J Occup Health 2007; 49: 224–234

Journal of Occupational Health

"How Fatigued Do You Currently Feel?" Convergent and Discriminant Validity of a Single-Item Fatigue Measure

Madelon L.M. VAN HOOFF, Sabine A.E. GEURTS, Michiel A.J. KOMPIER and Toon W. TARIS

Department of Work and Organizational Psychology, Radboud University Nijmegen, The Netherlands

Abstract: How Fatigued Do You Currently Feel? Convergent and Discriminant Validity of a Single-Item Fatigue Measure: Madelon L.M. VAN HOOFF, *et al*. Department of Work and Organizational Psychology, Radboud University Nijmegen, The Netherlands—The main aim of this study was to establish the convergent and discriminant validity of a single-item measure of daily fatigue ("How fatigued do you currently feel?") in a daily diary context. Convergent validity of our measure was examined by relating it to a validated multiple-item measure of fatigue (Profile of Mood States; McNair, Lorr, & Droppelman, 1971) and to other daily (work-home interference, sleep complaints, work-related effort) and global (fatigue, health complaints, work-home interference, job pressure) measures that are conceptually related to fatigue. Discriminant validity was assessed by relating the single-item fatigue measure to daily (work pleasure) and global (job control, social support, motivation to learn) measures that are conceptually distinct from fatigue. Data were collected among 120 academic staff members, who completed a general questionnaire (tapping the global measures under study) and who took part in a 9-d daily diary study (3 measurements daily). Correlation patterns and multilevel analyses revealed strong and significant associations between the single-item fatigue measure and the variables incorporated to assess convergent validity (especially with the POMS: r=0.80), thus supporting the convergent validity of our measure. Relations with variables included to examine discriminant validity were weak or insignificant, supporting the discriminant validity of the single-item fatigue measure. Despite this study's limitations (i.e., exclusive use of self-reporting, specific sample) we conclude that this single-item fatigue measure offers a valid way to assess daily fatigue.
(J Occup Health 2007; 49: 224–234)

Received Jun 29, 2006; Accepted Feb 26, 2007
Correspondence to: M.L.M. van Hooff, TNO Quality of Life, P.O. Box 718, 2130 AS Hoofddorp, The Netherlands (e-mail: Madelon.vanHooff@tno.nl)

Key words: Fatigue, Academics, Validity, Diary study

Fatigue is a central concept in occupational health psychology. Research on fatigue started with Angelo Mosso's pioneering work at the end of the 19th century[1] and gained momentum during and after both World Wars, when research focused on the development of performance standards and work and rest time schedules[2]. In recent years, many studies have examined fatigue (or its more extreme variant "exhaustion") in relation with, for example, high work pressure[3], work-home interference[4] or lack of recovery[5].

The aim of the present study was to validate a single-item measure of daily fatigue. The increasing usage of diary studies provided the inspiration for the present study[6]. Although such studies offer good opportunities to "capture life as it is lived"[7], participation in a diary study is rather demanding as it usually requires respondents' commitment over a period of several days (sometimes reporting several times per day). For that reason, it is important to ensure that potential respondents are not scared off in advance, and do not drop-out during the course of the study by keeping participants' effort investment within acceptable limits and by creating user-friendly diaries. One possible way to achieve this is to employ short, simple, and comprehensible questions to measure the constructs under study[8]. In this regard, single-item measures seem to offer important advantages over multiple-item measures.

Single-item measures

One of the advantages of single-item measures over multiple-item measures is their face validity[9]. It is immediately clear to the respondents which construct is being measured. A second and related advantage is that such a measure probably evokes less participant boredom, fatigue and frustration because there is no item redundancy[10] or repetition of comparable items. A third and more practical advantage is that single-item measures are quite convenient when space or time constraints limit

the number of items that can be incorporated in a (diary) survey[9, 10]. Finally, single-item measures may be more cost-effective than multiple-item measures, given that the costs of short questionnaires are lower than those of long questionnaires assessing the same concepts[9].

Despite these potential advantages of single-item measures, the common practice in academic research is to use multiple-item scales. Using single-item measures is generally discouraged, and reviewers often consider using such scales a "fatal error"[9] (p. 247). Indeed, single-item measures may suffer from at least two psychometric problems[10]. A first concern is that single-item measures assessing broad and multi-faceted constructs lack content validity, in that it is difficult to tap all aspects of such a construct with only one item. Conversely, if the construct under study is sufficiently unidimensional (as previous research has shown regarding fatigue[11]), and unambiguous to the respondents, single-item measures are not necessarily inferior to multiple-item measures[12]. A second issue is that, in cases of broad psychological constructs, multi-item measurement is needed to obtain reliable estimates of the participants' true scores on the phenomena of interest. The scores on the separate items of these measures may contain a large error component, but as these errors are presumed to be due to random factors, they should largely cancel each other out. Multi-item measures will therefore generally give a more reliable indication of the participants' true scores than single-item measures.

In sum, in the case of uni-dimensional constructs, single-item measures may be psychometrically acceptable substitutes for multiple-item measures, and may be preferred for reasons of efficiency and user-friendliness[10]. Indeed, previous research on single-item measures has already demonstrated that such measures can be valid alternatives for multiple-item scales tapping concepts such as job satisfaction[9, 13], self-esteem[10], the Big Five dimensions of personality[14], and stress symptoms[15]. With respect to the measurement of (chronic) fatigue, Rohland, Kruse and Rohrer[16] showed that a single item of burnout was a good alternative for the exhaustion subscale within a burnout questionnaire (Maslach Burnout Inventory). Furthermore, in their study on cancer-related fatigue, Kirsh, Passik, Holtsclaw, Donaghy and Theobald[17] concluded that a single fatigue item ("I get tired for no reason") could be a fast and accurate way to screen patients for fatigue. Schwartz *et al.*[18] concluded that a single-item 11-point fatigue scale ("What is your level of fatigue today?") is sensitive to moderate changes in fatigue and is easily administered in a clinical setting.

