ORD Carpet Study

Toxicology Report:

Evaluation of Off-Gassed Carpet Sample Atmospheres

Submitted: August 6, 1993

Documents Included		Pages
1.	Executive Summary	i - vi
2.	EPA Toxicology Study	EPA-1 - EPA-115
3.	Anderson Labs Toxicology Study	AL-1 - AL-84
4.	Peer Review Comments on EPA Study and Anderson Study	PR-1 - PR-27

EXECUTIVE SUMMARY

Toxicology Report: Evaluation of Off-Gassed Carpet Sample Atmospheres

August 6, 1993

Background

During the summer of 1992, through news releases, and in October, 1992 at a Senate hearing, Anderson Laboratories, Inc. released information describing neurotoxicity, pulmonary irritation and death in mice exposed to emissions from certain carpets. The information released described tests which had four primary features. First, an exposure feature, which involved placing carpet into an aquarium, heating the aquarium, and withdrawing air from the aquarium into a chamber from which four mice received their breathing air. The second feature involved modification of a standard test of sensory irritation (ASTM E 981). The standard test measures the rate at which the mice breathe during exposures. Research has shown that chemicals which humans find irritating (sensory irritation), when present in sufficient concentration, decrease the rate at which mice breathe. Further, certain chemicals, when present in sufficient concentration, change the pattern of breathing in a way that can be interpreted to indicate pulmonary irritation. Pulmonary irritation is considered to be more serious than sensory irritation. The standard ASTM method was modified to include multiple exposure periods rather than a single exposure. The third feature of the Anderson tests involves observation of the behavior of the animals after completion of the ASTM test. This behavior is scored by an observer and interpreted as an indicator of potential neurotoxicity. The fourth feature of the Anderson tests is the observation that death occurred in some mice. These descriptions of neurotoxic effects and death were particularly striking because effects of this magnitude are not consistent with our current knowledge of the emissions from carpets, and have not been described in the peer reviewed scientific literature.

The potential public health significance of these findings prompted a vigorous effort by EPA, with assistance from the Consumer Product Safety Commission, to evaluate their scientific reliability, by determining whether the findings could be replicated independently. Aside from being one of the cornerstones of science, independent replication in EPA laboratories is necessary before subsequent studies can be done to determine the causative agent(s) and the relevance of the observations for public health. This report describes the EPA effort to replicate independently the Anderson Laboratories' findings. The report also includes Anderson Labs' report of their data from the concurrently collaborative effort. Comments on these two efforts provided by four

independent peer reviewers are included with this document as well.

STUDIES by EPA's Office of Research and Development (ORD)

EPA is aware that carpet, along with other indoor products and some indoor environments in general has been associated with a variety of complaints, symptoms and signs. Based upon our knowledge to date, we have no reason to believe that the nature of complaints associated with carpet is different from the nature of complaints associated with other indoor sources. In order to establish a firm relationship between indoor sources and complaints, EPA must have available methods which are objective, sensitive, reliable and valid for this purpose. Methodologies which fit these criteria are not yet available. For this reason, the EPA's indoor air research program is identifying, developing, improving, and validating methodologies which can be used for this purpose. It is in the context of this research program that EPA began evaluating the ASTM E 981 method some time ago. Although we concluded, in a peer-reviewed paper published in the journal Indoor Environment last summer, that the value of ASTM E 981 for purposes of evaluating indoor air quality is currently limited, we nevertheless had planned to further evaluate the methodology as used by Anderson Laboratories prior to the reports of toxic effects of carpet. We were also in the process of working with other tests when the Anderson Laboratories' information came to our attention. Given the potential significance of Anderson's reports, we dedicated ourselves to investigating the four features of the Anderson Laboratories' testing described above.

ORD's goal was to understand the toxic effects reported by Anderson Laboratories. Before beginning to understand those effects, it was necessary to be able to reproduce them in our laboratory. The process by which this is done is called independent replication. Independent replication is also necessary to establish that a scientific finding is reliable. A scientific finding is reliable when it can be reproduced repeatedly by scientists who are given the protocol and the necessary equipment.

From the beginning, ORD took these studies very seriously. In addition, we remained highly confident throughout, and had every reason to believe that we would be successful in our effort to independently replicate the Anderson Laboratories' findings. Anderson Laboratories had indicated that their findings were reliable in their own laboratory, in that their testing of many different carpets revealed toxic effects. ORD assembled a large team of highly skilled scientists working on this project, and we maintained a collegial relationship with Dr. Anderson throughout. Even more to the point, part of the study team visited Anderson Laboratories in January, 1993, and not only witnessed a toxic exposure, but were able, using part of our apparatus and part of theirs, to produce lethality as well. There neither was, nor is, any doubt in the minds of the ORD team of scientists that these studies in Anderson Laboratories resulted in death to some mice.

The bottom line from our studies, however, is that despite our best efforts, which were considerable and which will be described below, we have not been able to independently replicate the severe toxicity described by Anderson Laboratories. In fact, we were not able to produce any convincing signs of even mild toxicity attributable to carpet in our tests. Our present

conclusion is that there must be an essential difference between the conditions of our experiments and those of Anderson Laboratories, which, despite our efforts, we have not been able to identify.

Following is a more detailed description of our efforts. First, the steps we took in our attempt to perform an independent replication of the Anderson Laboratories' observations will be outlined. Second, the scientific practices ORD followed during the course of these studies will be identified. Third, the results of our studies will be summarized and contrasted with the Anderson Laboratories' results. Finally, the results of an independent peer review of the ORD replication study will be presented.

The steps ORD took in an attempt to perform an independent replication of the Anderson Laboratories' observations included assembling a highly skilled scientific team, assembling the test apparatus, performing exploratory/pilot/shakedown studies, drafting a formal protocol, reciprocal visits with Anderson Laboratories' personnel, peer review of our study protocol, performing the formal replication study, analysis and peer review of the study results, performing additional exploratory studies, and initiation of a dialogue with industry scientists performing similar work.

The scientific practices ORD kept in focus during these studies were those of replication, blind testing when subjective measurements are involved, peer review, quality assurance, and inclusion of measurements which should help develop testable hypotheses to account for the observed findings. The importance of replication has already been mentioned. Scientists only accept cause-effect relationships when they can be demonstrated reliably by independent scientists.

High caliber science maintains high quality several different ways. Many scientific observations require subjective judgment or subjective scoring, particularly in the realm of biology. Because scientists recognize that whenever subjective evaluations are made, there is an unconscious tendency for expectations to color observations, good scientific practice requires that critical subjective observations be made by an individual who is not aware of the specific conditions of the test. This practice is called blinding, and in the case of these studies refers to the fact that the individual performing scoring should not know whether the data collected were from carpet-exposed mice or control mice. In ORD's formal replication, effort was made to ensure blinding of the testing laboratories by asking the Consumer Product Safety Commission to collect carpet and randomize shipment of carpet and no carpet to the test laboratories. At EPA we established elaborate procedures to ensure maintenance of this blinding.

Peer review refers to the process by which outside experts provide independent comment on the quality of a research design, the data collected and the interpretation placed upon the data. Most scientific journals send articles to peer reviewers before they will consider publication. Scientific progress is usually measured in terms of peer-reviewed publications, and EPA judges the performance of its scientific staff in large measure by the peer-reviewed publications they write. The process of publishing peer-reviewed papers is a lengthy one, and when EPA must act before important data can be published in the peer-reviewed literature, independent peer reviews are held. In the case of our carpet study, we believed that public concern was sufficient that we

could not wait for our studies to appear in the peer reviewed literature before they became known. Therefore we held independent peer reviews of both our research protocol, to ensure that the scientific design was of high quality, and of our research findings, to ensure that our data were appropriately collected and interpreted. In addition to planning for a peer review at the end of our study, we also arranged to have independent quality assurance audits of our procedures and our apparatus during the study. The purpose of these audits was to ensure that all systems were properly calibrated and operating as expected.

Finally, since ORD's expectation was that efforts would be successful in replicating the Anderson Laboratories' findings, some measurements were included which were designed to help develop testable hypotheses that would account for the findings. Since it was presumed that emissions from the tested carpets would account for the findings, evaluation of the emissions from the tested carpets was included. In addition, pieces from the carpets were analyzed directly, to determine the presence of pesticides and microbiological contaminants.

A major feature of the ORD study was collaboration with Anderson Laboratories in its design and execution. The protocol for performing the study was agreed to by both ORD and Anderson Laboratories. Each Lab received pieces from the same carpet to test, and the toxicology testing done routinely by Anderson Laboratories was included. Scoring procedures were discussed and modified to accommodate the desires of both sets of investigators. In addition, ORD performed a large number of other toxicological and analytical measurements. The analytical measurements characterizing carpet emissions and contaminants will be detailed in a separate document.

In this collaborative study, there were three treatments. Two treatments were different carpets, and one treatment was a control. Each of these three treatments was tested twice, so there was a total of 6 experiments performed. CPSC collected the carpets from sources which had previously supplied carpets to Anderson Laboratories, and which had been associated with severe toxicity and death when tested at Anderson Laboratories. CPSC randomized the 6 different sample sets, and sent a set simultaneously to EPA and to Anderson Labs for blind testing. After each laboratory had completed all 6 experiments, the code was broken for data analysis purposes. On May 26, 1993, a peer review was held of the two data sets (EPA's and Anderson Laboratories'). The peer reviewers had received draft reports from EPA and from Anderson Laboratories several days previously. At the peer review, each laboratory described their study and results, and the peer reviewers asked questions and discussed findings with study participants. The two reports, one from EPA and one from Anderson Laboratories, make clear that there was virtually nothing in common between the two sets of findings. EPA found no deaths in 24 tested animals, no severe or moderate sensory irritation, no severe or moderate pulmonary irritation, and no clear evidence of neurotoxicity. By contrast, Anderson Laboratories' findings included a total of 8 deaths in 24 tested animals, severe pulmonary irritation, and neurotoxicity.

During the course of preparing for and trying to understand these studies, EPA performed many exploratory studies (over 30) in which ORD examined a large number of variables and tested carpet samples supplied to us by Anderson Laboratories. Some of these studies were performed with as much as 10 square feet of carpet in a specially designed source chamber. Some studies

were done with very dry air, some with normal laboratory air, and some with partially humidified air. Some of the studies allowed for observation of animals for many days after exposure, and some of studies involved extensive heating of the carpet samples. In all, over 140 mice have been tested by ORD. With all of these studies, the only evidence of carpet-related neurotoxicity which we were able to produce occurred during our visit to Anderson Labs.

The science performed by EPA was characterized independently by each peer reviewers as of very high quality, yet we were not able to replicate the Anderson Laboratories' findings. The conclusion with which we are left is that very subtle but very important differences exist between the studies done by Anderson Laboratories and those done by us.

NEXT STEPS

We do not believe that the failure to replicate the Anderson Laboratories' findings proves that carpet emissions do not pose adverse health effects. At the same time, however, we do not have a sound basis for concluding that exposure to carpet emissions presents a health risk. Rather, we see two important issues. One is the meaning of the Anderson Laboratories' findings, and the other is the potential health effects of exposure to carpet emissions. These may be independent issues.

EPA will continue to follow up on the Anderson Laboratories' findings. The next step we will take is to hold a workshop at which data and hypotheses from all of the laboratories working on this problem will be presented and discussed. Based upon the outcome of this workshop, EPA will determine its next steps.

In addition to following up the Anderson Laboratories findings, EPA/ORD intends to continue its efforts to develop better methodologies for detecting and studying potential health effects of indoor sources. While many toxicology tests are available, most of them were designed to detect health effects which are very different from the type associated with most indoor air complaints. Our general strategy is to develop methods which can be used in humans, to be sure that they correlate with the symptoms and signs reported following exposure to some indoor environments. Once we have good methods for use in humans, we will develop animal models of those methods. Such models might be useful for screening purposes. Methodologies we are currently working on include tear film breakup for eye irritation, and trigeminal evoked potentials for sensory irritation. As we develop appropriate animal models, we can apply them to answer questions about the neurobehavioral effects that have been suggested by various scientists as related to indoor air complaints.

