



Available at www.sciencedirect.com



journal homepage: www.elsevier.com/locate/psyneuen



Olfactory influences on mood and autonomic, endocrine, and immune function

Janice K. Kiecolt-Glaser^{a,b,*}, Jennifer E. Graham^c, William B. Malarkey^{b,d,e}, Kyle Porter^f, Stanley Lemeshow^{b,f,g}, Ronald Glaser^{b,d,e}

^aDepartment of Psychiatry, The Ohio State University, 1670 Upham Drive, Columbus, OH 43210, USA

^bThe Ohio State Institute for Behavioral Medicine Research, 2175 Graves Hall, 333 West 10th Avenue, Columbus, OH 43210, USA

^cDepartment of Biobehavioral Health, The Pennsylvania State University, 315 East Health and Human Development Building, University Park, PA 16802, USA

^dDepartment of Internal Medicine, Davis Medical Clinic, The Ohio State University, 480 Medical Center Drive, Columbus, OH 43210, USA

^eMolecular Virology, Immunology, and Medical Genetics, 2078 Graves Hall, 333 West 10th Avenue, Columbus, OH 43210, USA

^fThe Ohio State University Center for Biostatistics, M200 Starling Loving, 320 West 10th Avenue, Columbus, OH 43210, USA

^gThe Ohio State University College of Public Health, M006 Starling Loving, 320 West 10th Avenue, Columbus, OH 43210, USA

Received 3 August 2007; received in revised form 8 October 2007; accepted 29 November 2007

KEYWORDS

Psychoneuro-immunology;
Aromatherapy;
Odor;
Complementary medicine

Summary

Despite aromatherapy's popularity, efficacy data are scant, and potential mechanisms are controversial. This randomized controlled trial examined the psychological, autonomic, endocrine, and immune consequences of one purported relaxant odor (lavender), one stimulant odor (lemon), and a no-odor control (water), before and after a stressor (cold pressor); 56 healthy men and women were exposed to each of the odors during three separate visits. To assess the effects of expectancies, participants randomized to the "blind" condition were given no information about the odors they would smell; "primed" individuals were told what odors they would smell during the session, and what changes to expect. Experimenters were blind.

Self-report and unobtrusive mood measures provided robust evidence that lemon oil reliably enhances positive mood compared to water and lavender regardless of expectancies or previous use of aromatherapy. Moreover, norepinephrine levels following the cold pressor remained elevated when subjects smelled lemon, compared to water or lavender. DTH responses to *Candida* were larger following inhalation of water than lemon

*Corresponding author at: Department of Psychiatry, The Ohio State University, 1670 Upham Drive, Columbus, OH 43210, USA.
Tel.: +1 614 292 0033; fax: +1 614 292 0038.

E-mail address: Kiecolt-Glaser.1@osu.edu (J.K. Kiecolt-Glaser).

or lavender. Odors did not reliably alter IL-6 and IL-10 production, salivary cortisol, heart rate or blood pressure, skin barrier repair following tape stripping, or pain ratings following the cold pressor.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

Widely utilized, aromatherapy is employed for relief of pain, relaxation and anxiety reduction, and enhanced energy; essential oils have been used to help women cope with labor pain, to relieve chemotherapy side effects, to enhance the rehabilitation of cardiac patients, to promote restful sleep, and to reduce post-surgical discomfort (Price and Price, 1999). However, efficacy data are scant, and potential mechanisms of action are controversial.

The “lock and key” or systemic effect theory posits that essential oils act like a drug or enzyme, such that particular odors should have very specific effects (Hirsch, 2001), a popular view among aromatherapy practitioners who prescribe certain odors for distinct health problems (Hirsch, 2001). For example, short-term inhalation of lavender oil (typically over the space of a few hours) is described as therapeutic for insomnia, influenza, headaches, migraines, anxiety, nervousness, and melancholy (Price and Price, 1999). Others have argued that lavender enhances immune function, as well as treating lung and sinus infections, laryngitis, and asthma (Keville and Green, 1995). These and other authors have emphasized lavender’s relaxant properties (Grace, 1999; Tisserand and Balacs, 1995).

Following a long history in folklore, lavender has been used as a sleep aid, and one small study suggested that the ambient odor of lavender can significantly enhance the amount of time asleep after withdrawal of medication for insomnia (Hardy et al., 1995). In another sleep study, lavender oil presented the first 2 min of every 10 min period for 40 min increased deep or slow-wave sleep compared to a control (distilled water) stimulus (Goel et al., 2005). Contingent negative variation (CNV), an EEG shift that occurs when individuals are expecting an event (e.g., a light that will signal a tone), is diminished by sedatives and enhanced by stimulants; lavender decreased CNV, but it did not affect reaction time or heart rate as sedatives do (Torii et al., 1988). Further, a comparison of rosemary and lavender oils using EEG and math computations showed that lavender increased patterns consistent with drowsiness, and subjects reported greater relaxation, while rosemary produced EEG patterns interpreted as increased alertness, consistent with faster and more accurate math test results (Diego et al., 1998).

In contrast to lavender’s sedative characterization, lemon oil is described as activating, immunomodulatory, and mood enhancing (Buchbauer et al., 1993; Keville and Green, 1995; Lis-Balchin and Hart, 2002; Price and Price, 1999); it has also been touted as an inhalation remedy for respiratory tract infections (Grace, 1999). Lemon oil has been associated with increased heart rate and enhanced mental and physical task performance in human studies (Jellinek, 1997). In addition, in one study participants exposed to lemon oil reported

fewer physical symptoms than individuals exposed to the unpleasant smell of dimethyl sulfide (Knasko, 1992).

Although some essential oils like lavender and lemon have been classified as either sedative or stimulating and these categories are broadly related to their putative CNS actions, there does not seem to be general agreement on mechanisms of action, and the health benefits are unclear; autonomic and self-report data have typically been used as a surrogate for health outcomes. In fact, despite the clear presumption that aromatherapy has immune consequences, we found only one human study that included immunological measures (Komori et al., 1995). Indeed, more broadly, most studies lack physiological or objectively assessed health endpoints.

Placebo and expectancy effects are central problems in human olfactory research, and interpretation of a number of studies is difficult for this reason; few studies are double or even single blind, and many assessed only a single odor without any control conditions (Martin, 2006). The paucity of these key controls is important because of what Jellinek terms the “placebo mechanism” (Jellinek, 1997); unlike the systemic effects theory described above, this theory holds that the characteristics of the odorant are irrelevant, and individuals’ expectancies determine the pattern of responses. In support of this perspective, participants to whom it was suggested that an odor would affect performance showed an improvement in math calculations, even when they were in fact exposed to no odor (Knasko et al., 1990). The general affective theory or reflectorial effect theory (Hirsch, 2001) provides yet another conceptual framework; it suggests that odors perceived as positive may induce positive moods, and these mood changes may enhance both physical and psychological well being.

To compare and contrast the diverse perspectives about whether and how odors affect health, we examined the autonomic, endocrine, and immune consequences of one purported sedating or relaxant odor, lavender, one activating or stimulant odor, lemon, and distilled water as a no-odor control during both resting and “challenge” or stress conditions in a mixed or between-within repeated measures design; each subject served as his or her own control during three separate 6 h visits. Depending on their random assignment, participants were either given no information about what odors they would be smelling or what to expect (the “blind” group), or they were told what odors they would smell and what changes to expect from the relaxant, stimulant, or no odor exposures (the “primed” group).

