Taylor & Francis
Taylor & Francis Group

Are there causal relationships between the dimensions of the Maslach Burnout Inventory? A review and two longitudinal tests

TOON W. TARIS[1], PASCALE M. LE BLANC[2], WILMAR B. SCHAUFELI[2], & PAUL J. G. SCHREURS[2,3]

[1]*Department of Work and Organizational Psychology, Radboud University Nijmegen, The Netherlands;* [2]*Utrecht University, Utrecht, The Netherlands; and* [3]*Institute of Work & Stress, Amstelveen, The Netherlands*

Abstract
Arising from interest concerning the possibility of causal relationships among the three components of the Maslach Burnout Inventory, several process models have been proposed for the development of burnout. The present paper first reviews the evidence in favour of the three most influential of these (Leiter and Maslach's model (1988); Golembiewski, Boudreau, Munzenrider, & Luo's (1996) phase model; and Lee and Ashforth's model (1993)). These three models, and our own model (which integrates of two of them, and includes feedback effects of depersonalization on emotional exhaustion) are then compared with each other using structural equation modelling, drawing on longitudinal data from two Dutch samples (total $N = 1185$). The review revealed that none of the seven previous studies on this issue provided any convincing support for any particular causal order proposed so far. In contrast, our own study showed that high levels of exhaustion were associated with high levels of depersonalization over time across both samples. Further, higher levels of depersonalization led to higher levels of emotional exhaustion and lower levels of personal accomplishment. To our knowledge, the present research is the first to provide reliable longitudinal evidence for the conceptualization of burnout as a developmental process, although the effects are not large enough to be of practical use in the recognition of burnout.

Keywords: *Burnout, Maslach Burnout Inventory, process model, longitudinal research, structural equation modelling, causality, teachers, care providers*

Introduction

One issue that has continued to interest burnout researchers over the past two decades concerns the relationships among the components of the widely-used Maslach Burnout Inventory (MBI, Maslach, Jackson, & Leiter, 1996). Based on pioneering work by Freudenberger (1975) and Maslach (1976), the MBI construes burnout as a psychological syndrome that occurs in response to chronic (especially interpersonal) stressors on the job (Maslach, Schaufeli, & Leiter, 2001, for an historical overview). The three key dimensions of this response are (1) emotional exhaustion (EE, the feeling of being overextended and

depleted of resources, representing the basic individual stress dimension of burnout), (2) depersonalization (DP, a cynical and distant attitude towards one's work and the people one works with, representing the interpersonal context in which burnout occurs), and (3) reduced personal accomplishment (PA, the tendency to evaluate one's achievements at work negatively, thus referring to the self-evaluation dimension of burnout). EE is usually considered to be the central quality of burnout and the most obvious manifestation of the syndrome. This does not mean that the other two dimensions are incidental or unnecessary (cf. Shirom, 1989). Theoretically, DP and PA capture critical aspects of the relationship people have with their work, thus covering different aspects of the concept of work-related well-being (cf. Van Horn, Taris, Schaufeli, & Schreurs, 2004). This is *not* to say that the MBI was grounded in firm clinical observation or based on sound theorizing. Rather, it has been inductively developed by factor-analysing a rather arbitrary set of items (Schaufeli, 2003).

Instead of considering the three MBI dimensions as more or less co-occurring phenomena, several researchers have interpreted the associations among these dimensions as resulting from an underlying causal process, in spite of the fact that the MBI was not *a priori* designed to capture such a process (see among others, Söderfeldt (1997) for a discussion). However, their views regarding the causal nature of this process have diverged strongly. The present research has two aims. The first is to review the empirical support for the three most prominent models for the causal development of burnout, i.e. Golembiewski, Munzenrider, and Stevenson's (1986) phase model, Leiter and Maslach's (1988) model, and Lee and Ashforth's (1993) model. The second aim is to test and compare these models as well as our own integration of two of these, using two Dutch longitudinal data sets.

The issue concerning the 'correct' causal order of the three dimensions of the MBI is not entirely of academic interest. It has been suggested that a good understanding of the aetiology and development of burnout could facilitate the early recognition of burnout (Lee & Ashforth, 1993; Van Dierendonck, Schaufeli, & Buunk, 2001). For instance, a decisive issue centres around the question of whether a high level of emotional exhaustion is a first symptom of a developing burnout (as proposed by Leiter & Maslach, 1988) or whether it is the end-point of it (Golembiewski et al., 1986). In the first case, interventions should be geared towards preventing relatively minor problems from getting worse, whereas in the latter case serious problems calling for curative interventions are present.

Since Golembiewski et al. (1986) and Leiter and Maslach (1988) proposed their respective models, a small body of research has addressed the causal relationships among the three components of the MBI. Below we first review the evidence for various causal orderings of the three burnout components. Then the most promising and theoretically plausible candidates for the relationships among the burnout dimensions are tested and compared using structural equation modelling, drawing on data from two longitudinal samples of oncology care providers and teachers.

From emotional exhaustion to personal accomplishment— and vice versa

Introducing the contenders. To date, three more or less influential models have been proposed for the causal relationships among the three dimensions of the Maslach Burnout Inventory. Perhaps the best known of these is the Leiter and Maslach (1988) model, that draws on the idea that burnout occurs in response to chronic job stress. That is, chronic high (quantitative and/or qualitative) job demands are presumed to trigger emotional exhaustion

as an individual stress response. In turn, high levels of EE would lead workers in both contactual professions (i.e. professions in which contact with other people constitutes a major part of the tasks) and non-contactual professions to withdraw themselves psychologically from the people they work with (for contactual professions) and their work (for non-contactual professions) in an attempt to cope with these stressors. For example, in the context of contactual professions, moderating one's compassion for clients by distancing oneself emotionally from them (i.e. developing a sense of 'detached concern', Lief & Fox, 1963) may be one way of protecting oneself from intense emotional arousal that could interfere with functioning effectively at the job (Maslach et al., 2001). However, excessive detachment combined with little concern may lead staff to respond to clients in negative, callous, dehumanized, and *depersonalized* ways. Finally, as feelings of depersonalization and cynicism towards the job and one's clients persist, the workers' feelings of self-efficacy and achievement diminish, since depersonalization precludes the achievement of work goals. Thus, Leiter and Maslach (1988) argued that high levels of EE would lead to high levels of DP, in turn leading to low levels of PA.

