ACUTE TOXICITY OF MARKING PEN EMISSIONS

Rosalind C. Anderson, Julius H. Anderson

Anderson Laboratories, Inc., West Hartford, Vermont, USA

To evaluate complaints of adverse reactions to marking pen emissions, groups of mice were exposed for 1 h to the emissions of 8 brands of felt-tip markers or white-board cleaner. Pneumotachographs and a computerized version of ASTM E-981 test method were used to measure changes in respiration. Sensory irritation (SI), pulmonary irritation (PI), and/or air flow limitation (AFL) of differing intensities were documented with each of the eight brands tested. At the peak of the effects, the largest SI was observed with pen F (72% of the breaths); the largest PI occurred with pen D (13% of the breaths), and the largest AFL was seen with pen F (25% of the breaths). Pens G and H produced minimal SI, PI, or AFL. A functional observational battery was used to screen for signs of neurotoxicity. Emissions from all eight of the pens produced behavioral abnormalities such as altered posture and gait, tremors, falling, and hyperactivity.

The exposure concentrations were similar to the total volatile organic compounds (TVOC) values near marking pens in actual use. Gas chromatography identified mixtures of alcohols, acetates, and/or ketones. Exposures to white-board cleaner solution resulted in similar toxicity (SI, PI, AFL, and neurotoxicity). These results document that some marking pens and white-board cleaner emit mixtures of chemicals that can produce acute respiratory toxicity and acute behavioral abnormalities in normal mice. These results provide a toxicological explanation for some of the human complaints concerning respiratory and neurological reactions to marking pen emissions.

Several complaints have been received by our laboratory concerning adverse reactions to the emissions of marking pens or air in the vicinity of white (dry-erase) boards. After sitting near students using marking pens to highlight lecture notes, one professional student experienced blurred vision, severe nasal congestion, ringing in the ears, headache, dizziness, difficulty concentrating, drowsiness, memory deterioration, decreased fine motor coordination, difficulty breathing, nausea, and swollen ankles. Other individuals have complained of burning face and eyes, difficulty breathing, mild mental confusion, and/or lightheadedness when near white boards or marking pen emissions. Although many marking pens are labeled “nontoxic,” marking pen emissions are deliberately inhaled by some drug abusers to induce solvent intoxication (McGarvey et al., 1999). Some of

Received 11 June 2002; sent for revision 9 July and 13 September 2002; accepted 29 October 2002.

Scott Hopkinson provided excellent technical assistance. Gary Cude of Matrix Analytical Labs in Addison, TX, provided excellent chemical analyses. Dr. Yves Alarie made many helpful suggestions in the preparation of the manuscript. There was no financial support for this research from outside of Anderson Laboratories (a private corporation). A preliminary report of this research appeared in Indoor Air 2002. Proceedings of the 9th International Conference on Indoor Air Quality and Climate, 30 June–5 July 2002, Monterey, CA.

Address correspondence to Rosalind Anderson, Anderson Laboratories, Inc., PO Box 323, West Hartford, VT 05084, USA. E-mail: JHARCA@hotmail.com
the complaints about marking pens might be due to solvent effects (Cornish, 1980), but other airborne chemicals might also be involved.

The discrepancy between the labels (“nontoxic”) and the human symptoms (which sound “toxic”) led us to systematically study the acute toxic potential of marking pen emissions using laboratory mice as objective test subjects. The study design was to expose mice for 1 h to the mixtures of volatile chemicals emitted by various marking pens. The mice were monitored with pneumotachographs to determine whether there were any effects on breathing rate, pattern, or air flow velocity. This is a sensitive way to detect toxic effects on the trigeminal nerve endings in the face and upper airway (sensory irritation), vagus nerve endings in the lower airway (pulmonary irritation), and bronchoconstriction effects (Vijayaraghavan et al., 1993; Anderson & Anderson, 1997, 1998, 1999a, 1999b, 1999c, 1999d, 2000). After exposures mice were examined and scored using a functional observational battery (FOB) to screen for evidence of neurotoxicity as previously done with carpet emissions (Anderson, 1995) and fragrance products (Anderson & Anderson, 1997, 1998). The emissions were subjected to gas chromatography to identify the major chemical components. Analogous studies were performed with white-board cleaner solution.