The present study

The present study contributes to previous research on the validity and utility of single-item measures of fatigue in five ways: 1) we examined daily fatigue in a sample of the general working population instead of a clinical sample; 2) we tried to minimize retrospection bias by assessing current fatigue three times a day rather than requesting a global rating reflecting fatigue during a whole day[18]; 3) we employed a nine-day daily diary study, which enabled us to obtain more reliable estimates of the relationships between the fatigue measure and other variables, and to examine the robustness of these relationships across time; 4) we extended validity information by investigating relationships between daily fatigue and daily constructs closely related to fatigue (i.e. alternative daily measures of fatigue, and daily measures of work-home interference, work-related effort, and sleep complaints), as well as with constructs conceptually different from fatigue (i.e. daily work pleasure); and 5) we examined relationships between the single-item daily fatigue measure and more habitual or global indicators of closely related (e.g., global fatigue, health complaints, work-home interference and work pressure) and different constructs (such as job control, social support, and motivation to learn). Such global measures are usually employed in survey research and may thus be considered the gold standard. As these measures should reflect an aggregate of day-to-day experiences, it is important to relate the single-item daily fatigue measure to these global measures as well.

The validity of a measure refers to the extent to which it actually measures what it claims to measure. One important source of validity evidence stems from relationships with other measures. Convergent validity evidence is obtained when a measure is positively related to questionnaires that tap similar constructs[19]. The absence of relationships with measures that tap different constructs provides discriminant validity evidence[19].

In order to provide convergent and discriminant validity evidence for a single-item measure of daily fatigue, we related a single-item report mark of fatigue to various other daily and global scales. Convergent validity evidence was investigated by relating the single-item report mark to the fatigue subscale of a well-validated instrument to measure fatigue, the Profile of Mood States (POMS)[20]. Both measures were assessed three times daily, for nine consecutive days. Convergent validity was further assessed by relating the single-item fatigue measure to other stress-related constructs, which are supposed to be associated with, but not identical to, fatigue. These related constructs comprised day-to-day measures on the one hand, and global measures on the other. To establish discriminant validity evidence, we investigated whether our single-item report mark exhibits weak or negative associations with external variables that are conceptually different from fatigue and that are, thus, supposed to be not or only weakly related to fatigue. Again, both daily and global measures are addressed.

In summary, in order to validate the daily single-item

report mark of fatigue, we addressed three research questions:

1) How does the single-item report mark of daily fatigue relate to daily fatigue as measured by the POMS? (convergent validity evidence);
2) How does the single-item report mark of fatigue relate to measures assessing other stress-related constructs, that is, to a) daily measures (i.e., work-home interference, work-related effort and sleep complaints) and b) global measures (i.e., fatigue, health complaints, work pressure, and work-home interference)? (convergent validity evidence); and
3) How does the single-item report mark of fatigue relate to measures that are conceptually different from fatigue, that is, to a) daily measures (i.e., work pleasure) and b) global measures (i.e., job control, social support, and motivation to learn)? (discriminant validity evidence).

Method

Participants and procedure

The study was conducted in two stages among academic staff members of a Dutch university. As the data of the present study were collected as part of a study focusing on the work-nonwork interface of academics, of 696 tenured employees who worked at least three days a week the participants were only those who (i) did not have a second job outside this university (to keep variation in work activities within acceptable limits), and (ii) lived together with a partner who worked at least 2.5 d a week (to increase the likelihood that the participants fulfilled at least some home obligations). A total of 146 employees agreed to participate. Of these, 133 completed a general questionnaire (91% response), assessing demographical information and global measures of interest for this study (see the "Measures" section).

Data of 13 participants were removed as they apparently did not meet the second inclusion criterion. The final sample therefore comprised 120 participants (62% male; 67% had at least one child living in the household; M_{age} was 45.2 yr, SD=7.8; they worked on average 34.2 (SD=5.5) contractual hours weekly; 46% worked as an assistant professor, 17% as an associate professor, 11% as a full professor, and the remaining 26% had other jobs, such as researcher or lecturer). Due to strict privacy regulations, it was unknown how many of the employees who were approached for participation in the study actually met our inclusion criteria (i.e. had no job outside the university and lived together with a partner who worked at least 2.5 d a week). Therefore, we do not know how many employees were in fact eligible for participation in the study, meaning that the overall response rate and the representativeness of our sample are not known.

In the second stage of the study (starting approximately ten days after the completion of the general questionnaire), the daily variables of interest were assessed by means of short questionnaires that were completed during two weekend days (1st Saturday and 1st Sunday), followed by five weekdays (Monday to Friday) and again ending with two weekend days (2nd Saturday and 2nd Sunday). On each of these nine consecutive days, three questionnaires were completed: (1) a morning questionnaire (to be completed after waking in the morning, i.e., between 7.30 and 8.30 a.m.), (2) an afternoon questionnaire (to be completed around 6 p.m.), and (3) an evening questionnaire (to be completed before bedtime, i.e., between 10 and 11 p.m.). Only diaries that were completed within an acceptable time range around the requested time were included in the final database. We removed morning questionnaires that were completed more than 2 h after waking, afternoon questionnaires that were completed before 4.30 p.m. or after 8 p.m., or less than 3 h after the morning questionnaires, and evening questionnaires that were filled in less than 2 h after the afternoon questionnaire or after 3 a.m.. The remaining questionnaires represented 72.1% of the morning, 72.6% of the afternoon, and 78.5% of the evening questionnaires.

Measures

Daily measures: The *single-item fatigue report mark* was obtained each day in the morning, afternoon and evening questionnaires. Participants rated their current state of fatigue ("How fatigued do you currently feel?") with a report mark varying from "1" ("not at all") to "10" ("extremely").