Table 1.
Summary of Toxicological Findings

Endpoint	EPA	Anderson
Deaths	none	8 of 24 (5 carpet, 3 accidental)
Severe Neurotoxicity	none	none
Neurotoxicity (extreme change from control)	none	carpet: many variables affected; control: only vocalization
Pulmonary Irritation	none	carpet: 4/16 severe; control: none
Sensory Irritation	carpet: 3/16 slight; control: 1/8 slight	carpet: 13/16 slight; control: none
General Appearance	carpet: facial swelling, lacrimation, hemorrhaging of pinna vessels, red tears; control: similar effects	carpet: facial swelling, lacrimation and bleeding, ear petechiae; control: ear petechiae

Table 2.
Summary of Analytical Chemistry/Microbiology

VOCs	- Unremarkable; most compounds in ppb range
4-PC	- Detected, but too little there to quantitate
Pesticides	- One carpet had high levels of chlorpyrifos, but not high enough to produce acute toxicity
Microbiol.	- Unremarkable

Toxicological Evaluation of Off-Gassed Carpet Sample Atmospheres

Submitted: August 6, 1993

Disclaimer

The research described in this article has been funded wholly by the U.S. Environmental Protection Agency (U.S. EPA) under contract 68-D2-0056 to ManTech Environmental Technology, Inc. It has been reviewed by the Health Effects Research Laboratory, U.S. EPA and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Acknowledgements

The authors wish to thank the following participants in the study for their effort and dedication to completing this task: Mette Jackson for exposure and pulmonary function testing; Pam Phillips for behavioral testing; Sean Dowd for computer/bioengineering support; Mary Daniels for heart puncture and blood preparation; Denise Sailstad for performing the orbital bleeds and blood preparation; Judy Richards and Rick Jaskot for clinical chemistries on serum and lavage; James Lehmann for peripheral smear, lavage, lung fixation, and cell differentials; Terisita Gabriel for hemoglobin and carboxy/methemoglobin measurements; Joel Norwood for nasal lavage and fixation; Elias Gaillard (EPL) for necropsy and histopathology evaluations; Lynne Cates (EPL) for performing the necropsy dissections; and Donald Doerfler for biostatistical analysis.

Contributors listed in alphabetical order:

Daniel L. Costa

Pulmonary Toxicology Branch, Environmental Toxicology Division, Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711

Robert S. Dyer

Office of the Associate Laboratory Director, Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711

Mark A. Mason

Indoor Air Branch, Pollution Control Division, Air and Energy Engineering Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711

Virginia C. Moser

ManTech Environmental Technology, Inc., P.O. Box 12313, Research Triangle Park, North Carolina 27709

Jeffrey S. Tepper

ManTech Environmental Technology, Inc., P.O. Box 12313, Research Triangle Park, North Carolina 27709

Table of Contents

Tables Titles 4

Figure Titles 5

Abstract 6

Introduction 8
- Background 8
- Goals 12
- Caveats 12

Methods 14
- Definitions 14
- Study Design 16
- Experimental Test Procedure 17
- Exposure System 19
- Animals 24
- Functional Observational Battery 25
- Frequency of Breathing Measurements 27
- Postmortem Evaluation 30
- General Data Analysis Strategy 32
- Quality Assurance 38

Results 42
- Exposure System 42
- Body Weight 46
- Irritancy Measurements 46
- Functional Observational Battery 51
- Postmortem Evaluation 56

Discussion 65
- Interpretation of the Results 65
- Potential Confounds 74
- Conclusion 77

References 80

Figure Legends 84

Figures 87

Appendix A: Summary of FOB Data for EPA Formal Replication Study 104

Tables Titles

Table 1 Test Matrix Using Aquarium Source Chamber

Table 2 Mouse Pulmonary Function Checklist

Table 3 Temperature (°C) - Bottom of Aquarium (S_2)

Table 4 Temperature (°C) - Inside of Aquarium (A_1)

Table 5 Temperature (°C) - Inside of Mouse Chamber (A_3)

Table 6 Percent Relative Humidity of Air Entering the Source Chamber

Table 7 Percent Decrease in Frequency of Breathing for the Mean of Four Mice

Table 8 Summary of Statistical Outcomes for Functional Observational Battery Data

Table 9 Organ Weight Data by Treatment Group

Table 10. Peripheral Blood White Cell Differential Counts

Table 11. Serum Clinical Chemistry Values

Table 12 Lung Lavage Chemistry

Table 13 Lung Lavage White Blood Cell Count Differentials ($x10^4$) /mL of Bronchoalveolar Lavage

Table 14. Incidence of Selected Histopathological Lesions

Table 15. Potential Causes of Clinical Chemistry and Differential Values

Table 16. Summary of Toxicological Findings

Figure Titles

Figure 1 Design of the Exposure System for Evaluating the Emissions

Figure 2 Change in Body Weight with Four Sequential Exposures

Figure 3 Control Period Frequency of Breathing

Figure 4 Exposure Period (a) Individual Animal Data, (b) Exposure effect

Figure 5. Recovery Period (a) Treatment Effect, (b) Exposure Effect

Figure 6. Sensory Irritation, Exposure Period

Figure 7. Pulmonary Irritation, Exposure Period

Figure 8. Disruption Index, Periods Collapsed

Figure 9. Lacrimation

Figure 10. Dilated Pinna

Figure 11. Jar Task

Figure 12. Tail Pinch

Figure 13. Click Response

Figure 14. Alertness

Figure 15. Hemoglobin

Figure 16. Clinical Chemistries: (a) protein, (b) albumin, (c) cholesterol

Figure 17. Glutathione from Nasal Lavage and Lavage Fluid Protein

Abstract

Anderson Laboratories, Inc. of Dedham, Mass. reported that the off-gassing of certain "complaint" carpets caused sensory and pulmonary irritation, changes in a checklist of neurobehavioral signs and, in the most extreme cases, death in exposed mice. This study was performed in an attempt to replicate and further investigate these findings primarily using two *in vivo* biological tests. One of these tests examines the reflex change in the breathing pattern of mice exposed to irritant airborne chemicals (American Standard Test Method for estimating sensory irritancy of airborne chemical [ASTM E-981-84]), and the other is an adaptation of a standard test method for evaluating the neurotoxic potential of chemicals (functional observational battery, Subdivision F Pesticide Assessment Guidelines-FIFRA). These two primary measures were coupled with several ancillary measures and a postmortem assessment in an attempt to ascertain the mechanism of toxicity, if observed. The postmortem evaluation included measurements of hemoglobin, serum clinical chemistries (liver, kidney and cardiac enzymes), blood and lung lavage white cell counts and differential, organ weights, and a gross necropsy with a microscopic evaluation of all major organs. The study evaluated three treatment groups comprised of two carpet exposures and one zero-grade air exposure. The investigators were blinded to the treatment group. Matched carpet samples were sent to the EPA laboratory and to Anderson Laboratories by the Consumer Product Safety Commission (CPSC). Eight mice were tested with each of the three treatment groups. There appeared to be no severe toxic effects associated with exposure to the off-gassing of the two tested carpets. Incidental findings of statistical significance, but unlikely to be of biological significance, were observed. Of all the effects that were observed, only the number of falls in the functional observational battery would be suggestive of an "adverse" effect. Several "possibly adverse" effects in each of the three categories of measurements

(irritancy, postexposure neurobehavior, and postmortem general toxicity screen) seemed to be randomly distributed across the three treatment groups. Most likely, these effects are false positives because of the numerous uncorrected comparisons performed. In conclusion, based on this assessment of irritation, neurobehavioral effects, and measures from a general toxicity screen, there is no indication that exposure to off-gassing from these two carpets poses a serious toxicological threat.

Introduction

Background Recently, Anderson Laboratories of Dedham, Mass, distributed data indicating irritancy and toxicity to mice exposed to emissions from carpets collected from sites with a history of carpet-related complaints. An *in vivo* test (ASTM E-981-84) was used to estimate irritancy in mice during exposure to air passed over warmed carpet samples in a test chamber. Briefly, carpet pieces were sealed in an aquarium that was then heated until the air temperature reached 37 °C. This air temperature was maintained for about an hour, after which room air was pulled through the aquarium and then through the mouse exposure chamber for one hour.

Measurements of changes in breathing rates and patterns were used to determine irritancy of emissions. Postexposure observations of the behavior and general condition of the mice using a list of clinical signs were used to assess correlated neurotoxicological effects. Death was observed in some exposed mice after multiple (up to four) one-hour exposures. No chemical or microbiological characterizations of the emissions from the exposure system were conducted during the tests.

These observations raise concern that some carpet emissions may adversely affect the health of exposed humans. Therefore, the EPA considered it important to (1) systematically replicate Anderson Laboratories' experiments to provide independent corroboration of test results; (2) apply the resources of the Health Effects Research Laboratory to perform a comprehensive toxicity screen to learn, if possible, what factors were responsible for the irritancy, neurobehavioral changes, and mortality observed in the test animals; and (3)

conduct a thorough chemical and microbial characterization of emissions from carpets.

This report deals only with the toxicology portion of the EPA evaluation. The EPA and Anderson Laboratories have completed a study in which the protocol required each laboratory to test three treatment conditions: subsamples of the same two carpets and an air control exposure. For these experiments, both the EPA and Anderson Laboratories were blind to the contents of the source chamber (carpet or air control). Upon completion of the study the EPA and Anderson Laboratories will exchange reports and compare results (the comparison will be the subject of a separate report). The two carpets were independently collected from their source by a third party (Consumer Product Safety Commission [CPSC]) for distribution to the test laboratories. Both carpets have previously been tested by Anderson Laboratories and were shown to produce severe toxicity in some test animals and death in others. The EPA tested subsamples of these same carpets and a zero-grade air control using methods specified by Anderson Laboratories. Additionally, postmortem evidence of toxicity in mice was evaluated. The EPA also attempted to characterize the carpet chemical and microbiological contaminants and emissions under test conditions (see companion chemistry/analytical report).

Test Method (ASTM E 981-84) Initial reports from Anderson Laboratories indicated that exposure to carpets could produce either sensory or pulmonary irritation, or a combination of both. This determination was made using an adaptation of the American Standard Test Method for Estimating the Sensory Irritancy of Airborne Chemical (ASTM E-981-84). This test method provides a quantitative estimate of the sensory irritant potential of an inhaled chemical and has been recently reviewed (Boz et al., 1992; Tepper and Costa, 1992). Irritancy

is detected by a characteristic change in the breathing pattern of mice, which results in a reduction in the breathing rate, during exposure to a test atmosphere. This characteristic response in mice has been demonstrated to qualitatively predict nose, throat, and eye irritation in humans for 51 chemicals (Alarie, 1973). A quantitative relationship has been established between published Threshold Limit Values and Time Weighted Averages (TLVs-TWAs) for 26 irritant chemicals and the concentrations at which the same chemicals reduce respiratory rate by 50% (RD_{50}) in mice. This statistical relationship indicates that 3% of the RD_{50} can be used to establish interim TLVs-TWAs, if toxicity is primarily based on sensory irritation (Alarie, 1981).

Although several types of irritant responses have been identified using ASTM E 981-84, they are broadly classified as two types, sensory and pulmonary (Alarie, 1973). "Sensory" (upper airway) irritation is usually produced by highly water-soluble chemicals such as ammonia, acrolein, and formaldehyde. Sensory irritation is caused by stimulation of the trigeminal nerve, which innervates the areas of the nose, throat, and eyes. In humans, exposures to sensory irritants readily produce a burning sensation to the eyes, often precipitating lacrimation. If the exposure is sufficient, the larynx also can be stimulated by these chemicals, producing cough. Analogous sensory stimulation can be readily identified in small laboratory mammals by a decrease in breathing rate (Alarie, 1973).