**MATERIALS AND METHODS**

**Samples**

Eight brands of marking pens and one type of white-board cleaner solution were purchased in local stores between January 1999 and December 2001. All the marking pens were labeled “nontoxic” and “conforms to ASTM D 4236”—an industry standard for labeling potentially toxic art materials (ASTM web site, www.ASTM.org, 2001). Pens A and B were made for use on white boards; the label on Pen B recommended use in “a well-ventilated space.” Pens C, D, E, F, and G were sold for general use, Pen C was permanent, Pen D contained a fragrance, and Pen E was washable. Pen H was sold for use on transparencies. The white-board cleaner was made for use as a spray; its label indicated it contained 2-butoxy ethanol/acetate and isopropanol (proportions not stated) and that it was irritating to the eyes and skin and should be used in a well-ventilated space away from children.

**Animals**

Male Swiss-Webster mice were obtained from Taconic Farms in Germantown, NY, and housed on corn-cob chips in polypropylene cages with 12 h light–12 h dark cycles according to published guidelines (Anonymous, 1996). Purina lab chow and bottled water were available except during exposures. The animal weights were between 25 and 28 g after a 1-wk acclimation. Extensive reports of source colony animal health were provided regularly by the animal supplier. Histological evaluation of lungs from our unexposed colony mice showed no signs of bacterial, viral, or parasitic infections.
Test Atmospheres

One or more marking pens were uncapped and placed in an all-glass chamber (40L; 0.25×0.30×0.50 m), which was sealed, warmed to 26°C, and allowed to equilibrate for one hr before use. The temperature in the animal chamber averaged 23°C (Cole Palmer thermister model 8402-00). The relative humidity in the animal chamber was 35% (range 20 to 50%). To study the white-board cleaner, 20 or 40g was placed in an open Petri dish inside the glass chamber and warmed as already described; about 10% of the board cleaner evaporated during these experiments.

Total Volatile Organic Compounds

The total volatile organic compounds (TVOC) in the exposure chamber was determined at 1- or 5-min intervals using flame ionization detection (Beckman Industrial, model 400A) with methane (100 ppm) as calibration gas. Additional TVOC values were determined after 1 or 4 markers were uncapped and used to draw 10-cm circles on a piece of tablet paper for 2 min. The sampling probe was positioned 26 cm above this drawing or above a sheet of glass being cleaned with white-board cleaner; 26 cm was an arbitrary choice and other distances were not examined.

Chemical Analysis of the Emissions

A 50-mg piece of felt tip was placed in a 1-L glass vessel. After 12 h the headspace was analyzed on 2 gas chromatography columns of differing polarities: (a) 30 m of 100% dimethyl polysiloxane, and (b) 30 m of 95% dimethyl–5% diphenyl polysiloxane. Peaks were detected by flame ionization and identified by retention times and referencing the material safety data (MSD) sheets related to these pens. This work was performed by Matrix Analytical Laboratories in Addison, TX.

Animal Exposures

For each test four mice were positioned in the glass exposure chamber as previously described (Vijayaraghavan et al., 1993). The head of each mouse extended into the central exposure area with the body in a side arm, which served as a whole-body plethysmograph. During a 15-min “baseline period” the animal exposure chamber was continuously ventilated with charcoal-filtered air. Five minutes later the animals began to breathe marking pen emissions carried by charcoal-filtered air, passed at 6L/min through the sample holding chamber to the animal exposure chamber and then out of the building. After 60 min the animals again inhaled charcoal-filtered air for a 15-min “recovery period.” The animals were then removed from the exposure chamber and allowed to rest for 15 min. The mice were then scored over a 20-min period using the FOB.

Over 200 mice were exposed to marking pen emissions, 4 mice per experiment. As controls, over 200 mice received sham exposures (the same insertion into the restraining tube for 95 min of monitoring while breathing
charcoal-filtered air), 4 mice per experiment. Each laboratory day involved a control experiment and three experiments with marking pen emissions. Each mouse served as its own control for statistically diagnosing adverse effects; group statistics involve pooled data from different mice tested on different days.