Our protocol for each session included a cold pressor, a laboratory stressor that elevates stress-related hormones, heart rate, and blood pressure (Blandini et al., 1995; Hirsch and Liebert, 1998). Both before and after the cold pressor we performed tape stripping, a common dermatological paradigm for studying restoration of the skin barrier, a process mediated by both endocrine and immune systems

(Choi et al., 2005). Our design thus provided a way to examine the ability of lemon and lavender odors to modulate stress and pain responses to the cold pressor, as well as wound healing via the speed of skin barrier repair.

Specific predictions can be derived from the various theories posited to explain the effects of essential oils. For example, if the systemic effect theory is correct, even relatively short-term exposure to lavender would be expected to produce larger declines in the production of cortisol and catecholamines, faster skin barrier repair, lower pain ratings in response to the cold pressor, and smaller stress-related immunological changes compared to lemon and the no-odor control; short-term exposure to lemon oil should produce greater transient increases in positive affect, heart rate, blood pressure, and catecholamines than either lavender oil or the no-odor control. If expectancies determine the pattern of responses (Jellinek, 1997), then the primed group's mood and physiological responses to lemon and lavender odors would be greater than the blind group; similarly, those with positive expectancies about aromatherapy in advance of participation would be expected to show greater changes. By assessing olfactory influences on mood and autonomic, endocrine, and immune function, our design allowed us to contrast these diverse conceptual perspectives, clarify mechanisms, and assess possible clinical efficacy.

2. Method

2.1. Participants

Participants recruited through ads were told that the study involved the assessment of responses to both strong and weak fruit and floral odors. We excluded individuals who were taking cardiovascular medications (statins, beta blockers, etc.) or medications or health problems with obvious immunological or endocrinological consequences. Additional exclusion criteria included perfume allergies, smoking, problems with smell or taste, respiratory symptoms or problems, asthma, and excessive alcohol or caffeine use.

The average age of the final sample of 21 men and 35 women was 24.41 ($SD = 6.05$, range = 18–43); 36 were white, 10 were African American, and 10 were Asian, and 54 had at least some college education. The Ohio State University Biomedical Research Review Committee approved the project; all subjects gave written informed consent prior to participation.

2.2. Screening session

We screened for general anosmia using phenylethyl alcohol and amyl acetate with a forced-choice staircase procedure (Coward, 1989). To assure that participants did not have a specific anosmia for lavender or lemon oils, they were given three sets of three bottles—two with distilled water, and the third which contained the essential oil—and asked to choose the one that differed from the other two. To be eligible for the study, subjects had to choose the correct response in all three trials for both odors.

We also conducted a structured interview to assess prior experiences with aromatherapy. Questions elicited participants' evaluation of and previous experience with aromatherapy, and their expectancies about the extent to which their own psychological and physiological responses would be influenced by odors.

Finally, in order to assess affective responses, cotton balls with each of the three stimuli were taped under each subject's nose in a randomly determined order. During exposure to each odor, participants rated their responses to a standardized series of positive and neutral slides (Lang et al., 1999). On completion of each series, subjects rated the odor's pleasantness, familiarity, and intensity using a 1–10 scale (Doty, 1986).

2.3. General Clinical Research Center (GCRC) visit timeline

After the screening session, subjects scheduled three visits to the GCRC, a hospital research unit. At least 2 weeks were scheduled between visits, and the average time to complete all three visits was 64.46 days ($SD = 48.40$). All visits followed the same timeline, illustrated in Figure 1, differing only in the odor condition.

On arrival a heparin well was placed in one arm, and participants then received a standard breakfast (after fasting since midnight). Subjects next completed a battery of questionnaires including a baseline PANAS to assess mood and the first of three emotional Stroop tests, both of which were repeated at the times shown in Figure 1. Following a 10-min baseline, we began automated blood pressure and heart rate assessments using a Dynamap/Critikon 1846SX/P. Baseline blood samples were drawn prior to odor exposure, after subjects had remained in a sitting position for 20–30 min.

Immediately prior to odor exposure, participants opened a sealed letter with blind or primed instructions (manipulation described below), and then a cotton ball with the essential oil or distilled water (depending on the day's randomized odor assignment) was taped under their nose (as described in more detail below). Next participants rated the extent to which they expected that the odor would affect their mood and physiological responses, as well as the odor's pleasantness, familiarity, and intensity using a 1–10 scale (Doty, 1986). Following about 1.25 h of odor exposure with periodic endocrine blood/saliva samples, they completed a second immune blood draw and the first of two tape-stripping sessions (Denda et al., 2000).

After undergoing the cold pressor test, participants had a third (and final) blood draw for immunological assays, and the second of two tape-stripping sessions. Following these tasks, subjects provided another salivary cortisol sample, and completed a thought-listing task. Subjects remained in the GCRC until 1:00 PM, with additional measurements of the tape-stripped sites, as well as salivary and blood sampling for hormone measurements.

2.4. Essential oils

A yellow-tinted cotton ball containing 100 μ l of the essential oil or distilled water was taped between the nose and upper

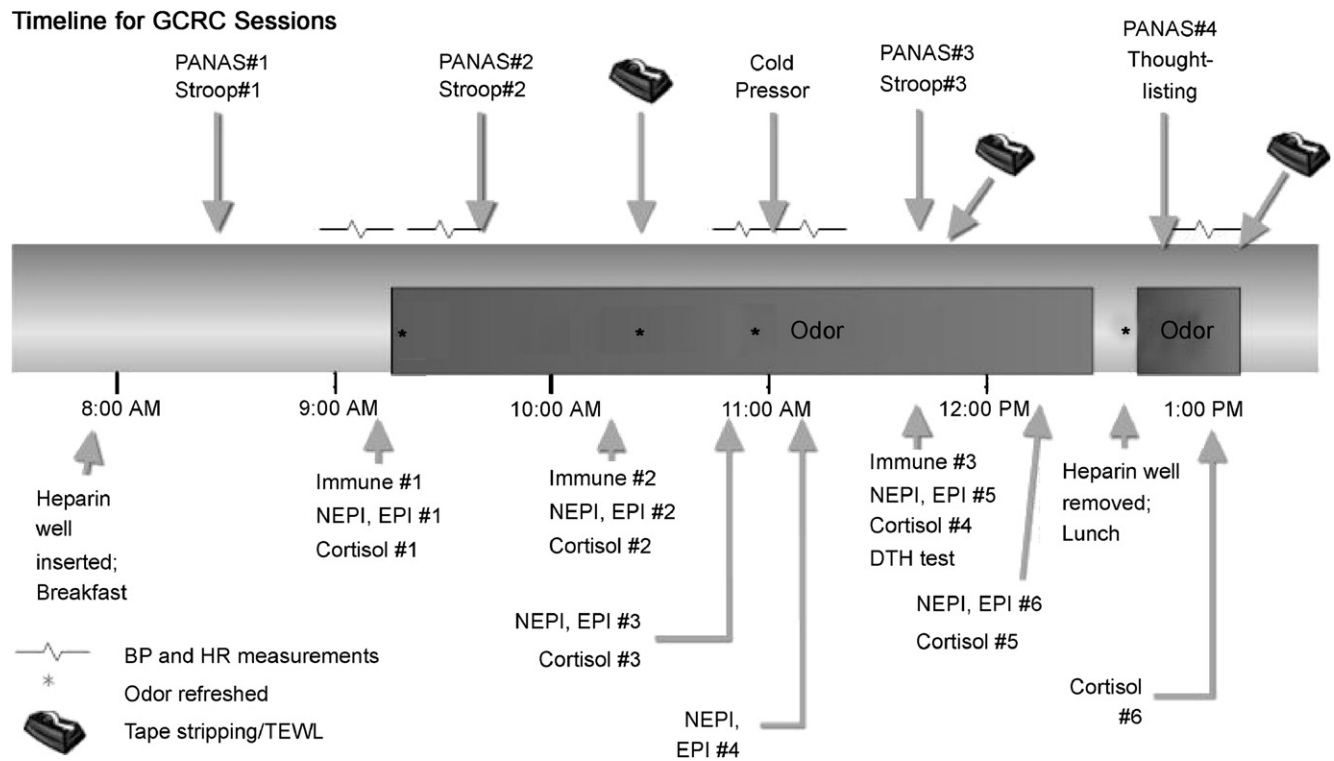


Figure 1 Timeline for experimental participation during each of the three GCRC sessions.

lip on top of a piece of surgical tape; use of the barrier tape avoided percutaneous absorption (Denda et al., 2000). This method provided continuous and uniform exposure across subjects that would not have been possible with ambient room inhalation, and helped maintain experimenter blindness. Figure 1 shows the intervals when the cotton was replaced throughout the session to maintain odor strength.

