Scandinavian Journal of Psychology, 2008, 49, 591–601 DOI: 10.1111/j.1467-9450.2008.00684.x

Health and Disability

Negative affect and 24-hour ambulatory physiological recordings as predictors of spontaneous improvement of medically unexplained symptoms

JAN H. HOUTVEEN and LORENZ J. P. VAN DOORNEN

Department of Clinical and Health Psychology, Utrecht University, Utrecht, The Netherlands

Houtveen, J. H. & van Doornen, L. J. P. (2008). Negative affect and 24-hour ambulatory physiological recordings as predictors of spontaneous improvement of medically unexplained symptoms. *Scandinavian Journal of Psychology*, *49*, 591–601.

The predictive value for spontaneous improvement in individuals suffering from medically unexplained symptoms (MUS) was explored of (1) anxiety and depression obtained from questionnaires, (2) negative affective states obtained from experience-sampling, and (3) ambulatory-assessed real-life physiological recordings. Sixty-seven individuals with MUS and 61 healthy controls were included. Twenty-four hour ambulatory recordings of cardiac autonomic activity, respiration, end-tidal CO_2 and saliva cortisol were combined with experience-sampling of somatic complaints and mood. Complaints were assessed again after one year. Although a reduction in symptoms (25%) was found, this could not be predicted from initial anxiety and depression. Improvement was somewhat related to relatively low diary reports of fatigue, especially in the late-afternoon and evening (3% variance explained). From the physiological measures only relatively high $PetCO_2$ values in the morning predicted improvement (5% explained). It was concluded that spontaneous recovery from MUS is hard to predict from self-reported distress and ambulatory physiological recordings.

Key words: Medically unexplained symptoms, prognostic factors, anxiety, ambulatory physiological recordings.

Jan Houtveen, Utrecht University, Faculty of Social and Behavioural Sciences, Clinical and Health Psychology, PO Box 80.140, 3508 TC Utrecht, The Netherlands. Tel: +31 (0)30 2539303; fax: +31 (0)30 2534718; e-mail: j.h.houtveen@uu.nl

INTRODUCTION

Medically unexplained or "functional somatic" syndromes (like chronic fatigue syndrome, multiple chemical sensitivity, irritable bowel syndrome, hyperventilation syndrome, etc.) are syndromes with heterogeneous somatic complaints without signs of organic pathology (Barsky & Borus, 1999; Costa Jr & McCrae, 1985; Da Costa, 1871; Watson & Pennebaker, 1989; Wessely, Nimnuan & Sharpe, 1999). These syndromes have a high prevalence, they share many symptoms and non-symptom characteristics (e.g., heterogeneous symptom reports) and treatments show only moderate success (Nimnuan, Hotopf & Wessely, 2001; Wessely *et al.*, 1999). Medically unexplained syndromes are a burdening problem in primary and secondary health care and more research is needed on the underlying mechanisms, effectiveness of treatments, and on the prognostic factors.

The starting point for the current prospective study on the prognostic factors of heterogeneous medically unexplained symptoms is the relation between reported somatic complaints and negative affect. Between-group differences in medically unexplained symptoms are closely tied to group differences in negative affective traits, like anxiety, depression and neuroticism (Barsky & Borus, 1999; Hotopf, 2002; Houtveen, Rietveld & De Geus, 2003; Pennebaker, 2000; Van Diest, De Peuter, Eertmans, Bogaerts, Victoir & Van Den Bergh, 2005; Watson & Pennebaker, 1989; Wessely *et al.*, 1999; Wientjes & Grossman, 1994). Experimentally induced negative affective states also increase self-reports of somatic complaints (Houtveen *et al.*, 2003; Wientjes & Grossman, 1994). The relation between negative affect and medically unexplained symptoms is supported by recent functional anatomical work indicating that bodily symptoms result from a synergistic interaction between higher order psychological factors and a lower order sensitization of the brain to bodily signals (Craig, 2002; Craig, 2003). Self-reported medically unexplained bodily symptoms are hypothesized to result from sensitization of the brain to interoceptive signals, which is modulated by negative-affect (Price, Zhou, Moshiree, Robinson & Verne, 2006; Verne, Himes, Robinson *et al.*, 2003; Verne, Robinson & Price, 2004).

Negative affect also has a predictive value for new-onset medically unexplained symptoms. Leiknes, Finset, Moum and Sandanger (2007) studied the course (1990 to 2001) and psychosocial predictors of medically unexplained pain symptoms in the general population. Stressful life events, anxiety and depression assessed in 1990 predicted medically unexplained pain in 2001. Surprisingly, initial report of pain was not related to future complaints. It remains, however, unclear whether negative affect also has a prognostic value for (impaired) *recovery* of medically unexplained symptoms. Rimes *et al.* studied new-onset and recovery of fatigue with time (4–6 months) in adolescents (Rimes, Goodman, Hotopf, Wessely, Meltzer & Chalder 2007). New cases of fatigue, but not recovery from it, were predicted by initial anxiety or depression disorder. Schmaling *et al.* studied prognostic factors over a 1.5-year period in patients suffering

from unexplained chronic fatigue (Schmaling, Fiedelak, Katon, Bader & Buchwald, 2003). Depression was not significantly related to a change in symptom severity over time, whereas mental health, antidepressant use, sedating medication use and somatic attributions were. Van der Werf *et al.* studied the natural course of chronic fatigue syndrome complaints (Van Der Werf, De Vree, Alberts, Van Der Meer & Bleijenberg, 2002). Persistence of complaints after one year follow-up was associated with illness duration, initial concentration problems, fatigue, and less psychosocial causal explanation. Again, no effect was found for psychological distress (anxiety and depression). Based on the studies mentioned above, it may be concluded that negative affect seems to predict new-onset unexplained symptoms, but its prognostic value for recovery from complaints remains to be established. The current study is focused on the prognostic value of negative affect on the natural course (i.e., spontaneous improvement) of medically unexplained symptoms using an improved methodology.

Individual differences in somatic complaints and negative affective traits (e.g., anxiety and depression) are traditionally assessed using retrospective questionnaires. However, recalled reports of symptoms and feelings can be biased by memory heuristics (Houtveen & Oei, 2007; Robinson & Clore, 2002). Retrospective reports of symptoms may for example be biased by identity-related beliefs leading to increased symptom reports with a longer recall period (Houtveen & Oei, 2007). It has been demonstrated that the relationship between negative affective traits and bodily symptoms is lower when participants rated their symptoms over a relatively short period (e.g., one day) than when they rated symptoms retrospectively over a longer period (e.g., weeks) (Larsen, 1992; Watson & Pennebaker, 1989). It has, therefore, been suggested that a negative affective trait like neuroticism may be associated with a tendency to over-recall symptoms, i.e., to remember them as being worse than they really were (Larsen, 1992). New methods to obtain measures of momentary-experienced affect and somatic complaints (i.e., assessed without recall bias) are experience-sampling and event-sampling (Bolger, Davis & Rafaeli, 2003; Stone & Shiffman, 2002). Brown and Moskowitz (2007) tested the effects of the personality trait neuroticism using traditional questionnaires versus concurrent affective state using a daily event-sampling design on momentary-assessed common physical symptoms. They showed that although neuroticism was correlated with momentary-assessed unpleasant affect, only unpleasant affective state influenced the report of physical symptoms, with no contribution from the trait neuroticism. Thus, retrospectively-assessed negative affective traits and momentary-assessed unpleasant affective states are two different aspects of emotional self-report. It may be speculated that neuroticism as a retrospectively assessed negative affective trait may be more associated with retrospective reports of somatic symptoms, whereas momentary-assessed negative affect rather may be associated with momentary-assessed symptoms. In the current study we will compare momentary reports of negative affective states and retrospectively assessed negative affective traits with respect to their predictive value for spontaneous improvement of medically unexplained symptoms.