American Society for Testing and Materials (ASTM) E-981

ASTM E-981 is a standardized toxicological test method for measuring acute biological effects of airborne irritant chemicals (ASTM, 1984; Alarie et al., 2000). Alarie and associates (Vijayaraghavan et al., 1993, 1994; Boylstein et al., 1995) added pneumotachographs to continuously measure air flow velocity during each breath from each mouse. Digital computer programs (Alarie, 2000) sample the analog data 500 times per second and integrate flow rates and calculate volume changes during each respiratory cycle of each mouse. The program determines the duration of pause after inspiration (TB, time of break), the duration of pause after expiration (TP, time of pause), the mid-expiratory air flow velocity (VD), the respiratory rate, and the tidal volume. The software determines the mean for each parameter for each 15-s period. During the first 15 min of each experiment, approximately 3500 breaths are used to define the baseline values for each mouse.

In each experimental hour, approximately 60,000 breaths are analyzed and averaged in 15-s units. To summarize this data, the digital programs provide two complementary approaches. First, adverse effects can be directly measured as changes in TB, TP, VD, respiratory rate, and/or tidal volume. These values can be compared to baseline values or compared to values for sham-exposed (control) mice. Second, one can compare each of these parameters to the baseline values for that mouse and only report (as diagnostic abnormalities) those that reflect statistically significant deviations from the baseline mean values. A diagnosis of sensory irritation (SI) requires an increase of TB by more than 2 times the standard deviation (SD) from the baseline mean value for that mouse that day. A diagnosis of pulmonary irritation (PI) requires an increase in the TP to greater than 2 SD from the baseline mean for that mouse that day. Air flow limitations (AFL) are diagnosed when VD falls 1.5 SD below the baseline mean value for that mouse that day. Both forms of data presentation were used: the second approach yields only those changes that were statistically significant, while the first approach more directly indicates the magnitude of the adverse effects.

The computer programs allow filtering and smoothing of the data using a polynomial spline procedure named maximum likelihood estimate–generalized cross validation (MEGCVM) (Boylstein et al., 1995). To generate tables comparing brands or doses, plots of the time course for each mouse were prepared, and the peak value for each parameter or diagnosis was manually extracted and tabulated.

FOB

Each mouse was individually observed in an open field and scored for 25 parameters related to appearance and behavior (Moser & Padilla, 1998):
posture; gait; overall activity; tremors; balance; climbing ability; reach reflex; grip strength; righting reflex; response to touch, click, and tail pinch; abdominal muscle tone; and orientation and foot placement while walking on a vertical or horizontal grid. We noted (when present) twitching of an eye or ear, gasping, severe facial swelling, severe lacrimation, piloerection, petechiae in ears, bleeding from eyes or ears, body tortion, weakness or paralysis of one or more limbs, repetitive stereotyped movements (such as lip smacking or hand washing), aggressive behavior, stupor, coma, and convulsions. These evaluations were made by a single technician who was aware of the previous treatment of the mice (not blinded). Videotapes were used to document severe reactions.

Data were tabulated as either the percentage of animals showing a specific abnormality or an FOB score reflecting the mean number of abnormalities observed per mouse. Individual components of the FOB were assessed used chi-square analysis of 2×2 contingency tables with the Yates correction factor (Zar, 1984). This analysis was applied while the data were in the form of proportions; the data were later changed to percent values for clarity of presentation.

Statistics and Graphics

SigmaPlot, Version 3 (Jandel Corporation, San Raphael, CA), was used to graph the data and perform t-tests on tabulated data. The criterion for significance was set at \( p < .05 \).

RESULTS

Exposure TVOCs

The mouse exposures involved TVOC values (68 to 56,300 ppm) that overlapped the range of TVOC values measured near simulated art projects (Table 1). While some values appear to be high numbers (e.g., 56,300 ppm), none of these exposure concentrations exceeded four times the TVOC concentrations found in the simulated art project (Table 1). The TVOC values in the exposure chamber varied 10% during the 60-min exposures. The control mice inhaled charcoal-filtered air with a TVOC of 12 ppm.