Our *Lavandula angustifolia* (true lavender) and *Citrus limonum* were essential oils steam extracted from plants (Aldrich, Milwaukee, WI). Both oils were initially characterized using mass spectrometry and gas chromatography by the vendor and confirmed by the Ohio State Chemical Instrumentation Center every 8–10 months to test variability in thawed aliquots (all from the same original lot, frozen at -80°C), focusing on the key active components of each oil.

2.5. Expectancies: blind vs. primed group information

Before being exposed to the odor for the first time in the session, all participants were given a letter in a sealed envelope. The blind group's letter thanked them for their participation and reminded them not to share any odor-related thoughts with staff members. Primed group letters included the same information given to blind group participants; in addition, they named the odor to which the participant would be exposed and provided specific physiological and psychological priming information. The lavender condition letter said that lavender is a calming odor which is sometimes used as a sleep aid for its relaxing qualities; further, it suggested that the scent might lower

heart rate and production of stress-related hormones, and that it might provoke positive or warm memories. The lemon condition letter said that lemon is a stimulating and activating odor and its scent might make participants feel more alert and energized; this letter also suggested that the scent might lift mood, increase heart rate, help them think more clearly, and perhaps also provoke positive memories and emotions. The water condition letter said that water was being used as a no-odor control to see how participants responded in the absence of an odor.

2.6. Experimenter blinding

Only the researcher who added the odor to the cotton ball knew the order in which participants were being exposed to the different odors, and she did not have any subject contact. She prepared the stimulus just prior to the scheduled time for use, and she gave the GCRC nurse the cotton balls in sealed glass vials. The nurses were told never to sniff the vials or the cotton as they taped the cotton under the subjects' noses.

A cream, Odor Perception Inhibitor (O-P-I)[®], was painted under the nose of all nurses and research assistants who had subject contact during the screening session and the admission to help keep personnel blind regarding the visit's odor assignment. The research assistants also wore a surgical mask, which enhanced the product's efficacy by trapping the O-P-I[®] odor. Further, sessions were run in a special GCRC room with reverse air ventilation, a Smok-Eter[®] system; use of the fan on high speed after each odor application minimized any traces of ambient odor.

2.7. Self-report measures

Health-related behaviors assessed at screening and each visit included recent medication use, exercise, caffeine and alcohol intake, sleep, and recent weight changes (Kiecolt-Glaser and Glaser, 1988). Physical activity was assessed with questions from Baecke et al. (1982) which have adequate reliability (.74–.88).

State mood was assessed with the two 10-item scales of the positive and negative affect schedule (PANAS) (Watson et al., 1988) which are largely uncorrelated. The scales show good convergent and discriminant validity with other state mood measures (Watson et al., 1988).

2.8. Indirect mood assessments

Picture ratings: At the screening session we used computer-presented positive and neutral pictures from the International Affective Picture System (IAPS) (Lang et al., 1999). Using IAPS normative data, we constructed three picture series, each lasting 12 min, with equivalent standardized valence and arousal data. Our participants rated the slides using the original normative study pictographs; for valence ratings participants continuously modulated a figure's face from an extreme frown to a broad smile, while for arousal they modulated a figure that showed inactivity to rapid movement.

Stroop: For the emotional Stroop, participants name the color in which negative or threatening words are printed (Williams et al., 1996). Following work by Mogg et al. (1993) we used 60 words from their lists from each of the following categories: anxiety-relevant negative words, depression-relevant negative words, positive words, and neutral household words; the word types were matched for length and frequency based on published norms (Carroll et al., 1971). The 60 words were divided into three sets of 20, with the three sets matched for length and frequency, and words were presented in random order.

The depression word interference scores for each subject and assessment time point were calculated by subtracting the mean latency for the neutral household words from those for depression-related words, and the same procedure was used for anxiety and positive words (Mogg et al., 1993). Thus, larger values indicate slower responses to emotion-relevant words than neutral household words (Mogg et al., 1993).

Thought-listing: Using a variation of a common thought-listing procedure (Cacioppo and Petty, 1981), the experimenter left the room as subjects spoke into a tape recorder for 2 min, describing their thoughts during their experimental participation. Participants were asked to describe any of their thoughts, including those about the odor, the setting, or the experimenter, and were told that we wanted an honest and complete a listing of what they actually thought about. The Linguistic Inquiry and Word Count (LIWC) program provided frequencies of positive and negative emotion word use from the tape transcripts (Pennebaker and Francis, 1999).

2.9. Cold pressor

Widely used in behavioral and psychophysiological research, the cold pressor provided a way to assess the impact of odor

on both acute pain and physiological recovery after a stressor (Blandini et al., 1995; Hirsch and Liebert, 1998). After sitting quietly for a 15-min adaptation period, participants immersed their right foot up to their ankle for 1 min in warm (37°C) water, and then immediately immersed their foot in a pan of 4°C water for 1 min (Hirsch and Liebert, 1998). Participants rated pain intensity on a 1–10 scale at the end of the cold pressor.

2.10. Immunological data

Delayed hypersensitivity to Candida: DTH memory responses to a common infectious agent provided a measure of T-cell immunity (Sokal, 1975). Nurses inoculated subjects' arm with .1 ml *Candida* (stock solution diluted 1:20 in saline, Greer Labs, NC) intradermally. The wheal diameter (two dimensions) was self-assessed at 24, 48, and 72 h by participants given detailed instructions and templates for measurement.

Blastogenesis: In all, 2×10^6 PBLs isolated from whole blood preparations were exposed to 2.5, 5.0, and 10.0 µg/ml Con A and PHA. The cells were incubated at 37°C for 65 h and then pulsed with MTS/PMS using the Celltiter 96 Aqueous nonradioactive cell proliferation assay (Promega). Cells incubated in media alone were used as controls.

Stimulated cytokine production: In all, 4×10^6 PBLs were isolated from whole blood preparations and incubated at 37°C for 72 h in 2 ml of RPMI-1640 media (Gibco) supplemented with 10% human male serum (Sigma) and 5 µg/ml of Con A (Sigma). Control cells were prepared in media without Con A. Supernatants were harvested and frozen at –86°C until assayed for IL-6 and IL-10 by ELISA using OptEIA kits from BD Biosciences.

2.11. Endocrine data

Saliva was collected using a salivette (Sarstedt, Newton, North Carolina), an untreated sterile cotton roll that was placed in the subject's mouth for ~2 min to ensure saturation, and assayed using the Cortisol Coat-A-Count RIA (Siemens Medical Solutions Diagnostics, Los Angeles, CA). Plasma catecholamine samples were frozen at –70°C and assayed by HPLC with ElectroChemical Detection using Standards and Chemistry (alumina extraction) from Thermo-Alko (Beverly, MA). All cortisol and catecholamine samples for a subject were frozen after collection and analyzed within the same assay run once the participant had completed the study.

