' (e.g., Agius, Blenkin, Deary, Zeally, & Wood, 1996; Swanson, Power, & Simpson, 1998). Because of its higher prevalence, WHI might have played a more crucial intervening pathway in the relationship between workload and work-related negative affect among medical residents than among both other samples.

Not only the basic hypothesized pattern of relationships between workload, WHI, and the two indicators of well-being (i.e., work-related negative affect and health complaints) was invariant across the three occupational groups, but the same was true for the parameter estimates of the relationships involved. High similarity in both the type and the strength of relationships in three different occupational groups can be considered a powerful validation of the basic hypothesized pattern of relationships (Byrne, 1991).

In general, the findings of Part I showed that whereas workload and health complaints were related only indirectly through WHI, a direct relationship existed between workload and work-related negative affect. This finding is in line with our assumption that WHI would play a full intervening role in the relationship of workload with general (i.e., context-free) outcome measures, and a partial mediating role in the relationship of workload with work-related outcome measures (Geurts et al., 1999). Another explanation for the direct link of workload with work-related negative affect (and the absence of this link with health complaints) is that negative affect, such as feeling angry, frustrated or irritated (either or not work-related) might be an acute and direct response to workload (or daily hassles in general) that appears and disappears more easily. Health complaints (e.g., fatigue, backache, and chest pain), on the other hand, might be more chronic by nature and the result of a long-term process in which WHI, and thus recovery during non-working hours, might play a more crucial role.

In Part II, we cross-validated the most robust model, that is, the basic model extended with a direct path from workload to affective well-being in an independent large-sized heterogeneous sample. The results showed that WHI played a *full* mediating role in the relationship between workload and both indicators of well-being (i.e., depressive mood and health complaints). Also in the heterogeneous sample, CFA showed that the two indicators of well-being were indeed distinct constructs and that also the association of workload with WHI was not based on one common underlying construct. The results of Part II are in line with the results of Part I, suggesting that workload, as such, may not cause depressive symptoms and health complaints, but only through a process of 'spillover' of load-effects that impede recovery during the non-working hours.

4.1. General conclusions and implications

Our results provided support for the E–R Model in that WHI played a significant role in mediating the impact of workload on workers' well-being. In the context of the E–R Model, our findings suggest that WHI (both time-based and strain-based) may compromise opportunities to recover (both quantitatively and qualitatively) from work demands, which in turn increase the chances that work demands erode affective well-being and subjective health. This mechanism for explaining the relationship between workload and well-being seemed to be differentially relevant for different indicators of well-being, with WHI playing a full mediating role with respect to depressive mood and health complaints, and a partial mediating role with respect to work-related negative affect. This differential role of WHI can be explained in at least two ways.

First, and in line with our expectations, WHI seemed to have played the most crucial intervening role in the relationship of workload with context-free indicators of well-being (i.e., depressive mood and health complaints). With respect to indicators of well-being that were not context-free (i.e., work-related negative affect), workload might have exerted its negative impact directly as well as indirectly (through WHI). Second, the different indicators of well-being might have developed in different stages of the stress process. Hereby, negative affect (either work-related or not) was likely to be an acute and direct response to workload that partly developed independent of WHI. Depressive mood and health complaints, on the other hand, might have reflected more chronic indicators of well-being that were likely the result of a long-term process in which WHI (and thus recovery during non-working hours) played a significant role. One might argue that depressive mood, as measured in the current study, also reflected a rather acute mood state because participants was asked how frequently they experienced depressive feelings during the past two weeks. However, taking a closer look at the specific items (e.g., "I felt happy"), we can assume that this mood state did not develop during the last two weeks, but existed longer than the period we asked for, and thus might be more chronic by nature.