Subjective experience of stress and negative emotional states clearly go together with changes in stress-physiological parameters related to the autonomic nervous system, respiration and cortisol (Allen & Crowell, 1989; Berntson, Cacioppo & Quigley, 1993; Berntson, Cacioppo, Quigley & Fabro, 1994; Boiten, Frijda & Wientjes, 1994; Grossman, 1983; Han, Schepers, Stegen, Van Den Bergh & Van De Woestijne, 2000; Kamphuis & Frowein, 1985; Kirschbaum & Hellhammer, 1989; Kirschbaum & Hellhammer, 1994; Langewitz & Ruddel, 1989; Ley & Yelich, 1998; Sherwood, Allen, Obrist & Langer, 1986; Suess, Alexander & Smith, 1980). However, self-reported negative affect and stress-physiological indicators have been demonstrated to show relatively low concordance on between-subject basis (Lang, 1994; Mauss, McCarter, Levenson, Wilhelm & Gross, 2005). Thus, the verbal report and the physiological signs of emotions and affect seem to be different representations of the same (central) state. This implicates that subjective report of emotional states and their peripheral stress-physiological expression confer relatively independent information. In the context of the current study we will explore whether peripheral physiological activity of three important stress-reactivity systems (autonomic, respiratory and cortisol) have an additional value in predicting spontaneous improvement of medically unexplained symptoms over and above self-reported psychological predictors. The physiological recordings observed in the current study are considered to be independent markers (in addition to self-reports) of previous and/or ongoing stress and negative affect.

A non-clinical group of individuals high on heterogeneous medically unexplained symptoms was examined for 24 hours using electronic diaries and real-life ambulatory physiological measurement devices. Their complaints were assessed again after a year. The predictive value of: (1) retrospectively reported anxiety and depression, (2) mean momentary-reported negative affective states obtained from diary assessment, and (3) the 24-hour stress-physiological indices obtained from real-life ambulatory monitoring was explored on the spontaneous improvement of complaints. As stress-physiological indices we opted for the activity of the cardiac autonomic (sympathetic and parasympathetic) nervous system, the respiratory system, and the hypothalamic-pituitary-adrenal (HPA) axis (i.e., cortisol). A control group low on symptoms was included as a reference group to allow comparison of the measures with a "normal" set-point.

METHOD

Participants

A non-clinical group of participants (n = 74) with heterogeneous somatic symptoms was recruited who, despite a history of medical tests, never received a medical diagnosis for their complaints. They were recruited by linking national websites on medically unexplained syndromes (e.g., hyperventilation syndrome, chronic fatigue syndrome, clinical burnout syndrome) to research information provided through the university website (www.fss.uu.nl\gp\stressprofielen). People who reacted (by electronic mail) received additional information and the inclusion questionnaires. Participants with complaints were included on the basis of high scores on the somatic subscale of the

Table 1. *Participant information*

	HSS-all[1] (n = 67)		LSS-LL (n = 51)		HSS-HI (n = 33)		HSS-LI (n = 34)		Sign.
Male (%)	22		27		18		26		n.s.
Smoking (%)	24		27		27		21		n.s.
Engaged in sports (%)	61		70		67		56		n.s.
Employed (%)	61		90		61		62		<0.001[a,b]
Social benefits (%)[2]	21		0		18		24		<0.001[a,b]
On sick leave (%)[3]	31		0		33		29		<0.001[a,b]
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Sign.
Age (years)	37.27	8.34	34.55	8.72	38.33	8.94	36.24	7.70	n.s.
BMI (kg/m^2)	23.87	4.61	22.70	2.30	23.61	3.96	24.12	5.21	n.s.
Screening SCL90R Som	34.85	5.92	13.98	1.43	36.94	5.91	32.82	5.27	<0.001[a,b,c]
Screening SCL90R Anx	23.22	7.67	11.16	1.35	24.55	8.44	21.94	6.71	<0.001[a,b]
Screening SCL90R Dep	37.40	9.88	18.25	2.09	39.12	9.04	35.74	10.50	<0.001[a,b]
T3-T1 (weeks)	57.6	14.7	68.49	24.69	54.45	13.56	60.94	15.30	<0.01[b]
T3-T2 (weeks)	53.1	14.2	66.49	21.91	50.50	13.48	55.94	14.61	<0.001[a,b]
Follow-up SCL90R Som	26.09	8.32	13.88	1.64	22.00	6.44	30.06	8.08	<0.001[a,b,c]
Follow-up SCL90R Anx	18.19	7.38	11.22	1.72	15.82	5.01	20.50	8.57	<0.001[a,b,c]
Follow-up SCL90R Dep	29.43	11.76	19.33	4.58	25.51	7.13	33.24	14.03	<0.001[a,b,c]
Change SCL90R Som	–8.93	1.66	–0.10	1.81	–14.94	3.81	–2.91	4.49	<0.001[a,b,c]
Bed time (hh:mm)	23:27	:58	23:41	:59	23:30	:60	23:26	:59	n.s.
Time awakening (hh:mm)	7:48	1:06	7:37	:53	7:47	1:09	7:49	1:05	n.s.
Hours of sleep	8:20	1:14	7:56	1:07	8:18	1:13	8:23	1:17	n.s.

Notes: [1] Used for the regression analysis. [2] Receive benefits from social security administration, unemployed because of chronic (mental) disability as a result of medical unexplained symptoms. [3] Employed but on sick leave whilst being assessed; HSS-all = all participants high on somatic symptoms; LSS-LL = low on somatic symptoms both at screening and follow-up; HSS-HI = high on somatic symptoms high improvement; HSS-LI = high on somatic symptoms low improvement; T1 = date of screening; T2 = date of assessment; T3 = date of follow-up; n.s. = not significant; [abc] significant post-hoc tests: [a] = LL-HI; [b] = LL-LI; [c] = HI-LI.

Symptom Check List (SCL-90-R; see appendix) in combination with (self reports of) absence of a medically diagnosed physical disease (see below). The control participants (n = 71) were recruited by the internet, advertisements in local papers, or were nearby volunteers. See Table 1 for further participant information. We used the one-year follow-up information from 67 participants high on symptoms (15 male, 52 female) and 61 control participants (16 male, 45 female). For the participants included, no intervention or treatment took place between assessments.

The somatic complaints, anxiety and depression subscales of the SCL-90-R were used for the selection of participants. The cut-off point used for the somatic score for the group high on symptoms (i.e., ≥27) is the median of the norm-distribution of Dutch psychiatric outpatients and the 95th percentile of the upper part of the norm-distribution of Dutch normal controls (Arrindell & Ettema, 1986). Participants high on symptoms were allowed to have "psychiatric outpatients above average" anxiety scores (≤33) and depression scores (≤55), while high and very high scorers (i.e., above respectively 33 and 55) were excluded. Further inclusion criteria for the group high on symptoms were age 25–50, a self-report of having consulted one or more physicians for their somatic complaints, and a self-report of never having got a physical diagnosis from a physician for their somatic complaints. With these criteria, we aimed to include only candidates with heterogeneous somatic complaints without organic cause, and without a primary anxiety or mood disorder.

The SCL-90-R cut-off points used for the inclusion of controls were average or below average scores on the somatic subscale (≤17), anxiety (≤14) and depression (≤23). These cut-off points were based on the norm-distribution of Dutch normal controls. Further inclusion criteria for the control group were age 25–50 and a self-report of not having a physical disease.

Exclusion criteria for all participants were the use of medication known to affect our physiological measures (e.g., corticosteroids, anti-inflammatory agents, anti-depressives, beta-blockers), a medically diagnosed physical disease, a body-mass index (BMI) > 30, and excessive use of alcohol or drugs. See Table 1 for further participant information.

Participants high on symptoms received a small present afterwards; control participants low on symptoms received a small monetary reward (10 Euro). All participants received an annotated summary of their recordings. The Ethics Committee of the University Medical Centrum Utrecht (UMCU) approved the study protocol and all subjects gave written consent before entering the study.

Inclusion and follow-up questionnaires

A Dutch translation of the SCL-90-R was used (Arrindell & Ettema, 1981). Distress ratings of a list of items experienced last week could be rated on a five-point Likert scale, ranging from 1 "*not at all*" to 5 "*extremely*". The subscales used were somatic complaints (n = 12 items, reliability α = 0.86) anxiety (n = 10 items, α = 0.88) and depression (n = 16 items, α = 0.90). The Somatic subscale was the main inclusion and outcome measure.