2.12. Skin barrier assessment

Cellophane tape stripping, a common dermatological paradigm for studying restoration of the skin barrier, allowed us to examine whether the time necessary for recovery from minor physical insults varied by odor exposure. Measurement of the rate of transepidermal water loss (TEWL) through human skin provides a noninvasive method to monitor changes in the stratum corneum barrier function of the skin (Choi et al., 2005). After obtaining baseline measurements on the volar forearm, cellophane tape (3M Scotch-type; St. Paul, MN) was applied repeatedly

(6–50 times) to remove the superficial layer of cornified skin cells. Tape stripping stopped when the TEWL was elevated from the basal level of 5–7 g/h/m² to at least 20 g/h/m². The number of strips required to reach TEWL \geq 20 g/h/m² was the measure of “barrier integrity” (Choi et al., 2005). TEWL was measured several times during the session using a computerized evaporimetry instrument, the DermaLab[®] (CyberDERM, Media, PA).

2.13. Statistical methods

Unless otherwise noted, analyses consisted of repeated-measures linear models. These models accounted for the correlation in measurements from the same subject across time and visits. Independent variables assessed in each model were odor, time, gender, and expectancy group (primed vs. blind), as well as their interactions; the baseline level of the dependent variable was included as a covariate. Log-transformations were used for right skewed distributions. Post hoc tests were performed to investigate significant interactions and pairwise differences, using Holm’s or Tukey’s procedure as appropriate to adjust for multiple comparisons. A two-sided significance level of $\alpha = .05$ was used. All analyses were performed in SAS[®] version 9.1. Nonsignificant data are summarized briefly because of space constraints.

3. Results

3.1. Blind vs. primed

Subjects’ ratings of expected mood and expected physiological change were highly correlated ($r = .84$), and thus we used the mean of these two ratings to evaluate expectancies following each odor application. The significant interaction between expectancy group and odor reflected the success of the priming manipulation, $F(2, 53) = 52.34$, $p < .001$; primed subjects had relatively higher expectations for lavender (4.82, SEM = .36) and lemon (4.94, SEM = .37) than blind subjects (lavender = 3.99, SEM = .33; lemon = 4.69, SEM = .35); ratings for water showed the expected reversal of this pattern, with primed subjects reporting lower expectations (1.46, SEM = .26) than blind subjects (2.35, SEM = .25).

Participants randomized to the blind or primed expectancy conditions did not differ on age, proportion of men and women, education, sense of smell, baseline negative or positive affect, or health behaviors.

3.2. Gender and health behaviors

The men and women who participated did not differ on age, education, sense of smell, health behaviors, or negative or positive affect. There were not reliable differences in any health behaviors related to odor, with the exception of alcohol intake over the prior 48 h, $F(2, 102) = 3.48$, $p = .03$. Subjects reported fewer drinks before sessions in which they were exposed to lavender (.36, SEM = .30) compared to lemon (1.12, SEM = .28) and water (1.01, SEM = .28).

3.3. Mood

PANAS: As shown in Figure 2a, the significant time by odor interaction for the PANAS-positive affect scores, $F(6, 240) = 4.43$, $p < .001$, reflected a greater increase from baseline to post-odor when subjects smelled lemon than when they received lavender (adjusted $p = .001$) or water (adjusted $p < .004$). After the cold pressor, positive affect increased for all odors relative to the post-odor assessment, but remained highest for lemon.

PANAS-negative affect scores were not normally distributed, and thus were categorized into scores of 10, 11, and greater than 11. A repeated-measures cumulative logistic regression model fit to this three-level outcome showed a marginal expectancy condition by odor interaction, $\chi^2(2) = 5.16$, $p = .076$, as well as significant change over time, $\chi^2(3) = 12.70$, $p = .005$. For the time effect, negative affect was higher at baseline than at the three later time points combined, OR = 2.80 (1.62, 4.85), $p = .001$.

Emotional Stroop: Analysis of Stroop interference scores with all negative emotion words (both anxiety and depression related) demonstrated a significant interaction between time and odor, $F(2, 188) = 5.07$, $p = .007$. The decreased reaction time from post-odor to post-stressor for lemon contrasted with the increased reaction times for both lavender (adjusted $p = .01$) and water (adjusted $p = .04$), shown in Figure 2b.

Secondary analyses demonstrated a significant time by odor interaction for depression-related words, $F(2, 188) = 4.25$, $p = .015$, but not for anxiety-related words, $F(2, 188) = .88$, $p = .42$. The decrease in reaction time to depressive words from post-odor to post-stressor was greater for lemon than for lavender (adjusted $p = .04$) or water (adjusted $p = .02$). The time by odor interaction was marginal for positive words, $F(2, 188) = 2.83$, $p = .062$.

Positive and negative emotion word use, via LIWC: The percentage of positive and negative words used in thought listings was modeled by replacing time with emotion word use (positive or negative) in the general model described previously. The significant odor by word-type interaction, $F(2, 103) = 4.84$, $p = .01$, reflected the fact that although the percentage of positive words used by participants was higher than that of negative words at each visit, the difference was greater on the days when subjects smelled lemon (adjusted $p = .02$) or lavender (adjusted $p = .03$) than the day they smelled water (Figure 2c).

IAPS picture ratings: Emotional valence ratings from the screening session differed by odor, $F(2, 108) = 8.13$, $p < .001$, as shown in Figure 2d. Compared to lemon, subjects’ slide ratings were significantly less positive when subjects smelled lavender (adjusted $p < .001$) and water (adjusted $p = .04$). Arousal ratings did not differ by odor, $F(2, 108) = .96$, $p = .39$.

3.4. Catecholamines and cortisol

A significant time by odor interaction for norepinephrine is illustrated in Figure 3, $F(10, 456) = 2.35$, $p = .01$. The decrease in norepinephrine following the cold pressor to the first recovery time point was smaller when subjects