An alternative view regarding the development of burnout was expressed by Golembiewski and his co-workers in their *phase model* (e.g. Golembiewski, Munzenrider, & Stevenson, 1986). This model proposed that each of the three MBI dimensions be split into low and high levels. By crossing these three dimensions, eight patterns (or phases) of burnout emerge. According to Golembiewski et al. (1986), workers may pass through several 'flight paths' (i.e. sequences of phases) while becoming burnt out. Chief among these is the path that starts with high levels of depersonalization in attempts to cope with the stresses of the job. Similar to Leiter and Maslach (1988), Golembiewski et al. (1986) argued that a certain degree of professional detachment is often functional and perhaps even necessary for performance in some occupations. However, beyond some point detachment becomes depersonalization, impairing the ability to form necessary relationships with others and undermining performance, thus affecting one's perception of accomplishment. Note that Golembiewski et al. (1986) distinguished between chronic and acute burnout; the latter form of burnout may be triggered by personal trauma (e.g. the death of a loved one), while chronic burnout is likely to derive from factors in the work setting. The flight path described above refers to the development of chronic (i.e. work-related) burnout; thus, depersonalization occurs first, in time leading to a reduced sense of personal accomplishment (as was also hypothesized by Leiter and Maslach, 1988). Emotional exhaustion would then occur in response to increasing depersonalization and decreasing personal accomplishment. Thus, whereas in Leiter and Maslach's conceptualization elevated levels of emotional exhaustion signify the onset of burnout, Golembiewski et al. (1986) consider emotional exhaustion to be the final stage of burnout.

Finally, Lee and Ashforth (1993) set out to compare the models of Leiter and Maslach (1988) and Golembiewski et al. (1986). Their exploratory analysis resulted in a variation on the Leiter and Maslach model in which emotional exhaustion was positively related to depersonalization (as in the Leiter and Maslach, and the Golembiewski et al. models), but where personal accomplishment developed independently from depersonalization; rather, they proposed that elevated levels of emotional exhaustion directly evoked decreases in personal accomplishment rather than indirectly through depersonalization. As Maslach et al. (2001) argue, it is difficult to gain a sense of accomplishment when feeling exhausted.

Previous research: What do we know? The above review shows that there are several models that describe the development of burnout in terms of a temporal sequence. The question,

then, is which model represents the underlying process best. Below we review the empirical evidence for each model. Studies that were possibly relevant to the issue of the causal relationships among the three burnout components were identified through a systematic literature search in the *PsychInfo* and *Medline* data bases (updated up to May 2005). Key words used were 'burnout', 'longitudinal', 'panel', 'development', 'causal', 'Maslach Burnout Inventory', 'MBI' and 'phase model', in various combinations. Further, the abstracts of studies in which the Leiter and Maslach (1988), the Golembiewski et al. (1986), or the Lee and Ashforth (1993) studies were cited were scrutinized for possible relevance. In order to be included, studies should (1) include at least two of the three dimensions of the MBI; (2) interpret the relationships among these dimensions causally; and (3) employ a longitudinal design. Although researchers have attempted to shed light on the causal order of the MBI components using cross-sectional designs (among others, Büssing & Glaser, 2000; Cordes, Dougherty, & Blum, 1997), only longitudinal designs can provide strong evidence on the temporal (and, thus, causal) order of variables (Taris, 2000, for an overview). Further, the reference lists to these studies were scrutinized for possibly relevant studies, as were the publications of several well-known burnout researchers. Table I summarizes the findings of the seven resulting studies.

Table I shows that studies 1–6 employed a complete panel design, i.e. a design in which all three MBI dimensions were measured at both waves of the study (no multi-wave studies were identified). Study 7 employed an incomplete panel design, i.e. a design in which not all three burnout dimensions were measured on all occasions. The time intervals between the waves of these studies varied from 3 months (Study 4) to 8 years (Study 7). Three studies (Studies 1, 2 and 7) were explicitly designed to *compare* alternative models for the causal order of the three burnout components. The other studies primarily explored the relationships among these components and other concepts, providing evidence on the causal relations among the three burnout components in passing.

To what degree does previous research provide longitudinal evidence for any of the three models described above? If anything, Table I reveals that such evidence is virtually absent. The longitudinal associations reported in these studies usually refer to 'stability effects' or 'autocorrelations', that is, associations between the measure of a particular variable at time 1 and the measure of that same variable at time 2. The studies described in Table I show that workers who report relatively high levels of exhaustion (depersonalization/personal accomplishment) report high levels of exhaustion (depersonalization/personal accomplishment) at a later point in time as well, seemingly irrespective of the length of the time interval between the study waves. Whereas it may be interesting to note that the temporal stability of the three MBI components is high, from our perspective these findings are largely irrelevant.

Four studies reported a longitudinal association between two different components of the MBI. Lee and Ashforth (1993) found that high levels of EE were associated with low levels of DP at a later point in time; this finding is inconsistent with any of the three models described above (note that the correlation underlying this association was positive, suggesting a suppressor effect; high levels of EE thus coincided with high levels of DP, which is consistent with the Leiter & Maslach, 1988, and Lee & Ashforth, 1993, models). Leiter (1990) reported that low levels of personal accomplishment were longitudinally associated with high levels of emotional exhaustion, thus partially supporting the Golembiewski et al. (1986) model. Leiter and Durup (1996) found that high levels of PA were longitudinally associated with low levels of DP; none of the three models compared here proposed such a relationship. The same applies to the finding of McManus, Winder,

Table I. Description of longitudinal studies reporting data relevant to the sequential development of burnout.