Time Course of a Representative Experiment

Figure 1 shows an experiment with Pen F. Pen F emissions were present from 20 until 80 min. Effects began within seconds after introduction of the test atmosphere (TVOC=4900 ppm). The value for TB rapidly increased to 170% of baseline and then gradually returned toward baseline and remained there for the rest of the exposure. The value of TP slowly rose about 10% over the course of the exposure. VD rapidly declined to about 80% of baseline value, then fell to about 70% of baseline value toward the end of the exposure (Figure 1). The results of the statistical diagnostic program are shown in Figure 2. At the peak of the effects the TB elevations reached diagnostic levels for SI in
TABLE 1. TVOC Values

<table>
<thead>
<tr>
<th>TVOC values (ppm)</th>
<th>Mouse chamber</th>
<th>Room air</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 pen</td>
<td>4 pens</td>
</tr>
<tr>
<td>Pen A</td>
<td>3800</td>
<td>17,300</td>
</tr>
<tr>
<td>Pen B</td>
<td>5700</td>
<td>26,000</td>
</tr>
<tr>
<td>Pen C</td>
<td>8900</td>
<td>32,000</td>
</tr>
<tr>
<td>Pen D</td>
<td>1800</td>
<td>4000</td>
</tr>
<tr>
<td>Pen E</td>
<td>530</td>
<td>2400</td>
</tr>
<tr>
<td>Pen F</td>
<td>5100</td>
<td>14,900</td>
</tr>
<tr>
<td>Pen G</td>
<td>70</td>
<td>380</td>
</tr>
<tr>
<td>Pen H</td>
<td>68</td>
<td>250</td>
</tr>
</tbody>
</table>

Note. Data represent the means of two or four experiments. The mouse chamber blank (no pens present) was 12 ppm, and the room air blank was 13 ppm.

FIGURE 1. Time course of representative experiment. Emissions of Pen F (TVOC = 4900 ppm) were introduced at 20 min; at 80 min the mice again breathed charcoal-filtered air. During the baseline period (0 to 15 min) the average values (SD) were TB = 0.029 (0.006) s, TP = 0.025 (0.003) s, VD = 2.563 (.387) ml/s, tidal volume = 0.205 (0.020) ml, and respiratory frequency = 266 (35) breaths/min. The thick lines show the means of 4 mice; the thin lines show the 95% confidence intervals.

47% of the breaths; the fall in VD at the end of the exposure allowed an AFL diagnosis in 30% of the breaths (Figure 2). The small elevation in TP resulted in diagnosis of PI in less than 5% of the breaths, so this line has been omitted from Figure 2.
Variability

Swiss Webster mice are outbred and demonstrate genetic variability in their physiological responses to various stimuli; their responses to felt tip marker emissions reflected this variability. For example, the TB elevation produced by Pen F peaked at 110, 165, 210, and 240% of baseline in the 4 mice that simultaneously received this exposure (average shown in Figure 1). These peaks for TB elevations occurred between 1 and 4 min after the onset of the exposure. Subsequent statements refer in general to the averages for groups of 12 or 16 mice.

Peak Effects

The peak effects for SI were markedly different with the eight brands of pens (Table 2). Pen F induced the most frequent SI (72% of the breaths), while Pens G and H produced no statistically significant effect (Table 2). The largest elevation of TB (TB = 264% of baseline) occurred with 4 units of Pen B. In the sham exposures 7% of the breaths were classified as SI (due to 12% random fluctuations in TB values).

All pens except C produced small amounts of PI, but the results were not markedly dose related (Table 3). The maximum TP value observed was 133% of baseline with nine units of Pen D.

AFL occurred with all pens except G (Table 3). Pen E was the most potent when tested as a single unit (AFL = 17% of the breaths), while Pen F produced the largest AFL (25% of the breaths, peak effect from 9 units of Pen F).
TABLE 2. Sensory irritation

<table>
<thead>
<tr>
<th></th>
<th>1 pen</th>
<th>4 pens</th>
<th>9 pens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pen A</td>
<td>4</td>
<td>35&lt;sup&gt;a&lt;/sup&gt;</td>
<td>46&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen B</td>
<td>63&lt;sup&gt;a&lt;/sup&gt;</td>
<td>69&lt;sup&gt;a&lt;/sup&gt;</td>
<td>57&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen C</td>
<td>59&lt;sup&gt;a&lt;/sup&gt;</td>
<td>63&lt;sup&gt;a&lt;/sup&gt;</td>
<td>33&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen D</td>
<td>13&lt;sup&gt;a&lt;/sup&gt;</td>
<td>30&lt;sup&gt;a&lt;/sup&gt;</td>
<td>70&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen E</td>
<td>5</td>
<td>—</td>
<td>21&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen F</td>
<td>35&lt;sup&gt;a&lt;/sup&gt;</td>
<td>50&lt;sup&gt;a&lt;/sup&gt;</td>
<td>72&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen G</td>
<td>4</td>
<td>—</td>
<td>6</td>
</tr>
<tr>
<td>Pen H</td>
<td>4</td>
<td>7</td>
<td>12</td>
</tr>
</tbody>
</table>

Note. Summary of peak effects during experiments with one, four, or nine pens open in the sample chamber. There were 220 shams and 12 or 16 mice in the other groups. Values are means (SE mean varied from 1 to 10). The sham value for SI was 7.