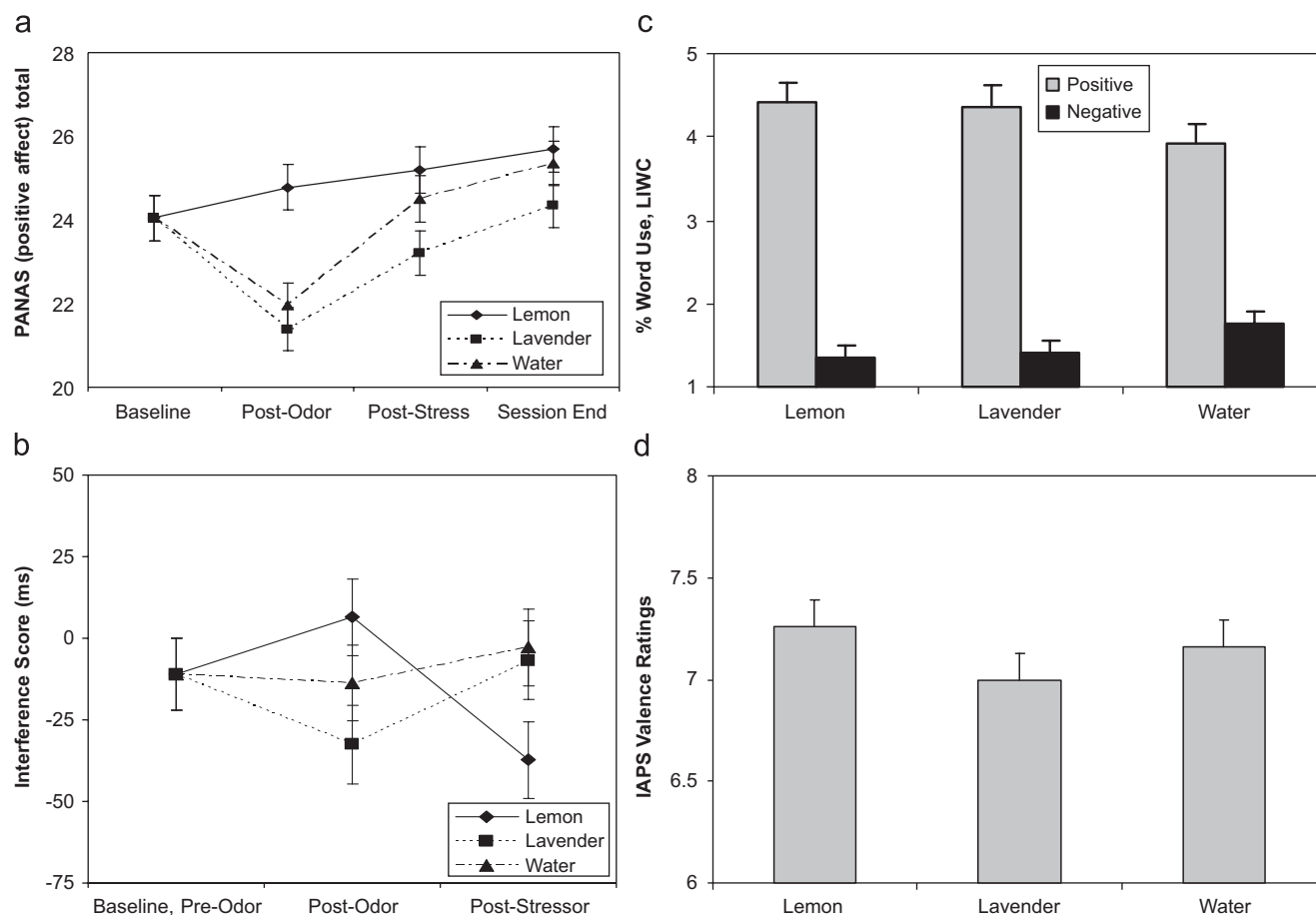


Figure 2 (a–d) Differences in one self-report and three unobtrusive mood measures as a function of odor condition. Baseline-adjusted mean (\pm SEM) changes in self-reported positive affect on the PANAS as a function of time and odor (a) showed the greatest difference immediately following the first odor application. Subjects showed greater recovery following the stressor when smelling lemon compared to lavender or water as reflected in baseline-adjusted mean (\pm SEM) interference in responses to negative emotion words on the Stroop (b). The differences in mean (\pm SEM) positive and negative emotion word use in thought listings were greater on the days when subjects smelled lemon or lavender than the day they smelled water (c). IAPS picture valence ratings from the screening session were significantly lower for lavender and water compared to lemon (d).

smelled lemon compared to water (adjusted $p = .019$). Norepinephrine returned to pre-stressor levels when subjects smelled water and lavender, but remained elevated when they smelled lemon. Analysis of epinephrine showed nonsignificant effects for odor, $F(2, 101) = 1.89$, $p = .16$, expectancy group, $F(1, 54) = .09$, $p = .76$, and the time by odor interaction, $F(10, 450) = .61$, $p = .80$.

For salivary cortisol, odor, $F(2, 98) = .19$, $p = .83$, expectancy group, $F(1, 53) = .59$, $p = .45$, and the time by odor interaction, $F(8, 380) = .64$, $p = .74$, were nonsignificant.

3.5. Heart rate and blood pressure

There was a marginal time by odor interaction for diastolic blood pressure, $F(12, 577) = 1.67$, $p = .07$, shown in Figure 4. Diastolic blood pressure decreased from post-tape stripping to post-cold pressor in subjects receiving lavender and increased in subjects receiving lemon or water. For systolic blood pressure, the odor effect, $F < 1$, $p = .50$, the

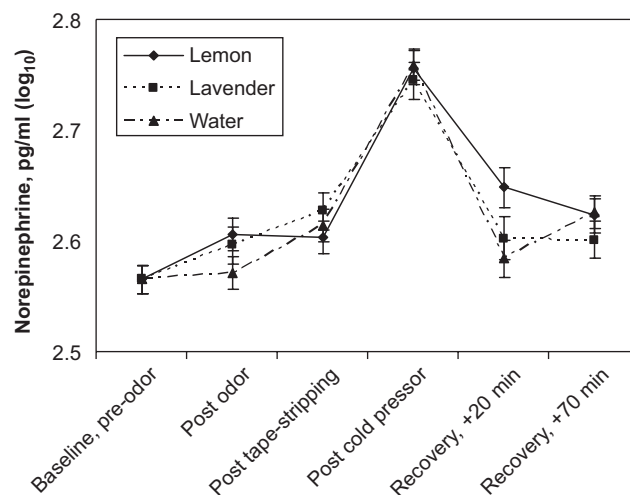


Figure 3 Baseline-adjusted mean (\pm SEM) norepinephrine levels from pre-odor exposure through the end of the admission as a function of odor exposure.

odor by time interaction, $F < 1$, $p = .99$, and expectancy group, $F(1, 53) = 1.70$, $p = .20$, were nonsignificant.

There was a significant expectancy group by time interaction for heart rate, $F(5, 266) = 2.48$, $p = .03$. Tests for the contrast for the quadratic trends across the pre-stressor, post-stressor, and recovery time points (times 3–5 in Figure 5) showed the quadratic trends were significantly different ($p = .002$) between the primed and blind groups. Primed subjects showed higher heart rate increases following administration of the cold water stressor than blind subjects, but the former also had larger decreases during recovery.

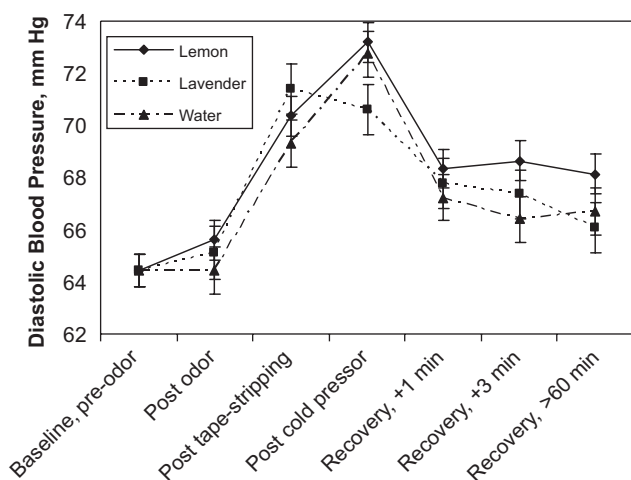


Figure 4 Baseline-adjusted mean (\pm SEM) diastolic blood pressure from pre-odor exposure through the end of the admission as a function of odor exposure. This marginal time by odor interaction, $p = .07$, shows that blood pressure decreased from post-tape stripping to post-cold pressor when subjects received lavender and increased when they receiving lemon or water.