Study		Sample	Design	Study goal	Longitudinal associations among burnout components
1	Bakker et al. (2000)	207 General Practitioners	2-wave complete panel design, 5 years in between	Compare Leiter & Maslach (1988) and Lee & Ashforth (1993), examine feedback loop	Stability effects: EE .50, DP .35, PA .54 Lagged effects: none
2	Lee & Ashforth (1993)	169 Supervisors and managers of a public welfare setting	2-wave complete panel design, 8 months in between	Compare Leiter & Maslach (1988) and Golembiewski et al. (1986) models	Stability effects: EE .52, DP .49, PA .63 Lagged effects: $+EE \rightarrow -DP$[a]
3	Leiter (1990)	122 Mental health workers	2-wave complete panel design, 6 months in between	Examine development of burnout	Stability effects: PA: .49, DP .42, EE .49 Lagged effects: $+PA \rightarrow -EE$
4	Leiter & Durup (1996)	151 hospital-based health care professionals	2-wave complete panel design, 3 months in between	Examine consequences of work-nonwork interference	Stability effects: EE .70, DP .48, PA .62 Lagged effects: $+PA \rightarrow -DP$
5	McManus et al. (2002)	331 Medical Doctors	2-wave complete panel design, 3 years in between	Explore relationships between stress and burnout	Stability effects: not reported. Lagged effects: $+PA \rightarrow +EE$
6	Savicki & Cooley (1994)	64 Child protective service workers	2-wave complete panel design, 18 months in between	Explore effects work characteristics on burnout: replicate Leiter & Maslach (1988) model	Stability effects: not reported. Lagged effects: none
7	Toppinen-Tanner et al. (2002)	556 Blue and white collar workers	2-wave incomplete panel design (only EE measured at both waves), 8 years in between	Compare models of Leiter & Maslach (1988), Golembiewski et al. (1986), Van Dierendonck et al. (2001), Lee & Ashforth (1993)	Stability effects: EE .62. Lagged effects: none

EE = Emotional exhaustion, DP = Depersonalization, PA = Personal accomplishment.
'+' = 'high levels of', '−' = 'low levels of', '→' = 'lead to'. For example, '$+EE \rightarrow -DP$' denotes 'high levels of EE lead to low levels of DP'.
[a]Suppressor effect: the underlying correlation has a reverse sign.

and Gordon (2000) that high levels of PA were longitudinally associated with higher levels of EE; again, this finding does not fit any of the models compared here.

Evaluation of previous research. One conclusion that can be drawn on the basis of previous research is that longitudinal evidence in favour of any of the three models compared here is largely absent. Indeed, insofar as longitudinal associations between two different burnout dimensions were found, these usually did not fit *any* of the relationships proposed by *any* of the three models. Thus, conclusions such as 'In general, the research on burnout has established the sequential link from exhaustion to cynicism [depersonalization]' (Maslach et al., 2001, p. 405) seem to be as yet unwarranted. Further, several studies merely explored

the data or aimed to replicate specific models, without comparing these models to their competitors. Thus, these competing models lacked the chance to prove their value; we do not know whether these would have fitted the data as well (and perhaps even better) than the model being tested.

Based on these considerations, it appears that none of the three models can be preferred to its competitors. This is not solely due to methodological limitations (failure to consider, test and compare alternative models) or lack of empirical support; we also believe that there is a conceptual reason for not accepting these models. In most of the studies reviewed above, depersonalization is considered to be a coping strategy. For instance, Leiter and Maslach (1988) consider depersonalization as 'an attempt to cope with working with difficult and demanding clients and the resulting feelings of exhaustion' (Van Dierendonck et al., 2001, p. 42). Lee and Ashforth (1996, p. 123) propose that depersonalization is a form of 'defensive coping' that should be interpreted in terms of Lazarus and Folkman's (1984) stress-strain-coping-self-evaluation process. Golembiewski, Boudreau, Munzenrider, and Luo (1996) argue that depersonalization develops as a (dysfunctional) means of coping in stressful situations. As most researchers agree that depersonalization should be considered to be a coping strategy, it is remarkable that none of the three models discussed so far explicitly treats depersonalization as such. That is, if depersonalization is an *effective* coping strategy, one would expect longitudinal effects of depersonalization on the two other burnout dimensions, such that elevated levels of depersonalization lead to lower levels of emotional exhaustion and higher levels of personal accomplishment. If depersonalization is a *dys*functional coping strategy (as argued by Golembiewski et al., 1996), elevated levels of depersonalization might lead to *negative* outcomes in terms of increased emotional exhaustion and lower personal accomplishment. Consistent with the latter reasoning, Bakker, Schaufeli, Sixma, Bosveld, and van Dierendonck (2000) found that General Practitioners reporting high levels of depersonalization at the first wave of their study experienced higher job demands at the second wave. In turn, high demands were cross-sectionally associated with elevated levels of emotional exhaustion. Bakker et al. (2000) suggest that depersonalized GPs treat their patients indifferently and that such indifferent behaviour triggers demanding and threatening behaviours from these patients, thus increasing their work load. Similarly, it would seem possible that depersonalized GPs tend to perceive their demands as higher. In a four-wave study among 668 workers, De Lange, Taris, Kompier, Houtman, and Bongers (2004) found that emotionally exhausted participants tended to consider their job demands as higher at a later point in time, even if nothing had changed in their jobs.

All in all, it appears that current models for the causal effects among the burnout dimensions should be extended to include effects of depersonalization on the other two burnout dimensions. Most researchers seem to agree that depersonalization is a dysfunctional coping strategy; whereas those who distance themselves from their work and the people they work with attempt to lower the emotional impact of their work, employing this strategy could well result in (perceptions of) higher job demands and, consequently, in even higher levels of exhaustion.

The present research. The current research was designed to circumvent the limitations of previous research in examining the causal relationships among the burnout components. We present findings from two studies. The first of these is a three-wave study among 218 Dutch oncology care providers; the second study employs data from a two-wave study among 967 Dutch teachers. By systematically comparing the three main models proposed

so far, each of these will have a fair chance of being accepted as the best approximation of the causal process underlying the associations among the three burnout dimensions. Further, by examining a model that includes 'feedback effects' of depersonalization on other components in these models, we can test whether depersonalization serves as a coping strategy that buffers the adverse effects of high strain. Finally, by replicating our results in two different occupational contexts we increase the external validity of our findings and decrease the chances of capitalization on chance.