4.2. Methodological limitations and future research

One serious limitation of our study, as well as of many other studies in this field, was the use of a cross-sectional design. This implies that one should exercise caution in drawing firm conclusions regarding causation, and thus regarding the mediating role of WHI. In fact, each model that is empirically tested on cross-sectional data is equivalent to the one where all arrows are reversed. There is simply no method or technique to determine which model would fit better as they both hypothesize identical relationships because no time element is involved. We did perform a post hoc analysis on the large-sized heterogeneous sample in which an alternative (and theoretically plausible) model was tested. In this model, we hypothesized that workload 'leads to' impaired well-being (i.e., depressive mood and health complaints) that in turn 'leads to' WHI (thus, WHI as outcome rather than as mediating variable). This model, however, provided a very bad fit ($\chi^2_{(3)} = 562.84$, $p = .00$, AGFI = .26, RMSEA = .34, and CFI = .45), indicating that well-being as mediating variable in the relationship of workload with WHI had no empirical support. In spite of the limitations of cross-sectional designs, the consistency of our findings across the four samples underlined their stability and robustness. However, a suggestion for future research is to study the mediating role of WHI in a longitudinal design in which workload, WHI and indicators of well-being are measured at various (preferably theoretically chosen) points in time, and to let mediating models concur with models in which WHI has different positions in the stressor–strain relationship (i.e., as stressor or as outcome measure).

A second limitation was the use of self-report measures. A study on this topic benefits from the use of more 'objective' measures of well-being, for example, physiological (hormonal) indicators concerning effort and recovery (e.g., (nor)adrenaline, see Sluiter, 1999), or from the use of other sources of information, for instance one's spouse or adolescent children (Jones & Fletcher, 1996a,b; Morrison & Clements, 1997). Another suggestion would be to obtain a more objective and detailed analysis of one's activities during non-work hours, and what consequences this might have in terms of energy consumption, recovery, and accumulation of load effects (see, for instance, the diary study of Sonnentag, 2001).

A third limitation involved the use of different measures in Part I and Part II concerning workload and the indicator of affective well-being. However, this might also be regarded as a strong point. Despite the use of different measures, the basic pattern of hypothesized relationships was successfully cross-validated. In future studies, we suggest to include both general (i.e., context-free) and domain-related (i.e., work-related or non-work-related) indicators of well-being, as well as acute and chronic indicators of well-being, in order to find out whether WHI indeed plays the most crucial mediating role with respect to context-free outcome measures that are more chronic by nature.

An additional suggestion for future research is to study not only how work might interfere with non-work but also how it might facilitate one's functioning at home, as well as the other way around. Employees may not just suffer from excessive work demands, but may as well benefit from a work situation that contains supports or resources that are relevant to meeting demands, for instance being part of a happy cohesive workforce (Jones & Fletcher, 1996a,b). From the perspective of the E–R Model, this situation might result in the spillover of positive load-effects and the mobilization of energy, through which one's functioning in the private situation is facilitated and one's well-being is positively influenced (see Geurts & Demerouti, 2003). Similar processes may also take place in the other direction, that is, interference and/or facilitation from the home situation. Recent studies (Demerouti & Geurts, in press; Grzywacz & Marks, 2000; Wagena & Geurts, 2000) suggest a four-dimensional structure of the interaction between work and home, that is, interference from work to home, facilitation from work to home, as well as interference and facilitation in the other direction (from home to work) (see also Frone, in press). Important is to find out whether or not various types of work–home interaction play similar mediating roles in the relationship between, on the one hand, potential stressors and supports in the work (and home) situation, and on the other, negative and positive indicators of well-being.

Despite these limitations, we feel our study contributed to previous research on WHI both theoretically and methodologically. By studying WHI from an original and relevant theoretical perspective and in multiple samples simultaneously, we provided support for a *mediating* role of WHI in the relationship of workload with various indicators of well-being. The question 'Does Work–Home Interference mediate the relationship between workload and well-being?' can thus be answered positively. Our study contributed to more general research in the field of OHP by shedding light on the underlying mechanism through which workload (or possible other demanding work characteristics) might exert its negative impact on psychological well-being and subjective health. In an attempt to open the 'black box' in the relationship between workload and well-being, our findings suggest a process of spillover of negative workload-effects that impede recovery during non-working hours through which particularly general well-being is negatively affected.

Acknowledgments

This study has been carried out within the framework of two programs that are financially supported by The Netherlands Organization for Scientific Research (NWO): The Priority Program 'Fatigue at Work' (580-02-104) and the program 'ASPASIA' (015.000.027). The authors thank Chantal IJland for providing the data among child care workers (Sample 2) that were used in Part I of this study. Data that were used in Part II of this study (Sample 4) were collected within the framework of the Study on Musculoskeletal Disorders, Absenteeism, Stress, and Health (SMASH), which was financially supported by the Dutch Ministry of Social Affairs and Employment, the Dutch Ministry of Public Health, Welfare and Sports, and the Dutch National Institute for Social Security. The authors thank Paulien Bongers from TNO Work and Employment for providing the SMASH-data. The authors thank Dr. Toon Taris and Dr. Jan de Jonge for their advice concerning the statistical analyses. Finally, the authors thank both anonymous reviewers for their fruitful suggestions to further improve the paper.

