

Odors and Sounds as Triggers for Medically Unexplained Symptoms: A Fixed-Occasion Diary Study of Gulf War Veterans

Eamonn Ferguson, Ph.D., Helen J Cassaday, Ph.D., and Peter A Bibby, Ph.D.

University of Nottingham

ABSTRACT

Background: Both laboratory studies on healthy volunteers and epidemiological evidence from patient samples indicate that odor can act as a trigger for the reporting of medically unexplained symptoms (MUSs). **Purpose:** The relationship between concurrent experiences of odor and MUSs has not been explored in a patient sample. **Methods:** This study used an 8-day fixed-occasion diary study, in which 17 veterans of the Persian Gulf War completed diary assessments of (a) the intensity and duration of odor and sound and (b) MUS severity. **Results:** The results showed that the intensity of odor was positively associated with the severity on the same day and subsequent days' symptoms, whereas the duration of odor was negatively related to the severity of MUS reporting on the same day. **Conclusions:** These results are consistent with an associative mechanism underlying symptom reporting in veterans. By contrast, the duration, but not the intensity, of sound was related to the severity of MUS reporting on the same day.

(Ann Behav Med 2004, 27(3):205–214)

INTRODUCTION

There is now a substantial body of empirical laboratory evidence indicating that odor can come to act as an associative trigger for symptom reporting in previously healthy participants (1–3).¹ Similarly, there is epidemiological evidence indicating that medically unexplained symptoms (MUSs) (4,5)² are related to retrospective reports of odorous environmental exposures (6–10). Empirical data exploring the relationship between concurrent experiences of odor and MUSs in a sample of patients who have been previously exposed to environmental odors are needed to help to tie the experimental and epidemiological findings together. This article, therefore, reports on a fixed-occasion diary study examining the relationships between daily experiences of odor, sound, and MUSs in a sample of ill veterans of the Persian Gulf War. Veterans were chosen for two reasons: (a) They have prior experience of odorous environmental hazards in

the Gulf (6–11) and (b) there are theoretical reasons to believe that an odor-based associative account can explain the persistence of MUS reporting in this group (11,12).

Odor and MUS Reporting

Van den Bergh and colleagues (1–3) have produced a substantial body of evidence showing that odor can act as a conditioned stimulus (CS) for symptom reporting. A series of experimental studies in humans have shown that an odor (e.g., ammonia CS), previously paired with illness induced by carbon-dioxide-enriched air (the unconditioned stimulus [UCS]), will on later presentation elicit the same pattern of symptoms (the conditioned response [CR]) as seen for carbon dioxide (e.g., dizziness). This effect is known to generalize to other perceptually similar odors (e.g., unpleasant odors, not pleasant odors) and to be supportable over long durations (1).

A similar bio-associative hypothesis has been offered for the persistence of symptom reporting observed in veterans of the Gulf War (11,12). It is proposed that odors present in the Persian Gulf (CS) became associated with a set of symptoms known as the sickness response (the unconditioned response [UCR]) (cf. 13). Subsequently, the same odors in the home environment (CS) could invoke the same responses (CR) and so result in MUS reporting.

Although odors may act in this way to influence symptom reporting, other environmental triggers may also have an effect. One such trigger that is pertinent to such a patient group is sound. In terms of accounts based on post-traumatic stress disorder (PTSD), the experience of intense sounds should lead to an increase in MUS reporting (e.g., 14,15).

Furthermore, there is evidence that general perceptions of disease severity can moderate the relationship between daily stress and symptom reporting, such that this relationship is stronger for those who perceive their symptoms to be more severe (16). Therefore, it is also proposed that veterans' general perceptions of their health may moderate, in a similar manner, the relationship between odor (or sound) and symptom reporting.

Summary of This Study

This study, therefore, explored the links between daily reports of symptoms, odors, and sounds based on a small but representative sample of veterans. In this study, both odor and sound ratings were taken and used as a within-subjects comparison. Both daily reports of odors and of sounds should act to trigger increased levels of MUS reporting (1–3,11–14). Therefore, the study design used here means that the relative efficacy of both of these two triggers can be examined. The idea of a trigger implies a temporal ordering with the trigger preceding the out-

¹Carbon dioxide (used in these studies) is odorless and acts as a trigeminal stimulant. Ammonia could act as a trigeminal stimulant or an odor. Although the pathways are different, the same associative principles should apply.

²MUSs are symptoms that are not diagnostic of any particular illness or disease (4,5).

Reprint Address: E. Ferguson, Ph.D., School of Psychology, University of Nottingham, Nottingham, NG7 2RD, United Kingdom. E-mail: eamonn.ferguson@nottingham.ac.uk

© 2004 by The Society of Behavioral Medicine.

come. To examine this, we explored 1-day lagged associations between reports of both odor and sound with respect to the subsequent day's symptom levels. One-day lags were used, as there is an extensive existing literature suggesting that symptoms and their predictors can vary over this interval (17–19). Thus, this study sampled occasions and examined temporal variability. To date, research in the area of illness associated with the Gulf War has done an excellent job of sampling people as well as sampling a wide variety of variables including stress, psychiatric morbidity, and personality (for reviews, see 11,20,21), but there are no data on short-term temporal variation in reported symptoms and environmental triggers by veterans of the Gulf War.