<sup>a</sup>Significant at p < .05 compared to shams.

TABLE 3. Pulmonary irritation and air flow limitation

<table>
<thead>
<tr>
<th></th>
<th>PI</th>
<th>AFL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 pen</td>
<td>4 pens</td>
</tr>
<tr>
<td>Pen A</td>
<td>9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3</td>
</tr>
<tr>
<td>Pen B</td>
<td>8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Pen C</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Pen D</td>
<td>6</td>
<td>11&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen E</td>
<td>4</td>
<td>n.a.</td>
</tr>
<tr>
<td>Pen F</td>
<td>3</td>
<td>7&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pen G</td>
<td>12&lt;sup&gt;a&lt;/sup&gt;</td>
<td>n.a.</td>
</tr>
<tr>
<td>Pen H</td>
<td>8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5</td>
</tr>
</tbody>
</table>

Note. Summary of peak effects during experiments with one, four, or nine pens open in the sample chamber. There were 220 shams and 12 or 16 mice in the other groups. Values are means (SE mean varied from 1 to 10). The sham value for PI was 4; the sham value for AFL was 4; n.a., not assessed.

<sup>a</sup>Significant at p < .05 compared to shams.

The lowest VD recorded was 31% below baseline in experiments with one unit of Pen D.

**Observational Abnormalities**

High frequencies of behavioral abnormalities occurred with posture, gait, tremors, hyperactivity, facial swelling, and gasping (Table 4). Severe lacrimation was seen with Pens A, B, and E. To preserve readability not all observations are
TABLE 4. Behavioral abnormalities

<table>
<thead>
<tr>
<th>Percent of mice with abnormality</th>
<th>Sham</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>5</td>
<td>67a</td>
<td>75a</td>
<td>42a</td>
<td>33a</td>
<td>100a</td>
<td>83a</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Gait</td>
<td>8</td>
<td>42a</td>
<td>75a</td>
<td>58a</td>
<td>75a</td>
<td>56a</td>
<td>42a</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tremor</td>
<td>17</td>
<td>92a</td>
<td>94a</td>
<td>50f</td>
<td>67a</td>
<td>75a</td>
<td>75a</td>
<td>50a</td>
<td>13</td>
</tr>
<tr>
<td>Disorientation</td>
<td>2</td>
<td>8</td>
<td>6</td>
<td>17</td>
<td>33a</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Falling</td>
<td>0</td>
<td>25a</td>
<td>6</td>
<td>0</td>
<td>33a</td>
<td>31a</td>
<td>8</td>
<td>0</td>
<td>13a</td>
</tr>
<tr>
<td>Tremor</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>83a</td>
<td>56a</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Swelling</td>
<td>0</td>
<td>67a</td>
<td>31a</td>
<td>17a</td>
<td>33a</td>
<td>19</td>
<td>67a</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lacrimation</td>
<td>0</td>
<td>17a</td>
<td>37a</td>
<td>17a</td>
<td>0</td>
<td>33a</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gasping</td>
<td>0</td>
<td>83a</td>
<td>88a</td>
<td>58a</td>
<td>42a</td>
<td>44a</td>
<td>92a</td>
<td>67a</td>
<td>0</td>
</tr>
<tr>
<td>FOB score</td>
<td>4</td>
<td>12a</td>
<td>12a</td>
<td>10a</td>
<td>9a</td>
<td>9a</td>
<td>11a</td>
<td>7a</td>
<td>4</td>
</tr>
</tbody>
</table>

Note. Data indicate percent of mice with the abnormality in experiments with one pen open in sample chamber. The last line gives the FOB score, the sum of the number of abnormalities noted. The different columns represent data from different pens as indicated. There were 64 shams and 12 or 16 mice in the other groups.