Pulmonary irritation, a second type of irritant response, is caused by stimulation of vagal nerve endings and usually occurs with less water-soluble chemicals. Prototypic examples are ozone (O_3), phosgene ($COCl_2$), and nitrogen dioxide (NO_2), chemicals that have greater peripheral lung deposition and thus, cause more damage in the deep lung.

Subjectively in humans, these chemicals may produce cough and substernal soreness without the burning sensation of the upper airways and eyes produced by the sensory irritants. In mice exposed to these irritants, breathing rate again decreases, but the breathing pattern is qualitatively different from the pattern of response seen with sensory irritation (Alarie, 1973).

Functional Observational Battery The initial results from Anderson Laboratories also included many signs of toxicity that implicated involvement of the nervous system (e.g., tremors, hindlimb weakness or paralysis, or changes in activity level). It was therefore appropriate to include neurological testing in this study. The test chosen was a functional observational battery that has been validated and is used by many organizations to screen chemicals for their neurotoxic potential. This first-tier evaluation has been recommended by several expert panels and is now required by the U.S. EPA for testing many chemicals (U.S. EPA,1991; for review, see Tilson and Moser, 1992). The purpose of this test battery is to identify potential neurotoxicants, which can then be studied in more depth to characterize any neurotoxic effects detected. The functional observational battery includes many functional indicators of the autonomic, sensory, motor, and behavioral status of the test subject. Much research has been conducted on the sensitivity, validity, and predictability of the functional observational battery using rats (for example see, Moser, 1989, 1990a, 1990b, 1991; Tilson and Moser, 1992). To a lesser extent, these specific tests have also been used in mice; however, the functional observational battery, used in rats, is based on the original screen for neurological signs in mice described by Irwin (1968). More recently, others have used the functional observational battery for mice with little or no modification (Tegeris, 1991; Beaton et al., 1990). Thus, this test procedure is well-suited to assess the possible neurotoxic effects of carpet vapors.

Postmortem Measurements There is no information currently available suggesting a mechanism for the toxicity observed by Anderson Laboratories. In an effort to obtain preliminary information pertaining to mechanism of toxic effect, a comprehensive toxicity screen was performed on each mouse in the study. The screen consisted of measurements on the blood, lung and nasal lavage, organ weights, gross necropsy and microscopic evaluation of several tissues. Analysis of the blood included determination of normal and abnormal hemoglobins, evaluation of the relative percent of different types of white blood cells (differential) and clinical chemistries to examine liver, kidney, heart, muscle and stress responses. Lung lavage was examined for evidence of inflammation (lavage differential), cell death and changes in lung permeability.

Goals The were two primary goals for this initial study: (1) Can we obtain comparable results from testing the same carpets using essentially the same equipment and procedures as those used at Anderson Laboratories? (2) If toxicity is observed in test animals, can the EPA gather information about the biological responses of the test animals and source emissions to determine the cause or develop plausible hypotheses to account for the findings?

Caveats This study must be considered **exploratory**. Due to time constraints, key variables associated with the test apparatus, preexposure sample conditioning, environmental conditions during testing and variability among samples have not been systematically examined. The effects observed on test animals at Anderson Labs reportedly have ranged from no observable effect to death within a single test that consists of four sequential one-hour exposures of four mice over a two-day period. Abnormal behavior, defined using Anderson's list of clinical observations, has ranged from transient to persistent and has not

occurred at predictable times after an exposure or exposures. Variability between emissions from subsamples of the same carpet is unknown. Quantitative and qualitative variability of emissions over the sequence of repeated heating/exposure cycles that constitute a single test is not known. Air temperature in the test chamber has been monitored during exposures at Anderson Labs; however, chamber surface and carpet temperatures have not been routinely monitored. These temperatures may or may not be critical to producing the reported effects.

Finally many biological end points, with the exception of death, used in these test procedures rely on subjective judgments by a trained observer. The functional observational battery, generally regarded as a screening tool that can be used to identify exposures that warrant further study, has been modified and adapted for these experiments. The battery requires a trained observer to make subjective evaluations of specific pre- and postexposure behaviors and appearance of the test animals. These judgments form the basis from which a determination is made as to severity of the effect of exposure. Furthermore, interpretation of pulmonary irritancy requires subjective evaluations of pulmonary waveform. Criteria were established, based on the ASTM E-981-84 for a numerical rating system to evaluate pulmonary waveforms. However, despite attempts at objectivity, these measurements are based on some subjective assumptions, and thus, interpretation between investigators may vary.

Methods

Definitions To aid the reader in understanding this document, we have attempted to use certain words in a specific and consistent manner. The following definitions will be used throughout this text.

Study The collection of six experiments using three treatment groups as defined below.

Treatment The study evaluated three treatment groups (A-C). Two of these treatment groups were exposed to carpet samples, whereas the third treatment group was exposed to humidified, zero-grade air.

Experiment The study included six experiments; each experiment used four mice. A single treatment group consisted of two replicate experiments.

Exposure Within each experiment, four one-hour exposures were performed over two days.

Period Within each exposure there were three periods: (1) a 15-minute period during which the mice were exposed in the plethysmographs to zero-grade, humidified air; (2) a 60-minute period during which mice were exposed to the airborne contents of the source chamber; and (3) and a 15-minute recovery period in zero-grade, humidified air.

Control Group Two types of controls were used in this study; each group having a unique designation. One control is the treatment group of mice exposed to only zero-grade

humidified air from the source chamber. This group, however, is always referred to as Treatment B. The second control used in the study refers to the non-exposed, unrestrained cage-control group that is used for comparison purposes in all of the postmortem evaluations. This group is referred to as the "control" group or variant of this term (e.g., cage control, unrestrained control, non-exposed control)

Source Chamber The source chamber was constructed from an aquarium and contained the carpet samples. In the zero-grade air treatment group, the source chamber was empty.

Animal Exposure Chamber The heads of the mice protruded into a glass animal exposure chamber that was connected to the source chamber during the 60-minute exposure period.

Plethysmograph Attached to the animal exposure chamber were four plethysmographs. A plethysmograph is a device for the measurement of changes in volume. Expansion of the chest (proportional to the volume inhaled) was sensed as a pressure change in the sealed plethysmograph tube because the mouse inspired from air external to the plethysmograph (i.e., from the animal exposure chamber). Frequency of breathing was instantaneously computed by calculating the time between pressure swings.

Sensory and Pulmonary Irritation A characteristic change in the pattern and frequency of breathing that occurs in mice exposed to irritant chemicals that stimulate the trigeminal nerve endings of the upper respiratory tract (Sensory Irritation) or stimulate the vagal nerve endings in the lower respiratory tract (Pulmonary Irritation).

Functional Observational Battery The functional observational battery is a test battery that has been primarily used to identify potential neurotoxicants that may require further, more in-depth evaluation. The battery includes many functional indicators of the autonomic, sensory, motor, and behavioral status of the test subject.

Study Design A total of six separate experiments were conducted, consisting of two experiments with no carpet in the source chamber (Treatment B), and four experiments using two different subsamples of carpets previously tested at Anderson Labs (Treatments A and C). Because some toxicological indicators used in this experiment rely on subjective judgments or interpretations of test animal behavior and appearance, and subjective interpretation of pulmonary waveform data, a blinding procedure was used to minimize the potential impact of negative or positive expectations of the experimenters on subjective observations. To ensure experimental blind, CPSC decided the order of testing by randomizing shipments one through six. Randomly selected subsamples of the same carpet were tested simultaneously at each laboratory. The actual test order is listed in Table 1.

Table 1. Test Matrix Using Aquarium Source Chamber

Treatment	Experiment #	Test Dates	Shipment
A-1	1	3/8 - 3/10	Carpet 1
B-1	2	3/10 - 3/12	Empty
B-2	3	3/22 - 3/24	Empty
C-1	4	3/24 - 3/26	Carpet 2
C-2	5	3/29 - 3/31	Carpet 2
A-2	6	3/31 - 4/2	Carpet 1

The CPSC shipped the appropriate subsamples (sealed in Tedlar bags) to the sample custodian at each laboratory, who received and logged the samples, loaded the source chamber, covered the sides and top with duct tape and placed the chamber in position for testing. At the completion of the four, one-hour exposures, the custodian removed the source chamber from the exposure system, removed the subsamples, and packaged and returned them to CPSC. For control tests, CPSC shipped a package containing an empty Tedlar bag. It was the responsibility of the experimenters to maintain the integrity of the blinding procedure.

Experimental Test Procedure The following summary briefly describes how the series of four, one-hour exposure tests were conducted. Details for the functional observational

battery, the method of assessing irritation (ASTM E-981-84), and the postmortem evaluations are described below. On the day before testing, three 1-ft^2 subsamples of a carpet were rolled and stapled together and then placed fiber-side down on the bottom of the source chamber. The source chamber was sealed and moved to the biology testing laboratory and the vinyl insulating blanket was placed on top of the source chamber. Meanwhile, six mice meeting the selection criteria were moved to the biology laboratory and the preexperiment functional observational battery was performed.

On the first exposure test day, the heating blankets were turned on high to achieve the inside air target temperature of 37±3 °C and outside bottom temperature 70±5 °C. Once these target temperatures were achieved, the carpet was allowed to bake under these static air flow conditions for one hour. During this time, all nylon bulkhead fittings were checked: the O_2 and CO_2 monitors were adjusted to 20.9 and 0.03%, respectively; the flow controllers, humidity sensor, and bottled air were turned on; and the vacuum rotameter was checked (Gilibrator Digital Flowmeter, NIST traceable) to ensure that the flow rate was 7 ± 0.2 LPM. The vacuum was then connected to the exhaust end of the animal exposure chamber.

Four test animals were placed in the animal exposure chamber per ASTM E-981-84 and provided with humidified clean air at 7 LPM for 15 minutes to establish baseline frequency of breathing rates and control waveform morphology (period 1). The source chamber was then connected to the exposure chamber and clean, zero-grade, humidified air was pulled through the source chamber across the carpets and into the animal exposure chamber for one hour (period 2). After one hour of exposure to effluent from the source chamber, the source chamber was again sealed and test animals were again exposed to humidified, zero-grade air

for a 15-minute period to evaluate if the animals recovered from the exposure (period 3). During these three evaluation periods (control, exposure, and recovery), magnehelic pressure, temperature from four thermocouples, humidity, and O_2 and CO_2 were monitored.

After the recovery period, the mice were immediately removed from their plethysmographs and were returned to their home cage, where their behavior and appearance was observed for 15 minutes. Two exposures were completed in one day with two hours separating the exposure periods. During this two-hour period, the heating blankets were left on and the temperature was monitored and adjusted to maintain, if possible, the temperature inside the source chamber. Following the second and fourth exposures (days 1 and 2), the mice were observed for 15 minutes, after which, each mouse was removed from the home cage and was examined and scored using the functional observational battery. This pattern was repeated on the following day such that one experiment included four, one-hour exposures conducted over a two-day period. Between the second and third exposures, carpet samples remain sealed in the source chamber overnight without additional heating.

Upon completion of the last (fourth) exposure and functional observational battery, all mice were sacrificed for the postmortem evaluation. Additionally, the flow through the entire exposure system was checked. This procedure was initiated to evaluate leaks into the exposure system from sources other than the supply cylinder air (e.g., loose fittings, holes in the duct tape front seal, etc.). The system leak check could only be performed after the fourth exposure, otherwise carpet emissions would be potentially lost.