Daily fatigue POMS was measured each day in the morning, afternoon and evening questionnaires with the six-item fatigue subscale of a shortened version of the Dutch translation of the Profile of Moods States (POMS)[21]. The POMS[20] is a questionnaire for the measurement of moods, and since its development, the instrument has been used in almost 3,000 scholarly publications[22]. There is ample evidence for the validity of this instrument[23–26].

Based on factor and item analyses, a shortened version of the Dutch translation of the POMS was developed by Wald en Mellenbergh[21], in which the fatigue subscale comprises six items. In a previous study examining the factor structure of the 65-item version of the POMS, these six items showed the highest loadings on the fatigue factor[25]. Wicherts and Vorst[27] found support for the factor structure of the shortened Dutch POMS in a sample of 5,880 psychology freshmen and reported measurement invariance across gender for the fatigue subscale as well.

Items were scored on a five-point scale (1="not at all", 2="a little", 3="moderately", 4="quite a bit", 5="extremely"), and scale scores were computed as the mean of the six items. Three exemplary items are "Right now, I feel exhausted", "Right now, I feel worn out",

"Right now, I feel bushed" (morning: α=0.87, afternoon: α=0.89, evening: α=0.89).

Daily work-home interference was measured with eight items during weekdays in the evening questionnaire. To this purpose the work-home interference subscale of the Survey Work-Home Interaction NijmeGen (SWING)[28] was slightly adapted to make the items suitable for day-to-day measurement. Two exemplary items are "Today, my work took up time that I would have liked to spend with my spouse/family/friends" and "Today I found it difficult to fulfil my domestic obligations, because I was constantly thinking about my work" (1="no", 2="a little" and 3="yes", α=0.82).

Work-related effort was measured during weekdays in the afternoon questionnaire. Participants were requested to indicate with a report mark the extent to which they considered the preceding workday as effortful (1="not at all", 10="extremely").

To assess daily *sleep complaints* (each morning questionnaire), a sum score was computed of five items from a sleep quality scale[29], slightly adapted to make them suitable for day-to-day measurement. Two exemplary items are: "Last night I woke up several times" and "I slept well last night" (reversed) (1="yes", 0="no", α=0.71). Each day's value for this variable refers to the previous night.

Daily work pleasure was measured with one item in each weekday afternoon questionnaire. Participants were asked to indicate with a report mark the extent to which they considered the preceding workday as pleasurable (1="not at all", 10="extremely").

Measures derived from the general questionnaire: *General fatigue* was assessed with the 10-item Fatigue Assessment Scale[30], which addresses mental as well as physical aspects of fatigue. Two exemplary items are "I am bothered by fatigue" and "Mentally, I feel exhausted" (1="(almost) never", 5="(almost) always"), with higher scores reflecting higher levels of fatigue (α=0.86).

Health complaints were measured with a Dutch questionnaire on subjective health developed by Dirken[31], the so-called VOEG. In this study the 13-item version (VOEG13) was used[32], which has been extensively validated in Dutch samples. Participants were asked whether or not they experienced any of 13 health complaints. Two exemplary items are: "Do you fairly often suffer from headache?" and "Do you fairly often feel dizzy?" ("yes"=1; "no"=0). For each participant, a sum score was computed reflecting the reported number of health complaints (α=0.71).

Global work-home interference was measured with an eight-item subscale from the SWING[28]. Two exemplary items are "How often does it happen that you find it difficult to fulfil your domestic obligations because you are constantly thinking about your work?" and "How often does it happen that you have to work so hard that you do not have time for any of your hobbies?", (0="(almost) never", 1="sometimes", 2="often", 3="(almost) always"). Higher scores reflect higher levels of work-home interference (α=0.73).

Job pressure was measured with five items adapted from the Job Content Questionnaire[33], that were rephrased as questions (e.g., "Do you have to work very fast?" and "Do you have enough time to get the job done?" (reversed) ; 1="(almost) never", 4="(almost) always"; α=0.73).

We used six items from Van Veldhoven *et al.*[29] to measure *Job control*. Two exemplary items are: "Can you take a short break if you feel this is necessary?" and "Can you decide for yourself how to do your job?" (1="(almost) never", 4="(almost) always"; α=0.68).

Social support was measured with eight items adapted from Geurts, Rutte and Peeters[34]. Four items assess support received from colleagues (e.g., "My colleagues show their appreciation for the way I do my job"), and also four items measure support received from a supervisor (e.g., "My supervisor shows her/his appreciation for the way I do my job", 1="(almost) never", 4="(almost) always"; α=0.90).

To measure employees' motivation to learn in their job, we used seven items from the *Motivation to learn* scale[35]. Two exemplary items are: "In my job, I feel encouraged to learn new things" and "In my job, I can develop myself" (1="(almost) never", 4="(almost) always"; α=0.76).

Demographic variables: Sex (0=male, 1=female), age (in years) and job class (1=assistant professor (both teaching and research) , 2=associate professor (both teaching and research), 2=full professor (both teaching and research), 4=researcher (no teaching), 5=lecturer (no research), 6=other) were included as demographic variables.

Statistical analyses

Global and daily variables: *Descriptive analysis*. Correlations were computed to obtain insight into the associations between the single-item fatigue report mark and the alternative fatigue measure (POMS, *Research Question 1*), and between the single-item fatigue report mark and the other (daily and global) variables that were incorporated in order to assess convergent (*Research Question 2*) and discriminant (*Research Question 3*) validity. For all daily variables (including both the fatigue report mark and the multiple-item alternative, POMS), the correlations were based on their mean values across all measurement points.