Method

Participants

The present research employed two data sets. The first of these concerned a three-wave sample of 218 Dutch oncology care providers; the second data set included data from a two-wave sample of 967 Dutch teachers.

Oncology care providers. The data were collected as part of a three-wave panel study among a national sample of 484 oncology care providers working in 18 different general and university hospitals. Oncology care providers may be faced with a multitude of psychosocial problems in their work. As cancer patients are confronted with a life-threatening disease, its treatment, and severe physical side effects, they may experience feelings of uncertainty and depressive symptoms. These and other patient reactions to cancer are difficult to handle with professional demeanour. Further, professionals in the field of oncology often must acknowledge the limits of modern medicine: they may feel incompetent, helpless helpers when confronted with an irresistibly advancing disease. Oncology care providers also face a variety of complicated treatment regimens and a rapidly growing, but not easy to assimilate, amount of knowledge in their subspeciality. Decision-making has become more compli-cated too, as treatment modalities have multiplied. In sum, oncology care providers must deal with a high level of patient-related, emotional demands in their jobs, next to stressors that are common to other health care workers, such as high quantitative job demands. All in all, it is not surprising that incumbents of this profession are a high-risk group for burnout (Maslach et al., 2001).

   In the first wave of the study, our participants completed a written questionnaire that addressed psychological and physical well-being, job demands and job control, and several biographical variables. The majority of the sample also co-operated in the second wave of the study (324 participants; 67% response rate), which was conducted after 1 year. The majority of those who participated in the second wave again participated in the third wave of the study (225 participants; 69% response rate), which was conducted after a further year. Multivariate analysis of variance revealed that the mean scores on the Time 1 study variables of those who dropped out of the study did not differ from those who remained in the study, $F(5, 478) = 1.18$, *ns*. Thus, dropout was not selective. After listwise deletion of missing values, the final sample included 218 oncology care providers (mean age $= 37.0$ years, $SD = 8.5$ years, 76% female, mean number of years of oncology experience was 10.3 years, $SD = 6.9$ years).

Teachers. The data were collected as part of a two-wave panel study among a nationally representative sample of 1309 Dutch primary and secondary school teachers. Although not being as intensely emotionally demanding as working at an oncology care unit, the interaction with pupils and their parents may be quite difficult. Interaction episodes with

single pupils are rather short and might not even exist at all. Thus, teachers have good chances to perform their working tasks even if they completely detach from several difficult pupils. However, they have almost no chance to avoid interaction with a difficult class. Further, in the Netherlands governmental regulations regarding the form and content of the curriculum change frequently, meaning that teachers find it difficult to routinize their tasks. This combines with the fact that classes in the Netherlands are on average quite large; especially in cities with a culturally diverse population, teachers may find it difficult to give all their pupils the attention they need. Finally, salaries, job security and career prospects are notoriously low in this sector. Next to the caring professions, teaching is a second major risk profession for burnout (Maslach et al., 2001).

In the first wave of the study, the participants filled out a questionnaire that addressed well-being, job characteristics, and biographical variables (Van Horn et al., 2004, for more information). The majority of the sample (998 participants) also co-operated in the second wave of the study that was conducted one year later, yielding a 76.2% response rate. MANOVA demonstrated that the mean scores on the Time 1 study variables of those who dropped out of the study did not differ significantly from the scores of those who remained in the study, $F(5, 1303) = .14$, *ns*. After listwise deletion of missing values, the final sample included 967 teachers (mean age $= 43.6$ years, $SD = 7.9$ years; 51% female; mean number of years of teaching experience was 19.5 years, $SD = 8.2$ years).

Measures

Oncology care providers. Burnout was measured using the 20-item Maslach Burnout Inventory—Human Services Survey (Schaufeli & Van Dierendonck, 2000). The reliability of the eight-item *Emotional exhaustion* scale (Cronbach's α) was .85 at Time 1, .87 at Time 2, and .90 at Time 3. *Depersonalization* was measured with five items; Time 1 α was .64, Time 2 α was .68, and Time 3 α was .62. *Personal accomplishment* was measured with seven items; α was .75 at Time 1, .77 at Time 2, and .79 at Time 3, respectively. For all three dimensions, scale scores could range from 0 (low exhaustion, depersonalization or personal accomplishment) to 6 (high exhaustion, depersonalization or personal accomplishment).

Teachers. Burnout was measured using the 22-item Maslach Burnout Inventory—Educators Survey (Schaufeli & Van Dierendonck, 2000). The reliability of the eight-item *Emotional exhaustion* scale (Cronbach's α) was .92 at Time 1 and .93 at Time 2. *Depersonalization* was measured with seven items; Time 1 α was .73 and Time 2 α was .75. *Personal accomplishment* was measured with seven items; α was .86 at Time 1 and .87 at Time 2, respectively. Again, scale scores could range from 0 (low exhaustion, depersonalization or personal accomplishment) to 6 (high scores on these dimensions).

Table II presents the means, standard deviations, and intercorrelations for the study variables.

Statistical analysis

The data were analysed using structural equation modelling as implemented in the LISREL 8.30 package (Jöreskog & Sörbom, 1999). Inspection of the univariate distributions of the variables revealed no substantive deviations from normality. Based on the theoretical considerations discussed above, four *a priori* models for the relationships among the three burnout dimensions were specified. For simplicity all constructs were included in the models at the scale level, i.e. we did not specify latent factor models for these constructs.

Table II. Means, standard deviations and correlations for the study variables (upper triangle: oncology care providers, $N=218$; lower triangle: teachers, $N=967$).

		Time 3			Time 2			Time 1		
		Emotional exhaustion	Depers.	Personal accompl.	Emotional exhaustion	Depers.	Personal accompl.	Emotional exhaustion	Depers.	Personal accompl.
Time 3										
Emotional exhaustion			.59	−.22	.73	.40	−.23	.65	.33	−.13
Depersonalization				−.32	.54	.64	−.28	.46	.55	−.22
Personal accomplishment					−.18	−.16	.70	−.18	−.17	.56
Time 2										
Emotional exhaustion						.59	−.26	.65	.35	−.07
Depersonalization					.46		−.21	.40	.54	−.15
Personal accomplishment					−.37	−.35		−.23	−.17	.64
Time 1										
Emotional exhaustion					.77	.38	−.30		.41	−.21
Depersonalization					.39	.61	−.35	.40		−.23
Personal accomplishment					−.33	−.27	.66	−.34	−.35	
Means and standard deviations										
Oncology care providers	M	1.61	.94	4.33	1.61	.98	4.28	1.48	.90	4.37
	SD	.97	.63	.83	.98	.70	.80	.84	.64	.81
Teachers	M				2.04	1.30	4.10	1.94	1.21	4.12
	SD				1.24	.82	.85	1.21	.80	.85

Correlations of .06 (teachers) and .13 (oncology care providers) are significant at $p<.05$.