A second questionnaire was created to collect demographic information (age, height, weight, work status, regularly engaged in sports, etc.) and self-reported health status. Recent experienced illness, medical history

and previous and current medication use had to be reported and a checklist had to be filled out on 18 medical problems (e.g., pain, inflammations, heart problems, asthma, etc.). Questions on each medical problem had to be completed regarding experiences in the last 6 months: whether a physician was ever seen for this complaint, a medical diagnosis was ever received, or a treatment (e.g., medication) for this complaint was prescribed.

Ambulatory measurement devices

The ambulatory electro cardiogram (ECG) and impedance cardiogram (ICG) were measured from a six Ag/AgCl electrode configuration using the VU-AMS (version 4.3, TD-FPP, Vrije Universiteit, Amsterdam, The Netherlands; www.psy.vu.nl/vu-ams). Electrode resistance was kept low by cleaning the skin with alcohol and rubbing. Vertical acceleration of the torso – integrated over 30-second periods – was additionally monitored and stored throughout the 24-hour recording time. This was used as a proxy for gross body movement (motility). By reason of memory limitations, continuous registration was set to 5 minutes per 15 minutes period, resulting in four data segments per hour. Details on the recording methodology, reliability and validity of the VU-AMS can be found elsewhere (De Geus & Van Doornen, 1996; De Geus, Willemsen, Klaver & Van Doornen, 1995; Houtveen, Groot & De Geus, 2006; Riese, Groot, Van den Berg *et al.*, 2003; Willemsen, De Geus, Klaver, Van Doornen & Carroll, 1996).

Palm™ M130 (www.PalmOne.com) Personal Digital Assistants (PDA) were used for (electronic diary) experience-sampling of symptoms and mood during the measurement day. Special software was developed for the generation of alarms (a 10-second auditory signal) during the day, and for the assessments of somatic complaints and mood. A fixed sampling protocol with an interval of 1.5 hours (i.e., without a random factor) was used that continued from awakening till bedtime. Alarms without response were repeated (max 3 times with 3 minutes time interval) resulting in a sampling interval of 90 ± 12 minutes. This sampling scheme leads to approximately 11 alarms per day (based on a 16-hour awakening period). Diary prompting was only disabled during sleep, initiated by a button on the PDA. Next, the PDA could be used as a morning alarm. Prompting continued after awakening. All unused buttons were blocked. The alarm software generated a log-file containing alarm and response times to be used for determination of the compliance. The questionnaire was launched by a start button that was visible for 12 minutes after prompting (the time-period during which repetitions of alarms were given). All questions were forced-choice, and they were displayed as sequential screens on the PDA. Participants were not allowed to leaf through the present or previous diaries. The self-reported dimensions measured on the PDA were: activity and posture, somatic complaints (measured by the 12 somatic complaint items used for selection based on the somatic complaints subscale of the SCL-90-R; see appendix) and mood (depression, vitality, anger, fatigue and tension; each measured by 3 items adapted from the shortened version of the POMS (Shacham, 1983); see appendix). Each item could be rated on a seven-point scale ranging from 1 "*not at all*" to 7 "*very much*".

The TG-951T CO_2 quantitative sensor Kit (Nihon Kohden Corporation, Tokyo, Japan) was used to measure the partial pressure of CO_2 at the end of a normal expiration ($PetCO_2$). $PetCO_2$ was measured during the day using a mainstream adapter (based on four respiratory cycles assessed once every 1.5 hours automatically initiated just after completing the diary questions) and during sleep using an air pump device in combination with a sidestream (nostrils) adapter (based on ten respiratory cycles assessed once every 15 minutes; we anticipated more loss of data as a result of mouth breathing during sleep as compared to day-time assessment). Digitized $PetCO_2$ values (in mmHg) from the capnometer were sent to the PDA computer for storage through an RS-232 interface cable.

Saliva was collected by salivettes, plastic tubes with cotton roles (Sarstedt, Etten-Leur, the Netherlands). Saliva was collected at 0, 15, 30 minutes after awakening for the cortisol awakening response and at noon, 6 pm and 11 pm (or sleeping time) for the day-curve. Saliva collection times were prompted by the PDA. Using the PDA allowed us to assess cortisol-influencing parameters as food intake, smoking, teeth brushing and medication. The cortisol alarm was repeated after recent food intake, smoking, or teeth brushing (i.e., to postpone saliva collection). The samples were kept in the refrigerator after collection and (after being recollected 24-hours later) stored at −20ºC. Samples were analyzed in a lab in Dusseldorf (Germany), by a time-resolved immunoassay with fluorescence detection as described elsewhere (Dressendorfer, Kirschbaum, Rohde, Stahl & Strasburger, 1992).

Procedure

Participants who met the inclusion criteria and who agreed to participate were contacted by telephone to receive more information on the study and to make an appointment for the 24-hour ambulatory recordings. They were visited at home where they completed additional questionnaires. Next, they were supplied with the VU-AMS, the PDA, the capnometer, and the salivettes. They received detailed verbal and written instructions how to use all equipment. Participants practiced how to complete the questions on the PDA. Next, they were instructed on the CO_2 assessment. Specific instructions were given how to breath as normal as possible during mainstream CO_2 assessments, which was also practiced under supervision. They got further instructions on how to regularly check the "all clear" signal of the VU-AMS device (a small blinking light on the side of the device) and how to respond to measurement alarms (e.g., a loose electrode contact) and electronic diary promptings. Twenty-four hour telephone assistance was available. Participants followed their normal day (e.g., working) routines. The following day, the researcher came back to collect the devices.

After a year they received the follow-up questionnaires (by e-mail or by regular mail). The follow-up questionnaires were similar to the questionnaires used for inclusion. When participants forgot to complete and post them, the request to complete these questionnaires was repeated after three weeks by e-mail, and again after two months by telephone. Participants who sent back the questionnaires received a small present.

Physiological data analyses

The recorded heart period time series (inter-beat intervals, IBI) together with the motility signal were displayed as a cardiotachogram for visual inspection by two independent raters. The (5-minute) segments containing too many artifacts (i.e., >5) were reduced in length (sub-segments low in artefacts should be at least 2.5 minutes) or rejected (no sub-segments could be selected). Because individual differences in motility could be responsible for group differences in the physiological measures, only segments (or sub-segments 2.5–5 minutes) low on motility were selected. "Low on motility" was based on the mean participant-specific motility value obtained during sleep. In 89% of the cases a (sub)segment low in motility could be found. Take note that low motility (sub)segments were only found during sleep (by definition) and (while awake) during relaxation in sitting or supine position. These (sub)segments were used for segment-specific analysis of the IBI and thoracic impedance (ICG) signals.

An artifact pre-processing was performed on the selected IBI data. Artifacts were detected automatically when a value in a particular segment deviated a specific number of *SD*s from the moving mean (e.g., value < M–3*SD* or value > M+3*SD*), or when a value in a particular segment was smaller or larger than a specific percentage of this mean (e.g., value < 0.5 M or value > 2 M). Corrections were verified by visual inspection and the default settings (i.e., 3, 0.5, 2) were adjusted for each participant. IBI artifacts were found and corrected for 0.07% (*SD* = 0.06) of the IBI values. Since artifacts cannot simply be deleted, because the continuity of time would be lost, spuriously short IBIs were

summed and missing beats were "created" by splitting spuriously long IBIs. The IBI mean values were computed from these corrected data for each segment.

Next, uniformly spaced samples were created by interpolation of the corrected IBI data using a Wavelet interpolation algorithm. Discrete Wavelet Transformation (DWT) was performed using a cardinal cubic spline function as base. This method results in identical power values for stationary relatively short data segments as compared to Fourier transformation (Houtveen & Molenaar, 2001), but it is superior for non-stationary data segments. Since the DWT (like Fourier) suffers from aliasing effects at both ends, the first and last 2.5 seconds of the time series were excluded from the derivation of the variances. The High Frequency (HF) power (the variance in the 0.125–0.5 Hz respiratory frequency-band window) was computed as main measure of respiratory sinus arrhythmia (RSA), which indicates cardiac parasympathetic (vagal) control (Berntson, Thomas Bigger Jr, Eckberg *et al.*, 1997).