Daily variables: Multilevel analyses. Correlations offer basic insight into the associations among the fatigue report mark, the POMS fatigue measure and the other daily variables under study. However, our rich diary data allow

us to examine the stability of the relationships of interest across all days of the observation period. To investigate this issue, multilevel analysis has to be used[36, 37], as our day-level data (level 1) are not statistically independent as they are nested within persons (level 2). Multilevel analysis takes into account that the data at the lowest level (in this case, the day level) are nested within a higher-order level (i.e., the participants), effectively resolving the statistical dependencies and the bias this may create. Using multilevel analysis it is possible to specify and compare models with each other. In the present case, the first model to be compared included only an intercept and in the following models predictors (both on the person and on the day level) were added consecutively. The improvement of one model above a previous one can be tested using a likelihood ratio statistic (following a χ^2-distribution with the number of additional predictors as df)[36, 37].

We used the MLWiN 2.0 software package[38] and all variables were standardized based on their grand mean. As there was no reason to expect relationships between independent and dependent variables to differ between the study's participants, we chose not to model random slopes, but only a random intercept.

To study the relationships between the fatigue report mark and the alternative fatigue measure (POMS; *Research Question 1*), a series of analyses was conducted, in which the POMS served as the dependent variable. We started with a Null model, in which only a random intercept was specified. In Model 1, the fatigue report mark was included to obtain insight into the relationships between this measure and the POMS. Model 2 additionally included sex, age and job class (entered as five dummy variables). Time and Day were subsequently added as covariates in Model 3, because fatigue may vary across the time of day as well as across day of the week. Time was entered as a continuous variable (0=morning, 1=afternoon, and 2=evening). The nine days of the study were represented by eight dummy variables, with Monday as the reference category. In Model 4, two Time × Fatigue report mark and eight Day × Fatigue report mark interactions were incorporated to examine whether the strength of the relationship between the two fatigue measures (report mark and POMS) was stable across the time of day and/or the days of the study.

Four series of analyses were conducted to investigate the relationships between the fatigue report mark and each of the four daily variables that were assessed to obtain convergent (*Research Question 2*: sleep complaints, work-related effort and work-home interference) and discriminant validity evidence (*Research Question 3*: work pleasure). In each case, the respective daily measure served as the dependent variable. As these measures were assessed only once a day, we chose to calculate and include mean daily levels of the fatigue report mark in these analyses, instead of the original three values for each day. Similar to *Research Question 1*, for each of the four daily variables, a Null model was computed that contained only a random intercept. In Model 1, the fatigue report mark was entered as a predictor. Model 2 additionally included sex, age and job class as covariates. The Day covariates were entered as dummy variables (again with Monday as a reference category) in Model 3. Finally, Model 4 incorporated eight (or four, if a measure was only assessed during week days) Day × Fatigue report mark interactions, to examine whether the strength of the association between fatigue and the daily dependent variable under study varied across the days of the study.

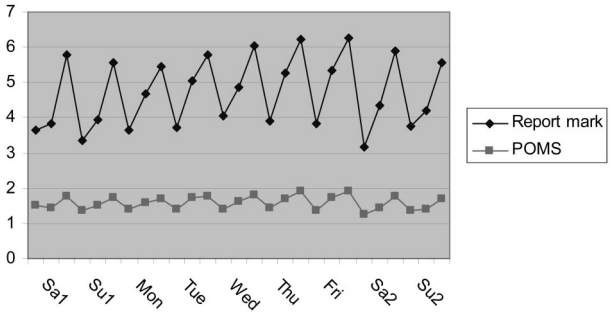

Fig. 1. Mean levels of the single-item fatigue report mark ("Report mark") and the POMS on all measurement occasions. Number of observations between 80 and 98 depending on missing values.

Results

Global and daily variables: Descriptive analysis

For each of the 27 measurement points the mean levels of both fatigue measures are presented in Fig. 1. This figure shows that both measures followed similar patterns during the research period (although the report mark's amplitude was higher due to its wider range, that is, 1–10 vs. 1–3 for the POMS measure).

Table 1 presents the means, standard deviations and correlations of the variables under study.

Levels of global fatigue in this sample (M=1.89, SD=0.59) did not significantly differ from those in a heterogeneous sample of 1,123 employees[28] [M=1.97, SD=0.57; T(1241)=–1.46]. Also, levels of health complaints (M=2.56, SD=2.33) did not differ from those in a heterogeneous sample of 1,421 employees[39] [M=2.62, SD=2.67; T(1539)=–0.24]. The levels of global WHI (M=1.02, SD=0.42) were higher than those in a heterogeneous sample of 1,857 workers[28] [M=0.86, SD=0.48; T(1975)=3.56], and levels of job control (M=3.23, SD=0.43) were higher than in a heterogeneous sample of 1,740 workers[28] [M=2.54, SD=0.63; T(1858)=11.81] as well.

Regarding convergent validity, Table 1 shows a high

Table 1. Means, standard deviations and correlations of the variables under study

	Measure	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12
Convergent validity															
	1 fatigue report mark	1.60	0.44	1											
Daily	2 POMS	4.73	1.59	.80**	1										
	3 daily WHI	1.38	0.36	.45**	.53**	1									
	4 daily sleep complaints	1.40	0.88	.45**	.45**	.32**	1								
	5 daily work-related effort	5.25	1.75	.47**	.28**	.34**	.22*	1							
Global	6 global fatigue	1.89	0.59	.51**	.52**	.33**	.36**	.11	1						
	7 global health complaints	2.56	2.33	.35**	.43**	.29**	.30**	–.03	.66**	1					
	8 global WHI	1.02	0.42	.55**	.58**	.66**	.37**	.34**	.63**	.48**	1				
	9 global job pressure	2.44	0.51	.16	.19	.20*	.10	.27**	.06	.03	.26**	1			
Discriminant validity															
Daily	10 daily work pleasure	6.76	0.95	–.02	–.08	–.24*	–.06	–.08	–.23*	–.15	–.21*	–.09	1		
Global	11 global job control	3.23	0.43	–.12	–.11	–.28**	–.20*	–.05	–.17	–.13	–.28**	–.25**	.25*	1	
	12 global social support	2.46	0.69	.02	.08	–.17	–.10	–.07	–.12	–.21*	–.18*	–.07	.15	.18	1
	13 global motivation to learn	2.60	0.43	–.02	.04	–.04	–.07	.00	–.18*	.03	–.04	.21*	.26**	.13	.27**