This approach seemed warranted, as our measures were generally reasonably reliable (αs ranging from .62 to .93, median α was .77).

The first model, which we will refer to as *model A*, corresponds with Golembiewski et al.'s (1986) conceptualization of burnout as starting with elevated levels of depersonalization, leading to decreases in feelings of personal accomplishment (denoted as $+DP \rightarrow -PA$); in turn, low levels of personal accomplishment lead to high levels of emotional exhaustion (i.e. $-PA \rightarrow +EE$). Model A can therefore be shown as:

$$+DP \rightarrow -PA \rightarrow +EE$$

These relationships were specified both within waves (in the form of correlated error terms for the endogenous—dependent—variables and correlations for the exogenous—in-dependent—variables; thus, no direction was specified for these associations) and across-waves (i.e. longitudinally). Thus, cross-sectionally we specified associations between DP and PA, and PA and EE; for the first wave these associations were specified as correlations among the independent variables, whereas for the next wave(s) these associations were specified as correlated errors. Further, each of the three MBI subscales was assumed to be longitudinally related to itself (i.e. EE as measured at Time 1 was presumed to be associated with EE as measured at Time 2, and so on). These effects were included in all other models as well.

The second model (*model B*) is based on the Leiter and Maslach (1988) model of burnout that starts with emotional exhaustion. In this model, high levels of exhaustion lead to elevated levels of depersonalization, in turn leading to low levels of personal accomplishment. Model B can therefore be shown as:

$$+EE \rightarrow +DP \rightarrow -PA$$

Again, these relationships were specified both within and across time points, with only the longitudinal effects specifying a particular causal direction.

The third model (*model C*) is Lee and Ashforth's (1993) adaptation of the 1988 Leiter and Maslach model. Compared to the Leiter and Maslach model, the Lee and Ashforth model replaces the effect of depersonalization on personal accomplishment with a direct effect of emotional exhaustion on personal accomplishment. Thus model C can be shown as:

$$+EE \rightarrow +DP, \text{ and } +EE \rightarrow -PA$$

The fourth model (*model D*) is our own integration of the very similar models B (Leiter and Maslach, 1988) and C (Lee and Ashforth, 1993), i.e.

$$+EE \rightarrow +DP \rightarrow -PA, \text{ and } +EE \rightarrow -PA.$$

Further, this model was extended with feedback-effects of Time 1 Depersonalization on all Time 2 burnout dimensions.

In an attempt to increase statistical power, for the relatively small three-wave oncology care providers to sample all corresponding across-wave effects among the three burnout dimensions were constrained to be equal (cf. Jöreskog & Sörbom, 1999). For example, in testing the Leiter and Maslach and Lee and Ashforth models, the lagged effect of Time 1 emotional exhaustion on Time 2 depersonalization was constrained to equal the effects of Time 2 emotional exhaustion on Time 3 depersonalization; similar constraints were made for the other effects implied by the models to be compared. This procedure means that the estimates of the respective constrained parameters are based on more information than would otherwise be the case, thus reducing their standard errors. Inspection of the chi-square value and modification indices provided by the LISREL programme reveals whether

the assumption that these constrained parameters are indeed the same across time is warranted.

Finally, model fit was judged using the chi-square (χ^2) test, as well as the Non-Normed Fit Index (NNFI), the Comparative Fit Index (CFI), and the Root Mean Squared Residual (RMSEA). Values of .90 and over (for the NNFI and the CFI) and .10 and lower (for RMSEA) indicate acceptable model fit (Byrne, 2001). Further, values for the Expected Cross-Validation Index (ECV1) are presented. For this index no cutoff values are available, but lower values indicate a higher likelihood that findings will replicate to other samples. Models were compared using the χ^2 difference test as well as the fit indexes mentioned above.

Results

Comparison of models

Table III presents the fit indices for the four models. In the oncology sample, the model of Golembiewski et al. (1986) fitted the data considerably less well than the models of Lee and Ashforth (1993) and Leiter and Maslach (1988), as evidenced by the fact that the χ^2 for the Golembiewski et al. model is at least 97.28 points higher than that for the other two models, with the same number of degrees of freedom. The other fit indexes confirm this impression.

Table III. Comparison of the fit of five models, for oncology care providers and teachers separately.

Model	X^2	df	RMSEA	NNFI	CFI	ECVI
Oncology care providers, $N=218$						
A. The Golembiewski et al (1986) model: +DP → −PA → +EE	167.77	19	.15	.80	.90	.65
B. The Leiter & Maslach (1988) model: +EE → +DP → −PA	52.73	19	.07	.96	.93	.34
C. The Lee & Ashforth (1993) model: +EE → +DP, and +EE → −PA	70.49	19	.09	.93	.96	.39
D. Integration of models B and C plus feedback effects: +EE → +DP → −PA, and +EE → −PA[a]	23.11	14	.05	.98	.99	.28
E. Model D revised: insignificant effects omitted [a, b]	23.30	17	.03	.99	1.00	.26
Teachers, $N=967$						
A. The Golembiewski et al (1986) model: +DP → −PA → +EE	119.2	6	.13	.89	.95	.14
B. The Leiter & Maslach (1988) model: +EE → +DP → −PA	62.2	6	.09	.94	.98	.09
C. The Lee & Ashforth (1993) model: +EE → +DP, and +EE → −PA	68.8	6	.10	.94	.97	.10
D. Integration of models B and C plus feedback effects: +EE → +DP → −PA, and +EE → −PA[a]	4.12	2	.03	.99	1.00	.04
E. Model D revised: insignificant effects omitted [a, b]	6.63	3	.03	.99	1.00	.04

EE = Emotional exhaustion, DP = Depersonalization, PA = Personal accomplishment.
'+' = 'high levels of', '−' = 'low levels of', '→' = 'lead to'. For example, '+EE → −DP' denotes 'high levels of EE lead to low levels of DP'.
[a] = chi-square difference tests indicate that this model fits the data significantly better ($p < .05$) than models A–C.
[b] = chi-square difference test indicates that this model fits the data not significantly worse ($p > .05$) than model D.