Exposure System The design of the exposure system for evaluating the emissions of

carpets is shown in Figure 1. Basically, the system consists of a humidified clean air system, a modified aquarium as the source exposure chamber, carpet samples, a heating and insulation system for the source chamber (not shown), an animal exposure chamber, and several physical and biological monitoring systems.

Air and humidity system During the carpet exposure period (see protocol), the emissions from the source chamber were pulled through the animal exposure chamber at 7 LPM under approximately -0.06" of water pressure. Certified zero-grade air (National Welders) from a compressed gas cylinder was used to supply excess air flow (approximately 14 LPM) to the source chamber. The airflow from the cylinder was divided into two streams, with the flow rate of each stream controlled by a mass flow controller (one 10 SLPM and one 20 SLPM mass flow controller electronically controlled by a Tylan RO-28 Flow controller). The mass flow controllers were calibrated just before the first experiment (Gilibrator Digital Flowmeter, NIST traceable). One air stream passed across room-temperature distilled water held in a one liter glass impinger, the other air stream remained dry. The two flows were recombined and monitored for relative humidity (Omega RH411, Digital Thermo-Hygrometer) just before entering the source chamber. Alteration of the ratio of the wet and dry flows was used to adjust the humidity of the inlet stream. A relative humidity of 50 ± 10% was set as the target concentration. Excess airflow (approximately 7 LPM = 14 LPM cylinder - 7 LPM vacuum) was allowed to escape into the laboratory before entering the source chamber.

Source Chamber A commercially purchased 10 gallon (38 L) fish tank (Fish Pros, Raleigh NC) was used as the carpet source chamber. All aquariums were first prepared by removing the plastic rim using a hot air gun and a knife. Excess silicone adhesive was then removed

with razor blades and precision knives (X-acto). The aquarium was turned on its side so that the opening was facing outward (toward the animal exposure chamber) and the long side panels (10"x 20") became the top and the bottom. The top and sides were covered in duct tape to conceal the contents. A front cover was constructed from a 10" x 20" double thick (0.1875") plate glass. In diagonal corners of the front cover, 3" from each side, two 0.5"ID holes were fitted with nylon bulkhead fittings. This front cover was attached to the aquarium with duct tape and covered with duct tape to conceal the contents of the source chamber. A 28/12 ground glass female socket was placed on the upper 0.5"ID nylon fitting, and a 28/12 ground glass male ball joint was connected to the lower fitting. Supply air entered through this upper hole and the source chamber atmosphere exited through the lower hole into the animal exposure chamber. An additional hole was cut in the front plate to accommodate a teflon coated 6" thermocouple probe placed in a 0.25"ID nylon fitting located 2" below the inlet fitting. A heating pad (Sunbeam model E12107) was used to heat the bottom of the chamber, while the top of the chamber was heated with a wool heating throw blanket (Sunbeam model HT-1). The top, bottom, and sides of the chamber were insulated with a vinyl coated fiberglass blanket that was adjusted (opened or sealed) as necessary to achieve the desired target temperatures.

Prior to use, the aquariums were baked overnight at test temperature conditions using laboratory air at 7 LPM to flush the excess adhesive vapors. The aquariums were then washed with an Liqui-nox^{R} detergent solution and rinsed with deionized water. The aquariums were air dried or dried with a clean tissue and considered suitable for use. Just prior to the toxicology experiments, the clean dry tank was heated for one hour under test conditions and duplicate 3L ST032 sorbents were collected from the heated aquarium to establish a system

background before the carpet was installed. The carpet was installed in three 1-ft^2 sections, rolled fiber-side out and stacked in pyramid fashion diagonally on the aquarium bottom surface. During control experiments, the aquarium was sealed empty (i.e., Treatment B).

Carpet source and sample collection Two carpets were selected for evaluation from sites that have been tested at Anderson Labs and were shown to produce severe toxicity or death in mice under test conditions. Carpets were collected by the CPSC, according to protocols described briefly below. The collection, transportation, and storage processes were standardized to minimize potential differences among subsamples. Carpet samples were collected by CPSC personnel as soon as was practical after source sites were selected and access had been obtained.

Sample Collection Protocol Site data were collected using CPSC's Investigation Guideline (appendix G, Carpet Test Plan). Approximately 50-100 ft^2 of carpet was collected by CPSC from the site and stored. Each sample or portion collected was tagged with an identification number and wrapped in Tedlar. Labels with the official sample collection number were attached to the bag. A collection report was included with the sample describing the collection procedures, location, date of collection, and signature of the collecting agent. Packaged samples were placed in an appropriate shipping container and shipped to CPSC within 48 hours of collection. Samples were logged on receipt by the CPSC, inspected for shipping damage, subdivided into 1-ft^2 pieces that were labeled, randomized, and then packaged in groups of three in sealed Tedlar bags. Bags were stored at room temperature in a sample storage room. The CPSC maintained a bound log of all samples and subsamples. Subsamples were shipped via air express to EPA and Anderson labs approximately 48 hours

before testing.

Animal exposure chamber A 2.3-L glass chamber with four attached plethysmographs was purchased (Crown Glass, Somerville, NJ) to conform to ASTM E-981-84 standards. Prior to use of the chamber and between experiments, it was washed with analytical grade glass cleaner (Liqui-nox®) and thoroughly dried in a drying oven at 180° F. During exposure, mice were restrained in head-out plethysmograph tubes, according to ASTM E 981-84. Mice were loaded into the plethysmographs using a rubber stopper with a plunger that sealed the back end of the plethysmograph and pushed the hind quarter of the animal forward forcing the head through a latex collar. The neck was sealed around a 5/16" hole in a latex dam (thin gauge green #02146, Hygienic, Inc.). The latex dam was attached to the animal exposure chamber using duct tape. A 1/2" hole in the duct tape aligned with the 5/16" hole in the latex to accommodate the animal's neck. A cotton-tipped applicator was placed in front of the collar during loading to prevent the animal from biting the collars. Between exposures, collars were checked for potential leaks and changed as necessary. Plethysmographs were wiped clean with tissue and distilled water after each exposure. Latex collars were always replaced between experiments.

Exposure measurement system Three additional systems were used to monitor conditions before and during carpet emissions. Four thermocouple probes and an electronic thermometer (Omega HH21 Thermocouple Thermometer) were used to monitor and regulate the source and animal exposure chambers. The probes were located in the following positions (Figure 1): (1) S_2 (chamber bottom, outside surface, target temperature = 70 ± 5 °C); (2) S_7 (chamber top, outside surface, expected temperature = 40 ± 3 °C); (3) A_1

(source chamber air, target temperature = 37 ± 3 °C); and (4) A_3 (animal exposure chamber, expected temperature = 24 ± 2 °C). All thermocouples were calibrated before the initiation of the experiment against a NIST thermometer.

Percent oxygen (O_2) (Beckman Labman Oxygen analyzer) and carbon dioxide (CO_2) (SensorMedics Medical Gas analyzer LB-2) were monitored and recorded every 15 minutes during the experiment. A single point calibration was performed daily and multipoint calibrations using certified gas sources were performed just prior to and midway through the study. Additionally, static pressure was measured inside the animal exposure chamber using both positive and negative magnehelics (Dwyer Instrument Co.) that were capable of reading between 0 and 0.5" water pressure.

Animals Weekly, 20 male Swiss-Webster mice (viral antibody free), weighing 18 to 22 grams were delivered from Taconic Farms, NY. From each shipment of mice, sentinel animals (four to five mice per experiment, total number of sentinel animals = 27) were sent for serology, microbiology, and parasitology. The sentinel mice were found free of all common infections, including murine virus antibody, respiratory tract pathogens, endoparasites, ectoparasites, and fecal *Pseudomonas*. Mice were group housed (10 per cage) on corn cob bedding in a climate- and light- (12 hours light/12 hours dark) controlled AAALAC-accredited vivarium for at least seven days before testing. Mouse chow (Tech Lab Agway) and tap water were available *ad libitum*. Mice were considered acceptable for use after this waiting period if they appeared healthy and their weight was between 25.5 and 28.0 grams. Prior to weighing mice, the mouse scale accuracy was assessed using a 100 gram transfer standard weight. Particulate filter tops were used when transporting the animals between the vivarium and the

laboratory for exposure.

Initially, six animals that met the weight criteria were tail-marked and evaluated using the functional observational battery (see description below). Those that were acceptable (general appearance normal and could perform the tests in the neurobehavioral screen) were candidates for use in the exposure on the following day. On the day of exposure, four mice were loaded into the plethysmographs. The two remaining mice were not exposed and served as nonrestrained cage controls for postmortem evaluations. During the entire study, three mice that were initially loaded into the plethysmograph on the first exposure day were eliminated, according to ASTM E-981-84 guidelines, because of abnormally low baseline rates or misshaped respiratory waveform patterns. When this occurred, the mouse was swapped with one of the cage controls. Between exposures, the four mice in an experiment were returned to an acrylic cage with corn cob bedding located in the laboratory. This cage was covered with a filter top through which charcoal-purified, HEPA-filtered air was forced so as to maintain a chemical-free environment during the two days in which exposures were performed. While the mice lived in the laboratory, they received food and water *ad libitum*, except during exposure. The two non-exposed (unrestrained controls) mice from the original six were housed similarly, but separately from the exposed group. Besides the screening body weight measurements, body weights were also measured before each of the four exposures. For the first 15 minutes after each exposure, general appearance and activity in their home cage were also noted.

Functional Observational Battery

Test procedure The functional observational battery was conducted according to the

protocol (Appendix B, U.S. EPA Carpet Test Plan), which describes the progression of tests and standardized scoring criteria for each measure. This protocol was originally based on that used in this laboratory for rats; modifications and additions were made to stress the possible neuromuscular component of this toxicity syndrome, and to include measures which Dr. Anderson, in her experience, felt were particularly sensitive.

The tests began with the trained observer hanging the mouse by one hind leg onto the lip of a 1-gallon glass jar. The mouse had to pull up and stand upright on the jar rim, and this was repeated three times. Next, the mouse was placed horizontally across a screen, held at a 45° angle, and was given 20 seconds to either rotate and walk up the screen, walk straight across, or rotate and move down the screen; this was also repeated three times. While on the screen, the observer scored how often the mouse slipped a paw down between the wire mesh. The screen was then placed on a table edge, and the mouse, held by the tail, was lowered towards it. A positive forelimb placing response was noted when the mouse reached toward the screen. The mouse was then allowed to grab the screen and was raised up until the screen was perpendicular. The number of times the mouse dropped the screen was counted. The mouse was then held in the observer's hand, during which time its reactivity to being handled, lacrimation, palpebral closure, and salivation were ranked. In addition, the presence of piloerection; exophthalmus; cyanosis; gasping; facial swelling; or blood around the eyes, ears, or nose were noted. Body tone was scored by assessing the resistance of the abdominal muscle to light finger pressure. Holding the mouse between the palms of the hands, the observer flipped it over and scored the ease with which the mouse regained normal posture.

Open-field observations took place on the top of a laboratory cart (60 x 90 cm) with a 3" rim around the perimeter and covered with plastic-backed paper (which was changed before each test). The mouse was observed for exactly two minutes. During this time, the number of rears were counted, and gait characteristics, ataxia, degree of body tilt, alertness, and overall level of activity were scored. Descriptions were recorded for body posture and any clonic or tonic motor movements. The observer also recorded any diarrhea, excessive vocalizations, stereotypic movements, or any other atypical behaviors. The mouse's reactions to a puff of air delivered to the face, the sound of a metal clicker, or a tail pinch using forceps were then scored. Finally, the mouse was placed on the screen, which was then inverted. The time required for the mouse to climb over the edge, back to the top, was recorded, with a 60-second cut-off. During the preexposure test only, mice were given up to three tries, along with some prodding in some cases, to learn this task.