Number (%) of missing values between 0 (0%) and 22 (18,3%); mean number of missing values=13 (10.8%).
*p<0.05; **p<0.01.

correlation between the POMS and the report mark (r=0.80, p<0.01). Furthermore, the fatigue report mark was substantially related to global fatigue (r=0.51, p<0.01) and global work-home interference (r=0.55, p<0.01). A somewhat lower association was observed with global health complaints (r=0.35, p<0.01). The fatigue report mark showed no significant association with global job pressure (r=0.16, p>0.05). As to the daily measures, the fatigue report mark was significantly related to all three measures incorporated to address convergent validity ($r_{\text{daily WHI}}$=0.45, p<0.01; $r_{\text{daily sleep complaints}}$=0.45, p<0.01; $r_{\text{dailywork-related effort}}$=0.47, p<0.01). In sum, these results provide convergent validity evidence for the fatigue report mark. Furthermore, the single-item report mark was not significantly related to the global ($r_{\text{job control}}$=–0.12, p>0.05; $r_{\text{social support}}$=0.02, p>0.05; $r_{\text{motivation to learn}}$=–0.02, p>0.05) and daily ($r_{\text{work pleasure}}$=–0.02, p>0.05) variables incorporated to provide discriminant validity evidence. As discriminant validity can be established by showing the construct under study not to be related to measures that tap conceptually different constructs[19], these results offer support for the discriminant validity of the fatigue report mark.

Daily variables: Multilevel analyses

Research question 1. Table 2 presents the multilevel estimates for the models predicting the POMS fatigue measure from the fatigue report mark. Model 1, which includes the fatigue report mark, provided a significantly better fit than the Null model and revealed a strong and positive association between both fatigue measures (β=0.70, p<0.01). The covariates are added in Model 2, but this model did not improve upon Model 1. This indicates that sex, age and job class are not related to fatigue as measured with the POMS. The inclusion of Time and Day in Model 3 did provide a better fit than Model 2, but none of the separate time or day effects were significant. Model 4, which fitted better than Model 3, included the Time × Fatigue report mark and Day × Fatigue report mark interactions. These interactions show that the strength of the relationship between the fatigue report mark and the POMS increased slightly from morning to evening (Time × Fatigue report mark interaction: β=0.09, p<0.01), and during the course of the working week (Day × Fatigue report mark interaction: Wednesday: β=0.10, p<0.05; Thursday: β=0.16, p<0.01; Friday: β=0.17, p<0.01). Although statistically significant, the relevance of these variations can be questioned, as the model including these relationships explained only one percent more variance than the model not including them. In sum, these results provide convergent validity evidence for the fatigue report mark.

Research questions 2 and 3. Multilevel estimates for models relating the daily variables to the single-item measure of fatigue are presented in Table 3.

Daily sleep complaints

Model 1, in which the single-item fatigue measure is included, provided a significant improvement over the Null model. The model shows the report mark and sleep complaints to be positively related (β=0.51, p<0.01).

Table 2. Multilevel estimates for models predicting fatigue POMS from the fatigue report mark

Model and variables	–2*LL	Diff –2*LL (df)	Level 1 intercept variance (SE)	Level 2 intercept variance (SE)	R^2
Null model Intercept	5891.92		.65 (.02)	.38 (.06)	
Model 1 Intercept, report mark	4267.06	1624.86 (1) ***	.33 (.01)	.14 (.02)	.53
Model 2 Intercept, report mark, covariates	4261.99	5.07 (7)	.33 (.01)	.14 (.02)	.54
Model 3 Intercept, report mark, covariates, Time, Day	4243.87	18.12 (9)*	.32 (.01)	.14 (.02)	.54
Model 4 Intercept, report mark, covariates, Time, Day, Time* report mark, Day* report mark	4174.76	69.10 (9)***	.31 (.02)	.14 (.02)	.55

*p<0.05; ***p<0.001. LL=log likelihood; Diff=difference.

Table 3. Multilevel estimates for models relating the fatigue report mark to the daily variables

Daily variable	Model and variables#	–2*LL	Diff –2*LL(df)	Level 1 intercept variance (SE)	Level 2 intercept variance (SE)	R^2
Daily sleep complaints	Null model	2090.52		.80 (.04)	.19 (.04)	
	Model 1	1930.53	159.99 (1)***	.65 (.04)	.14 (.03)	.20
	Model 2	1927.03	3.5 (7)	.65 (.04)	.13 (.03)	.21
	Model 3	1890.56	36.47 (4)***	.62 (.03)	.14 (.03)	.23
	Model 4	1884.50	6.06 (4)	.61 (.03)	.14 (61)	.24
Daily WHI	Null model	985.58		.51 (.04)	.50 (.09)	
	Model 1	947.48	38.10 (1) ***	.49 (.04)	.35 (.07)	.16
	Model 2	939.47	8.01 (7)	.48 (.04)	.32 (.07)	.20
	Model 3	928.40	11.07 (4)*	.47 (.04)	.32 (.06)	.21
	Model 4	928.02	0.37 (4)	.47 (.04)	.31 (.06)	.22
Daily work-related effort	Null model	962.26		.48 (.04)	.53 (.10)	
	Model 1	922.60	39.66 (1)***	.45 (.04)	.38 (.07)	.19
	Model 2	918.69	3.91 (7)	.45 (.04)	.36 (.07)	.21
	Model 3	913.43	5.26 (4)	.45 (.04)	.35 (.07)	.22
	Model 4	910.41	3.02 (4)	.45 (.04)	.34 (.07)	.22
Daily work pleasure	Null model	1028.29		.66 (.06)	.32 (.07)	
	Model 1	1019.17	9.11 (1)**	.64 (.05)	.33 (.07)	.02
	Model 2	1013.65	5.52 (7)	.64 (.05)	.30 (.07)	.05
	Model 3	1007.86	5.79 (4)	.62 (.05)	.30 (.07)	.06
	Model 4	1001.09	6.77 (4)	.61 (.05)	.30 (.07)	.07