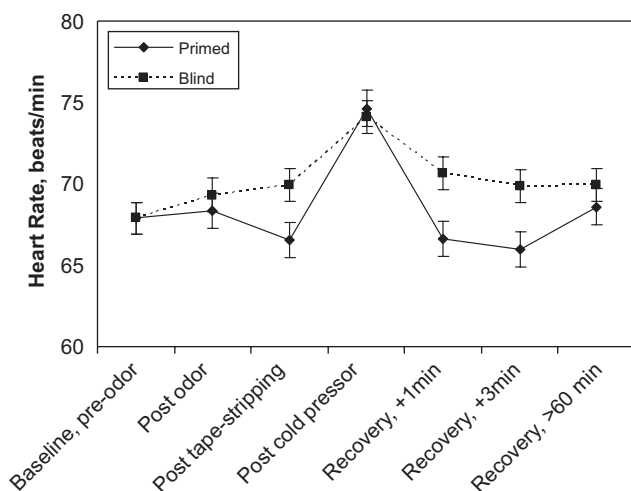


Figure 5 Baseline-adjusted mean (\pm SEM) heart rate from pre-odor exposure through the end of the admission as a function of expectancy group assignment.

3.6. Immunological data

DTH: Using the maximum DTH measurement obtained for each odor at 24, 48, or 72 h produced a significant difference in mean DTH wheal size across odors, $F(2, 92) = 4.47$, $p = .014$, reflecting the fact that subjects had larger DTH responses when they smelled water compared to lavender (adjusted $p = .02$) or lemon (adjusted $p = .06$), shown in Figure 6; blind and primed groups did not differ, $p = .35$. Two participants were excluded from DTH analyses because they never showed a positive DTH reaction to *Candida*, i.e., wheals < 5 mm at all time points for each session (Sokal, 1975). Using only the 72 h measurements produced the same significant odor effect with water the highest, $F(2, 80) = 4.71$, $p = .01$.

Production of IL-6 and IL-10: There was a marginal expectancy group effect for stimulated IL-10 production, $F(1, 192) = 3.20$, $p = .08$, with a trend toward higher production in the primed group compared to blind subjects. None of odor, the odor by time, or the odor by expectancy interactions were significant for IL-6 ($ps > .10$) or IL-10 ($ps > .50$).

Blastogenesis: There were significant odor by expectancy group by time interactions for proliferative responses to PHA, $F(2, 81) = 7.67$, $p < .001$, and Con A, $F(2, 95) = 4.37$, $p = .015$, shown in Figure 7. Controlling for pre-odor baseline, for both mitogens the effect of expectancy group on change from post-odor to post-cold pressor was significantly different for lavender than for both water (adj $p = .001$ for PHA, $p = .016$ for Con A) and lemon (adj $p = .009$ for PHA, $p = .094$ for Con A). When subjects smelled lavender, the blastogenic responses for both Con A and PHA increased in the primed group from pre- to post-stressor and decreased in the blind group, while the effect by primed group was in the opposite direction for water and lemon.

3.7. Skin barrier repair

Using the method to calculate percent recovery as described by Denda (Denda et al., 2000), recovery from the first tape stripping series improved with time, as would be expected, $F(1, 55) = 43.40$, $p < .001$, but neither odor nor the odor by time interaction was significant, $ps > .60$. Comparisons of the speed of recovery at 75 min using both the first and second tape stripping series showed no difference between

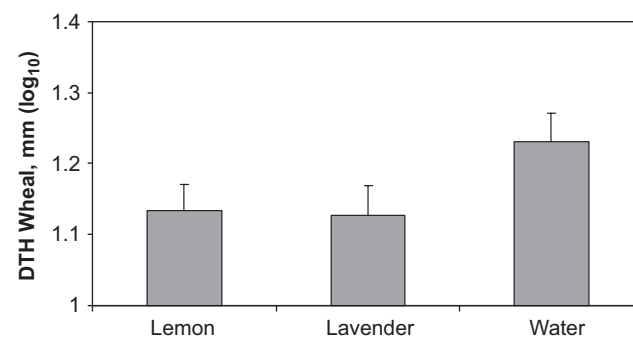


Figure 6 Mean (\pm SEM) of the maximum DTH measurement for each odor at 24, 48, or 72 h after the GCRC session.

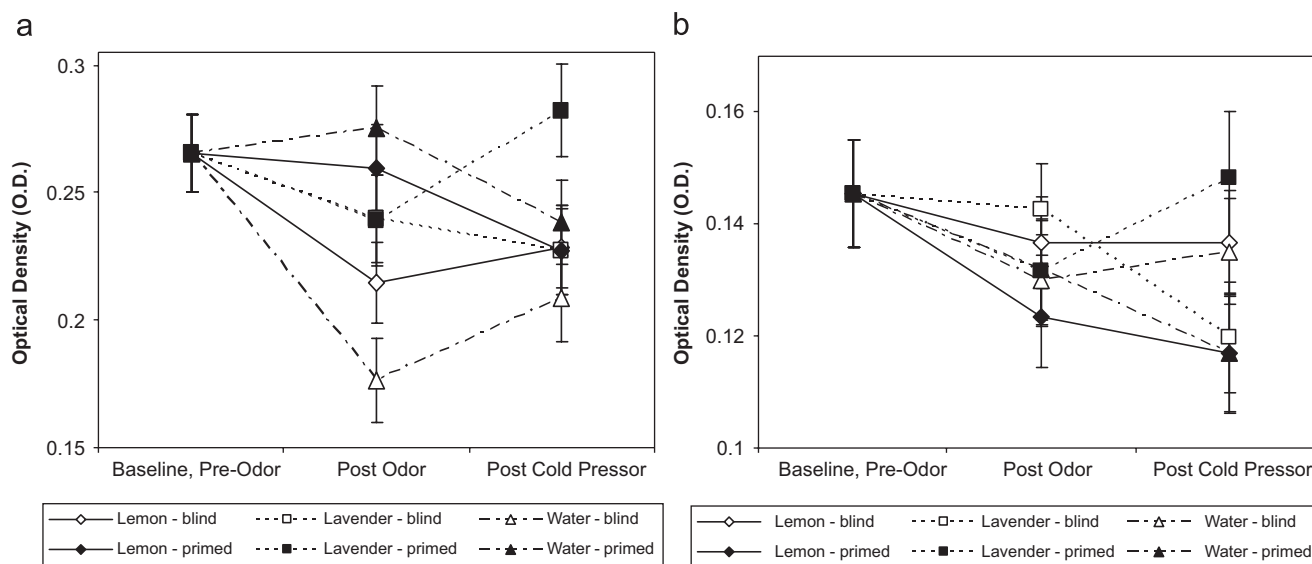


Figure 7 (a, b) Baseline-adjusted mean (\pm SEM) changes in PHA (6a) and Con A (6b) related to both odor exposure and expectancy group assignment. The values shown are means of responses to the 2.5, 5.0, and 10.0 $\mu\text{g}/\text{ml}$ concentrations.

the two, and no odor or odor by time interactions, $ps > .20$. Skin barrier integrity was poorer after the stressor than before, reflected in the fewer tape strips required to meet the TEWL criterion post-stressor, $F(1, 54) = 36.18$, $p < .001$, but neither the odor nor the odor by time interaction was significant, $p's > .14$.

3.8. Pain

Pain intensity ratings following the cold pressor did not differ among the three odors, $F < 1$, $p = .58$.

3.9. Experimenter blindness

We initially did not evaluate experimenter blindness (i.e., whether or not the experimenter guessed the day's odor correctly), but later assessed the variable and evaluated results with 30 subjects. Analyses were performed using only these 30 subjects, with one model equivalent to the primary analysis and one model with an indicator for experimenter blindness added. The results of these two models were consistent for affect and the physiological data; the results were not influenced by whether or not the experimenter correctly guessed the day's odor.