The fit of the Leiter and Maslach (1988) and the Lee and Ashforth (1993) models is quite comparable, with the Leiter and Maslach model yielding a slightly better fit than the Lee and Ashforth model. However, our synthesis/extension of the Lee and Ashforth/Leiter and Maslach models (model D) fitted the data significantly better than its three contenders. Compared to the Leiter and Maslach (1988) model, our model gains 29.62 χ^2 points at the loss of only 5 degrees of freedom, $p < .001$, whereas the other indexes also show that Model D fits the data better than Models A-C. The results for the teacher sample were very similar. Whereas the Golembiewski et al. (1986) model fitted the data badly, the Lee and Ashforth (1993) and Leiter and Maslach (1988) models fitted the data about equally well. Our model fitted the data better than the other models (compared to the best-fitting Leiter and Maslach model, model D gains 64.7 χ^2 points at the loss of 4 degrees of freedom, $p < .001$) and is thus accepted as the best *a priori* model for the relationships among the dimensions of the MBI.

Closer inspection of Model D for both samples revealed that not all effects included in these models were statistically significant. These effects were omitted. Specifically, for the teacher sample we omitted the direct effect of emotional exhaustion (EE) on personal accomplishment (PA). This effect was also omitted for the oncology care providers. For the latter sample we also excluded the effect of depersonalization (DP) on PA and the feedback effect of DP on EE. After omitting these non-significant associations, the final models (one for each of the two samples) still fitted the data acceptably well (models E in Table III).

Substantive implications

Table IV presents the standardized maximum likelihood estimates for both samples for the final models. Figure 1 presents the main findings graphically, omitting stability effects and R^2 values. As Table IV shows, the scores on EE, DP and PA were fairly stable across time; the magnitude of these coefficients compares well to the findings reported in previous research (cf. Table I). For the oncology care providers we also found positive Time 1–Time 3 effects; although the magnitude of these effects was lower than that of the one-year effects, they underline the temporal stability of EE, DP and PA, respectively.

Of more interest are the lagged effects shown in Figure 1. Consistent with the Leiter and Maslach (1988) and Lee and Ashforth (1993) models, we found that high levels of EE were associated with high levels of DP at a later point in time (standardized effects of .18 for the oncology care providers and .15 for the teachers, both effects $p < .001$). Further, if psychological withdrawal in the form of depersonalization is a dysfunctional coping strategy, one would expect elevated levels of DP at one wave to be associated with higher levels of EE and lower levels of PA at a later wave. Both effects were partly confirmed; in the teacher sample we found that high levels of depersonalization were associated with higher levels of EE and lower levels of PA across time. Both effects support the notion that DP is a dysfunctional coping strategy; the first effect is consistent with previous cross-sectional findings by Bakker et al. (2000), whereas the latter effect was already proposed by Leiter and Maslach (1988) and Golembiewski et al. (1986). Note, however, that these effects were not replicated in the smaller oncology care sample.

Discussion

The present research aimed to provide more insight in the possible causal relationships among the three dimensions of the Maslach Burnout Inventory. Our literature review revealed that up until now three major models for the causal order among these dimensions

Table IV. Standardized maximum likelihood estimates for the lagged effects in the final models (upper estimate: oncology care providers, lower estimate: teachers).

Variables	Time 2			Time 3		
	Emotional exhaustion	Deperson-alization	Personal accomplishment	Emotional exhaustion	Depersonalization	Personal accomplishment
T1 Emotional exhaustion	.67**	.18**[a]		.29**		
	.73**	.15**				
T1 Depersonalization	−	.50**			.22**	
	.10*	.55**	−.14**			
T1 Personal accomplishment			.64**			.20**
			.60**			
T2 Emotional exhaustion				.51**	.18**[a]	
T2 Depersonalization					.45**	
T2 Personal accomplishment						.56**
R^2	.44	.36	.41	.55	.50	.51
	.60	.40	.44			

All effects significant at $p < .05$, except * $p < .01$ and ** $p < .001$.
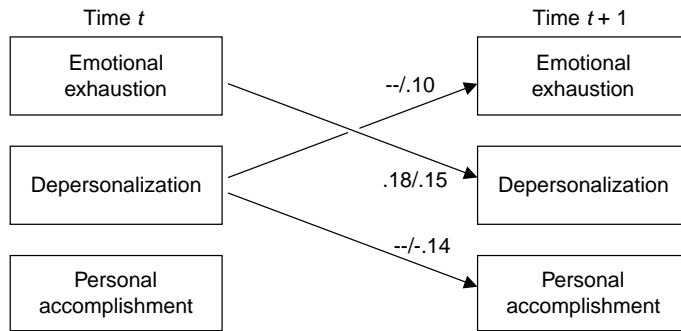[a]Effect was constrained to be equal across waves.

Figure 1. Lagged effects among the three burnout dimensions in the present study (models E), all effects significant at $p < .05$ or better. Of the two estimates given for each effect, the first is for the sample of oncology care providers and the second for the teachers.

have emerged. The Leiter and Maslach (1988) and Lee and Ashforth (1993) approaches were rather similar, proposing that burnout starts with emotional exhaustion and ends with reduced personal accomplishment. In contrast, Golembiewski et al. (1986) argued that depersonalization marks the onset of a burnout and that emotional exhaustion reflects its completion.