*p<0.05; **p<0.01; ***p<0.001. LL=log likelihood; Diff=difference

Null model: Intercept only; Model 1: Intercept, report mark; Model 2: Intercept, report mark, covariates; Model 3: Intercept, report mark, covariates, Day; Model 4: Intercept, report mark, covariates, Day, Day * report mark.

Adding sex, age and job class in Model 2 did not significantly improve the fit compared to Model 1. However, including Day in Model 3 did result in a better fitting model. Model 3 shows that, generally, compared to Monday, sleep complaints were lower in the weekends (1st Sunday: β=–0.28, $p<0.05$; 2nd Saturday: β=–0.59, $p<0.01$; 2nd Sunday: β=–0.25, $p<0.05$) and in the second half of the working week (Thursday: β=–0.37, $p<0.01$; Friday: β=–0.50, $p<0.01$). Finally, Model 4, including the Fatigue report mark × Day interactions, did not provide a better fit, indicating that the strength of the relationship between sleep complaints and the fatigue report mark was constant during the observation period. Altogether, these results offer support for the validity of the fatigue report mark.

Daily work-home interference

Model 1, which fitted the data significantly better than the Null model, showed a positive association between the fatigue report mark and daily work-home interference (β=0.35, $p<0.01$). Model 2, in which the covariates are modelled, did not improve significantly upon Model 1, indicating that sex, age and job class are not related to daily work-home interference. Day was included in Model 3, and this model fitted better than Model 2. This model revealed that work-home interference was generally lower on Friday (β=–0.27, $p<0.05$) compared to Monday. Note that weekend-days were not included in the model, as work-home interference was only assessed on weekdays. As including interactions with Day did not improve the model fit, it can be concluded that the strength of the association between fatigue and daily work-home interference was invariant across the week days. Thus, in sum, these findings provide support for the convergent validity of the fatigue report mark.

Daily work-related effort

Model 1 showed a positive association between the fatigue report mark and daily work-related effort (β=0.35, $p<0.01$) and provided a significantly better fit than the Null model. The covariates were incorporated in Model 2, but this model did not fit better than Model 1, indicating that sex, age and job class are not related to daily work-related effort. Model 3, in which Day is included, did not fit better than Model 2, indicating that, generally, levels of work-related effort were stable during the week. Note that, again, weekend-days were not included in the model, as work-related effort was only assessed during weekdays. Including Fatigue report mark × Day interactions in Model 4 did not result in a better fitting model. Therefore, it can be concluded that the strength of the association between the fatigue report mark and work-related effort did not differ from day to day. Overall, these findings provide support for the convergent validity of the fatigue report mark.

Daily work pleasure

Model 1 provided a significant improvement over the Null model and showed that the fatigue report mark was negatively, albeit relatively weakly, related to daily work pleasure (β=–0.18, $p<0.01$). Including the covariates in Model 2 did not result in a better fitting model; thus, age, sex and job class were not related to daily work pleasure. Model 3, in which Day is included, did not provide a better fit either. This indicates that levels of work pleasure were stable across the days of the week. As work pleasure was only assessed during weekdays, weekend days were not included in the model. Incorporating Fatigue report mark × Day interactions in Model 4 did not improve the model fit, pointing out that the strength of the association between the fatigue report mark and daily work pleasure was stable during the observation period. As work pleasure is conceptually different from fatigue, its weak (negative) association with fatigue supports the discriminant validity of the fatigue report mark.

Discussion

The present study was designed to establish convergent and discriminant validity evidence for a single-item report mark of fatigue in the context of a daily diary study. To this purpose, we related this report mark to other daily diary measures and to more habitual or global measures derived from a general questionnaire.

Convergent validity

The results provided evidence for the convergent validity of the single-item fatigue report mark. First, crude correlations revealed a very strong association between the report mark and the alternative multiple item measure (POMS). This result was confirmed using multilevel analysis. Although this analysis also revealed some statistically significant variations in the strength of this association across the time of the day and days of the week, the relevance of these variations can be questioned, as they only explained one percent of additional variance.

Second, the fatigue report mark was substantially correlated with other, supposedly related, daily variables: daily work-home interference, daily sleep complaints and daily work-related effort. These findings were confirmed by means of multilevel analysis. Moreover, this latter analysis showed that the associations between the report mark and these daily variables were stable across the observation period, and, thus, did not depend on the day they were measured.

Finally, the fatigue report mark was related to three out of the four global variables included to investigate its convergent validity. It was substantially correlated with global fatigue, global health complaints and global work-home interference. No significant association was

found with global job pressure.

Discriminant validity evidence

The results also support the discriminant validity of the single-item fatigue report mark, as it revealed only non-significant or weak relationships with measures supposed to tap constructs other than fatigue. Correlations show that this measure is not significantly related to daily work pleasure and multilevel analysis revealed only a weak negative association with this variable. The report mark was also unrelated to any of the global measures incorporated to examine discriminant validity (i.e., global job control, global social support, global motivation to learn).