References

Agius, R. M., Blenkin, H., Deary, I. J., Zeally, H., & Wood, R. A. (1996). Survey of perceived stress and work demands of consultant doctors. *Occupational and Environmental Medicine, 53*, 217–224.

Allen, T. D., Herst, D. E., Bruck, C. S., & Sutton, M. (2000). Consequences associated with work-to-family conflict: A review and agenda for future research. *Journal of Occupational Health Psychology, 5*(2), 278–308.

Bagozzi, R. P., & Heatherton, T. F. (1994). A general approach to representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling, 1*, 35–67.

Barnett, R. C. (1998). Toward a review and reconceptualization of the work/family literature. *Genetic, Social, and General Psychology Monographs, 124*, 125–182.

Bentler, P. M. (1985). *Theory and implementation of EQS: A structural equations program*. Los Angeles, CA: BMDP Statistical software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 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., & Wolff, Ch. J., de (1986). *Vos-D, vragenlijst organisatiestress-D. Handleiding* [Vos-D, Questionnaire Organizational Stress-D. Guideline]. Nijmegen: Katholieke Universiteit Nijmegen.

Bond, T., Galinsky, E., & Swanberg, J. E. (1998). *The 1997 National study of the changing workforce*. New York: Families and Work Institute.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. Scott Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage Publications.

Burke, R. J. (1988). Some antecedents and consequences of work–family conflict. *Journal of Social Behavior and Personality, 3*, 287–302.

Burke, R. J. (1993). Work–family stress, conflict, coping and burnout in police officers. *Stress Medicine, 9*, 171–180.

Burke, R. J., & Greenglass, E. R. (1999). Work–family conflict, spouse support, and nursing staff well-being during organizational restructuring. *Journal of Occupational Health Psychology (Special issue: Relationship between work and family life), 4*, 327–336.

Byrne, B. M. (1991). The Maslach Burnout Inventory: Validating factorial structure and invariance across intermediate, secondary, and university educators'. *Multivariate Behavioral Research, 26*, 583–605.

Demerouti, E., & Geurts (in press). Towards a typology of work–home interaction. *Community, work & family*. Manuscript submitted for publication.

De Jonge, J., Reuvers, M. E. N., Houtman, I. L. D., & Kompier, M. A. J. (2000). Linear and nonlinear relations between psychosocial job characteristics, subjective outcomes, and sickness absence: Baseline results from SMASH. *Journal of Occupational Health Psychology, 5*(2), 256–268.

Dirken, J. M. (1969). *Arbeid en stress. [work and stress]*. Groningen: Wolters Noordhoff.

Frone, M. R. (in press). Work–family balance. In J. C. Quick, & L. E. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 00–00). Washington, DC: American Psychological Association.

Frone, M. R., Barnes, G. M., & Farrell, M. P. (1994). Relationship of work–family conflict to substance use among mothers: The role of negative affect. *Journal of Marriage and the Family, 56*, 1019–1030.

Frone, M. R., Russell, M., & Barnes, G. M. (1996). Work–family conflict, gender and health-related outcomes: A study of employed parents in two community samples. *Journal of Occupational Health Psychology, 1*, 57–69.

Frone, M. R., Russell, M., & Cooper, M. L. (1991). Relationship of work and family stressors to psychological distress: The independent moderating influence of social support, mastery, active coping and self-focused attention. *Journal of Social Behavior and Personality, 6*, 227–250.

Frone, M. R., Russell, M., & Cooper, M. L. (1992). Antecedents and outcomes of work–family conflict: Testing a model of the work–family interface. *Journal of Applied Psychology, 77*, 65–78.

Frone, M. R., Russell, M., & Cooper, M. L. (1997a). Relation of work–family conflict to health outcomes: A four-year longitudinal study of employed parents. *Journal of Occupational and Organizational Psychology, 70*, 325–335.