METHODS

Sampling and Participants

Veterans were sampled through two veterans support charities (the National Gulf Veterans and Families Association and the Gulf Veterans Association). Advertisements were placed in their newsletters and magazines requesting veterans to participate. Data were gathered on 21 veterans who agreed to participate (M age = 43.3 years, SD = 9.1 years). They were 95% male, they had served on average 111 days in the Gulf (SD = 44 days), and 56% were smokers or had smoked. The demographics for the veteran sample are representative of those reported for large-scale epidemiological studies (e.g., 6,7,22,23).

Design

The 8-day fixed-occasion diary (Monday to the following Monday) study was completed, following the same procedures as used in other diary studies (24,25). Variables were assessed at two levels. Baseline symptoms (estimated average degree of symptom severity) constituted the Level 2 (between-subjects) variable and were assessed prior to the diaries being completed. The experiences of daily symptoms, odors, and sounds constituted the Level 1 (within-subjects) variables. The assessment of the Level 1 variables conformed to a fixed-occasion design (see 26). That is, there was a fixed set of $t = 1, \dots, m$ measurement occasions. Daily assessments were chosen on the basis of the extensive existing literature suggesting that symptoms and their predictors can vary over this interval (see 17–19).

Eight days of assessment were chosen for three reasons. First, applying the Snijders and Bosker (26) formula for the required number of Level 1 units to pilot data ($N = 23$ undergraduates) indicated that 7 days were necessary. Second, initial pilot data from 6 veterans (different from the participants in the study reported here), using a 14-day fixed-occasion diary study, resulted in zero returns (so it was inadvisable to greatly exceed the minimum required). Third, 8 days is a time span that has been used previously (24,25). The requirements of the diary study were discussed with workers from the veterans' charities, so that they could explain the procedure to the participants.

Measures

Baseline symptom levels. Veterans indicated, in general, how severely they experienced each of 19 symptoms with re-

spect to the course of their illness (median year of onset was 1992, range = 1991–1996). Judgments were made using 6-point Likert-type scales ranging from 0 (*never experienced the symptom*) to 5 (*experienced very severely*). These 19 symptoms were derived from the six cluster definitions provided by Haley, Kurt, and Hom (8) and so covered most of the common symptoms reported by Gulf War veterans. Example symptoms used to map Haley et al.'s (8) Syndrome Cluster 1 of "impaired cognition" were "problems remembering things," "feeling sleepy in the daytime," and "migraine type headaches." The following are example symptoms designed to index the remaining five symptom clusters: (a) Cluster 2 ("confusion ataxia") was indexed by symptoms including "problems stumbling or losing things," "sexual problems"; (b) Cluster 3 ("arthro-myo-neuropathy") was indexed by symptoms including "muscle fatigue after exertion," "tingling in the hands, feet, arms or legs"; (c) Cluster 4 ("phobia-apraxia") was indexed by symptoms including "trembling during activities like cleaning your teeth"; (d) Cluster 5 ("fever-adenopathy") was indexed by "swollen gland in any bodily location"; and (e) Cluster 6 ("weakness-incontinence") was indexed by symptoms including "problems controlling your bowels or bladder," "problems standing from a chair." Recent evidence suggests that a single factor can account for the covariation among MUSs such as these (4,5,11). Therefore, these 19 symptoms were summed to give a single symptom index ($\alpha = .84$).

Daily symptoms. These were measured with the same 19-symptom checklist used to assess baseline symptoms. However, in this case participants indicated, at the end of each day, how severely they had experienced each separate symptom that day. Symptoms were scored on a 6-point Likert-type scale ranging from 0 (*did not experience the symptom*) to 5 (*experienced very severely*). Coefficient alphas showed that the sum of these 19 symptoms was reliable for each day of assessment for the veterans (range = .90–.93, $M = .91$).

The Chemical Intolerance Index. The Chemical Intolerance Index (CII) is a five-item scale designed to assess the extent to which each of five chemicals (e.g., perfume) are perceived to cause illness (27). Ratings were made on 5-point Likert-type scales ranging from 1 (*almost never*) to 5 (*almost always*). A cutoff of 22 (out of a maximum of 25) indicates risk for multiple chemical sensitivity (MCS) (27) and was used to distinguish associative from nonassociative (directly sensitizing) effects of odor. For the veterans the CII had a coefficient alpha of .69. One veteran scored as being at risk for MCS on the CII. The analyses were, therefore, conducted twice with this veteran included and excluded. Including this veteran in the analyses did not alter the results. Therefore, the results reported are based on the full sample.

Intensity and duration of odors. On a daily basis, if participants noticed a smell, they were instructed to write down the odor's name and to indicate its intensity, using a 5-point Likert-type scale ranging from 1 (*very weak [only just noticed]*)

to 5 (*very strong [overbearing]*). Participants also provided an estimate of odor duration (in minutes). Participants indicated the single odor that had been most obvious rather than all odors. Where responses indicated the absence of a distinctive odor, a score of zero was returned. Participants were not instructed to describe smells that made them feel ill, just to describe any smell they noticed. To avoid setting up any expectancy effects or confounds, we did not ask veterans to report on odors and exposures they experienced in the Gulf. Rather, the large body of epidemiological evidence was taken to indicate the general types of exposure experienced (9,10,23).

Intensity and duration of sounds. Sounds were assessed the same way as odors. Participants provided a brief description of any sounds on a daily basis and indicated loudness (intensity) on a 5-point Likert-type scale ranging from 1 (*very quiet [only just noticeable]*) to 5 (*very loud [overbearing]*). Participants also indicated the duration (as before, in minutes). Again, participants indicated the single sound they most noticed rather than all sounds. Where responses indicated the absence of a distinctive sound, a score of zero was returned. Participants were not instructed to describe sounds that made them feel ill, just to describe any sound they noticed.