Changes in the thoracic impedance (dZ) were used to assess respiration. Respiratory-related changes in dZ were obtained by band-pass filtering (0.125–0.5 Hz) of the ICG signal using a DWT filter with a cubic spline function as base. Next, the respiratory power values were computed as the variances of the filtered time series for each (selected) data segment. Changes in the respiratory power values were used as a (raw) estimation of changes in respiratory depth as described elsewhere (Houtveen *et al.*, 2006). Respiratory frequencies were obtained from the band-pass filtered thoracic impedance (dZ) signal by counting the number of up-going zero crossings and dividing this value by the time of a segment as described elsewhere (De Geus *et al.*, 1995; Houtveen *et al.*, 2006).

Pre-ejection period was manually scored using the VU-AMS interactive software which graphically displays the large-scale ensemble averages (i.e., averaged per segment). Pre-ejection period reflects the time interval between the onset of the electromechanical systole (Q-wave onset) in the ECG and the onset of left ventricular ejection at the opening of the aortic valves (B-point) in the ICG. The B-points were manually determined for each ensemble averaged (selected) segment, and the pre-ejection period values were determined by summing a fixed Q-to-R interval of 48 ms to the R-B interval time. Pre-ejection period was computed as a measure of cardiac sympathetic control (inotropic control over cardiac contractility) (Sherwood, Allen, Fahrenberg, Kelsey, Lovallo & Van Doornen, 1990).

$PetCO_2$ values were assessed with 1.5 hour intervals during daytime (based on four respiratory cycles) and 15 minutes intervals during sleep (based on ten respiratory cycles). The largest value of the respiratory cycles (four during the day; ten during the night) was taken as the most reliable value (i.e., based on the idea that an end-tidal value could not exceed the true arterial partial pressure of CO_2). Visual inspection was performed on the sleep data to reject incidental low values due to mouth breathing. Suspicious values were evaluated in combination with respiratory depth and frequency.

Statistical analyses

Electronic diary data were mean-aggregated over four 4-hour time periods (8–12 hr, 12–16 hr, 16–20 hr and 20–24 hr); physiological data (except for cortisol) were mean-aggregated over six 4-hour time periods (including 0–4 hr and 4–8 hr). The electronic diary data, IBI, HF heart period variability power, respiratory power, pre-ejection period and cortisol values were ^{10}log transformed to obtain normal distributions. A change in self-reported somatic complaints score was computed for each participant by subtracting the SCL-90-R somatic subscale score obtained during screening from the one-year follow-up score.

Linear regression analyses (SPSS 14) were performed to predict one-year change in complaints from questionnaire data, electronic diary data and 24-hour physiology. Because of multicollinearity (i.e., high correlations between the predictors) the within-subject time-of-day information could not be used in the regression analyses: the repeated measures were therefore aggregated over the day (explained in more detail in the results section). The "Stepwise" variable selection method (P-entry = 0.05; P-removal = 0.10) was used to predict one-year change in somatic complaints by: (model I) gender, age and BMI, (model II) SCL-90-R screening levels of self-reported somatic complaints, (model III) the SCL-90-R screening levels of self-reported anxiety and depression (i.e., our trait measures of negative affect), (model IV) the day-means of momentary reported complaints obtained from experience-sampling, (model V) the day-means of momentary reported negative affective states (mood) obtained from experience-sampling, and (model VI) the day-means of the real-life ambulatory physiological measures.

Next, to test for specific interactions with time-of-day (i.e., 6 periods for the physiological measures; 4 periods for the diary measures), three subgroups were created for further repeated-measures analysis of variance: (1) control participants without complaints that scored below the inclusion threshold on the SCL-90-R somatic subscale for controls both during screening and follow-up, (2) participants high on improvement who scored high on the somatic subscale during screening but relatively low at follow-up, and (3) participants low on improvement who scored high both during screening and follow-up. The median value of the change score (i.e., 10 points improvement on the SCL-90-R subscale somatic complaints) was used to separate the subgroups with high versus low improvement. Repeated-measures Group (3) by Time-Period (4 or 6) MANOVA tests (using SPSS 14) were performed for all measures. Age and BMI were not related to group and could, therefore, be used in an analysis of covariance to improve the power of the tests (see Miller & Chapman, 2001). Finally, repeated-measures tests comparing only the HSS-HI and HSS-LI participants including the SCL-90-R screening scores of somatic complaints as a covariate were performed for all measures. *P*-values below 0.05 were considered significant.

RESULTS

Improvement of somatic complaints

The mean (and *SD*) scores on the SCL-90-R subscales for somatic complaints, anxiety and depression obtained during screening and after one year are shown in Table 1. For the participants high on somatic symptoms (HSS-all, left column), the mean somatic complaints score was significantly reduced (on average 8.75 points; 25% improvement) after a year ($t(66) = 9.57$, $p < 0.001$). The correlations between the one-year change in the SCL-90-R subscale somatic complaints (follow-up – screening) and the potential predictors for HSS-all participants are shown in Table 2. Because a table containing so many correlations may produce Type I errors, significancies in this table should be interpreted with care. More reliable are the results from the overall regression and repeated measures analysis described below.

Regression analysis

The model I regression analysis to predict the one-year change in complaints from gender, age and BMI was not significant ($F(3, 63) = 0.93$, $p = 0.43$). Model II, predicting one-year change in complaints from the screening level of SCL-90-R somatic complaints was significant ($F(1, 65) = 4.99$, $p < 0.05$, adjusted $R^2 = 0.06$): a larger improvement was related to higher initial somatic complaints scores. This model was used as the base model in the further regression analyses. Model III, adding the

Table 2. *Correlations between one-year change in the SCL-90 subscale somatic complaints (follow-up – screening) and the potential predictors for the participants high on somatic symptoms (*n = *67)*

Screening SCL90 Som	–0.29*	Fatigue 16–20	0.20	PEP 0–4	0.00
Screening SCL90 Anx	–0.20	Fatigue 20–24	0.26*	PEP 4–8	–0.12
Screening SCL90 Dep	–0.08	SCL90 Som 8–12	0.12	PEP 8–12	–0.08
Bed time	0.04	SCL90 Som 12–16	0.18	PEP 12–16	–0.08
Get up time	0.19	SCL90 Som 16–20	0.12	PEP 16–20	–0.02
Hours of sleep	0.15	SCL90 Som 20–24	0.08	PEP 20–24	0.04
Depression 8–12	0.25	Cortisol awake	0.14	$PetCO_2$ 0–4	–0.17
Depression 12–16	0.21	Cortisol +15 min	0.15	$PetCO_2$ 4–8	–0.13
Depression 16–20	0.20	Cortisol +30 min	0.14	$PetCO_2$ 8–12	–0.26*
Depression 20–24	0.28*	Cortisol 12:00	0.15	$PetCO_2$ 12–16	–0.17
Vitality 8–12	–0.07	Cortisol 18:00	–0.01	$PetCO_2$ 16–20	–0.11
Vitality 12–16	–0.17	Cortisol bedtime	0.08	$PetCO_2$ 20–24	–0.21
Vitality 16–20	–0.16	IBI 0–4	–0.11	R-freq 0–4	0.14
Vitality 20–24	–0.17	IBI 4–8	–0.12	R-freq 4–8	0.14
Anger 8–12	0.11	IBI 8–12	–0.09	R-freq 8–12	0.14
Anger 12–16	0.18	IBI 12–16	–0.10	R-freq 12–16	0.03
Anger 16–20	0.04	IBI 16–20	–0.12	R-freq 16–20	0.03
Anger 20–14	0.13	IBI 20-24	–0.17	R-freq 20–24	0.06
Tension 8–12	0.11	HF-pow 0–4	0.18	R-pow 0–4	–0.21
Tension 12–16	0.20	HF-pow 4–8	0.20	R-pow 4–8	–0.15
Tension 16–20	0.07	HF-pow 8–12	0.26*	R-pow 8–12	–0.07
Tension 20–24	0.14	HF-pow 12–16	0.22	R-pow 12–16	–0.21
Fatigue 8–12	0.10	HF-pow 16–20	0.24	R-pow 16–20	–0.26*
Fatigue 12–16	0.20	HF-pow 20–24	0.23	R-pow 20–24	–0.34**

Notes: Values were ^{10}log transformed when needed to obtain a normal distribution. The computed correlations are partial correlations with age and BMI as control variables. IBI = inter beat interval (ms); HF-pow = heart period variability high frequency power; PEP = pre-ejection period; cortisol = saliva cortisol; $PetCO_2$ = end-tidal partial pressure of CO_2; R-freq = respiratory frequency; R-pow = respiratory power; * $p < 0.05$; ** $p < 0.01$.