Single-item vs. multiple-item measures of fatigue

Whereas these results support the convergent and discriminant validity of our single-item measure of fatigue, they also raise the question of how our single-item measure performs compared to the six-item fatigue scale of the POMS. To address this issue, post hoc analyses were conducted. Regarding the correlations with the other daily measures and with the global measures in used the study, the results show that both fatigue measures were equally strongly related to all measures except to daily work-related effort, with which the report mark showed a somewhat stronger relationship (see Table 1). We also repeated the multilevel analyses for the daily measures that were included to examine convergent and discriminant validity, but for this we included the POMS (instead of the report mark) as the independent variable (results can be obtained from the first author on request). For each daily variable, we compared the fit of two models (one the POMS and the other including the report mark as an independent variable) using Schwarz's Bayesian Information Criterion (BIC)[40]. Following this procedure, it became clear that the POMS and the single-item fatigue measure were equally related to the other daily measures (daily sleep complaints, daily WHI, daily work-related effort, and daily work pleasure). Thus, based on these additional analyses, it can be concluded that the report mark is equivalent to the well-validated six-item measure of fatigue.

Limitations and suggestions for future research

We believe three limitations and suggestions for future research deserve to be mentioned. First, the present research employed a specific sample: academic staff members who worked at least 3 d a week and who lived together with a partner who worked at least 2.5 d a week. Although there seems no reason to assume that our main findings on the relations between the single-item fatigue measure and the other daily and global variables are unique to this sample, it is desirable that this study is replicated among employees in other professions, in other family situations and with other working hours.

Second, our study exclusively employed self-report measures. This may have led to an overestimation of the associations among the variables under study due to common method variance[41]. However, common method variance should have inflated all associations studied, and not just part of them. Thus, the fact that some relationships were found in this study while others were not, argues against this possibility. Moreover, alternative measures such as physiological measures should not by definition be considered superior to self-reported measures, because these are not free of error variance either[42–44]. Furthermore, by demonstrating that a) using self-reporting does not guarantee finding significant results, b) potential biasing variables (social desirability, negative affectivity and acquiescence) do not generally inflate correlations among study variables, and c) monomethod correlations are not by definition higher than multimethod correlations, Spector[41] concluded that "the popular position suggesting common method variance automatically affects variables measured with the same method is a distortion and oversimplification of the true state of affairs" (p221). In sum, we do not believe that common method bias severely biased our findings. In this respect, the use of physiological and performance measures in addition to self-reporting could provide interesting insights in future research.

Third, although the present study sheds light on the associations between a single-item fatigue measure and variables supposed to provide an indication of convergent and discriminant validity evidence, it did not examine the ability of this measure to capture differences in fatigue due to interventions such as vacations or overtime reduction programs. Future research should therefore examine whether this single-item fatigue measure is sensitive to the supposedly beneficial effects of these interventions (cf. Schwartz *et al.*'s study[18] on minimally important clinical differences).

Contributions and implications of this study

In addition to its limitations, we believe the present study's assets should be acknowledged as well. First, we employed a daily diary design with 27 repeated measurements (i.e., nine consecutive days and three measurements daily). This design enabled us to obtain reliable estimates of the relationships between the single-item fatigue measure and the other daily variables included, and made it possible to investigate, and confirm, the stability of these relationships across the research period.

Second, we included relationships between our daily single-item fatigue measure and both daily and global measures assessing constructs that could provide convergent and discriminant validity evidence.

Finally, and most importantly, the results of our study

provide substantial and convincing evidence for the validity of a report mark as a daily measure of fatigue. As the single item measure is psychometrically equivalent to the well-validated six-item fatigue measure, our study has important practical implications for research in this area. It implies that, in contexts where it is important to ask participants as few questions as possible (e.g., in daily diary studies), a single-item suffices to measure fatigue.

Acknowledgments: This study was supported by ASPASIA Grant 015.000.027 from the Netherlands Organization for Scientific Research (NWO) to Sabine Geurts. We thank Sabine Sonnentag for her valuable input in the design of the study and Jessica Rehe for her assistance with the data collection.