Data Analysis

The diary data were analyzed using hierarchical multivariate linear modeling (HMLM) using HLM5 (28). Snijders and Bosker (26) and Raudenbush, Bryk, Cheong, and Congdon (28) argued that this is a suitable modeling strategy for a fixed-occasion design, as it allows various models of the temporal variance-covariance structure to be assessed (see also 29). In all the models tested here, those that assumed unrestricted Level 1 variance-covariance structures were the best fits to the data, and the parameters from these models are reported (for more detail on these procedures, see 28,29).

The data hierarchy in this case consists of two levels. Level 1 assesses the within-person variation (i.e., daily patterns in the ratings of physical symptoms, smells, and sounds), and Level 2 assesses the between-persons variability (i.e., baseline average symptoms). The Level 1 predictor variables were not centered, as they were on a meaningful metric. That is, not experiencing a smell or sound provided a meaningful score of zero.³ The Level 2 variable was centered around the grand mean (29,30).

Data Exclusions

Participants were excluded from the HMLM analyses if they had missing data. Data were classed as missing if the participant had omitted a whole section of the diary (e.g., no record of any daily symptoms for the 8 days or no detail on odors or sounds over the 8 days). If only 1 or 2 days' ratings were missing, the remaining data could be included.

³The same pattern of results emerged when the Level 1 data were group centered.

RESULTS

Number of Level 1 Units

Applying Snijders and Bosker's (26) formula to these data indicated that for reliable estimates 4 days' data were required. Therefore, the 8 days collected were sufficient for the analyses.

HMLM of the Diary Data

Of the 21 veterans who returned diaries, 17 (81%) provided usable data sets for the multilevel modeling of the intensity and duration of odors and sound as predictors of daily symptoms. These 17 participants provided 133 days of data. The results for initial models exploring the within-subjects relationships for odor and sound with respect to symptoms are presented first. These are followed by an examination of how the between-subjects factor (general symptom severity) modifies these relationships.

Initial Level 1 models (effects of odor and sound on symptoms). Two models were examined initially. The first model explored the relative strength of odor and sound (in each case, both their intensity and duration) to predict symptom reporting on the same day (Lag 0). The second model explored the power of the previous day's odor and sound (in each case, both their intensity and duration) to predict the subsequent day's symptom severity (Lag 1).

The Lag 0 model was specified for the following theoretical reasons: (a) Animal learning theory data show that flavor aversion is readily acquired (31), and (b) there is evidence for concurrent associations between daily events and mood (32).

This initial zero lag model is expressed by the following equations.

Level 1:

$$Y_{ij} = B_{0j} + B_1(\text{Intensity of odor}) + B_2(\text{Duration of odor}) + B_3(\text{Intensity of sound}) + B_4(\text{Duration of sound}) + r_{ij}$$

Level 2:

$$\begin{aligned} B_0 &= \gamma_{00}, \\ B_1 &= \gamma_{10}, \\ B_2 &= \gamma_{20}, \\ B_3 &= \gamma_{30}, \\ B_4 &= \gamma_{40}, \end{aligned}$$

where Y_{ij} = the within-person variation in daily symptoms, B_{0j} = the intercept term, B_1 to B_4 = the slope estimates for the intensity and duration of smell and sound, and r_{ij} = the error term; γ = the structural coefficients expressing the Level 1 variation.

The 1-day lag model was explored because previous work has shown that psychological factors (e.g., emotions and stress) can predict somatic symptoms with this order of delay (18,19). The equation for this analysis is of the same general form as the zero-lag equation, except that the previous day's odor and sound (intensity and duration) are used to predict the subsequent day's symptom severity.

The results for these two initial models are presented in Table 1. The results show that at zero and 1-day lags the level of symptom severity reported by the veterans was significantly different from zero (γ_{00}). There were also significant positive associations for the intensity of odor with symptom severity (γ_{10}) at lags zero and 1-day. That is, the greater the reported intensity of an odor, the greater the reported symptom severity on the same and subsequent days. There was also a significant negative association for the duration of odor with symptom severity (γ_{20}) at the zero lag. This means that short-duration odors were associated with the reporting of more severe symptoms on the same day. There was also a marginally significant ($p = .052$) positive effect for duration of sound at Lag 0 (γ_{40}). This means that long-duration sounds were associated with more severe same-day symptoms.

Modifying effects of Level 2 general symptoms. A series of further models examined whether significant Level 1 effects (e.g., significant lagged associations for odor) were conditional on (modified by) Level 2 variables (i.e., baseline symptoms). Although such a two-step procedure has been suggested to provide a useful way to model data (see 29), other researchers have cautioned that selecting only significant Level 1 effects to be entered into a conditional model should at best be viewed as a heuristic and that analysis should also be guided by theory (26). We therefore explored cross-level effects for all conventionally significant Level 1 effects ($p < .05$). In addition, it is theoretically plausible that veterans' perceptions of their general health may moderate the relationship between triggers and symptoms (16,33). Therefore, marginally significant effects were also explored.

Thus, the conditional model at Lag 0 was specified only for the significant effects (odor intensity and duration at Lag 0) and the marginally significant effect for the duration of sound. The conditional model at the 1-day lag contained one significant term for odor intensity.