*Significant at $p < .05$ compared to shams.

shown in Table 4. Avoidance behavior was observed in 50% of the mice exposed to Pen B. The hyperactivity with Pen D was so severe that 75% of these mice were recorded as “explosive.” Decreased grip strength was observed in several mice exposed to Pens D and F, and the righting reflex was slow in several mice exposed to Pen B. Higher concentrations produced similar types of abnormalities. The last line in Table 4 shows the FOB score (a simple counting of all abnormalities) for comparison.

FOB Scores

The FOB score was related to the exposure concentration as shown in Figure 3. Data are consistent with a linear relationship between FOB and log TVOC with $r^2 = .711$.

Chemical Evaluation of Pen Emissions

The pens with greatest sensory irritation (B, C, and F) emitted various different mixtures of methanol, ethanol, propanol, butanol, acetone, butyl acetate, and methyl isobutyl ketone. The least toxic pens (E, G, and H) emitted only traces of ethanol and isopropanol. Intermediate toxicity pens (A and D) emitted larger amounts of ethanol and isopropanol; D also emitted an ester. It was not possible to attribute all the irritancy to any individual chemical.

White-Board Cleaner

Mice were exposed to white-board cleaner at concentrations (Table 5) that were slightly higher than the TVOC values developed when using the cleaner on glass in an open room (4270 ppm). The lower concentration tested with mice (5400 ppm) produced SI and AFL, while the higher concentration tested
FIGURE 3. FOB score versus exposure TVOC. Each point represents data from 12 or 16 mice exposed for 1 h to marking pen emissions. Regression line has $r^2 = .711$.

TABLE 5. White-board cleaner

<table>
<thead>
<tr>
<th>Exposure TVOC (ppm)</th>
<th>SI</th>
<th>PI</th>
<th>AFL</th>
<th>FOB score</th>
</tr>
</thead>
<tbody>
<tr>
<td>5400</td>
<td>47a</td>
<td>5</td>
<td>24a</td>
<td>12a</td>
</tr>
<tr>
<td>14,700</td>
<td>54a</td>
<td>11a</td>
<td>11a</td>
<td>14a</td>
</tr>
</tbody>
</table>

Note. Sixteen mice in each test.

*Significant at $p < .05$ compared to 64 shams.

(14,700 ppm) affected all three (SI, PI, and AFL) end points (Table 5). TB maximum was 228% of baseline; TP maximum was 113% baseline, and VD minimum was 25% below baseline. The FOB scores were significantly elevated (12 to 14 abnormalities/mouse) at both exposure concentrations. The observational abnormalities with high frequency of occurrence were posture (100%), gait (83%), tremor (92%), falling (50%), hyperactivity (92%), facial swelling (42%), and gasping (83% of the mice).

**DISCUSSION**

Several brands of marking pens emitted mixtures of chemicals that produced SI, PI, AFL, and behavioral changes in normal mice. These adverse effects occurred at TVOC concentrations similar to those a human would encounter near an art project involving one to four marking pens. One can only speculate
whether there might be even higher TVOC concentrations in poorly ventilated classrooms with 30 students using marking pens continuously over 45 to 60 min. These studies involved only the mixtures of VOCs actually emitted by these pens; effects of individual components were not evaluated. Interactions between the effects of various chemical components of the emission mixtures may account for the fact that there are simple dose-response relationships present in some parts of the data but not others. For example, lines 1, 4, 6, and 8 in Table 2 show simple relationships, line 2 suggests a saturation phenomenon, and line 3 might represent a biphasic relationship. The presence of simple dose-response relationships (as in lines 1, 2, 4, 6, and 8) suggests that one or a few components of the emission mixture is dominant in producing the overall effect. In explaining a biphasic or complex dose-response relationship (as in line 3) one must consider both (a) interference between different physiological reflexes (e.g., SI reflex vs. PI reflex) and (b) interference between effects of different chemicals in the emission mixtures. For example, a mixture of a stimulant with a low ED50 and an inhibitor with a high ED50 might give a biphasic dose-response relationship. The data in Table 3 concerning PI and AFL involve smaller effects, and one must avoid overinterpretation of these numbers. Considering the potential complexities of evaluating multiple neurological effects of mixtures of chemicals, the relative simplicity of the dose-response relationship in Figure 3 is rather remarkable.