Observational testing required 5 to 10 minutes per mouse. All data were recorded on preprinted sheets and were later entered into a computer for further analysis. The observers (one performed the manipulations, the other recorded the data) were unaware of the treatment conditions. Testing was conducted in the same laboratory where exposures took place. Entry to the room was restricted during testing, and extraneous noises were held to a minimum.

Frequency of breathing measurements

Data acquisition system Four pressure transducers (Validyne DP-45) were connected to each plethysmograph via a 7" segment of thick-walled Tygon tubing. The transducers were connected to a chart recorder (Astro-Med, Dash-4) via preamplifiers (Hewlett-Packard, Model 8805B). A signal from the preamplifiers was also fed to a custom built frequency-to-voltage

converter that reported the instantaneous frequency between two peaks meeting its threshold requirements. The frequency data were then converted to a voltage and digitized by a personal computer every five seconds. A computer program converted the voltage back to frequency and stored and displayed the median value for the 12 samples as the representative frequency for each mouse during the one-minute sample. Data files for each experiment were named by date (e.g., MT921101.EXT) and stored on the hard disk.

Before each experiment, the data acquisition system was calibrated using several independent methods. The chart recorder and frequency-to-voltage converter were first tested using internal calibration circuits residing within the instruments. The preamplifiers were balanced without, and then with, the electrical load contributed by the transducer. A closed vial was loaded into the plethysmograph to simulate the displacement volume of a 25-gram mouse. A fixed volume syringe pump (approximately 0.25 mL) was oscillated at a fixed frequency (approximately 250 cps) into the constant volume plethysmograph and the amplitude of each of the four chambers was adjusted to 6 volts of the full 10-volt scale. The frequency was then noted on the frequency-to-voltage converter and hand checked on the chart recorder. Previous experiments had verified that the transducer/plethysmograph system was linear between 100 and 300 breaths per minute, and that the frequency-to-voltage converter was accurately digitized by the computer between rates of 0 and 600 cycles per second.

Once the mice were loaded into the plethysmographs, the chart recorder was run at 10 mm/second and tidal breathing of the animals was examined to see that the signal was at least 50%, but not more than 90%, of the width of that channel's chart and that the rates were

between 200 and 275 breaths per minute. The chart recorder was then set to run in dual speed mode (10 seconds at 10 mm/second, 50 seconds at 5 mm/second) and marked with the experiment and exposure information. Similar information (investigator name, animal species, strain, sex, date of birth, date of arrival, experiment number, exposure number and exposure information [up to 1 line of text], animal numbers, animal weights, and exposure protocol information [control, exposure, and recovery times]) was then entered in the computer. The chart recorder and computer program were started simultaneously so that samples would be time stamped and matched. At the conclusion of the exposure, the data files were backed up to floppy disk and stored.

Subjective evaluation of waveforms During exposure, the chart recorder was run at two speeds so that waveform morphology could be subjectively scored according to criteria established in the ASTM E-981-84 and subsequently refined in discussions between our laboratory and Anderson Laboratories. For each minute, for each mouse, a subjective score between 0 and 3 (0=none, 1=slight, 2=moderate, and 3=severe) was entered onto a preprinted data sheet in three different categories. Subjective scores were entered for sensory irritation, pulmonary irritation, and a disruption index based on the extent and severity of irregular waveforms not fitting into the above categories (e.g., movement artifacts). The disruption index primarily identified sighs (large tidal excursions) and movement artifacts. Two to three sighs or disruptions were scored as slight (1), while more frequent and prolonged disruptions were scored moderate (approximately 2 to 10 disruptions) or severe (>10 disruptions). For sensory and pulmonary irritation, if several breaths were of the characteristic shape, and those breaths did not precede or come immediately after a disruption, the entire minute period was scored by those breaths. This technique would tend to overestimate the

amount and severity of irritation compared to the decrease in frequency of breathing. All respiratory waveforms were scored by the same person, usually during each exposure.

The minute-by-minute data (90 minutes) for each animal (N=24) for each exposure (4) for the three different measures (sensory irritation, pulmonary irritation, and the disruption index) (total =25,920 hand measurements) were reduced to a single score for each period (control, exposure, and recovery) for each animal for each exposure for each measurement type. Previous analyses indicated that the highest (1 to 3) severity score that appeared four or more times would be significantly different from all lower scores. Thus, the highest severity score that occurred four times was used as the score for the entire period for that animal. The criteria for scoring a period as not zero were purposely set low (4 observations during the 60-minute exposure) to increase the sensitivity of detecting positive effects.

Postmortem Evaluation Immediately after the functional observational battery testing, mice were deeply anesthetized with halothane and the orbital plexus was tapped to obtain blood for analysis of %methemoglobin, %carboxyhemoglobin, and hemoglobin content (Operator's Manual IL282 CO-Oximeter, Instrumentation; Brown, 1980). The mouse was then exsanguinated via heart puncture, a peripheral smear was obtained, the remaining blood was centrifuged, and serum was collected and immediately frozen at -70 °C for analysis within the week. Peripheral blood cell differential counts were enumerated after applying Wright Giemsa stain.

Clinical Chemistries Clinical chemistries were performed on all available serum. For most

mice, 350 uL was obtained so that the full battery could be examined, which included bilirubin, lactate dehydrogenase (LDH), alanine aminotransferase (ALT), aspartate aminotransferase (AST), creatinine, blood urea nitrogen (BUN), 5' nucleotidase (5'-ND), glucose, alkaline phosphatase (ALP), total protein, isocitrate dehydrogenase (ICD), albumin, triglycerides, cholesterol, sorbitol dehydrogenase (SDH), and bile acids. However, for several mice with insufficient serum, clinical chemistries were not performed on SDH and bile acids. All clinical chemistries were performed using a centrifugal analyzer (Cobas Fara II clinical chemistry analyzer) using kits (Sigma) adapted for use with this analyzer.

Nasal and Lung Lavage After the mice were bled, a nasal lavage apparatus was inserted into the posterior pharynx and 0.5 mL of warmed 0.85% saline was injected through the nose and collected at the end of the nares. The mouse was then tracheostomized (20 gauge Luer adaptor) and lavaged (BAL) three times using the same aliquot of saline (0.8 mL). The nasal lavage and BAL were held on ice until centrifugation and the cell pellet and supernatant were prepared for further biochemical analyses or for evaluation of cell count and differential (Ghio et al., 1991). Cell count was obtained using a Coulter counter (Coulter Electronics, Inc.) and cell differential was enumerated after Wright Giemsa staining. The supernatant of the lavage was immediately analyzed for indicators of lung injury, protein, and LDH (Ghio et al., 1991; Smialowicz et al., 1991) using a centrifugal analyzer. Protein was measured using the Bio-Rad method for measurement of total protein using bovine serum albumin as a standard. Lactate dehydrogenase was measured using a purchased kit (Sigma). Perchloric acid (17.5 uL) was added to the supernatant from the nasal lavage prior to high speed centrifugation. The high speed supernatant was stored at -70 °C until it was analyzed for ascorbic and uric acids, and glutathione. The high speed pellet was analyzed for total protein using the

centrifugal analyzer as described above.

Necropsy and Histopathology Major organs (kidney, liver, brain, lung, nose, heart, thymus, spleen, adrenal, diaphragm, thigh muscle, stomach, duodenum, and colon) were grossly observed and removed under the supervision of a certified veterinary pathologist (Experimental Pathology Laboratories, Inc.). Organ weights for brain, heart, thymus, spleen, liver, and kidney were obtained prior to fixation. Lungs were infused with a syringe containing a volume of 10% neutral buffered formalin equivalent to the volume removed after BAL. The removed organs and remaining carcass were appropriately fixed in 10% neutral buffered formalin for histopathology and stored (Feldman and Seely, 1988). The following tissues were paraffin embedded, stained with hematoxylin and eosin, and microscopically examined for histopathological damage: kidney, liver, brain, lung, nose, heart, thymus and spleen, adrenal, diaphragm, thigh muscle, stomach, duodenum, colon, and any other tissue with gross abnormalities.

The relative degree of severity of inflammatory, degenerative, and proliferative changes were graded using the following scale: minimal (1), slight/mild (2), moderate (3), moderately severe (4), and severe/high (5). Congenital lesions were not graded, but instead were designated as present when observed.

General Data Analysis Strategy Due to the small number of carpet samples and test animals used in these exploratory experiments, detailed statistical analyses correlating variables of the test system with measures of toxicity were not feasible. However, the general data analysis strategy was designed to answer if the EPA could replicate Anderson's results

with carpets shown to produce toxicity to mice in her system. Positive answers to any of the four questions posed below would constitute a successful replication of toxicity observed at Anderson Laboratories.

1. Are deaths observed at both laboratories in test animals exposed to emissions from either carpet, but not observed from exposure to air?

2. In the absence of death, is unquestionable evidence of toxicity observed at both laboratories using the functional observational battery criteria (e.g., paralysis or seizure-like behaviors) from either carpet, but not air?

3. Do the detailed clinical chemistry, lung lavage analyses, and histopathologic evaluations suggest abnormalities in mice exposed to carpet, but not air?

4. Are severe alterations in breathing pattern and rate as defined in ASTM E-981-84 observed for carpets, but not air exposures at both laboratories?

For all biological end points, descriptive statistics (means, standard deviations, and plots) were obtained. No attempt was made to evaluate the equivalence of the replicate exposure groups because the number of animals was too small (N=4). Treatment group A was a carpet exposure, B was an exposure to zero-grade air, and C was an exposure to a different carpet sample than Treatment A. For most end points, one-way analysis of variance (ANOVA) was used as a first step to examine overall treatment effects on any single parameter. In the event of a significant treatment effect, statistical comparisons to Treatment

B (zero-grade air group) as well as the nonrestrained control group were made using appropriate parametric or nonparametric tests. The significance level for accepting false positives (Type I errors) was set at 5%. Because this study was exploratory in nature, corrections for multiple pairwise comparisons were not made, thus increasing the likelihood of calling a finding significant when it was not. This strategy was developed so that the results could be used to help define potentially important parameters for testing in later studies. Unless otherwise stated, data are presented as mean ± standard deviation for all tables and figures.

For some end points, the analysis examined the three treatment groups, A, B, and C, during or after four exposures. This type of data required a two-way ANOVA using treatment and exposure as the two factors. Treatment refers to the contents of the source chamber only, whereas exposure refers to the four consecutive exposures in any experiment regardless of the contents of the source chamber. Thus, a significant treatment group effect would indicate that a significant difference between treatment groups (A, B, or C) occurred, but it was unrelated to the number of exposures or exposure procedure (i.e., the same effect was seen after each exposure). A significant exposure effect would indicate that there were significant difference between exposures, but that it was unrelated to the treatment group (i.e., all treatment groups behaved the same way). If a significant treatment-by-exposure interaction occurred, then the effect at different exposures was different for different treatment groups. For example, if only on the last exposure day, Treatments A and C had significant decreases in the frequency of breathing, then a significant treatment-by-exposure interaction would occur. If Treatments A and C showed decreases in frequency compared to Treatment B after every exposure, there would be a significant treatment effect. If all treatment groups

increased the frequency of breathing with each exposure, then a significant exposure effect would be found.

Physical exposure data Mean values from serial measurements (temperature and humidity) and maximal deviation (magnehelic pressure, $\%O_2$, and $\%CO_2$) were reported on the data summary sheet. No formal analyses of these data were undertaken, except to note when excursions beyond the target conditions existed.

Subjective Appearance The incidence of occurrence (yes/no) for puffy faces, hemorrhaged ears, dilated ear vessels, and lacrimation were analyzed using a categorical modeling procedure for the linear analysis for nonparametric data (CATMOD; SAS 1990). Two factors, treatment and exposure, were evaluated. If no significant treatment-by-exposure effects were observed, the data were collapsed across exposure and reanalyzed to determine if any treatment effects could be detected.