The equations for the conditional model for the zero lag are as follows:

Level 1:

$$Y_{ij} = B_{0j} + B_1(\text{Intensity of odor}) + B_2(\text{Duration of odor}) + B_3(\text{Intensity of sound}) = B_4(\text{Duration of sound}) + r_{ij}.$$

Level 2:

$$\begin{aligned} B_0 &= \gamma_{00} + \gamma_{01} (\text{symptoms}), \\ B_1 &= \gamma_{10} + \gamma_{11} (\text{symptoms}), \\ B_2 &= \gamma_{20} + \gamma_{21} (\text{symptoms}), \\ B_3 &= \gamma_{30} + \gamma_{31} (\text{symptoms}). \end{aligned}$$

This model includes cross-level terms. For example γ_{00} indicates (as in the initial model) the mean level of symptoms, and γ_{01} (*symptoms*) indicates, in addition, the extent to which this average is modified by baseline symptom levels. Similarly, γ_{10} indicates the average size of the relationship between the intensity of odor and symptom severity, and γ_{11} (*symptoms*), the extent to which that relationship is moderated by or conditional on baseline symptom levels, and so on. The model for the 1-day lag would be the same, except it would not include the terms for the durations of odor or sound.

The results for these two conditional models are presented in Table 2. The mean levels of symptom severity (γ_{00}) were significantly different from zero. The significant positive association between the intensity of odor and symptom severity at the zero and 1-day lags (γ_{10}) remained. Also the significant negative association between the duration of odor and symptom severity (γ_{20}) and the positive effect for duration of sound (γ_{30}) remained.

Baseline symptom levels significantly influenced average symptom severity (γ_{01}) at zero and 1-day lags. The positive coefficients with baseline symptoms indicated that those veterans who reported higher average symptom severity also tended to report more severe daily symptom levels. Baseline symptoms

TABLE 1
Initial Model at Zero and 1-Day Lags

Effect	Symbol	Zero Lag			1-Day Lag		
		Coefficient	SE	p	Coefficient	SE	p
Intercepts							
Symptoms	γ_{00}	34.85	3.80	< .0001	41.81	3.65	< .0001
Slopes							
Intensity of odor	γ_{10}	0.63	0.17	< .01	0.52	0.20	< .05
Duration of odor	γ_{20}	-0.01	0.004	< .01	0.004	0.005	= .40
Intensity of sound	γ_{30}	0.21	0.19	= .281	-0.22	0.16	= .19
Duration of sound	γ_{40}	0.007	0.003	= .052	-0.0007	0.003	= .83
Deviance		811.53 (41)			698.59 (33)		

Note. Values in parentheses are degrees of freedom for the deviance statistics. Symbol = hierarchical multivariate linear modeling symbol; Coefficient = hierarchical multivariate linear modeling coefficient (these are not standardized).

TABLE 2
Conditional Model at Zero and 1-Day Lags

Effect	Zero Lag				1-Day Lag		
	Symbol	Coefficient	SE	p	Coefficient	SE	p
Intercepts							
Symptoms	γ_{00}	38.93	0.93	< .0001	43.54	0.64	< .0001
Slopes							
Intensity of odor	γ_{10}	1.40	0.11	< .0001	0.29	0.09	< .001
Duration of odor	γ_{20}	-0.03	0.004	< .0001			
Duration of sound	γ_{30}	0.017	0.002	< .05			
Cross-level interactions with general symptom levels							
Symptoms	γ_{01}	0.73	0.04	< .0001	0.82	0.03	< .0001
Intensity of odor	γ_{11}	0.05	0.006	< .0001	0.004	0.005	= .42
Duration of odor	γ_{21}	-0.0006	0.0002	< .05			
Duration of sound	γ_{31}	-0.0008	0.0001	< .0001			
Deviance		758.65 (44)			642.45 (32)		

Note. Values in parentheses are degrees of freedom for the deviance statistics. Symbol = hierarchical multivariate linear modeling symbol; coefficient = hierarchical multivariate linear modeling coefficient (these are not standardized).

also showed cross-level interactions with the slope coefficients at zero lag (γ_{11} , γ_{21} , and γ_{31}). The positive coefficient for odor intensity (γ_{11}) indicates that those veterans who reported (on average) more severe symptoms showed a significantly stronger positive association between odor intensity and symptoms. The negative cross-level interaction for odor duration (γ_{21}) indicates that the negative association between odor duration and symptoms was stronger for higher levels of baseline symptom reporting. Thus, for those with higher baseline symptom levels, shorter duration odors were associated with increased symptom reporting. Finally, the negative cross-level interaction for the duration of sound (γ_{31}) indicates that the positive association between the duration of sound and symptom reporting is reduced. Thus, shorter sound durations were more likely to be associated with increased symptom reporting for those reporting higher levels of baseline symptoms.

To check that the inclusion of the marginal significant effect (for the duration of sound) from the initial model added significantly to the conditional model, we compared the deviance statistic for the conditional model with the duration of sound included (785.6, $df = 44$) to the deviance for the conditional model when it was not included (815.9, $df = 42$). The difference between these two models (30.3, $df = 2$) was significant, indicating that the model that included the marginal effect was an improvement.

Summary of HMLM results. A summary of these results is presented in Table 3. The veterans showed consistent positive associations (both on the same day and lagged over a single day) between reported odor intensity and symptoms. This association was stronger for veterans reporting more severe baseline symptoms. Shorter duration odors were associated with more severe daily symptoms, and this effect was also stronger for veterans reporting more severe baseline symptoms. Longer duration

sounds were associated with reports of more severe daily symptoms. However, this effect was reduced for those with higher levels of baseline symptom severity, such that shorter sound durations were more likely to be associated with increased symptom reporting in those with higher levels of baseline symptoms.