Table 3. *Results of the principal component analysis and significances of the components to predict one-year change in complaints*

Diary components	Day-time (8–24 hr) means of	Variance	Sign.
Complaints and fatigue	somatic complaints, fatigue	12.9%	<0.10
Negative affective state	anger, tension, depression	16.6%	<0.10
Vitality	vitality	3.5%	n.s.
Physiological components	**Means of**	**Variance**	**Sign.**
Mean PEP	24 hr pre-ejection period	9.1%	n.s.
Mean HF-power	24 hr cardiac HF-power	7.7%	n.s.
Mean $PetCO_2$	24 hr $PetCO_2$	6.4%	n.s.
Mean IBI	24 hr inter-beat interval	5.4%	n.s.
Mean R-power	24 hr Respiratory power	5.1%	n.s.
Mean R-frequency	24 hr Respiratory frequency	4.6%	n.s.
Mean cortisol awakening	0, 15, 30 minutes	4.1%	n.s.
Mean daytime cortisol	noon, 6 pm, 11 pm	3.5%	n.s.

Notes: variance = percentage of variance explained in the principal component analysis; n.s. = not significant.

screening levels of SCL-90-R anxiety and depression, did not add a significant predictive value to the base model ($t_{anxiety} = -0.17$, $p > 0.10$; $t_{depression} = -0.29$, $p > 0.10$). Note that removing the screening level of SCL-90-R somatic complaints from this model did not change these results.

Before the subsequent regression analyses could be performed, all additional potential predictors were first tested for multicollinearity. Unacceptable high measures of collinearity (i.e., VIF values > 10) were found for the electronic diary and physiological measures. Principal component analysis (Varimax rotation with Kaiser normalization) showed the existence of 11 components explaining 79% of the variance from all these predictors (see Table 3). These components were step-wisely added to the base model to predict one-year change in complaints. Near significant predictions were found for model IV, adding day-mean momentary-reported "complaints and fatigue" (obtained from experience-sampling) to the base model ($t_{complaints\ and\ fatigue} = 1.75$, $p = 0.08$, adjusted R^2 change = 0.03), and for model V,

adding day-mean momentary-reported "negative affective state" (obtained from experience-sampling) ($t_{\text{negative affective state}} = 1.69$, $p = 0.06$, adjusted R^2 change = 0.03). No predictive value was found for day-mean "vitality" ($t_{\text{vitality}} = -1.09$, $p > 0.10$). Adding the six 24-hr means of the real-life ambulatory physiological measures and mean "cortisol awakening" and "cortisol day level" to the base model (model VI), did not significantly contribute to the prediction (all p's > 0.10).

Interactions with time-of-day

Table 2 suggest that some time-specific momentary-reported and physiological measures were significantly correlated with change in complaints. Because of multicollinearity, time-specific predictive values of these measures could not be tested with the regression analyses described above. Therefore, additional repeated measure analysis of variance tests were performed to test for interactions with time. To do so, the 67 participants high on complaints were divided into two subgroups with relatively high improvement (≥10 points; HSS-HI; $n = 33$) versus low improvement (<10 points; HSS-LI; $n = 34$) on the SCL-90-R subscale somatic complaints. Median duration of complaints was 78 months for the HSS-LI and 48 months for the HSS-HI subgroup; this difference was, however, not significant ($\chi^2(1) = 0.75$, $p > 0.10$). A subgroup of control participants low on symptoms who scored below the selection threshold both at screening and at follow-up was further selected (LSS-LL; $n = 51$). See Table 1 for participant information of the so-obtained three subgroups. The three subgroups were not different on demographics like gender, age, BMI, sports behavior, and smoking (p's > 0.05).

Group differences on the SCL-90-R screening measures. A significant difference was found between HSS-HI and HSS-LI participants on the mean SCL-90-R somatic subscale score used for screening ($t(65) = 3.01$, $p < 0.01$): the HSS-HI participants had a higher somatic complaints screening score than the HSS-LI participants. This is in line with the model II regression showing that the worse cases improved most. Note that their mean follow-up score was much lower as compared to the HSS-LI participants (see Fig. 1). In line with the results obtained from the model III regression analysis, the HSS-HI and HSS-LI participants did not differ on their SCL-90-R screening scores of anxiety and depression (p's > 0.05; see Table 1), demonstrating that initially retrospectively-reported anxiety and depression scores do not predict one-year improvement of somatic complaints.

Repeated measure analyses on the diary measures. The mean number of diary (and day-time $PetCO_2$) responses during a day was 10.70 (HSS-HI: $M = 10.76$, range 10–11; HSS-LI: $M = 10.79$, range 10–11; LSS-LL: $M = 10.61$, range 9–11). No significant group differences in compliance were found ($F(2, 117) = 1.52$, $p = 0.22$). The Group by Time-period means of momentary-reported somatic complaints and mood are depicted in Fig. 2. Main effects of Group were found for momentary somatic complaints ($F(2, 112) = 106.36$, $p < 0.001$), depression ($F(2, 112) = 12.89$, $p < 0.001$), vitality ($F(2, 112) = 40.62$, $p < 0.001$), anger ($F(2, 112) = 10.18$, $p < 0.001$), fatigue ($F(2, 112) = 88.11$, $p < 0.001$) and tension ($F(2, 112) = 19.33$, $p < 0.001$). Post-hoc group comparisons showed that the LSS-LL group scored lower on the negative emotional states and somatic complaints and higher on vitality than both HSS (HI and LI) groups (all p's < 0.001).

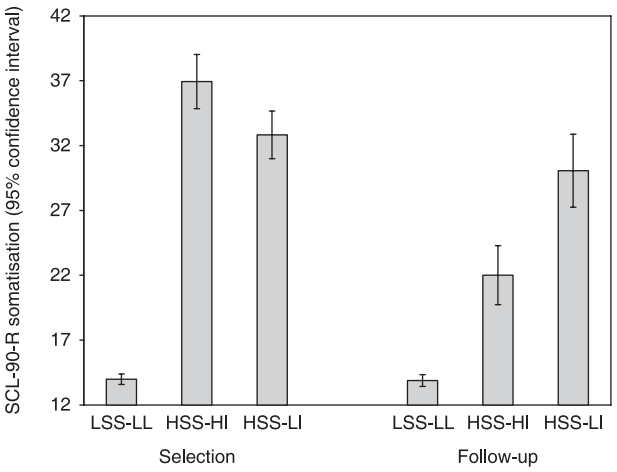

Fig. 1. Mean SCL-90-R somatic complaints subscale scores during screening and follow-up.

Post-hoc tests comparing the HI and LI groups yielded that despite the large difference between the HI and LI groups in retrospectively reported SCL-90-R somatic complaints (see above and illustrated in Fig. 1), a difference for momentary-reported somatic complaints was completely absent ($p = 0.49$). A near significant difference between the HI and LI groups was, however, found for momentary-reported depression and a significant difference was found for fatigue: the HSS-LI participants tended to report more momentary depression ($p = 0.056$) and they reported more fatigue ($p = 0.019$) as compared to the HSS-HI group. A significant Group by Time-period interaction effect was found for momentary fatigue ($F(6, 220) = 2.10$, $p < 0.05$): the HI-LI difference was specific for the time periods 16–20 and 20–24 hr (p's < 0.05). Thus, particularly relatively high levels of momentary-reported fatigue was found to be a predictor for less improvement after a year, and this effect was most pronounced for fatigue reported in the evening hours. Note that this specific effect could not be extracted from the regression analyses described above in which only a trend was found for the components "complaints and fatigue" and "negative affective state". Repeated measures test comparing only the HSS-HI and HSS-LI participants including the SCL-90-R screening scores of somatic complaints as a covariate yielded similar results. Thus the effects as reported are not attributable to the differences found at the moment of screening.