The pens involved in this study were all labeled “nontoxic.” These certification labels indicate that a toxicologist has reviewed “the complete formulas,” “each ingredient and its quantity,” “possible adverse interactions with other ingredients,” and “potential acute and chronic harm to any part of the human body” (Art & Creative Materials Institute [ACMI], 2002). These evaluations considered both children and adults with use and misuse (such as ingestion) by a small child (ACMI, 2002). The toxicity data presented in this article make one seriously question the adequacy of the approach outlined above. The anecdotal evidence reflected by human complaints about marker emissions suggests that laboratory testing allows more accurate predictions than does the review process.

These results demonstrate a toxicological basis for some of the complaints received concerning acute toxicity due to marking pen emissions. It is not known whether marking pen emissions are causing widespread health problems or whether the toxic effects are evident in only a few sensitive individuals.

SI Testing in Mice and Humans

SI results from chemical irritation of trigeminal nerves in the conjunctivae, facial skin, and nasal mucosa (Alarie, 1973, 2000; Neilsen, 1991). The microscopic details of the nerve terminations in the conjunctivae have been elaborated (MacIver & Tanelian, 1993). The receptors on the surface of these nerves have been named vanilloid receptors, because they have been extensively studied using capsaicin aerosols, and the vanilloid portion of the capsaicin molecule is
important in ligand binding (Carterina et al., 1997). The stereospecificity of these receptors has been documented (Kasanen et al., 1998; Larsen et al., 2000). When airborne irritants bind to these tetrameric receptors (Kuzhikandathil et al., 2001), there is a calcium flux, release of substance P, and local neurogenic inflammation (Bascom et al., 1997). Trigeminal nerve firing rate is proportional to the sum of the individual stimulations by various airborne chemicals (Carstens et al., 1998). The trigeminal nuclei in the pons (Chamberlin & Saper, 1998) process the afferent information and produce efferent signals in the phrenic nerves to change the respiratory pattern and rate (Alarie, 1973).

Data concerning SI in mice can be directly extrapolated to predict human experience and/or threshold limit values for occupational exposure for more than 80 volatile or organic chemicals (VOCs) (Shaper, 1993; Alarie, 1981). Because Swiss-Webster mice are less sensitive than humans to many airborne irritants (Y. Alarie, personal communication), the SI test is not likely to give “false positive results” when used as a screening tool (Tepper & Costa, 1992).

SI testing with mice has proved useful in studies of individual VOCs (Kasanen et al., 1999; Neilsen et al., 1999; Larsen & Neilsen, 2000) and mixtures of VOCs derived from mold cultures (Korpi et al., 1999) and consumer products (Anderson & Anderson, 1997, 1998, 1999a, 1999b, 1999c, 2000). The objectivity of the mouse tests has proved useful in evaluating samples of air taken from schools and offices with complaints of indoor air pollution (Anderson & Coogan, 1994; Anderson & Anderson, 1999d).

SI in humans has been studied with single VOCs and mixtures of VOCs (Hudnell et al., 1992; Cometto-Muniz et al., 1997, 1998a; Hemple-Jorgensen et al., 1999; Millqvist et al., 1999; Molhave et al., 2000). The conjunctivae of the eyes and the nasal mucosa are approximately equal in chemesthetic sensitivity (Cometto-Muniz et al., 1998b). Conjunctival hyperemia has been photographically documented (Hemple-Jorgensen et al., 1998), and mucosal inflammation (Buckley et al., 1984; Meggs et al., 1996) has been assessed by analyzing nasal exudate (Koren et al., 1992). The physical properties of volatile chemicals can be used to develop mathematical equations to predict the effects of mixtures of chemicals as sensory irritants in humans (Alarie et al., 2000).

The VOCs that produce SI in mice cause humans to experience burning of eyes, throat, skin, nose, or chest; conjunctivitis, lacrimation; coughing; and/or gagging (Alarie, 1973), plus fatigue and shortness of breath (Millqvist et al., 1999). Thus the SI observed in mice exposed to marking pen emissions and white-board cleaner probably explains the complaints of burning face and eyes in some humans exposed to marking pen emissions and air near white boards.