Confounds and Biases

The main focus of this study was to examine the data for associations (both same-day and lagged effects) between reported intensity and duration of odors and sounds with symptoms. The analyses identified patterns consistent with causal effects. However, it is necessary to consider potential confounds that provide alternative explanations of how these associations could arise. These are explored in terms of (a) reactivity, (b) the nature of the triggers, and (c) the frequency with which triggers are experienced.

TABLE 3
Summary of the Effects of Environmental Triggers on Symptoms in Veterans

Effect	Zero Lag	1-Day Lag
Intensity of odor	Positive ^a	Positive
Duration of odor	Negative ^a	No effect
Intensity of sound	No effect	No effect
Duration of sound	Positive ^a	No effect

Note. Positive = positive association between environmental triggers and symptoms; negative = negative association between environmental triggers and symptoms; no effect = no association between environmental triggers and symptoms.

^aEffects modified by reported levels of general symptoms.

Reactivity. A major methodological problem discussed in relation to daily diary studies is that of “reactivity,” or the notion that the mere act of completing the diary daily might affect relationships in the data (34). For example, over time, participants could become more aware both of their symptoms and of environmental odors and sounds. Thus, if a participant reported more severe symptoms over the 8-day diary period, or more intense or longer lasting odors and sounds, any lagged associations could be attributable to these general trends in the data.

A series of five separate repeated measures analyses of variance were conducted on the intensity and duration of odor and sound ratings and the symptom severity rating. The 8 days of assessment acted as the repeated measure. There were no significant effects for day of assessment (maximum $F = 1.4$). Although finding no trend suggests that reactivity is not present, it does not fully rule it out (34).

The nature of daily reported odors and sounds. Patterns in the data may be related to the different types of odors and sounds experienced (see Tables 4 and 5). With the exception of gunfire (reported by 1 veteran on 1 day only), the odors and sounds reported were typical of everyday life. Tinnitus was reported by 3 of the veterans. Given its chronic nature, having tinnitus could in principle influence the results for the sound

TABLE 4
Daily Reported Odors

Odor	Occurrences
Car exhaust	5
Diesel	4
Petrol	3
Sweat	3
Burning	3
Perfume	3
Vomit	3
Bleach	2
Fabric conditioner	2
Damp	2
Paint	2
Tar	1
Dog	1
Smoke	1
Coffee	1
Sea air	1
Urine	1
Cut grass	1
Wet canvas	1
Acid	1
Polythene	1
Oil	1
Aftershave	1
Lavender	1

Note. Occurrences refer to the number of days each odor was mentioned. The descriptive terms used are taken directly from the wording used in the diaries.

TABLE 5
Daily Reported Sounds

Sound	Occurrences
Tinnitus	15
Traffic	12
Alarms (car, house, radio)	8
Talking	4
Banging	3
Aircraft	3
Screaming child	3
Workmen's drills	2
Music	2
Bouncing ball	2
Buzzing	2
Humming	1
Gun fire	1
Quiet song	1
Knocking	1
Building work	1
Crowds	1
Hedge trimmer	1
TV	1

Note. Occurrences refer to the number of days each sound was mentioned. The descriptive terms used are taken directly from the wording used in the diaries.

data. The analyses were therefore repeated with these veterans removed. The results were unchanged by their exclusion.

Frequency of experiences. Veterans did not report the experience of odors and sounds on all days, and it may be that any reporting differences between odor and sound account for the differences seen in the data. On average, veterans' reports indicated the absence of distinctive odors on 5.3 days and of sounds on 3.8 days. A repeated measures t test indicated no difference, $t(17) = 1.7, p = .097$, two-tailed.

DISCUSSION

Laboratory-based associative models of symptom reporting have shown that after the appropriate pairings of odor and symptoms, healthy undergraduates will report symptoms in the presence of the conditioning odor (1). The pattern of results reported in this article similarly indicates that daily experiences of odor are linked to increased symptom reporting in a sample of veterans of the Persian Gulf War with chronic illness—a group previously exposed to odor and illness (6–11). For this group, odor, but not sound, is related to the same-day as well as subsequent-day's symptom reporting. The implications of these findings are discussed next.

Odor and Symptom Reporting

The pattern of same-day and lagged associations between reported odors and symptom severity supports the basic hypotheses arising from the bio-associative model of Gulf War-related

illness (11,12) and the experimental work of Van den Bergh and colleagues (1–3). To our knowledge, these are the first data that show in a patient population with MUSs that current experiences of odor are associated with subsequent MUS reporting. These findings, along with the experimental data of Van den Bergh's group, indicate that associative effects could provide a mechanism for persistent MUSs.

Although not based on experimental manipulations, diary data of this nature do allow for some degree of causal inferences to be drawn, with odor predicting the subsequent day's symptom reports. A number of the findings from this study are consistent with the hypothesis of an associative mechanism linking odor and symptoms. First, more intense odors were associated with increased symptom reporting, both on the same day and 24 hr later, whereas the shorter odor duration was associated with increased symptom reporting only on the same day. This differential patterning of response to intensity and duration is consistent with animal learning theory: Conditioning increases with stimulus intensity and is reduced with extended CSs through inhibition of delay (35). Second, the types of odors reported by the veterans were likely to have been experienced in the Gulf (6–10), where they would have provided likely CSs. Third, these odors are everyday odors that are not themselves likely to cause illness. Fourth, it may be the case that the data also support a sensitization account. However, the sample was screened for chemical intolerance, and only 1 veteran scored above the cutoff point for potential MCSs. When this veteran was excluded from the analyses, the pattern of results remained the same.