Repeated measure analyses on the physiological measures. The Group by Time-period means of the physiological measures are depicted in Fig. 3. No significant main effect of Group was found for cortisol, IBI, HF-power, PEP, $PetCO_2$, respiratory

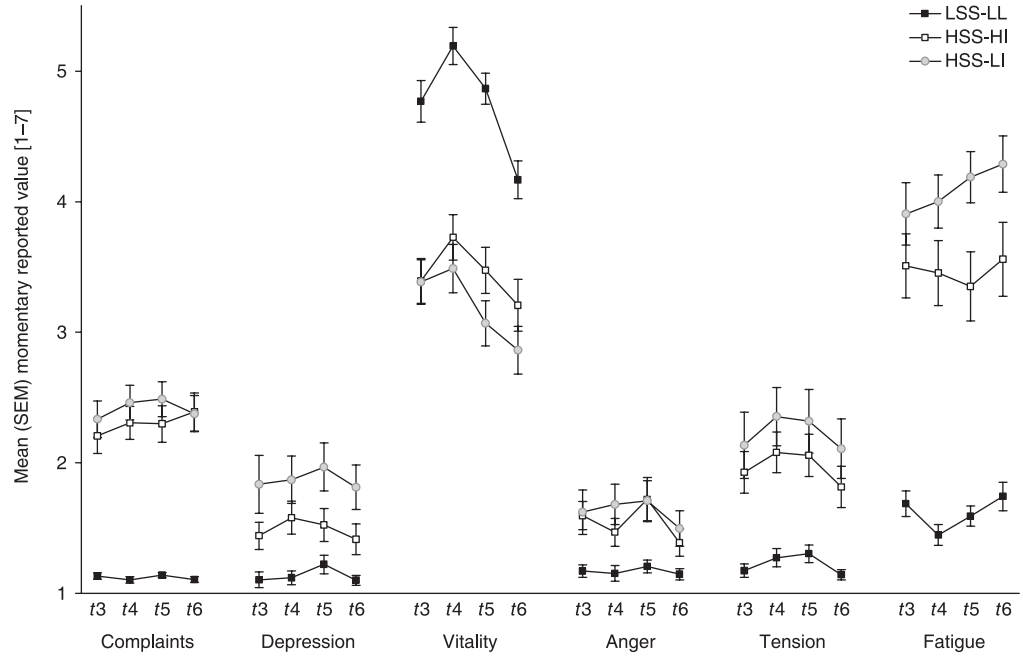

Fig. 2. Measures obtained from experience-sampling.
Notes: *t*3 = 8–12 am; *t*4 = 12–16 pm; *t*5 = 16–20 pm; *t*6 = 20–24 pm.

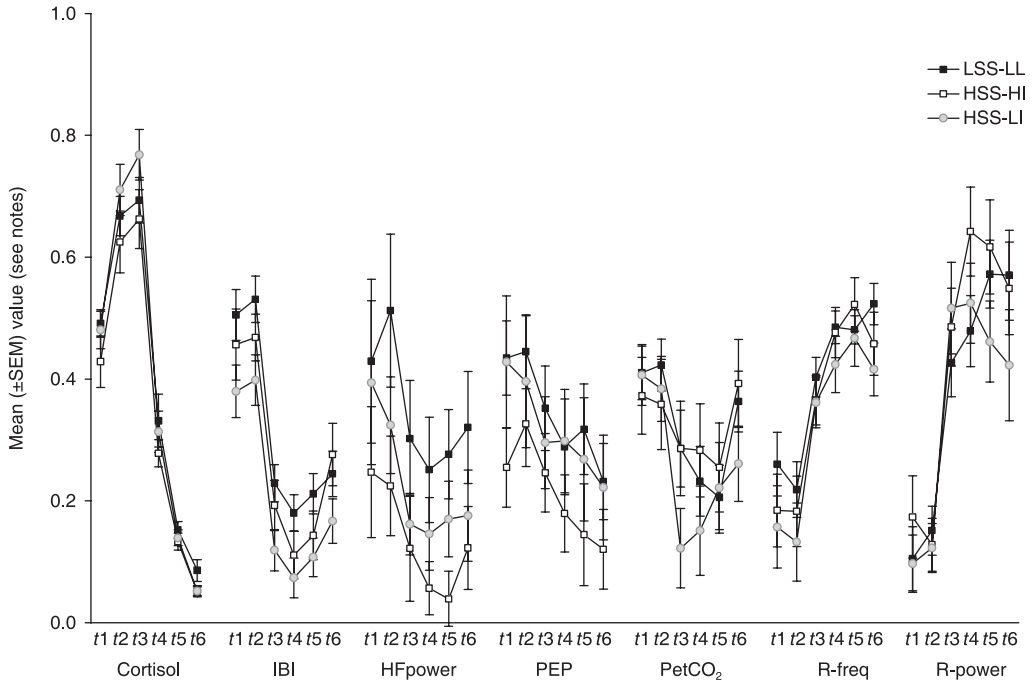

Fig. 3. Physiological measures.
Notes: *t*1 = 0–4 am; *t*2 = 4–8 am; *t*3 = 8–12 am; *t*4 = 12–16 pm; *t*5 = 16–20 pm; *t*6 = 20–24 pm; cortisol = value*30 (nmol/l); IBI = value*500 + 700 (ms); HFpower = value*1000 + 300 (ms^2); PEP = value*10 + 84 (ms); $PetCO_2$ = value*10 + 33 (mmHg); R-freq = value/10 + 0.25 (Hz); R-power = value/10000 + 3000 (arbitrary units).

frequency and respiratory power (all p's > 0.10). A Group by Time-period interaction effect was only found for respiratory power ($F(10, 218) = 2.81$, $p < 0.01$). However, post-hoc Tukey HSD tests did not yield specific time-period effects for respiratory power (all p's > 0.05). Repeated measures test comparing only the HSS-HI and HSS-LI participants including the SCL-90-R screening scores of somatic complaints as a covariate did not yield a significant effect for respiratory power anymore. Although

$PetCO_2$ did not show up as a main effect, a repeated measures test comparing only the HSS-HI and HSS-LI participants including the SCL-90-R screening scores of somatic complaints as covariate yielded a significant Group by Time-period interaction effect for $PetCO_2$ ($F(5, 53) = 2.87$, $p < 0.05$). Post-hoc tests indicated significantly lower $PetCO_2$ values in the morning (8–12 hr: $p < 0.05$) for the HSS-LI participants. Thus, of all physiological measures only the absence of lower $PetCO_2$ values in the morning was related to improvement after a year.

Explained variance

To get an impression of the variance that could be explained by the significant predictors found in the current study, an additional regression analysis was performed in which the base line model II (predicting one-year change in complaints by the screening level of SCL-90-R somatic complaints) was extended by adding fatigue in the evening and $PetCO_2$ in the morning. The adjusted R^2 of the model II prediction was 0.057. Adding fatigue in the evening increased the adjusted R^2 to 0.085 (i.e., additional 2.8% explained), while adding $PetCO_2$ in the morning increased the adjusted R^2 to 0.11 (i.e., additional 5.3% explained). A model adding both predictors to model II resulted in an adjusted R^2 of 0.133 (i.e., additional 7.6% explained).

DISCUSSION

Spontaneous improvement of complaints was observed in a non-clinical group of participants suffering from heterogeneous somatic symptoms that could not medically be explained. Although no intervention or treatment (e.g., psychotherapy or medication) took place, on average a 25% reduction of complaints was found after a year. Improvement of somatic complaints was related to relatively high initial (retrospectively-reported) SCL-90-R somatic complaints scores. Initial complaints predicted 6% in the reduction of complaints. In addition, relatively low ratings of fatigue in the evening obtained from experience-sampling and relatively high (but normal as compared to controls) $PetCO_2$ values in the morning obtained from 24 hr ambulatory monitoring could reliably be related to improvement of symptoms after a year. Momentary fatigue and $PetCO_2$, however, predicted less than 8% of the total variance observed in the reduction of symptoms. Thus, recovery from medically unexplained symptoms can only marginally be predicted from all measures used in the current study.