3.10. General affective or reflectorial effect theory

To address the general affective or reflectorial effect theory (Hirsch, 2001), we first examined the association between odor pleasantness and affect. Adjusting for odor, higher PANAS-positive affect scores were associated with higher odor pleasantness ratings, $F(1, 555) = 4.43$, $p = .04$. On average, PANAS-positive affect scores increased .41 points for each one unit increase in pleasantness rating. However, this association did not differ significantly between odors. Odor pleasantness ratings were not significantly associated

with negative affect on the PANAS, or with the physiological data.

3.11. Expectancies and odor awareness

In addition to differences related to the subject's blind/primed assignment, secondary analyses examined the extent to which individuals' ratings of expected mood and physiological change following the initial application of each odor were related to their subsequent pattern of responses. Participants' ratings of expected change were not significantly associated with changes in affect or physiological data.

We also asked subjects to name the odor they smelled, to see if odor awareness made a difference for "blind" subjects that might have influenced the expectancy manipulation. Lemon was correctly guessed by 77% of the blind group, compared to 7% for lavender and 46% for water ("no smell" was counted as correct for the latter). Participants' accuracy in identifying the odor was not significantly associated with changes in affect or physiological data.

3.12. Aromatherapy users vs. nonusers

Responses to the scripted interview administered during screening identified 21 subjects who had previously used aromatherapy for reasons other than enjoyment of the scent (to relax, to improve mood, etc.). Of the 21 users, 16 had used some form of aromatherapy at least monthly and rated their aromatherapy experience positively; these participants were considered enthusiastic users. Analyses of IAPS picture rating data showed a significant odor by user interaction, $F(2, 106) = 5.45$, $p = .006$; in the user group, arousal was higher in response to lemon than in response to water or lavender, while arousal did not differ by odor for non-users. For the remaining variables, inclusion of user or enthusiastic user groupings did not significantly alter

differences related to odor, primed group membership, time, or their interactions.

4. Discussion

Lemon oil's positive effects on mood were reflected in both self-reports (PANAS-positive mood scale) as well as three unobtrusive mood assessments (IAPS valence ratings, emotional Stroop responses, and emotion word use in thought listings). In contrast, lavender's effects on mood were no better than water (and sometimes more negative) on the PANAS, the Stroop, and IAPS valence ratings.

The immunological data did not support the purported clinical efficacy of lemon or lavender oil. In fact, DTH skin responses were larger following inhalation of water than lemon or lavender, and greater DTH responses are typically thought to indicate a more robust or better immune response to the antigen (Segerstrom, 2006). Odor exposure was unrelated to IL-6 and IL-10 production. The sole potentially supportive immunological finding was equivocal, as it also reflected the impact of expectancies: When subjects smelled lavender, the blastogenic response to mitogens increased in the primed group from pre-stressor to post-stressor and decreased in the blind group, while the priming effect was in the opposite direction for water and lemon.

If lemon oil inhalation had such reliably positive effects on mood, why not physiology? Marshaling evidence from experimental studies, a recent review suggested that inducing an activated state of positive affect may trigger short-term rises in various physiological systems that have transient effects on immune and autonomic function (Pressman and Cohen, 2005). For example, induction of positive affect via hypnosis or through a movie enhanced catecholamine secretion in two laboratory studies (Levi, 1965; Zachariae et al., 2001). In our study the catecholamine time by odor interaction reflected the fact that norepinephrine remained elevated 20 min after the stressor when subjects were smelling lemon, but returned to pre-stressor levels when they were smelling lavender or water. Indeed, lemon oil's effects on both norepinephrine and mood are consistent with the systemic effect theory which suggests that specific scents can evoke catecholamine changes (Feller, 1997; Price and Price, 1999). Moreover, catecholamines can inhibit the antigen-presenting capacity of Langerhans cells in the skin (Seiffert et al., 2002); thus, lemon's enhanced norepinephrine response could have contributed to the lower DTH responses observed following lemon compared to water.

Our study was designed to contrast diverse conceptual perspectives, clarify mechanisms, and assess possible clinical efficacy. In terms of conceptual perspectives, the lemon-enhanced norepinephrine and mood responses supported the systemic effect theory (Feller, 1997; Price and Price, 1999). In contrast, the only evidence that expectancies impacted responses to odors was the lymphocyte proliferation differences for lavender; we found no other significant odor-related expectancy effects in other domains. In accord with the general affective theory, lemon was perceived most positively and had the greatest positive effect on mood; however, positive mood did not translate into enhanced physical well being.

In a finding relevant to clinical efficacy, pain ratings following the cold pressor were not modulated by either lemon or lavender; by limiting the cold pressor time to a minute we had a very mild stressor that nonetheless reliably elevated catecholamines, heart rate and blood pressure, altered lymphocyte proliferation, and delayed skin barrier repair. The finding that odor exposure was unrelated to pain ratings is in accord with two well-controlled studies that used much more challenging pain stimuli than our brief cold pressor. For example, men and women who placed their hand and forearm in ice water for up to 15 min reported more pain when the ambient odor was either pleasant (lemon) or unpleasant (machine oil) compared to the no-odor condition (Martin, 2006). In another study researchers compared pain responses to contact heat, pressure, and ischemic pain in a randomized crossover design with more varied and taxing pain stimuli, e.g., ischemic pain was terminated either after the subject rated the pain as maximally unpleasant, after 15 min, or on the subject's request; their odors (lavender, rosemary, and distilled water) were not reliably related to pain ratings immediately following the stimuli (Gedney et al., 2004). The absence of any analgesic benefits across all three studies is notable, because pain reduction is a primary reason for aromatherapy's widespread use in health-related applications ranging from labor pain to post-surgical discomfort (Price and Price, 1999).

We did not find gender differences in odor responsiveness even though women have a more acute sense of smell than men as assessed by standardized tests of odor identification and detection (Bartoshuk and Beauchamp, 1994). However, our use of standardized detection tests as part of our selection criteria, an important control not found in most other aromatherapy studies, likely minimized subsequent gender differences.

Repeated or prolonged exposure to an odor can produce decrements in sensitivity to that odorant, and adaptation may also diminish behavioral responses; indeed, 30 min of exposure to lemon-smelling citral resulted in reductions in subsequent ratings of its pleasantness (Dalton and Wysocki, 1996). Accordingly, we may have had some adaptation over the interval of exposure in the GCRC. Nonetheless, in clinical applications of aromatherapy, the same odor is used over a period of several hours, and thus our paradigm provided relevant data on results that might be expected from normal aromatherapy practices. However, this study did not test the long-term effects of aromatherapy.

Many complementary/alternative therapies have not been subjected to well-controlled tests. The data from this randomized controlled trial are important because they directly address both potential mechanisms and clinical efficacy. We chose lemon and lavender because they are widely used purported stimulant and relaxant odors, and health benefits have been repeatedly ascribed to them, particularly lavender. Our sample included regular aromatherapy users as well as skeptics, so we also investigated the possibility that "true believers" might show greater benefits. We found clear and consistent evidence that lemon oil inhalation enhances positive mood and also boosts norepinephrine release (in line with its activating properties), but no other obvious physiological or health-related benefits from either lemon or lavender; indeed, the finding

that both odors appeared to depress DTH responses relative to water suggests that the immunomodulatory effects of these odors were negative, at least for this aspect of the immune response.

Role of funding sources

This research was supported by Grant AT002122 from the National Center for Complementary and Alternative Medicine (NCCAM) at the National Institutes of Health (NIH), NIH Training Grant AI55411, by General Clinical Research Center Grant MO1-RR-0034, and by Ohio State Comprehensive Cancer Center Core Grant CA16058; the NIH had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Conflict of interest

None declared.

Acknowledgments

We appreciate the helpful assistance of Cathie Atkinson, Michael DiGregorio, Bryon Laskowski, and Laura Von Hoene.