The data also add to the previous research showing that retrospective reports of exposures to odors such as petrol in the Gulf are associated with current symptom ratings. However, such studies can be affected by recall biases (20). Of importance, using a daily diary study method that reduces recall biases, the results of this study showed that currently experienced environmental odors are also associated with current reported symptoms. Indeed, Wessely (21) recently suggested that the most plausible exposures contributing to symptoms reported by Gulf War veterans are likely to be general everyday triggers such as the ones found here. Thus, odor would seem to be a strong contender for such a trigger.

Sound and Symptom Reporting

For these veterans there was also a positive effect for the duration of sound, with longer duration sounds associated with an increase in symptom reporting. A simple account based on PTSD might explain these results (14,15), and it may also be true the case that sound acts as an associative trigger. However, if this were the case, it might be expected that the intensity, as well as the duration, of sound would also be associated with symptom reporting; it was not. Moreover, in contrast to the duration of odor, for which the association with MUS reporting was negative, the association for the duration of sound was positive. This indicates that odor and sound operate via different mechanisms. It may be that long-duration sounds have uncondi-

tioned effects that make the veterans feel less well. Noise stress is in and of itself aversive (cf. 36). Alternatively, odor and sound stimuli could result in differential sensory adaptation and/or produce differential levels of inhibition of delay (35). However, such accounts are inevitably post hoc, and comparisons across modalities are difficult to make because there will also be confounded differences in stimulus salience. Thus, further research is needed to fully explore the mechanism(s) that might explain the effect for sound duration seen in veterans. For example, an increase in the unconditioned reaction to sound should show as a generally heightened sensitivity.

The Role of Baseline Symptom Severity

The within-subjects relationships between odor, sound, and daily symptom severity (described earlier) were moderated by veterans' general perceptions of their symptoms. This indicates that there are individual differences with respect to the involvement of odor and sound in the perception of symptoms. For veterans who reported more severe general symptoms, the relationship between odor intensity and symptom severity was stronger, and shorter duration odors and sounds were more likely to be associated with increased severity of symptom reporting. Thus, general perceptions of illness severity seem to enhance or sensitize the link between perceptions of environmental stimuli and symptom reporting. A similar picture has been reported with respect to stress and symptoms (16).

Alternatively, it may be the case that the influence of general perceptions of symptom severity can be explained by personality characteristics that increase sensitivity to and awareness of environmental stimuli and enhanced general symptom reporting (37). These could include neuroticism (N) and/or negative affectivity (NA), which are discussed later, as well as variables that are related to increased awareness and misinterpretation of bodily sensations such as somatosensory amplification (38,39), selective bodily attention (40), and symptom attributions (41). Future studies should consider these personality characteristics as potential Level 2 variables in this kind of research.

Generalization to Other Syndromes With MUSs

The relationships observed in these data pertain to a specific sample of Gulf War veterans. However, it has been argued that the range of syndromes that include MUSs (e.g., chronic fatigue syndrome, fibromyalgia, sick-building syndrome) may not represent separate clinical entities (1,4,5,11). As such, the basic theoretical model described here can be easily extended to symptom reporting in other clinical groups with MUSs. For example, a whole variety of factors could act as the UCS (e.g., stress, vaccination, having an infection) to produce symptoms. These symptoms will be experienced within particular environments where CSs (e.g., odors) are present, with the symptoms eventually becoming CRs supported by the associated CSs. In this way, a basic associative mechanism can explain persistent MUS reporting.

Limitations and Future Directions

It may be that potential sample bias and demand characteristics account for the relationship observed in this study. For example, use of a mailed survey means that diaries may be returned preferentially by veterans who chose to report odors and symptoms. However, there is no evidence in the data that this was in fact the case. For example, the respondents did not report odors every day, and there was no evidence that the sample was at risk for MCS. Therefore, although sample biases can never be completely removed, particular odor-based biases appear to be limited.

It is also possible that the observed relationship between daily odor and daily symptoms reflects the operation of a third variable, and as such inferences about causality are reduced. One prime candidate here is the personality dimension of N or NA. N-NA is known to be related to olfactory sensitivity (42), general MUS reporting (43,44) and the development of conditioned symptom reporting (1, but see 45). However, although the association between N-NA and symptom reporting is observed when data are cross-sectional, studies have shown that this association is not always observed when daily assessments of symptoms are recorded (46). Thus, N or NA may be less important to the explanation of daily variation in symptom reporting. Furthermore, this study took account of baseline differences in terms of the general level of symptom reporting (rated before the diary phase of the study), and the role of general symptom reporting was modeled. However, future studies would be improved by the inclusion of personality measures.

It is also possible that daily mood could mediate the relationship between odor and symptom reporting. There is evidence that daily negative mood is related to increased daily symptom reporting (47,48) and that unpleasant odors are related to reporting negative mood (49).

Both personality and mood would be particularly likely to play an important role in the early stages of acquiring associations between environmental events and symptoms (1). Once the links become established, their effects may be less important. All of these are essentially empirical questions that required further investigation.

CONCLUSIONS

This study provides an analysis of temporal patterning in MUSs and potential environmental triggers in a demographically representative sample of Gulf War veterans. These diary data show that odor intensity was the most prominent environmental trigger for the same and subsequent day's symptom reporting in veterans of the Gulf War. Effects of odor and sound duration were confined to same-day effects for the veterans. These naturalistic data provide support for the role of odor in symptom reporting that has previously been reported in healthy individuals in laboratory settings. However, to show causality, it is necessary to also test the possible contribution of additional variables such as personality and mood.