The subgroups high versus low on improvement were different on the retrospective reports of complaints used for screening but, surprisingly, not on momentary reports of the same complaints during the assessment day. It thus seems that group differences in illness-related semantic knowledge rather than truly experienced complaints predicted recovery. This speculation is based on the suggestion that semantic knowledge (e.g., illness-related beliefs and expectations) influences retrospective symptom reports more strongly than momentary reports (Houtveen & Oei, 2007). However, identity-related semantic knowledge like illness beliefs and somatic attribution of complaints has been related in the literature to *poorer* outcomes (Schmaling *et al.*, 2003; Wilson, Hickie, Lloyd *et al.*, 1994) whereas higher complaint scores during selection were related to a *better* prognosis in the current study. These contradictory results indicate that identity-related beliefs probably are not the explanation for a prognostic effect of relatively high initial (retrospective) reports of somatic complaints. Another more plausible explanation may be that some participants initially inflated their symptoms during screening to be included in the study. Remarkable, regarding the absolute level, is that less complaints were found after a year for the subgroup high on improvement as compared to the subgroup low on improvement while their initial scores were higher (see Fig. 1). This may be the result of a "regression towards the mean" phenomenon. The confounding effect, however, of high retrospectively reported somatic complaints during screening has been accounted for by including the screening scores in the base model of the subsequent regression analyses and as a covariate in the repeated measures analyses.

One-year change in heterogeneous medically unexplained symptoms could not be predicted from initial retrospective reports of anxiety and depression (i.e., using the SCL-90-R subscales anxiety and depression). This is in line with previous studies on this issue that also could not demonstrate a predictive effect of anxiety or depression on the recovery of chronic fatigue syndrome (Rimes *et al.*, 2007; Van Der Werf *et al.*, 2002). The only momentary-reported measure that predicted improvement in the current study was fatigue. This is in line with the results of van der Werf *et al.* who also indicated that persistence of chronic fatigue syndrome complaints after one year follow-up was associated with initial fatigue (Van Der Werf *et al.*, 2002). The current study found similar results for heterogeneous medically unexplained symptoms and, more specifically, it demonstrates that increased momentary-reported fatigue in the late afternoon and evening was related to less improvement.

As potential psychophysiological aspects of previous and/or ongoing negative affect we assessed the 24-hour real-life activity of the (cardiac) autonomic nervous system, the respiratory system, and cortisol. Note that only sub-segments low on physical activity (motility) were selected, excluding the confounding influence of activity-related variance in these measures. Changes in complaints were correlated with $PetCO_2$ and RSA (HF power) estimates in the morning and respiratory depth (respiratory power) in the evening (see Table 2, further illustrated in Fig. 3). Of these measures, the effects of RSA and respiratory depth did not emerge consistently from the subsequent regression and repeated measures analyses. Additionally, they could not clearly be interpreted. Comparison with a normal control group showed that RSA values were the lowest for the subgroup high on improvement (see Fig. 3), which is contrary to the prediction that negative affect (supposedly reflected in low RSA – Allen & Crowell, 1989; Berntson *et al.*, 1994; Langewitz & Ruddel, 1989) would be related to less improvement over time. It may further be speculated that participants low on improvement (who reported more fatigue in the evening) were also less active

in the evening, resulting in a lower respiratory activity as have been found for this subgroup. Therefore, the significant correlations found between change in complaints and RSA and respiratory depth (Table 2) may very well have resulted from statistical Type I errors, and it is concluded that these measures do not have prognostic value in predicting spontaneous improvement of medically unexplained symptoms.

Relatively low $PetCO_2$ was related to less improvement after a year. As depicted in Fig. 3, participants low on improvement had relatively lower $PetCO_2$ values in the morning (i.e., on average 2 mmHg reduction) as compared to normal controls and participants high on improvement. A reduction in $PetCO_2$ has been related in the literature to medically unexplained symptoms and to increased levels of negative affect (Han *et al.*, 2000; Ley & Yelich, 1998; Suess *et al.*, 1980). Note that participants in the current study were *not* hyperventilating; the $PetCO_2$ values observed were not low enough to reach a hypocapnic state. Inspection of individual $PetCO_2$ curves also did not yield cases of true hyperventilation.

To summarize, a mean reduction of complaints of 25% was found after a year in a non-clinical group suffering from heterogeneous medically unexplained symptoms. This reduction was spontaneous since no intervention or treatment took place. In line with previous studies, spontaneous improvement of unexplained somatic symptoms could not be predicted by initial anxiety or depression questionnaire scores. The improved methodology of experience-sampling used in the current study did not yield much improvement over and above the prediction by retrospective questionnaires. Even including a series of 24-hour real-life ambulatory assessed physiological measures (as independent markers of previous and/or ongoing negative affect) led to little additional explanation in the variance of improvement. Thus, although previous studies have shown that negative affect correlates with medically unexplained symptoms and may even predict the onset of symptoms (Leiknes *et al.*, 2007; Rimes *et al.*, 2007), the current study indicates that apparently other factors determine spontaneous improvement of medically unexplained symptoms. Other psychological factors that may be related to (reduced) spontaneous improvement are locus of control, somatic attribution and secondary gain. Physiological determinants may be found in genetic factors and sensitization of the physiological systems involved in interoception. Further research is needed for pinpointing these and other factors, and for determining predictive factors of improvement in specific clinical groups of medically unexplained symptoms.

This study was funded by a grant from the Netherlands Organization for Scientific Research (NWO No 452-02-011). The authors gratefully acknowledge the aid of Marte Kaan, Bregtje Gunter Moor and Marjolein Raaijmaakers.

REFERENCES

Allen, M. T. & Crowell, M. D. (1989). Patterns of autonomic response during laboratory stressors. *Psychophysiology*, *26*, 603–614.

Arrindell, W. A. & Ettema, J. H. M. (1981). *Dimensionele structuur, betrouwbaarheid en validiteit van de Nederlandse bewerking van de Symptom Checklist (SCL-90)* [Dimensional structure, reliability and validity of the Dutch translation of the Symptom Checklist (SCL-90)]. Nederlands Tijdschrift voor de Psychologie. Lisse, Swets & Zeitlinger, *36*, 77–108.

Arrindell, W. A. & Ettema, J. H. M. (1986). *SCL-90: Handleiding bij een multidimensionele psychopathologie indicator* [in Dutch]. Lisse, the Netherlands, Swets & Zeitlinger.

Barsky, A. J. & Borus, J. F. (1999). Functional somatic syndromes. *Annals of Internal Medicine*, *130*, 910–921.

Berntson, G. G., Cacioppo, J. T. & Quigley, K. S. (1993). Cardiac psychophysiology and autonomic space in humans: Empirical perspectives and conceptual implications. *Psychological Bulletin*, *114*, 296–322.

Berntson, G. G., Cacioppo, J. T., Quigley, K. S. & Fabro, V. T. (1994). Autonomic space and psychophysiological response. *Psychophysiology*, *31*, 44–61.

Berntson, G. G., Thomas Bigger Jr, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., *et al*. (1997). Heart rate variability: Origins methods, and interpretive caveats. *Psychophysiology*, *34*, 623–648.

Boiten, F. A., Frijda, N. H. & Wientjes, C. J. E. (1994). Emotions and respiratory patterns: Review and critical analysis. *International Journal of Psychophysiology*, *17*, 103–128.

Bolger, N., Davis, A. & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579–616.

Brown, K. W. & Moskowitz, D. S. (1997). Does unhappiness make you sick? The role of affect and neuroticism in the experience of common physical symptoms. *Journal of Personality and Social Psychology*, *72*, 907–917.

Costa Jr, P. T. & McCrae, R. R. (1985). Hypochondriasis, neuroticism, and aging: When are somatic complaints unfounded? *The American psychologist*, *40*, 19–28.

Craig, A. D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. *Nature Reviews Neuroscience*, *3*, 655–666.

Craig, A. D. (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology*, *13*, 500–505.

Da Costa, J. M. (1871). On irritable heart: A clinical study of functional cardiac disorder and its consequences *American Journal Medical Science*, *61*, 17–52.

De Geus, E. J. C. & Van Doornen, L. J. P. (1996). Ambulatory assessment of parasympathetic/sympathetic balance by impedance cardiography Ambulatory Assessment. In J. Fahrenberg & M. Myrtek (Eds.), *Computer assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 141–164). Berlin, Hogrefe & Huber.

De Geus, E. J. C., Willemsen, G. H. M., Klaver, C. H. A. M. & Van Doornen, L. J. P. (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. *Biological Psychology*, *41*, 205–227.