Frequency of breathing One-minute measurements of breathing frequency data were collected for each mouse for the three periods during each exposure. The one-minute measurement consisted of the median of 12 instantaneous samples from the frequency-to-voltage converter that were sampled by the computer every five seconds. The median was used rather than the mean to filter outliers in the measurement sample. A spreadsheet program (Excel, Microsoft Inc.) transformed the data into percent difference by dividing the mean of the last five minutes in the control period by each minute during the exposure and recovery period, subtracting one from the ratio and multiplying by 100. The program then highlighted the highest percent increase and decrease from control during the exposure

period. Thus, for each animal at each exposure, four values were reported, the mean of the last five minutes of the control period, the greatest percent increase and decrease during the exposure period, and the mean of the recovery period. All statistical analyses were done on these data. The mean frequency of breathing for the group of four mice and the mean percent difference from the control period were also calculated for three-minute periods as per ASTM E-981-84.

Because of the repeated nature of these data (i.e., the same animal was measured four times), two-way repeated measures ANOVAs were used for the analyses. The analyses modeled treatment with three levels (A, B, and C) as one factor, and exposure was the repeated measure with four levels (1 to 4). Four analyses were performed. The first analysis examined the mean control period value for each animal for each exposure, to evaluate if the animals from different treatment groups were initially the same. Additionally, the analysis examined if the control period from subsequent exposures was the same as the initial data or whether there was a carry-over effect. Carry-over effects could result either from the treatment or the exposure procedure, and could result in differences in subsequent control periods. The second and third analyses evaluated effects during the exposure period. These analyses examined whether the greatest increase or decrease percent difference from control was affected by either the treatment or the exposure procedure. The fourth analysis evaluated whether the recovery period was different from the control period and whether carry-over effects occurred due to treatment or the exposure procedure.

Subjective scores of pulmonary waveforms The incidence of severity score data were analyzed using a categorical modeling procedure for the linear analysis for nonparametric data

(CATMOD; SAS 1990). Two factors, treatment and exposure, were evaluated. If no significant treatment-by-exposure effects were observed, the data were collapsed across exposure and reanalyzed to determine if any treatment effects could be detected. If the treatment-by-exposure interaction was significant, one-way ANOVAs were conducted on the treatment effects after each individual exposure.

Functional Observational Battery The two test groups for each treatment (sample) were combined for all analyses, thus n=8 per group. Statistical analysis procedures for the data derived from these tests have been described (Creason, 1989). Descriptive and rank data were analyzed using a categorical modeling procedure for the linear analysis for nonparametric data (CATMOD; SAS 1990). Initially, preexposure data were examined using a one-way ANOVA to determine if significant differences existed between treatment groups. For those measures that showed a significant or equivocal difference in the preexposure test, the analyses were applied to the change from baseline (delta) values rather than actual values. Overall, ANOVAs were conducted using treatment as the grouping factor, and repeated-measures across time (days 1 and 2, corresponding to the tests after the second and fourth exposures). If there was a significant treatment-by-time interaction factor, one-way ANOVAs for each day of testing were conducted. If only the treatment effect was significant, treatment groups were collapsed across the days and subjected to ANOVA. Treatment groups were compared using mean contrasts (SAS 1990).

The number of rears, being the only continuous variable, was similarly analyzed except that the GLM procedure was used (SAS, 1990). Square-root transformation was included to more closely approach a normal distribution of the data.

In all cases, probability values <0.05 were considered significant. However, p-values that approached significance ($p<0.10$) were taken into account, along with time-course or severity of effect, when evaluating the data from each measure.

Postmortem variables The analysis of most of the postmortem variables was simplified because only one value for each measured parameter existed for each animal. One-way ANOVA was used to evaluate the pooled (N=8) three treatment groups (Treatments A, B, and C). If a significant treatment effect was observed ($p\leq0.05$), two post-hoc Students t-tests were performed comparing the mean of Treatment B with that of Treatment A and C. If non-exposed control animal measurements were available for the parameter, a separate ANOVA was performed, using the non-exposed control as a fourth treatment group in the ANOVA. In the event of a significant effect using this analysis, the nonrestrained control group was compared to the restrained control group (Treatment B) to see if there was a significant effect of the exposure regimen.

Quality Assurance Because of the high visibility of this study and the importance of this study to the EPA, an internal and external quality assurance program was instituted that was specifically tailored to this study. For the biology studies conducted at the Health Effects Research Laboratory, an interim Quality Assurance Project Plan, written protocols, and operating procedures for all techniques and end points were developed. Additionally, several internal quality control checklists were developed to ensure that all steps in all procedures were carefully defined and followed. Table 2 is an example of the exposure checklist form.

Table 2. Mouse Pulmonary Function Checklist

Experiment Name: Exposure Number:

Date: Name:

	Glass chamber is clean and all ports are connected or closed.
	Dams are not jagged or punctured and completely sealed against chamber.
	Create page in logbook with animal number/weights and comments as needed.
	Mark beginning of chart with experiment name, exposure and animal numbers.
	Check O_2 setting at 20.9% and CO_2 analyzer at 0.03%
	Calibrate Frequency to Voltage converter (1st exposure only)
	Tidal signals should be triangular and of similar bandwidth (≈2/3).
	Make sure corks are tight, but stoppers are pulled back behind pressure port.
	Before exposure, turn Tylan and humidity sensor on
	Turn air cylinder on
	Humidity should read between 40 and 50%
	Water vessel is appropriately filled
	Set Vacuum rotameter and check flow with Gilabrator
	Connect supply air and vacuum, tighten fittings
	Mark beginning of exposure period.
	Record pressure on report sheet and temperature and humidity every 15 minutes
	Mark end of exposure period.
	At end of experiment, copy data onto backup floppy.
	Write O2 and CO2 peak values on report sheet
	Each day analyze data and put analysis and this form in a notebook

Tm	Humidity	Inside Temp (°C)	Top Temp (°C)	BottomTemp (°C)	Mouse Temp (°C)	
0						
15						
30						
45						
60						
75						
90						

Besides internal quality control procedures, the EPA requested an external audit by the Research Triangle Institute, to review all procedures, ensure that all critical instruments were operating correctly, verify that procedures were being followed, and recommend improvements in existing procedures. Two types of quality assurance audits were performed. A technical systems audit and a performance evaluation audit of relative humidity and temperature sensors, positive and negative magnehelics, O_2 and CO_2 monitors, and indicated flow-rate measurement instruments were performed.

The objectives of the audits were (1) to assess this study for compliance with requirements documented in the study protocol, Quality Assurance Project Plan, and operating procedures for the EPA carpet emissions project; and (2) to evaluate the study record-keeping procedures for data completeness, accuracy, traceability, and defensibility.

Due to concern that the audit personnel and activities could influence the test animals, and thereby compromise the experiment, the technical system audit was not conducted while exposures and postexposure functional observational batteries were being performed. Consequently, the audit was based on interviews with project personnel and review of project records, and not on direct observation of experimental procedures. Interviews were conducted using a checklist, and review of project records included a data tracking exercise conducted for two animals

The objective of the performance audit was to assess the performance of instruments to provide an independent evaluation of the quality of the data generated by them. Performance evaluations of the carbon dioxide and oxygen analyzer, five temperature sensors, one relative humidity sensor, two pressure sensors, and two flowrate sensors were

conducted. The oxygen analyzer was evaluated at one concentration using direct sampling of a National Institute of Standards and Technology-Standard Reference Material (NIST-SRM). Because this was a primary standard, no verification was necessary. The flow rates were measured using a soap film flow meter, and the magnehelic gauges were audited using an inclined manometer. Both of these devices are primary standards and verifications were not necessary. Temperature and relative humidity sensors were conducted according to SOP comparing NIST traceable sensors with test system probes.

Results

Exposure system

Temperature Two temperatures were considered critical to replicate the temperature conditions used at Anderson Laboratories. The first critical measurement was the air temperature in the source chamber. Although it was clear that there was a nonuniform distribution of temperatures in the source chamber, the A_1 probe site was used to estimate average source chamber air temperature. Attempts to keep this temperature stable (37±2 °C), despite the crude method of heating, were successful for all 24 exposures (Table 3). Similarly, attempts to keep the outside bottom temperature (S_2) at 70±5 °C were successful; however, there were specific days when we could never achieve the target temperature of 70 °C (Table 4). Despite our best attempts, the temperature in the plethysmograph hovered at the outer limits of acceptability (24±2 °C) for 16 of the 24 exposures (Table 5). This was primarily due to changes in the laboratory ambient temperature. With the discovery and subsequent repair of a leaky room air conditioning unit, the final experiment more closely approached the target temperature.

Table 3. Temperature (°C) - Inside of Aquarium (A_1)

TREATMENT		A		B		C	
EXPERIMENT #		9317	9322	9318	9319	9320	9321
	EXP 1	37.7	36.0	37.4	37.3	36.9	38.3
	EXP 2	37.0	36.1	36.9	37.6	37.3	36.3
	EXP 3	37.3	37.6	39.3	37.6	37.6	36.4
	EXP 4	37.7	36.3	36.9	37.7	36.9	36.1

Table 4. Temperature (°C) - Bottom of Aquarium (S_2)

TREATMENT		A		B		C	
EXPERIMENT #		9317	9322	9318	9319	9320	9321
	EXP 1	69.9	71.9	64.0	65.4	71.7	72.6
	EXP 2	71.4	72.7	64.7	66.1	67.6	72.9
	EXP 3	71.3	73.7	64.4	65.9	70.7	73.6
	EXP 4	71.6	72.6	64.4	66.6	70.3	73.0

Table 5. Temperature (°C) - Inside of Mouse Exposure Chamber (A_3)

TREATMENT		A		B		C	
EXPERIMENT #		9317	9322	9318	9319	9320	9321
	EXP 1	25.6	22.1	25.4	25.6	26.1	24.7
	EXP 2	25.7	23.1	25.4	26.4	**30.0**	23.6
	EXP 3	25.3	23.1	25.4	25.9	26.1	23.4
	EXP 4	25.4	22.7	25.0	26.0	26.3	22.7

Bold number exceeded study limit

Humidity The target relative humidity of 50±10% was not achieved. Although the humidity probe was audited just prior to the study, and typically, such probes maintain their calibration for months, unbeknownst to us, at some time just before the study the sensor stopped working correctly. This error was discovered during the quality assurance audit after the study was complete. Using a recently calibrated sensor, we were able to reconstruct the ratio of dry and wet flows and calculate the actual relative humidity of air entering the source chamber. These values are listed in Table 6. The actual humidity range was between 18.7% and 29.2%. If experiments are examined in date order, the progressive failure of the relative humidity sensor can be observed.

Table 6. Percent Relative Humidity of Air Entering the Source Chamber

TREATMENT		A		B		C	
EXPERIMENT #		9317	9322	9318	9319	9320	9321
EXPOSURE DATE		3/9-10	4/1-2	3/11-12	3/23-24	3/25-26	3/30-31
	EXP 1	**29.2%**	**18.7%**	**27.6%**	**21.2%**	**19.9%**	**19.4%**
	EXP 2	**29.2%**	**19.9%**	**27.6%**	**24.2%**	**19.9%**	**21.0%**
	EXP 3	**27.6%**	**19.4%**	**27.2%**	**20.3%**	**20.8%**	**19.4%**
	EXP 4	**27.6%**	**18.7%**	**27.2%**	**20.3%**	**20.8%**	**20.3%**

Bold numbers are outside study target range

System Flow, Static Pressure, Oxygen and Carbon Dioxide These four measures were made to support the critical measurements described above. System flow, when measured at the vacuum source, was always within 3% of the target flow (7 LPM) and within the limitations of the normal house vacuum fluctuations. At the end of each experiment, flow through the entire system was checked to evaluate possible leaks in the system. Flow through the entire system indicated that no more than 4% leakage occurred, which is well within measurement error, indicating that there were no leaks. Static pressure was monitored to make sure that the animals were not exposed for prolonged periods to excessive negative or positive pressure (±0.3" water pressure). During all conditions, static pressure was negative (vacuum driven system) and ranged between -0.05 and -0.075" water pressure, except for very brief excursions (<10 seconds) when the source chamber was connected to the animal exposure

chamber. Larger excursions in pressure were tolerated at this time period to ensure that none of the contents of the source chamber escaped into the laboratory. Oxygen and carbon dioxide were monitored with no unusual findings, given the limits of accuracy and reliability of these instruments. The poststudy audit of the flow meter, the magnehelics and the oxygen and carbon dioxide monitors indicated that all instruments were functioning properly and were accurate within the limitations of the instrument.