REFERENCES

- (1) Devriese S, Winters W, Stegen K, et al.: Generalization of acquired somatic symptoms in response to odors: A Pavlovian perspective on multiple chemical sensitivity. *Psychosomatic Medicine*. 2000, 62:751–759.
- (2) Van den Bergh O, Devriese S, Winters W, et al.: Acquiring symptoms in response to odors: A learning perspective on multiple chemical sensitivity. *Annals of the New York Academy of Sciences*. 2001, 933:278–289.
- (3) Van den Bergh O, Stegen K, Van de Woestijne KP: Learning to have psychosomatic complaints: Conditioning of respiratory behavior and somatic complaints in psychosomatic patients. *Psychosomatic Medicine*. 1997, 59:12–23.
- (4) Deary IJ: A taxonomy of medically unexplained symptoms. *Journal of Psychosomatic Research*. 1999, 47:51–59.
- (5) Wessely S, Nimnuan C, Sharpe M: Functional somatic syndromes: One or many? *Lancet*. 1999, 354:936–939.
- (6) Unwin C, Blatchley N, Coker W, et al.: Health of UK servicemen who served in the Persian Gulf War. *Lancet*. 1999, 353:169–178.
- (7) McCauley L, Joos S, Lasarev M, Storzbach D, Bourdette D: Gulf War unexplained illnesses: Persistence and unexplained nature of self-report symptoms. *Environmental Research (Section A)*. 1999b, 81:215–223.
- (8) Haley RW, Kurt T, Hom J: Is there a Gulf War Syndrome? Searching for syndromes by factor analysis of symptoms. *Journal of the American Medical Association*. 1997, 277:215–222.
- (9) Haley RW, Kurt TL: Self-reported exposure to neurotoxic chemical combinations in the Gulf War. *Journal of the American Medical Association*. 1997, 277:231–237.
- (10) Proctor S, Heeren T, White R, et al.: Health status of Persian Gulf War veterans: Self-reported symptoms, environmental exposures and the effect of stress. *International Journal of Epidemiology*. 1998, 27:1000–1010.
- (11) Ferguson E, Cassaday HJ: Theoretical accounts of Gulf War syndrome: From environmental toxins to psychoneuro-immunology and neurodegeneration. *Behavioural Neurology*. 2001, 13:133–147.
- (12) Ferguson E, Cassaday HJ: The Gulf War and illness by association. *British Journal of Psychology*. 1999, 90:459–476.
- (13) Maier SF, Watkins LR: Cytokines for psychologists: Implications of bi-directional immune-to-brain communication for understanding behavior, mood and cognition. *Psychological Review*. 1998, 105:83–107.
- (14) Morgan CA, Hill S, Fox P, Kingham P, Southwick SM: Anniversary reactions in Gulf War veterans: A follow up inquiry 6 years after the war. *American Journal of Psychiatry*. 1999, 156:1075–1079.
- (15) Wolfe J, Erickson D, Sharkansky E, King D, King L: Course and predictors of post-traumatic stress disorder among Gulf War veterans: A prospective analysis. *Journal of Consulting and Clinical Psychology*. 1999, 67:520–528.
- (16) Rothrock NE, Lutgendorf SK, Kerder KJ, Ratliff T, Zimmerman B: Stress and symptoms in patients with interstitial cystitis: A life stress model. *Urology*. 2001, 57:422–427.
- (17) Feldman SJ, Downey G, Schaffer-Neitz R: Pain, negative mood, and perceived support in chronic pain patients: A daily diary study of people with reflex sympathetic dystrophy syndrome. *Journal of Consulting and Clinical Psychology*. 1999, 67:776–785.
- (18) Hokanson JE, Richard LT, Xufeng N, Stader SR, Flynn HA: Illustration of concomitant time series analyses in a case of somatoform disorder. *Cognitive Theory and Research*. 1994, 18:413–437.
- (19) Stader SR, Hokanson JE: Psychosocial antecedents of depressive symptoms: An evaluation using daily experience methodology. *Journal of Abnormal Psychology*. 1998, 107:17–26.