Frone, M. T., Yardley, J. K., & Markel, K. S. (1997b). Developing and testing an integrative model of the work–family interface. *Journal of Vocational Behavior, 50*, 145–167.

Galinsky, E., Bond, J. T., & Friedman, D. E. (1993). *The changing workforce: Highlights of the national study*. New York: Families and Work Institute.

Geurts, S. A. E., & Demerouti, E. (2003). Work-nonwork interface: A review of theories and findings. In M. Schabracq, J. Winnubst, & C. Cooper (Eds.), *Handbook of work and health psychology* (pp. 279–312). Chichester: Wiley.

Geurts, S., Rutte, C., & Peeters, M. (1999). Antecedents and consequences of work–home interference among medical residents. *Social Science & Medicine, 48*, 1135–1148.

Grandey, A., & Cropanzano, R. (1999). The conservation of resources model applied to work–family conflict and strain. *Journal of Vocational Behavior, 54*, 350–370.

Greenhaus, J. H., & Beutell, N. J. (1985). Sources of conflict between work and family roles. *Academy of Management Review, 10*, 76–88.

Grzywacz, J. G., & Marks, N. F. (2000). Reconceptualizing the work–family interface: An ecological perspective on the correlates of positive and negative spillover between work and family. *Journal of Occupational Health Psychology, 5*, 111–126.

Hoogendoorn, W. E., Bongers, P. M., De Vet, H. C. W., Douwes, M., Koes, B. W., Miedema, M. C., Ariens, G. A. M., & Bouter, L. M. (2000). Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: Results from a prospective cohort study. *Spine, 25*, 3087–3092.

Houtman, I., Goudswaard, A., Dhondt, S., Grinten, M., van der Hildebrandt, V., & Kompier, M. A. J. (1995). *Evaluatie van de monitorstudie naar stress en lichamelijke belasting*. Den Haag: SZW.

Jones, F., & Fletcher, B. (1996a). Taking work home: A study of daily fluctuations in work stressors, effects on moods and impacts on marital partners. *Journal of Occupational and Organizational Psychology, 69*, 89–106.

Jones, F., & Fletcher, B. (1996b). Job control and health. In M. Schabracq, J. Winnubst, & C. Cooper (Eds.), *Handbook of work and health psychology* (pp. 33–50). Chichester: Wiley.

Joosten, J., & Drop, M. J. (1987). De betrouwbaarheid en vergelijkbaarheid van drie versies van de VOEG. [Reliability and comparability of three versions of the VOEG]. *Gezondheid en Samenleving, 8*, 251–265.

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 Publications.

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

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). *Organizational stress: Studies in role conflict and ambiguity*. New York: Wiley.

Karasek, R. A. (1985). *Job content instrument: Questionnaire and user's guide*. Los Angeles: Department of Industrial and Systems Engineering, University of Southern California.

Kinnunen, U., & Mauno, S. (1998). Antecedents and outcomes of work–family conflict among employed women and men in Finland. *Human Relations, 51*, 157–177.

Kohout, F. J., Berkman, L. F., Evans, D. A., & Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D Depression Symptoms Index. *Journal of Aging and Health, 5*, 179–193.

Kompier, M. A. J. (1988). *Arbeid en gezondheid van stadsbuschauffeurs* [Work and health of city bus drivers]. Doctoral thesis, University of Groningen. Delft: Uitgeverij Eburon.

Kopelman, R. E., Greenhaus, J. H., & Connolly, T. F. (1983). A model of work, family, and interrole conflict: A construct validation study. *Organizational Behavior and Human Performance, 32*, 198–213.

Kossek, E. E., & Ozeki, C. (1998). Work–family conflict, policies, and the job-life satisfaction relationship: A review and directions for organizational behavior-human resources research. *Journal of Applied Psychology, 83*, 139–149.

Leiter, M. P., & Durup, M. J. (1996). Work, home and in-between: A longitudinal study of spillover. *Journal of Applied Behavioral Science, 32*(1), 29–47.

MacEwen, K. E., & Barling, J. (1994). Daily consequences of work interference with family and family interference with work. *Work and Stress, 8*, 244–254.