Dressendorfer, R. A., Kirschbaum, C., Rohde, W., Stahl, F. & Strasburger, C. J. (1992). Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. *The Journal of Steroid Biochemistry and Molecular Biology*, *43*, 683–692.

Grossman, P. (1983). Respiration, stress, and cardiovascular function. *Psychophysiology*, *20*, 284–300.

Han, J. N., Schepers, R., Stegen, K., Van Den Bergh, O. & Van De Woestijne, K. P. (2000). Psychosomatic symptoms and breathing pattern. *Journal of Psychosomatic Research*, *49*, 319–333.

Hotopf, M. (2002). Childhood experience of illness as a risk factor for medically unexplained symptoms. *Scandinavian Journal of Psychology*, *43*, 139–46.

Houtveen, J. H., Groot, P. F. C. & De Geus, E. J. C. (2006). Validation

of the thoracic impedance derived respiratory signal using multilevel analysis. *International Journal of Psychophysiology*, *59*, 97–106.

Houtveen, J. H. & Molenaar, P. C. (2001). Comparison between the Fourier and Wavelet methods of spectral analysis applied to stationary and nonstationary heart period data. *Psychophysiology*, *38*, 729–735.

Houtveen, J. H. & Oei, N. Y. L. (2007). Recall bias in reporting medically unexplained symptoms comes from semantic memory. *Journal of Psychosomatic Research*, *62*, 277–282.

Houtveen, J. H., Rietveld, S. & De Geus, E. J. C. (2003). Exaggerated perception of normal physiological responses to stress and hypercapnia in young women with numerous functional somatic symptoms. *Journal of Psychosomatic Research*, *55*, 481–490.

Kamphuis, A. & Frowein, H. W. (1985). Assessment of mental effort by means of heart rate spectral analysis. In J. F. Orlebeke, G. Mulder & L. J. P. Van Doornen (Eds.), *The psychophysiology of cardiovascular control* (pp. 841–853). New York: Plenum Press.

Kirschbaum, C. & Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: An overview. *Neuropsychobiology*, *22*, 150–169.

Kirschbaum, C. & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology*, *19*, 313–333.

Lang, P. J. (1994). The varieties of emotional experience: A meditation on James-Lange theory. *Psychological Review*, *101*, 211–221.

Langewitz, W. & Ruddel, H. (1989). Spectral analysis of heart rate variability under mental stress. *Journal of Hypertension*, *7*, S32–S33.

Larsen, R. J. (1992). Neuroticism and selective encoding and recall of symptoms: Evidence from a combined concurrent-retrospective study. *Journal of Personality and Social Psychology*, *62*, 480–488.

Leiknes, K. A., Finset, A., Moum, T. & Sandanger, I. (2007). Course and predictors of medically unexplained pain symptoms in the general population. *Journal of Psychosomatic Research*, *62*, 119–128.

Ley, R. & Yelich, G. (1998). Fractional end-tidal CO_2 as an index of the effects of stress on math performance and verbal memory of test-anxious adolescents. *Biological Psychology*, *49*, 83–94.

Mauss, I. B., McCarter, L., Levenson, R. W., Wilhelm, F. H. & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*, 175–190.

Miller, G. A. & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*, 40–48.

Nimnuan, C., Hotopf, M. & Wessely, S. (2001). Medically unexplained symptoms an epidemiological study in seven specialities. *Journal of Psychosomatic Research*, *51*, 361–367.

Pennebaker, J. W. (2000). Psychological factors influencing the reporting of physical symptoms. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman & V. S. Cain (Eds.), The science of self-report Implications for research and practice (pp. 301–315). London, LEA.

Price, D. D., Zhou, Q., Moshiree, B., Robinson, M. E. & Verne, G. N. (2006). Peripheral and central contributions to hyperalgesia in irritable bowel syndrome. *Journal of Pain*, *7*, 529–535.

Riese, H., Groot, P. F. C., Van den Berg, M., Kupper, N. H. M., Magnee, E. H. B., Rohaan, E. J., *et al*. (2003). Large-scale ensemble averaging of ambulatory impedance cardiograms. *Behavior Research Methods, Instruments, and Computers*, *35*, 467–477.

Rimes, K. A., Goodman, R., Hotopf, M., Wessely, S., Meltzer, H. & Chalder, T. (2007). Incidence, prognosis, and risk factors for fatigue and chronic fatigue syndrome in adolescents: A prospective community study. *Pediatrics*, *119*.

Robinson, M. D. & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*, 934–960.

Schmaling, K. B., Fiedelak, J. I., Katon, W. J., Bader, J. O. & Buchwald, D. S. (2003). Prospective study of the prognosis of unexplained chronic fatigue in a clinic-based cohort. *Psychosomatic Medicine*, *65*, 1047–1054.

Shacham, S. (1983). A shortened version of the Profile of Mood States. *Journal of Personality Assessment*, *47*, 305–306.

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R. & Van Doornen, L. J. P. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology*, *27*, 1–23.

Sherwood, A., Allen, M. T., Obrist, P. A. & Langer, A. W. (1986). Evaluation of beta-adrenergic influences on cardiovascular and metabolic adjustments to physical and psychological stress. *Psychophysiology*, *23*, 89–104.

Stone, A. A. & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine*, *24*, 236–243.

Suess, W. M., Alexander, A. B. & Smith, D. D. (1980). The effects of psychological stress on respiration: A preliminary study of anxiety and hyperventilation. *Psychophysiology*, *17*, 535–540.

Van Der Werf, S. P., De Vree, B., Alberts, M., Van Der Meer, J. W. M. & Bleijenberg, G. (2002). Natural course and predicting self-reported improvement in patients with chronic fatigue syndrome with a relatively short illness duration. *Journal of Psychosomatic Research*, *53*, 749–753.

Van Diest, I., De Peuter, S., Eertmans, A., Bogaerts, K., Victoir, A. & Van Den Bergh, O. (2005). Negative affectivity and enhanced symptom reports: Differentiating between symptoms in men and women. *Social Science and Medicine*, *61*, 1835–1845.

Verne, G. N., Himes, N. C., Robinson, M. E., Gopinath, K. S., Briggs, R. W., Crosson, B., *et al*. (2003). Central representation of visceral and cutaneous hypersensitivity in the irritable bowel syndrome. *Pain*, *103*, 99–110.

Verne, G. N., Robinson, M. E. & Price, D. D. (2004). Representations of pain in the brain. *Current Rheumatology Reports*, *6*, 261–265.

Watson, D. & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review*, *96*, 234–254.

Wessely, S., Nimnuan, C. & Sharpe, M. (1999). Functional somatic syndromes: One or many? *Lancet*, *354*, 936–939.

Wientjes, C. J. E. & Grossman, P. (1994). Overreactivity of the psyche or the soma? Interindividual associations between psychosomatic symptoms, anxiety, heart rate, and end-tidal partial carbon dioxide pressure. *Psychosomatic Medicine*, *56*, 533–540.

Willemsen, G. H. M., De Geus, E. J. C., Klaver, C. H. A. M., Van Doornen, L. J. P. & Carroll, D. (1996). Ambulatory monitoring of the impedance cardiogram. *Psychophysiology*, *33*, 184–193.

Wilson, A., Hickie, I., Lloyd, A., Hadzi-Pavlovic, D., Boughton, C., Dwyer, J., *et al*. (1994). Longitudinal study of outcome of chronic fatigue syndrome. *British Medical Journal*, *308*, 756–759.

Received 8 April 2008, accepted 10 June 2008

APPENDIX

The 12 items of the somatic complaints subscale of the SCL-90-R used for the current study are: (1) headache, (2) faintness or dizziness, (3) pain in heart or chest, (4) pain in lower back, (5) nausea or upset stomach, (6) soreness of muscles, (7) trouble getting your breath, (8) hot or cold spells, (9) numbness or tingling in part of the body, (10) lump in your throat, (11) heavy feeling in your arms or legs and (12) feeling weak in parts of your body. The used items of the shortened version of the Profile of Mood States questionnaire are: (1) depression: unhappy, sad, hopeless, (2) vitality: active, energetic, lively, (3) anger: angry, annoyed, moody, (4) fatigue: tired, weary, fatigued and (5) tension: tense, nervous, anxious.