Body Weight All mice used in the exposure studies met the weight requirement for inclusion in the study one day prior to testing (25.0 to 28.0 grams). Mice gained weight between the pretest day and the first exposure day. Body weight on the first day of exposure for all 24 animals ranged from 25.8 to 28.9 grams. Treatment (A, B, or C) had no significant effect on body weight. On the other hand, the experimental procedure caused a reduction in body weight ($p<0.001$), with each succeeding exposure, causing a further decrement in body weight (Figure 2). On average, animals lost 8.2% of their body weight between the beginning of the first and the end of the last exposure.

Irritancy Measurements

Rate changes Frequency of breathing was one of the primary measurements evaluated in this study. As described in the methods section, the data were analyzed in two ways. First, the data were analyzed according to the ASTM E-981-84 procedure using the group mean (N=4). To increase sensitivity, a second analysis was performed using individual animal data.

ASTM E-981-84 Analysis The ASTM method specifies that the greatest percent reduction in frequency for the mean of all four animals is to be used as "the" response for the test. Thus, statistics could not be performed because the raw data (Table 7) consisted of only six data

points per exposure (i.e., two per treatment group per exposure).

Table 7. Percent Decrease in Frequency of Breathing for the Mean of Four Mice

Treatment	Sample #	Exposure 1	Exposure 2	Exposure 3	Exposure 4
A-1	1	6.8	6.0	10.2	**13.3**
A-2	6	4.3	3.2	9.1	**13.3**
B-1	2	**12.3**	6.3	6.7	5.8
B-2	3	6.6	4.2	8.5	9.1
C-1	4	5.3	2.0	6.7	6.4
C-2	5	4.1	5.7	11.6	**14.2**

Bold numbers indicate those groups in which the percent decrease in frequency of breathing would be classified as slight (12-20% decrease) sensory irritation by ASTM E-981-84.

The data from Table 7 suggest that with an increased number of exposures, a decrease in frequency of breathing is observed. When the two replicate experiments were combined, there is some suggestion of a greater decrease in frequency of breathing during the fourth exposure for Treatments A (13.3%) and C (10.3%), compared to Treatment group B (7.5%).

Individual Animal Analyses Because there was some suggestion of an effect using the ASTM E-981-84 criteria, a more in-depth analysis of the individual animal's response to

treatment was examined. The analysis attempted to answer three questions related to the treatment response of individual animals with four successive exposures during the three measurement periods.

1. Was there a treatment-related difference in frequency of breathing during the clean zero-grade air control period? The ANOVA indicated that there was a treatment-related difference ($p=0.0129$) in the control periods, with the Treatment B group having a slightly lower frequency of breathing than Treatment A or C (Figure 3a). Additionally, there was a significant exposure-related effect ($p<0.001$) without an exposure-by-treatment related interaction ($p=0.759$). The lack of a significant interaction term would indicate that the treatment-related differences that occurred were maintained across the four exposures. Thus, when the data were collapsed across all treatment groups, frequency of breathing elevated with increasing numbers of exposures (Figure 3b). This result would suggest that there was a carry-over effect from the previous exposures, but it was unrelated to source chamber contents.

2. *Were differences in treatment groups observed during the exposure period?* Because Treatment group B had a significantly lower frequency of breathing initially, and because that difference was maintained across the four exposures, all subsequent analyses were performed on data adjusted for this difference. Percent difference from each animal's control response was used as the adjustment. Analysis of each animal's greatest one-minute percent decline in frequency (Figure 4a) indicated that there were significant exposure-related effects ($p<0.001$), but no significant treatment ($p=0.972$) effects or treatment-by-exposure interactions ($p=0.639$). The significant exposure effect occurred because the decrease in frequency was less after the second exposure and greater after the fourth exposure than the decrease observed during the first and third exposure (Figure 4b). This effect was significant only when

the data were collapsed across all treatment groups. A similar analysis of each animal's greatest one-minute percent increase in frequency indicated that there were no significant treatment or exposure-related effects. With all treatments and exposures combined, the mean greatest one-minute increase in frequency was 15.1%. Thus, overall, no treatment group related increases or decreases in frequency of breathing could be detected.

3. *Were differences in treatment groups observed during the recovery period?* Although no treatment-related differences were observed during the exposure period, it is possible that treatment-related differences may become manifest during the recovery period. Another two-way ANOVA was performed evaluating the percent difference from the control period. The analysis indicated that both significant exposure ($p=0.013$) and significant treatment ($p=0.044$) effects occurred; however, a significant treatment-by-exposure interaction was not observed ($p=0.529$). The significant treatment-related effect occurred because, when collapsed across all exposures, the Treatment A or B groups did not fully recover, whereas the Treatment C group did (Figure 5a). The exposure effect occurred because the second exposure recovery period was significantly ($p=0.029$) different than the percent recovery for exposures 1, 3, and 4 (Figure 5b).

Subjective Evaluations of Pulmonary Waveforms Alterations of the normal sinusoidal respiratory waveforms were classified as either looking like sensory or pulmonary irritation, or abnormal, but not either of the above (disruption index). The existence and severity of altered respiratory waveforms were evaluated for each of the three exposure periods. In general, between 70 and 90% of all periods in an exposure had less than four one-minute scores greater than zero (i.e., no evidence of sensory-like or pulmonary irritation). Analysis indicated that significantly more respiratory waveforms that looked like slight sensory irritation occurred

during the control period (p= 0.012) and during the second and fourth exposure periods (p=0.005) for the Treatment B group. Figure 6 shows the incidence of sensory irritation during the exposure period when the data are collapsed across the four exposures.

Although a significant treatment-by-exposure interaction was observed for pulmonary irritation during the exposure period, the interaction occurred because of differences between Treatment A and C during the fourth exposure. When collapsed across exposure periods, no significant differences between Treatment B (zero-grade air group) versus A or C were noted (Figure 7).

There was also a significant (p=0.001) treatment-by-exposure interaction for the control period when the disruption index was analyzed. This interaction occurred because Treatment A showed more disruptions than Treatment B during the third exposure. Although no differences in the disruption index were observed during the exposure period, a marginally significant (p=0.051) treatment effect was revealed during the recovery period because Treatment A had more disruptions than Treatment B (p=0.039) when the data were collapsed across all four exposures. The disruption index for all periods combined (control, exposure and recovery) and collapsed across all exposures (1 through 4) is shown in Figure 8.

Postexposure appearance Although not part of the formal functional observational battery, mice were observed for at least 15 minutes after each exposure. During this time, any unusual behavior or abnormal appearance was noted. During most of these postexposure observation periods, mice would initially be still and then after 2-3 minutes they would begin to groom and drink water. No abnormal behavioral observations were noted for any of the animals. However, their appearance was abnormal. These effects, which included

lacrimation (tearing), dilated or hemorrhaged blood vessels of the pinna (outer ear), and facial swelling, were observed in all experimental groups after most exposures. Analysis of the incidence of occurrence indicated that there were no differences in incidence among treatment groups for the facial swelling and hemorrhaged pinna. For lacrimation (Figure 9) and dilated pinna vessels (Figure 10), significant treatment differences were observed. Treatment C was found to have significantly ($p<0.001$) less lacrimation, but more dilated pinna vessels ($p<0.001$) than Treatment groups A or B. No treatment-by-exposure interactions were significant on these observations.

Functional Observational Battery A summary of the data for all of the measures of the functional observational battery is included in Appendix A. All tests were conducted at the prescheduled times. For ease of interpretation, the functional observational battery tests have been sorted into groups representing the neurological functions they most closely represent: (1) neuromuscular function, (2) sensorimotor function, (3) general activity and excitability, and (4) general appearance and other measures.

Some tests were not affected at any time; they showed no variability across groups and therefore were not subjected to statistical analysis. These were: Body Tilt, Forelimb Placing, Tonic Movements, Excessive Vocalizations, Diarrhea, Gasping, Cyanosis, Exophthalmus, and Salivation. The results of statistical analysis for the remaining behavioral measures are shown in Table 8. Significant or equivocal differences between groups were detected in 8 of 18 analyses.

Table 8. Summary of Statistical Outcomes for Functional Observational Battery Data[a]

FOB Measure	Treatment (TRTMT)	TRTMT-by-Time	Day 1	Day 2	TRTMT
Neuromuscular					
Jar Task	Sig.	Sig.	Sig. A≠B	Sig. A≠B C≠B	
Grip Strength	NS	NS			
Body Tone	NS	NS			
Righting Reflex	NS	NS			
Body Posture	NS	NS			
Ataxia Score	Equiv.	NS			NS
Gait Score	NS	NS			
Inverted Screen Test	NS	NS			
Missteps	NS	NS			
Sensorimotor					
Air Puff Response[b]	NS	NS			
Click Response	Sig.	NS			Sig. A≠B
Tail Pinch Response[b]	Equiv.	NS			Sig. A≠B C≠B
Activity and Excitability					
Activity	NS	Sig.	NS	Equiv.	
Alertness	NS	Sig.	NS	Sig. A≠B C≠B	
Handling Reactivity	NS	NS			
Rearing	NS	Equiv.	NS	NS	
Clonic Movements	NS	NS			
Other					
Tilted Screen	Sig.	NS			Sig. A≠B

[a] Probability of significant difference between groups
Sig.= $p<0.05$
NS = $p>.10$
Equiv. = $0.05<p<0.10$, may be slight effect
[b] Analysis conducted on change from preexposure data

Neuromuscular function Performance on the jar task was affected by treatment (Figure 11), showing a significant treatment-by-time interaction ($p=0.042$). Following the second exposure (day 1), mice in the treatment A group had more falls than did Treatment B (air-exposed group, $p=0.002$). This difference was apparent again after the fourth exposure, but the contrast was not quite significant ($p<0.052$). Group C was significantly different from Treatment B, but this was because the Treatment C mice had less falls than the air-exposed group ($p=0.017$). Thus, more falls occurred after exposure to Treatment A than with Treatment B, and better performance was observed with Treatment C after the last exposure.

The overall analysis of ataxia score produced a marginal treatment effect ($p=0.084$). However, when the treatment data were collapsed across days, the significant effect ($p=0.051$) was between Treatment A and C. Thus, exposure to carpets did not produce effects on this measure that could be differentiated from the air-exposed mice.

One mouse in the Treatment B group dropped the screen twice, after both the second and fourth exposures. There were some mice with slightly decreased body tone in all treatment groups. In addition, there was one instance of slightly slow righting in each of Treatments B and C, and one hunched posture in Treatment B. Four, five, and six mice displayed somewhat abnormal gaits (score of 3) in the Treatment A, B, and C groups, respectively; one mouse in Treatment B and one in Treatment C also showed marked gait abnormalities (score=4). For all these measures, however, there were no statistically significant differences in the distributions of these scores across treatment groups. Thus, although there were some measures that showed deviations from preexposure values in these mice, these differences were equally distributed across all treatment groups. When there were only one or very few mice affected, most commonly it occurred in the Treatment B

group.