- (20) Soetekouw PMMB, de Vries M, van Bergen L, et al.: Somatic hypotheses of war syndromes. *European Journal of Clinical Investigations*. 2000, 30:630–641.
- (21) Wessely S: Ten years on: What do we know about the Gulf War syndrome? *Clinical Medicine*. 2001, 1:28–37.
- (22) Coker W, Bhatt B, Graham, J: Clinical findings for the first 1000 Gulf War veterans in the Ministry of Defence's medical assessment programme. *British Medical Journal*. 1999, 318:290–294.
- (23) McCauley L, Joos S, Spence P, Lasarev M, Shuell DT: Strategies to assess validity of self-reported exposures during the Persian Gulf War. *Environmental Research (Section A)*. 1999, 81:195–205.
- (24) David JP, Green PJ, Martin R, Suls J: Differential roles of neuroticism, extraversion, and event desirability for mood in daily life: An integrative model of top-down and bottom-up influences. *Journal of Personality and Social Psychology*. 1999, 73:149–159.
- (25) David JP, Suls J: Coping efforts in daily life: Role of the Big Five traits and problem appraisals. *Journal of Personality*. 1999, 67:265–294.
- (26) Snijders T, Bosker R: *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling*. London: Sage, 1999.
- (27) Szarek MJ, Bell IR, Schwartz GE: Validation of a brief screening measure of environmental chemical sensitivity: The chemical odor intolerance index. *Journal of Environmental Psychology*. 1997, 17:345–351.
- (28) Raudenbush S, Bryk A, Cheong YF, Congdon R: *HLM5: Hierarchical Linear and Non-linear Modelling*. IL: Scientific Software International, 2000.
- (29) Raudenbush SW, Bryk AS: *Hierarchical Linear Models* (2nd Ed.). London: Sage, 2002.
- (30) Nezlek JB: Multilevel random coefficients analyses of event- and interval-contingent data in social and personality psychology research. *Personality and Social Psychological Bulletin*. 2001, 27:771–785.
- (31) Garcia J, Koelling RA: Relation of cue to consequence in avoidance learning. *Psychonomic Science*. 1966, 4:123–124.
- (32) Stone AA, Kessler RC, Haythornthwaite JA: Measuring daily events and experiences: Decisions for the researcher. *Journal of Personality*. 1991, 59:575–607.
- (33) Heady B, Wearing A: Personality, life events and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*. 1989, 57:731–739.
- (34) Affleck G, Zautra A, Tennen H, Armeli S: Multilevel daily process designs for consulting and clinical psychology: A preface for the perplexed. *Journal of Consulting and Clinical Psychology*. 1999, 67:746–754.
- (35) Hilgard ER, Marquis DC: *Conditioning and Learning*. New York: Appleton-Century-Crofts, 1940.
- (36) Pennebaker JW: *The Psychology of Physical Symptoms*. New York: Springer-Verlag, 1982.
- (37) Kolk AMM, Hanewald GJFP, Schagen S, Gijsbers van Wijk CMT: Predicting medically unexplained physical symptoms and health care utilization. A symptom perception approach. *Journal of Psychosomatic Research*. 2002, 52:35–44.
- (38) Barsky AJ, Goodson JD, Lane RS, Cleary PD: The amplification of somatic symptoms. *Psychosomatic Medicine*. 1988, 50:510–519.
- (39) Ferguson E: Hypochondriacal concerns and the five factor model of personality. *Journal of Personality*. 2000, 68:705–724.
- (40) Shields SA, Mallory ME, Simon A: The body awareness questionnaire: Reliability and validity. *Journal of Personality Assessment*. 1989, 53:802–815.
- (41) Robbins JM, Kirmayer LJ: Attributions of common somatic symptoms. *Psychological Medicine*. 1991, 21:1029–1045.
- (42) Pause BM, Ferstl R, Fehm-Wolfsdorf G: Personality and olfactory sensitivity. *Journal of Research in Personality*. 1998, 32:510–518.
- (43) Deaty IJ, Scott S, Wilson, JA: Neuroticism, alexithymia and medically unexplained symptoms. *Personality and Individual Differences*. 1997, 22:551–564.
- (44) Feldman PJ, Cohen S, Doyle WJ, et al.: The impact of personality on the reporting of unfounded symptoms and illness. *Journal of Personality and Social Psychology*. 1999, 77:370–378.
- (45) Van den Bergh O, Kempynck PJ, Van de Woestijne KP, Baeyens F, Ellen P: Respiratory learning and somatic complaints: A conditioning approach using CO₂-enriched air inhalation. *Behaviour Research and Therapy*. 1995, 33:517–527.
- (46) Brown KW, Moskowitz DS: Does unhappiness make you sick? The role of affect and neuroticism in the experience of common physical symptoms. *Journal of Personality and Social Psychology*. 1997, 72:907–917.
- (47) Schanberg LE, Sandstrom MJ, Starr K, et al.: The relationship of daily mood and stressful events to symptoms in juvenile rheumatic disease. *Arthritis Care and Research*. 2000, 13:33–41.
- (48) Larsen RJ, Kasimatis M: Day-to-day physical symptoms: Individual differences in the occurrence, duration, and emotional concomitants of minor daily illnesses. *Journal of Personality*. 1991, 59:387–423.
- (49) Marchand S, Arsenault P: Odors modulate pain perception: A gender-specific effect. *Physiology and Behavior*. 2002, 76:251–256.

APPENDIX

The strength of specifying these models using HMLM is the inclusion of the Δ matrix, which captures the variation and covariation among the T measurement periods. The Δ matrix can be used to examine hypotheses about different patterns of temporal variance and covariance in the data. The following variance–covariance structures were examined: (a) unrestricted Level 1 variance, (b) homogeneous Level 1 variance, and (c) a first-order autoregressive model at Level 1. The difference between the deviance statistic for the models run under the different variance–covariance assumptions was used to decide which variance–covariance pattern provided the best fit to the data (see 29). In all the models tested here, those that assumed unrestricted Level 1 variance–covariance structures were the best fits to the data and the parameters from these models are reported.

The two levels of the hierarchy in this study are represented by the following two equations. The first expresses the within-person Level 1 variation, and the second expresses the Level 2 between-persons variation.

Level 1 equation:

$$Y_{ij} = B_{0j} + \sum_{q=1}^{Q_j} B_{qj}(\text{triggers}) + r_{ij},$$

where Y_{ij} = the within-person variation in daily symptoms, B_{0j} = the intercept terms and indicates the mean level of symptom

within each person over the time series, B_{qj} = the slope estimates for the triggers (the intensity and duration of odors and sounds), and r_{ij} is the error term. The model stated above does not specify random variation between persons, as this is absorbed into the Δ matrix.

Level 2 equation:

$$B_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} W_{qsj} \gamma_{qsj},$$

where W represents the Level 2 predictor (in this case baseline symptoms) and γ represents the Level 2 coefficients.

Do Not Copy