Meijman, T. F., & Mulder, G. (1998). Psychological aspects of workload. In P. J. Drenth, H. Thierry, & C. J. de Wolff (Eds.), *Handbook of work and organizational psychology* (2nd ed., pp. 5–33). Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

Morrison, D. L., & Clements, R. (1997). The effect of one partner's job characteristics on the other partner's distress: A serendipitous, but naturalistic experiment. *Journal of Occupational and Organizational Psychology, 70*, 307–324.

Paoli, P. (1997). *Second European survey on working conditions 1996*. Dublin: European Foundation for the Improvement of Living and Working Conditions.

Parasuraman, S., Greenhaus, J. H., & Granrose, C. S. (1992). Role stressors, social support, and well-being among two-career couples. *Journal of Organizational Behavior, 13*, 339–356.

Parasuraman, S., Purohit, Y. S., Godshalk, V. M., & Beutell, N. J. (1996). Work and family variables, entrepreneurial career success, and psychological well-being. *Journal of Vocational Behavior, 48*, 275–300.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385.

Reskin, B. F., & Padavic, I. (1994). *Women and men at work*. Thousand Oaks, CA: Pine Forge Press.

Rice, R., Frone, M. R., & McFarlin, D. B. (1992). Work–nonwork conflict and the perceived quality of life. *Journal of Organizational Behavior, 13*, 155–168.

SCP (Sociaal Cultureel Planbureau 2000). *De kunst van het combineren. Taakverdeling onder partners* [The skill of combining. Division of tasks among partners]. Den Haag: SCP.

SCP (Sociaal Cultureel Planbureau 2001). *Trends in de tijd*. [Trends in time]. Den Haag: SCP.

Sluiter, J. K. (1999). *How about work demands, recovery, and health*. Doctoral thesis, University of Amsterdam.

Sonnentag, S. (2001). Work, recovery activities, and individual well-being: A diary study. *Journal of Occupational Health Psychology, 6*(3), 196–210.

Stephens, M. A. P., Franks, M. M., & Atziena, A. A. (1997). Where two roles intersect: Spillover between parent care and employment. *Psychology and Aging, 12*, 30–37.

Swanson, V., Power, K. G., & Simpson, R. J. (1998). Occupational stress and family life: A comparison of male and female doctors. *Journal of Occupational and Organizational Psychology, 71*, 237–260.

Thomas, L. T., & Ganster, D. C. (1985). Impact of family supportive work variables on work–family conflict and strain: A control perspective. *Journal of Applied Psychology, 80*, 6–15.

Ursin, H. (1980). Personality, activation and somatic health: A new psychosomatic theory. In S. Levine, & H. Ursin (Eds.), *Coping and health* (pp. 259–279). New York: Plenum Press.

Van Veldhoven, M., Meijman, T. F., Broersen, J. P. J., & Fortuin, R. J. (1997). *Handleiding VBBA: Onderzoek naar de beleving van psychosociale arbeidsbelasting en werkstress met behulp van de vragenlijst beleving en beoordeling van de arbeid* [Guideline VBBA: Research on the experience of psychosocial work load and job stress with the Questionnaire on the Experience and Evaluation of Work]. Amsterdam: SKB (Stichting Kwaliteitsbevordering BGZ Nederland).

Van Veldhoven, M., Broersen, J. P. J., & Fortuin, R. J. (1999). *Werkstress in beeld. Psychosociale arbeidsbelasting en werkstress in Nederland* [The picture of job stress. Psychosocial work load and job stress in the Netherlands]. Amsterdam: SKB (Stichting Kwaliteitsbevordering BGZ Nederland).

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

Wallace, J. E. (1997). It's about time: A study of hours worked and work spillover among law firm lawyers. *Journal of Vocational Behavior, 50*, 227–248.

Wagena, E., & Geurts, S. (2000). SWING* Ontwikkeling en validering van de 'Survey Wek-thuis Interferentie-Nijmegen' [SWING. Development and validation of the 'Survey Work–home Interference-Nijmegen']. *Gedrag & Gezondheid (themanummer Psychische vermoeidheid in de arbeidssituatie), 28*(3), 138–158.

Warr, P. B. (1987). *Work: Unemployment and mental Health*. Oxford: Oxford University Press.

Warr, P. B. (1990). The measurement of well-being and other aspects of mental health. *Journal of Occupational Psychology, 63*, 193–210.