All three models have repeatedly been tested in previous research, frequently leading to the conclusion that one of these models fitted the data best and should therefore be accepted as reflecting the causal process that generated the associations among the three MBI dimensions. However, our literature review identified major problems with this research. Previous work on this issue recurrently employed cross-sectional designs, meaning that causal claims are unwarranted. Further, the studies that did employ longitudinal designs failed to provide any *longitudinal* evidence for the causal models tested therein. Moreover, often only one particular model was tested; comparison with competing models is impossible in these cases. Finally, while most theorists argue that psychological withdrawal in the form of depersonalization should be considered a coping strategy, two of the three major models for the development of burnout do not consider possible 'feedback' effects of depersonalization on the variable marking the onset of burnout (i.e. emotional exhaustion in both the Leiter & Maslach (1988), and Lee & Ashforth (1993), approaches—according to Golembiewski et al. (1986), depersonalization is the start of the burnout process). All in all, it seems fair to conclude that previous research failed to provide convincing evidence for any of the three models for the development of burnout.

Using longitudinal data from two samples of 218 oncology care providers and 967 teachers, the three models discussed above were compared to each other and to a fourth model (our own), which integrated the ideas of Leiter and Maslach (1988) and Lee and Ashforth (1993), and that included effects of depersonalization on emotional exhaustion. As in previous research, the results largely failed to confirm the model of Golembiewski et al. (1986). The models of Lee and Ashforth (1993) and Leiter and Maslach (1988) were about equally strongly supported, with the Leiter and Maslach (1988) approach receiving slightly stronger support than its competitor. However, our synthesis/extension of the Lee and Ashforth (1993) and Leiter and Maslach (1988) models fitted the data even better in both samples. After omitting non-significant effects, our findings supported the idea that higher levels of exhaustion trigger higher levels of depersonalization (this effect was replicated across both samples and is consistent with the Leiter & Maslach (1988), and Lee & Ashforth (1993), models). Further, we found indications that higher levels of

depersonalization are associated with lower levels of personal accomplishment across time (this effect was only present for the teachers, and supports the ideas advanced by Golembiewski et al. (1986) and Leiter & Maslach (1988)). Finally, higher levels of depersonalization were associated with higher levels of exhaustion across time (teacher sample only; this effect is consistent with the notion that depersonalization is a dysfunctional coping strategy, cf. Bakker et al., 2000).

Study limitations

In order to judge the evidence for the models for the across-time development of burnout adequately, several limitations of the present research need to be discussed. First, it should be noted that the lagged effects reported in this research are relatively small, providing one explanation of why previous research (usually employing fairly small samples; the 7 samples discussed in Table I included on average only 229 participants) failed to provide longitudinal support for any of the three models for the development of burnout. These small lagged effects also mean that little variance in the across-time change in the MBI components is accounted for. While this is to be expected if the phenomenon under study is relatively stable (as seems to be the case for burnout, see Table I), it also suggests that from a practical point of view it is actually hard to predict the development of burnout. Whereas at any given time point there are strong associations among the burnout components, across-time *changes* in these components (e.g. changes in levels of depersonalization) hardly depend on the scores on the component that conceptually precedes this component (e.g. emotional exhaustion).

A second limitation concerns the length of the time interval used. In both samples we used one-year intervals between the study waves. While this is convenient for controlling seasonal influences (De Lange, Taris, Kompier, Houtman, & Bongers, 2003), it would seem possible that this particular time interval does not adequately reflect the causal interval that separates changes in, say, emotional exhaustion from corresponding changes in depersonalization (Zapf, Dormann, & Frese, 1996). Unfortunately, as yet we have little evidence regarding the correct length of the time between waves. As burnout is usually construed as the result of a chronic exposure to high job demands (e.g. Maslach et al. (2001)—but see Golembiewski et al. (1986), who also distinguishes an acute form of burnout), one might expect that relatively long time intervals would reflect the true causal interval best. However, our review revealed that *all* previous research failed to provide longitudinal evidence for any model in particular, irrespective of the time interval used. Nevertheless, if the underlying causal interval deviates from the interval that is employed in a particular study, the magnitude of the lagged effects will be estimated conservatively. Probably the best way to study the development of burnout would be to conduct a long-term longitudinal study consisting of many waves separated by short time intervals. Such a design would allow researchers to examine which time interval reflects the true causal interval best. In practice, however, such a design will often be unfeasible.

Finally, one should note that the present research focused on the across-time associations among the three MBI dimensions only; contextual information (e.g. job characteristics) was largely ignored. This could mean that the effect sizes reported in this research are biased because we failed to control for potentially relevant confounding variables; it is well known that all three burnout components are strongly associated with work characteristics such as job demands and job control (De Lange et al., 2003; Lee & Ashforth, 1993; Söderfeldt, 1997). Although the present data sets included measures of demands and control, for simplicity these were not included in the current analyses. However, to examine the

magnitude of the bias that possibly resulted from not controlling these contextual variables, we re-estimated the final models (models E) reported in Table III after partialling out the scores on job demands and job control. Although the sizes of the effects reported in Table IV/Figure 1 changed somewhat, our conclusions generally remained the same. Thus, our results do not seem to have been severely biased by not controlling for contextual variables.

Implications

In spite of these limitations, we believe that the present research extends and enhances current knowledge in at least three respects. First, our review of previous research revealed that to date, no convincing support for any particular causal order of the MBI dimensions was available, in spite of the fact that researchers frequently assume that there is such evidence (Kitaoka-Higashiguchi et al., 2004). Second, our results showed that all three models proposed to date were at least partly supported longitudinally. To our knowledge, our research is the first to provide replicable longitudinal evidence that high levels of exhaustion trigger high levels of depersonalization (as suggested by Leiter & Maslach, 1988 and by Lee & Ashforth, 1993); we also found longitudinal support for the assumption of Golembiewski et al. (1986) and Leiter and Maslach (1988) that high levels of depersonalization lead to low levels of personal accomplishment, but only for the teacher sample. Third, our results underlined the idea that psychological withdrawal in the form of depersonalization is a dysfunctional and not an effective coping strategy. Thus, while a little professional detachment may well be necessary and functional in dealing with workplace stressors (cf. Golembiewski et al., 1996; Lief & Fox, 1963), too much of a good thing seems to have adverse consequences for worker well-being. The precise structure of this process is as yet largely unknown. For instance, overly detached workers may find it impossible to form necessary relationships with others (cf. Golembiewski et al., 1986), their indifferent behaviour may trigger demanding behaviours from others (cf. Bakker et al., 2000), they may simply perceive their job to be more taxing, thus increasing levels of exhaustion (as suggested by De Lange et al., 2004), or perhaps all these explanations apply; only follow-up research can provide evidence regarding this issue.