References

- Baecke, J.A.H., Burema, J., Frijters, J.E.R., 1982. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am. J. Clin. Nutr.* 36, 936–942.
- Bartoshuk, L.M., Beauchamp, G.K., 1994. Chemical senses. *Annu. Rev. Psychol.* 45, 419–449.
- Blandini, F., Martignoni, E., Sances, E., Bono, G., Nappi, G., 1995. Combined response of plasma and platelet catecholamines to different types of short-term stress. *Life Sci.* 56, 1113–1120.
- Buchbauer, G., Jirovetz, L., Jager, W., Plank, C., Dietrich, H., 1993. Fragrance compounds and essential oils with sedative effects upon inhalation. *J. Pharm. Sci.* 82, 660–664.
- Cacioppo, J.T., Petty, R.E., 1981. Electromyographic specificity during covert information processing. *Psychophysiology* 18, 518–523.
- Carroll, J.B., Davies, P., Richman, B., 1971. *Word Frequency Book*. Duke University Press, Durham, NC.
- Choi, E.-H., Brown, B.E., Crumrine, D., Chang, S., Man, M.-Q., Elias, P.M., Feingold, K.R., 2005. Mechanisms by which psychologic stress alters cutaneous permeability barrier homeostasis and stratum corneum integrity. *J. Invest. Dermatol.* 124, 587–595.
- Cowart, B., 1989. Relationships between taste and smell across the life-span. In: Murphy, C., Cain, W.S., Hegsted, D.M. (Eds.), *Nutrition and the Chemical Senses in Aging: Recent Advances and Current Research Needs*. Annals of the New York Academy of Sciences, New York, pp. 39–55.
- Dalton, P., Wysocki, C.J., 1996. The nature and duration of adaptation following long-term odor exposure. *Percept. Psychophys.* 58, 781–792.
- Denda, M., Tsuchiya, T., Shoji, K., Tanida, M., 2000. Odorant inhalation affects skin barrier homeostasis in mice and humans. *Br. J. Dermatol.* 142, 1007–1010.
- Diego, M.A., Jones, N.A., Field, T., Hernandez-Reif, M., Schanberg, S., Kuhn, C., McAdam, V., Galamaga, R., Galamaga, M., 1998. Aromatherapy positively affects mood, EEG patterns of alertness and math computations. *Int. J. Neurosci.* 96, 217–224.
- Doty, R.L., 1986. Gender and endocrine-related influences on human olfactory perception. In: Meiselman, H.L., Rivlin, R.S. (Eds.), *Clinical Measurement of Taste and Smell*. Macmillan Publishing Company, New York, pp. 377–413.
- Feller, R.M., 1997. *Practical Aromatherapy: Understanding and Using Essential Oils to Heal the Mind and Body*. Berkley Books, New York.
- Gedney, J.J., Glover, T.L., Fillingim, R.B., 2004. Sensory and affective pain discrimination after inhalation of essential oils. *Psychosom. Med.* 66, 599–606.
- Goel, N., Kim, H., Lao, R.P., 2005. An olfactory stimulus modifies nighttime sleep in young men and women. *Chronobiol. Int.* 22, 889–904.
- Grace, K., 1999. *Aromatherapy Pocketbook*. Llewellyn Publications, St. Paul, MN.
- Hardy, M., Kirk-Smith, M.D., Stretch, D.D., 1995. Replacement of drug treatment for insomnia by ambient odour. *Lancet* 346, 701.
- Hirsch, A.R., 2001. Aromatherapy: art, science, or myth? In: Weintraub, M.I. (Ed.), *Alternative and Complementary Treatment in Neurologic Illness*. Churchill Livingstone, Philadelphia, PA, pp. 128–150.
- Hirsch, M.S., Liebert, R.M., 1998. The physical and psychological experience of pain: the effects of labeling and cold pressor temperature on three pain measures in college women. *Pain* 77, 41–48.
- Jellinek, J.S., 1997. Psychodynamic odor effects and their mechanisms. *Cosmet. Toilet.* 112, 61–71.
- Keville, K., Green, M., 1995. *Aromatherapy: A Complete Guide to the Healing Art*. The Crossing Press, Freedom, CA.
- Kiecolt-Glaser, J.K., Glaser, R., 1988. Methodological issues in behavioral immunology research with humans. *Brain. Behav. Immun.* 2, 67–78.
- Knasko, S.C., 1992. Ambient odor's effect on creativity, mood, and perceived health. *Chem. Senses* 17, 27–35.
- Knasko, S.C., Gilbert, A.N., Sabini, J., 1990. Emotional state, physical well-being, and performance in the presence of feigned ambient odor. *J. Appl. Soc. Psychol.* 20, 1345–1357.
- Komori, T., Fujiwara, H., Tanida, M., Nomura, J., Yokoyama, M.M., 1995. Effects of citrus fragrance on immune function and depressive states. *Neuroimmunomodulation* 2, 174–180.
- Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1999. *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. The Center for Research in Psychophysiology, University of Florida, Gainesville, FL.
- Levi, L., 1965. The urinary output of adrenalin and noradrenalin during pleasant and unpleasant emotional states: a preliminary report. *Psychosom. Med.* 27, 80–85.
- Lis-Balchin, M., Hart, S., 2002. Chemical profiles of lavender oils and pharmacology. In: Lis-Balchin, M. (Ed.), *Lavender: The Genus Lavandula*. Taylor & Francis, New York, pp. 243–250.
- Martin, G.N., 2006. The effect of exposure to odor on the perception of pain. *Psychosom. Med.* 68, 613–616.
- Mogg, K., Bradley, B.P., Williams, R., Mathews, A., 1993. Subliminal processing of emotional information in anxiety and depression. *J. Abnorm. Psychol.* 102, 304–311.
- Pennebaker, J.W., Francis, M.E., 1999. *Linguistic Inquiry and Word Count: LIWC*. Erlbaum, Mahwah, NJ.
- Pressman, S., Cohen, S., 2005. Does positive affect influence health? *Psychol. Bull.* 131, 925–971.
- Price, S., Price, L., 1999. *Aromatherapy for Health Professionals*. Churchill Livingstone, Edinburgh.
- Segerstrom, S.C., 2006. How does optimism suppress immunity? Evaluation of three affective pathways. *Health Psychol.* 25, 653–657.
- Seiffert, K., Hosoi, J., Torii, H., Ozawa, H., Ding, W.H., Campton, K., Wagner, J.A., Granstein, R.D., 2002. Catecholamines inhibit the antigen-presenting capability of epidermal Langerhans cells. *J. Immunol.* 168, 6128–6135.

- Sokal, J.E., 1975. Measurement of delayed skin-test responses. *N. Engl. J. Med.* 293, 501–502.
- Tisserand, R., Balacs, T., 1995. *Essential Oil Safety: A Guide for Health Care Professionals*. Churchill Livingstone, Edinburgh.
- Torii, S., Fukuda, H., Kanemoto, H., Miyachi, R., Hamazu, Y., Kawasaki, M., 1988. Contingent negative variation (CNV) and the psychological effects of odour. In: van Toller, S., Dodd, G.H. (Eds.), *Perfumery: The Psychology and Biology of Fragrance*. Chapman & Hall, London, pp. 107–120.
- Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070.
- Williams, J.M.G., Mathews, A., MacLeod, C., 1996. The emotional Stroop task and psychopathology. *Psychol. Bull.* 120, 3–24.
- Zachariae, R., Jorgensen, M.M., Egekvist, H., Bjerring, P., 2001. Skin reactions to histamine of healthy subjects after hypnotically induced emotions of sadness, anger, and happiness. *Allergy* 56, 734–740.