PI Reactions in Mice and Humans

Several of the pens and the white-board cleaner produced PI at high concentrations. Mice and humans react in a similar manner to agents that induce lower airway irritation and inflammation (Alarie, 1973), but extrapolation of PI from mice to humans is not as numerically precise as it is with SI (Schaper, 1993).
**AFL in Mice and Humans**

The VD measurement used with these mice and the FEV$_1$ measurement (forced expiratory volume in 1 s) employed to evaluate human asthmatics are both reflections of expiratory air flow velocity. The VD values in Figure 1 decreased by 30% during this exposure, similar to the decreases in FEV$_1$ seen in human asthmatics during bronchoconstriction episodes (O’Byrne et al., 1997). The procedure for measuring VD reactions in mice is relatively new, so there is not yet sufficient experience to allow direct extrapolation to humans. Although mice have fewer respiratory bronchioles (Mercer & Crapo, 1992) and faster breathing rates, both mice and humans decrease air flow velocity when challenged with cholinergic agents such as carbamylcholine and methacholine (Boylstein et al., 1995). Both mice and asthmatics react with air flow decreases on challenge with fragrance products (Kumar et al., 1995; Anderson & Anderson, 1998) and other mixtures of VOCs (Pappas et al., 2000). Until more information develops, it is prudent to believe that a chemical or mixture that produces AFL in mice is probably capable of causing air flow reduction in some humans.

Because AFL occurred during the first hour of encounter between mice and marking pen emissions, these air flow reductions probably represent acute toxic effects rather than allergic phenomena. Toxic changes in air flow are probably involved in production and/or exacerbation of some forms of occupational asthma (Chan-Yeung & Malo, 1995).

**Time Course of Respiratory Reactions**

The TB response was transient despite a constant TVOC value in the presented emissions; it is possible that the transience of the SI response reflects physiological adaptation. Similar transient responses of TB occurred in analogous experiments with fragrance products and mattresses (Anderson & Anderson, 1998, 1999c), while prolonged responses occur with individual chemicals such as propanolol and 2-chlorobenzylchloride (Vijayaraghavan et al., 1994) and with the mixtures of chemicals emitted by air fresheners, fabric softeners, and vinyl mattress covers (Anderson & Anderson, 1997, 1999a, 2000).

**FOB Scores**

The FOB scores are based on gross changes in behavior or appearance. Without more invasive experimentation one can only speculate about the mechanisms of these adverse effects. Some of the abnormalities (e.g., abnormal posture, abnormal gait, tremors, and falling) appear to reflect neurological or neuromuscular effects. Some of the other findings (e.g., facial swelling and severe lacrimation) might indicate local inflammatory responses or autonomic nervous system effects (Bascom et al., 1997). The difficulties involved in interpreting gross behavior effects of solvent toxicity have been extensively discussed (Warren et al., 2000).

The FOB results in the mice suggest a toxicological basis for some of the complaints of humans exposed to marker emissions. Data suggest some of the
VOCs entered the central nervous system and caused toxic changes that resulted in gross abnormalities in appearance or behavior of the mice. Because of the complexity of the central nervous system, one cannot make a one-to-one extrapolation of signs in mice to symptoms in humans.

**Chemistry of Emissions and Dose-Response Curve for FOB**

It is not known which of the chemicals present in the emission mixtures were most responsible for the adverse effects measured. It was assumed that the results are probably due to combined effects of several of the moieties in these mixtures. Individually, propanol, isopropanol, butanol, and methyl isobutyl ketone produce eye, nose, and throat irritation; central nervous system depression; liver and kidney toxicity; and muscle weakness or tenderness (Wilson, 1993; Anonymous, 2002; Noraberg & Arlien-Soborg, 2000; Neilsen et al., 1988). Combinations of solvents can act additively or synergistically in human toxicity studies (Cometto-Muniz et al., 1997; Noraberg & Arlien-Soborg, 2000). The apparent linearity of the data in Figure 3 would be consistent with nonspecific solvent effects as the primary mechanism of action for the neurological toxicity of these mixtures of VOCs. However, data are not sufficient for a critical evaluation of this hypothesis.

**White-Board Cleaner**

At concentrations similar to those that a human would encounter in actual use, white-board cleaner produced significant adverse effects in mice. The product label identified 2-butoxy ethanol/acetate and isopropanol as present, but the relative contributions of the two moieties to the toxic effects is not known. The label on the white-board cleaner clearly warns of irritant effects and instructs users to have a well-ventilated space. Unfortunately, many schools are using marking pens and white-board cleaner in rooms with low ventilation rates. Marking pens and white-board cleaner may be partially responsible for symptoms of indoor air pollution (eye irritation, pulmonary irritation, and neurotoxicity) in certain schools (Anderson & Coogan, 1994; Anderson & Anderson, 1999d).

**REFERENCES**