In conclusion, the main contribution of the present research lies in the fact that it supports the notion that burnout can indeed be conceptualized as a developmental process. However, the sizes of the lagged effects were rather low, suggesting that it is incorrect to assume that high levels of exhaustion automatically lead to high levels of depersonalization and low levels of personal accomplishment. Rather, while highly exhausted workers may have a tendency to become more depersonalized and to experience lower levels of achievement across time, these effects seem to be too weak to be of much practical use (e.g. for the early recognition and treatment of a case of developing burnout). Thus, whereas the practical value of the present set of findings may be quite limited, our research suggests that more research into the consequences of high levels of detachment and depersonalization for worker well-being, behaviour and performance could be worthwhile.

References

Bakker, A. B., Schaufeli, W. B., Sixma, H. J., Bosveld, W., & van Dierendonck, D. (2000). Patient demands, lack of reciprocity, and burnout: A five-year longitudinal study among general practitioners. *Journal of Organizational Behavior*, *21*, 425–441.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Hillsdale, NJ: Lawrence Erlbaum.

Büssing, A., & Glaser, J. (2000). Four-stage process model of the core factors of burnout: The role of work stressors and work-related resources. *Work and Stress*, *14*, 329–346.

Cordes, C. L., Dougherty, T. W., & Blum, M. (1997). Patterns of burnout among managers and professionals: A comparison of models. *Journal of Organizational Behavior*, *18*, 685–701.

De Lange, A. H., Taris, T. W., Kompier, M. A. J., Houtman, I. L. D., & Bongers, P. M. (2003). The *very* best of the millennium: Longitudinal research and the demand-control(-support) model. *Journal of Occupational Health Psychology*, *8*, 282–305.

De Lange, A. H., Taris, T. W., Kompier, M. A. J., Houtman, I. L. D., & Bongers, P. M. (2004). The relationships between job characteristics and mental health: Examining normal, reversed and reciprocal relationships in a 4-wave study. *Work & Stress*, *18*, 149–166.

Freudenberger, H. J. (1975). The staff burnout syndrome in alternative institutions. *Psychotherapy: Theory, Research and Practice*, *12*, 72–83.

Golembiewski, R. T., Boudreau, R. A., Munzenrider, R. F., & Luo, H. (1996). *Global burnout: A worldwide pandemic explored by the phase model*. Greenwich, CT: JAI Press.

Golembiewski, R. T., Munzenrider, R. F., & Stevenson, J. G. (1986). *Phases of burnout: Developments in concepts and applications*. New York: Praeger.

Jöreskog, K. G., & Sörbom, D. (1999). *LISREL-8.30* (computer manual). Chicago, IL: Scientific Software.

Kitaoka-Higashiguchi, K., Nakagawa, H., Morikawa, Y., Ishizaki, M., Miura, K., Naruse, Y., Kido, T., & Higashiyama, M. (2004). Construct validity of the Maslach Burnout Inventory—General Survey. *Stress & Health*, *20*, 255–260.

Lazarus, R.L., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lee, R. T., & Ashforth, B. E. (1993). A longitudinal study of burnout among supervisors and managers: Comparisons between the Leiter and Maslach (1988) and Golembiewski et al. (1986) models. *Organizational Behavior and Human Decision Processes*, *54*, 369–398.

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. (1990). The impact of family resources, control coping, and skill utilization on the development of burnout: A longitudinal study. *Human Relations*, *43*, 1067–1083.

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

Leiter, M. P., & Maslach, C. (1988). The impact of interpersonal environment of burnout and organizational commitment. *Journal of Organizational Behavior*, *9*, 297–308.

Lief, H. I., & Fox, R. C. (1963). Training for 'detached concern' in medical students. In H. I. Lief, V. F. Lief, & N. R. Lief (Eds), *The psychological basis of medical practice* (pp. 12–35). New York: Harper & Row.

Maslach, C. (1976). Burned-out. *Human Behavior*, *5*, 16–22.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *MBI: The Maslach Burnout Inventory: Manual research edition*. Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397–422.

McManus, I. C., Winder, B. C., & Gordon, D. (2002). The causal links between stress and burnout in a longitudinal study of UK doctors. *The Lancet*, *359*, 2089–2090.

Savicki, V., & Cooley, E. J. (1994). Burnout in child protective service workers: A longitudinal study. *Journal of Organizational Behavior*, *15*, 655–666.

Schaufeli, W. B. (2003). Past performance and future perspectives of burnout research. *South African Journal of Industrial Psychology*, *29*, 1–15.

Schaufeli, W. B., & Van Dierendonck, D. (2000). *Utrechtse Burnout Schaal* (Utrecht Burnout Inventory). Lisse: Swets Test Publishers.

Shirom, A. (1989). Burnout in work organizations. In C. L. Cooper, & I. T. Robertson (Eds), *International review of industrial and organizational psychology 1989* (pp. 25–48). New York: Wiley.

Söderfeldt, M. (1997). *Burnout?* Unpublished doctoral dissertation, Lunds Universitet, Lunds.

Taris, T. W. (2000). *A primer in longitudinal data analysis*. London: Sage.

Toppinen-Tanner, S., Kalimo, R., & Mutanen, P. (2002). The process of burnout in white-collar and blue-collar jobs: Eight-year prospective study of exhaustion. *Journal of Organizational Behavior*, *23*, 555–570.

Van Dierendonck, D., Schaufeli, W. B., & Buunk, B. P. (2001). Toward a process model of burnout: Results from a secondary analysis. *European Journal of Work and Organizational Psychology*, *10*, 41–52.

Van Horn, J. E., Taris, T. W., Schaufeli, W. B., & Schreurs, P. J. G. (2004). The structure of occupational well-being: A study among Dutch teachers. *Journal of Occupational and Organizational Psychology*, *77*, 365–375.

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