References

1) Mosso A. La fatigue intellectuelle et physique. Paris: Felix Alcan, 1894.
2) Meijman T and Schaufeli W: Mental fatigue and work: developments in work and organizational psychology. De Psycholoog 6, 236–241 (1996) (in French)
3) De Lange AH, Taris TW, Kompier MAJ, Houtman ILD and Bongers PM: The relationships between work characteristics and mental health: examining normal, reversed and reciprocal relationships in a 4-wave study. Work Stress 18, 149–166 (2004)
4) Van Hooff ML, Geurts SA, Taris TW, Kompier MA, Dikkers JS, Houtman IL and van den Heuvel FM: Disentangling the causal relationships between work-home interference and employee health. Scand J Work Environ Health 31, 15–29 (2005)
5) Sluiter JK, Frings-Dresen MH, Van der Beek AJ and Meijman TF: The relation between work-induced neuroendocrine reactivity and recovery, subjective need for recovery, and health status. J Psychosom Res 50, 29–37 (2001)
6) Van Eerde W, Holman D and Totterdell P: Special section: editorial. J Occup Organ Psych 78, 151–154 (2005)
7) Bolger N, Davis A and Rafaeli E: Diary methods: Capturing life as it is lived. Annu Rev Psychol 54, 579–616 (2003)
8) Taris TW. A primer in longitudinal data analysis. London: Sage Publications, 2000.
9) Wanous JP, Reichers AE and Hudy MJ: Overall job-satisfaction: how good are single-item measures? J Appl Psychol 82, 247–252 (1997)
10) Robins RW, Hendin HM and Trzesniewski KH: Measuring global self-esteem: construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. Pers Soc Psychol B 27, 151–161 (2001)
11) Michielsen HJ, De Vries J, Van Heck GL, Van de Vijver FJR and Sijtsma K: Examination of the dimensionality of fatigue: the construction of the Fatigue Assessment Scale (FAS). Eur J Psychol Assess 20, 39–48 (2004)
12) Sacket PR, Larson JR. Research strategies and tactics in industrial and organizational psychology. In: Dunnette MD, Hough LM, eds. Handbook of industrial and organizational psychology (2nd ed., Vol. 1). Palo Alto, CA: Consulting Psychologists Press, 1990: 419–489.
13) Nagy MS: Using a single-item approach to measure facet job satisfaction. J Occup Organ Psych 75, 77–86 (2002)
14) Woods SA and Hampson SE: Measuring the Big Five with single items using a bipolar response scale. Eur J Personality 19, 373–390 (2005)
15) Elo A, Leppänen A and Jahkola A: Validity of a single-item measure of stress symptoms. Scand J Work Environ Health 29, 444–451 (2003)
16) Rohland BM, Kruse GR and Rohrer JE: Validation of a single-item measure of burnout against the Maslach Burnout Inventory among physicians. Stress Health 20, 75–79 (2004)
17) Kirsh KL, Passik S, Holtsclaw E, Donaghy K and Theobald D: I get tired for no reason: a single item screening of cancer-related fatigue. J Pain Symptom Manag 22, 931–937 (2001)
18) Schwartz AL, Meek PM, Nail LM, Fargo J, Lundquist M, Donofrio M, Grainger M, Throckmorton T and Mateo M: Measurement of fatigue determining minimally important clinical differences. J Clin Epidemiol 55, 239–244 (2002)
19) American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington: American Educational Research Association, 1999.
20) McNair DM, Lorr M, Droppelman LF. Manual for the Profile of Mood States. San Diego (CA): Educational and Industrial Testing Service, 1971/1982/1992.
21) Wald FDM and Mellenbergh GJ: The shortened version of the Dutch translation of the Profile of Mood States (POMS). Nederlands Tijdschrift voor de Psychologie 45, 86–90 (1990) (in Dutch)
22) McNair DM, Heuchert JP, Shilony E. Profile of Mood States: Bibliography 1964–2002. North Tonawanda, NY (etc): Multi Health Systems Inc., 2003.
23) Boyle GJ: A cross-validation of the factor structure of the Profile of Mood States; were the factors correctly identified in the first instance? Psychol Rep 60, 343–354 (1987)
24) Jacobson AF, Weiss BL and Steinbook RM: The measurement of psychological states by use of factors derived from a combination of items from mood and symptom checklists. J Clin Psychol 34, 677–685 (1978)
25) Norcoss JC, Guadagnoli E and Prochaska JO: Factor structure of the Profile of Mood States (POMS): two partial replications. J Clin Psychol 40, 1270–1277 (1984)
26) Reddon JR, Marceau R and Holden RR: A confirmatory evaluation of the Profile of Mood States: convergent and discriminant item validity. J Psychopathol Behav 7, 243–259 (1985)
27) Wicherts JM and Vorst HCM: Model fit of the Dutch shortened Profile of Mood States and measurement invariance over gender. Nederlands Tijdschrift voor de

Pychologie 59, 12–21 (2004) (in Dutch)
28) Geurts SAE, Taris TW, Kompier MAJ, Dikkers JSE, Van Hooff MLM and Kinunnen UM: Work-home interaction from a work psychological perspective: development and validation of a new questionnaire. Work Stress 19, 319–339 (2005)
29) Van Veldhoven M, De Jonge J, Broersen S, Kompier M and Meijman T: Specific relationships between psychosocial job-conditions and job-related stress: a three level analytic approach. Work Stress 16, 207–228 (2002)
30) Michielsen HJ, De Vries J and Van Heck GL: Psychometric qualities of a brief self-rated fatigue measure: the Fatigue Assessment Scale. J Psychosom Res 54, 345–352 (2003)
31) Dirken JM. Work and Stress. Groningen: Wolters Noordhoff, 1969 (in Dutch).
32) Joosten J and Drop MJ: Reliability and comparability of three versions of the VOEG. Gezondheid en Samenleving 8, 251–265 (1987) (in Dutch)
33) Karasek RA. Job content instrument: Questionnaire and user's guide. Los Angeles: Dept. of Industrial and Systems Engineering, University of Southern California, 1985.
34) Geurts SAE, Rutte C and Peeters M: Antecedents and consequences of work-home interference among medical residents. Soc Sci Med 48, 1135–1148 (1999)
35) Van Mierlo H, Rutte CG, Seinen B and Kompier M: Autonomous teamwork and psychological well-being. Eur J Work Organ Psychol 10, 291–301 (2001)
36) Hox J. Multilevel analysis: Techniques and Applications. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.
37) Snijders TAB and Bosker RJ. Multilevel Analysis: An introduction to basic and advanced multilevel modeling. London: SAGE, 1999.
38) Centre for Multilevel Modelling. MLwiN 2.0. Bristol: University of Bristol, 2005.
39) Geurts SAE, Kompier MAJ, Roxburgh S and Houtman ILD: Does work-home interference mediate the relationship between workload and well-being? J Vocat Behav 63, 532–559 (2003)
40) Schwarz G: Estimating the dimension of a model. Annals of Statistics 6, 461–464 (1978)
41) Spector PE: Method variance in organizational research: truth or urban legend? Organizational Research Methods 9, 221–232 (2006)
42) Semmer NK, Grebner S and Elfering A. Beyond self-report: Using observational, physiological, and situation-based measures in research on occupational stress. In: Perrewé PL, Ganster DC, eds. Research in Occupational Stress and Well-being, Vol. 3. Amsterdam: JAI, 2004: 207–263.
43) Kompier M: Assessing the psychosocial work environment—"subjective" versus "objective" measurement (editorial). Scand J Work Environ Health 31, 405–408 (2005)
44) Sonnentag S: Burnout and HPA-axis functioning—There are no simple answers (editorial). Scand J Work Environ Health (in press)