Sensorimotor measures There was a clear preexposure difference between treatment groups (p=0.01) in the response to the tail pinch, and responses to the air-puff stimulus showed a marginal (p=0.061) differences on the day before testing. Therefore the data for these measures were analyzed using the change from preexposure values. A marginally significant overall treatment effect (p=0.066), but no treatment-by-time interaction, was detected in the tail pinch response; there were no changes in the air-puff response. Mice exposed to either carpet showed less increases in response to the tail-pinch than controls; that is, mice in Treatment B showed more increased responses than did mice in either Treatment A or C (see Figure 12). However, the data for mice in Treatments A and C were most like their pre-exposure values, and therefore the significance of this difference is questionable.

There was a significant overall treatment effect with the click response (p=0.004), but no treatment-by-time interaction. Further analysis revealed that mice in Treatment A showed less reactivity to the click stimulus than did mice in Treatment B. Figure 13 shows that following Treatment A, there were more responses rated as "slight" (rank=2), whereas with Treatment B, more mice showed "clear" (rank=3) responses. However, as with the tail pinch response, the scores in group A mice were more like their own preexposure data than were the Treatment B group's data.

Activity and Excitability Changes in the ranking of handling reactivity and the number of rears were the same in all treatment groups. One mouse in Treatment A and one in Treatment B showed slight quivers after the fourth exposure. None of these measures were

statistically significant.

The overall treatment-by-time interaction for activity level was significant, but there were no significant differences in rank distribution on any one day. A trend was detected after the fourth exposure ($p<.0772$), at which time the mice in Treatment A showed lower activity levels than did those in Treatment B.

The treatment-by-time interaction was also significant for the scoring of alertness ($p=0.024$). Univariate ANOVAs showed that on day 2, the mice in both Treatments A ($p=0.007$) and C ($p=0.029$) showed lower scores than did those in Treatment B. Review of the data (Figure 14) showed that this difference was due to more Treatment B mice appearing hyper-alert, which was only observed in one mouse before exposures began. Treatment C mice, in fact, showed no changes in level of arousal across days, whereas two Treatment A mice showed slightly lower values on day 2.

Other Measures There was a significant overall difference between treatment groups on the direction taken on the tilted screen ($p=0.03$), but no treatment-by-time interaction. Collapsing the data across days, it appeared that more mice went down the screen (2 or 3 times out of 3 trials) following exposure to Treatment A than did after exposure to Treatment B.

General Appearance After exposures, almost all mice showed facial swelling, lacrimation, dilation and hemorrhaging of pinna vessels, and some chromodacryorrhea (reddish tears). Although these signs were recorded during the behavioral tests as well as upon removal from the animal exposure chamber, the analysis of the data was conducted on

the observations made immediately after removal from all four exposures (see above).

Postmortem Evaluation Two separate one-way ANOVAs were used to analyze the postmortem data. The first analysis examined differences between the three treatment groups (A, B, and C), whereas the second analysis examined the three treatment groups and also included the non-exposed cage-control mice.

Gravimetric data Gravimetric data were obtained during the necropsy for six tissues: brain, heart, liver, kidneys, thymus, and spleen (Table 9).

Table 9. Organ Weight Data by Treatment Group

Treatment	Brain	Heart	Liver	Kidney	Thymus	Spleen
Control	0.43±0.02	0.14±0.01	1.46±0.13	0.43±0.04	0.08±0.02	0.10±0.02
A	0.42±0.03	0.13±0.01	1.25±0.08	0.38±0.03	0.06±0.02	0.08±0.02
B	0.44 0.02	0.13±0.01	1.29±0.11*	0.40±0.05	0.06±0.02*	0.09±0.03
C	0.43±0.02	0.13±0.01	**1.18±0.07**	0.41±0.04	0.06±0.01	0.09±0.01

Bold number indicates significant difference from Treatment B.

* Indicates significant difference from non-exposed unrestrained cage control group.

Significant treatment differences in organ weight were found only for liver weight ($p=0.033$). This difference occurred because Treatment C livers weighed less ($p=0.021$) than Treatment B (air-control group). When liver weights were corrected for individual differences in body weight, the difference was only marginally significant ($p=0.055$); however, there were no significant differences in body weight between treatment groups.

Analysis of treatment effects compared to the non-exposed control indicated that there were significant treatment effects for liver ($p < 0.001$) and thymus ($p=0.027$). Both liver and thymus weights of the non-exposed controls were larger than any of the treatment groups and were significantly different from Treatment B.

Hemoglobin Measurements Three parameters were analyzed to evaluate potential differences in the oxygen-carrying ability of blood from carpet-exposed mice. Comparison of the three treatment groups indicated a marginally significant ($p=0.059$) effect of treatment on hemoglobin concentration. When the non-exposed control group was included into the analysis as an additional treatment group, significant differences in hemoglobin concentration were observed ($p=0.026$). The analysis indicated that the non-exposed and Treatment B groups had significantly more hemoglobin than the Treatment A or C groups (Figure 15), but Treatment B was not different from the cage control. The other two parameters that were evaluated, %methemoglobin and %carboxyhemoglobin, were found not to be significantly different by treatment with or without the non-exposed controls added into the analysis.

White cell differential Analysis of the differential white cell percentages in peripheral blood indicated that there was an overall treatment effect for neutrophils ($p<0.001$) when examining the three treatment groups and the non-exposed controls (Table 10). Significantly more neutrophils were in the peripheral blood of the Treatment B mice ($p<0.001$) than for the non-exposed controls. There were no significant differences between Treatment B and Treatments A or C.

Table 10. Peripheral Blood White Cell Differential

	Control	Treatment A	Treatment B	Treatment C
%Monocyte	9.0 ± 12.0	10.7 ± 7.8	5.9 ± 3.0	4.3 ± 2.4
%Neutrophil	15.0 ± 5.1	49.7 ± 22.1	40.2 ± 9.0*	46.7 ± 8.9
%Lymphocyte	55.9 ± 15.1	31.0 ± 16.4	42.7 ± 12.2	41.7 ± 7.2
%Eosinophil	1.7 ± 2.3	0 ± 0	0.4 ± 0.7	0.1 ± 0.4
%Unknown	18.5 ± 8.6	8.3 ± 1.5	10.7 ± 4.7	7.1 ± 4.9

* Indicates significant difference from non-exposed unrestrained cage control group.

Serum Clinical Chemistry Sixteen measurements indicative of liver, kidney, heart, and stress-related responses were evaluated in the serum of treatment mice and non-exposed controls (Table 11). Of these 16 measurements, three were found to be significantly different in the three group treatment comparison. Total serum proteins ($p=0.003$), albumin ($p=0.002$) and cholesterol ($p=0.023$) were found to be lower in Treatment C compared to Treatment B (Figure 16). Treatment A was not different from Treatment B for any of these measurements. When the non-exposed control group was included into the analysis as an additional treatment group, eight of the sixteen measurements showed significant treatment differences. Of the three measurements that were significant in the three treatment group comparison, Treatment C was most similar to the non-exposed control group for all three measurements.

Table 11. Serum Clinical Chemistry Values

Parameter	Control	Treatment A	Treatment B	Treatment C
N	12	8	8	8
TBIL	0.429 ± 0.130	0.526 ± 0.359	0.446 ± 0.283	0.436 ± 0.107
LDH	541.7 ± 187.8	633.5 ± 127.3	773.3 ± 356.0	806.1 ± 332.6
ALT	37.1 ± 19.9	60.5 ± 24.3	47.5 ± 21.3	49.0 ± 15.2
AST	75.9 ± 24.6	123.0 ± 32.7	108.6 ± 37.4*	157.8 ± 61.5
CREA	0.583 ± 0.083	0.614 ± 0.177	0.603 ± 0.074	0.546 ± 0.153
BUN	19.7 ± 6.2	29.0 ± 7.0	30.8 ± 7.2*	33.7 ± 12.3
5-ND	22.3 ± 5.4	24.0 ± 4.5	31.4 ± 17.3	26.4 ± 9.0
GLUC	236.8 ± 23.5	207.9 ± 18.0	214.8 ± 22.9*	198.0 ± 18.8
ALP	138.6 ± 27.4	138.6 ± 29.2	157.8 ± 11.6	143.8 ± 29.3
PRO	4.90 ± 0.38	5.44 ± 0.46	5.48 ± 0.36*	**4.83 ± 0.27**
ICD	20.7 ± 11.1	33.0 ± 6.9	38.0 ± 9.0*	37.6 ± 13.7
ALB	2.93 ± 0.21	3.35 ± 0.40	3.45 ± 0.17*	**2.93 ± 0.20**
TRIG	89.1 ± 26.6	60.4 ± 12.3	56.7 ± 10.9*	45.7 ± 13.2
CHOL	111.5 ± 16.9	136.3 ± 24.0	146.1 ± 13.6*	**115.7 ± 22.7**
SDH	15.0 ± 3.1	17.1 ± 6.7	18.3 ± 2.0[a]	16.9 ± 3.1[a]
BILE	45.4 ± 9.6[a]	103.0 ± 103.8[b]	222.0 ± 222.0[c]	228.2 ± 295.8[d]

Bold numbers indicate a significant difference from Treatment B.
* Indicates significant difference from non-exposed unrestrained cage control group.
[a,b,c,d] Insufficent serum was available, N= 7, 5, 2, and 6, respectively

Nasal and Bronchoalveolar Lavage Chemistry Four parameters were measured from the nasal lavage - ascorbic acid, uric acid, glutathione, and total protein. Of these measurements, uric acid was not detectable in the nasal lavage and no significant treatment effects occurred for ascorbic acid or total protein. Glutathione obtained from the nasal lavage was significantly ($p=0.021$) greater in Treatment B than in Treatment A, but not different than Treatment C or the non-exposed control group (Figure 17). Only two parameters were examined in the bronchoalveolar lavage of mice - total lavageable protein and lactate dehydrogenase. A significant ($p<0.001$) treatment effect for total protein was observed, indicating that Treatment B had less protein in the lavage than Treatments A or C. However, the lavageable protein in Treatment B was also significantly ($p=0.027$) less than the non-exposed controls (Table 12). No differences in lactate dehydrogenase were noted.

Table 12. Lung Lavage Chemistry

	Total Protein	Lactate Dehydrogenase
Controls	**136 ± 36**	**33.5 ± 8**
Treatment A	**144 ± 17**	**46 ± 15**
Treatment B	**108 ± 15***	**41 ± 15**
Treatment C	**157 ± 26**	**43 ± 10**

Bold numbers indicate a significant difference from Treatment B.

* Indicates significant difference from non-exposed unrestrained cage control group.

Lung Lavage White Cell Differential Between 60,00 and 100,000 cells were removed from the lungs of both non-exposed and exposed mice, with the alveolar macrophages being the most prevalent (approximately 80% of cells, Table 13) . Analysis of variance indicated a significant treatment effect for the number of lymphocytes when either the three treatment groups were analyzed (p=0.012) or when the non-exposed control was an additional group in the analysis (p=0.003) . No significant difference was detected in a comparison of the number of lymphocytes in Treatment B versus the non-exposed control group. Treatment A, but not C, had significantly (p=0.027) more lymphocytes than Treatment B.

Table 13. Lung Lavage White Cell Count and Differential ($x10^4$) / mL of Bronchoalveolar Lavage

	Control	Treatment A	Treatment B	Treatment C
White Count	69.6 ± 33.8	102.8 ± 67.7	61.6 ± 20.4	100.0 ± 57.9
Macrophage	49.1 ± 18.8	81.9 ± 56.2	52.7 ± 18.6	78.2 ± 56.5
Neutrophil	0.51 ± 0.41	0.42 ± 0.51	0.21 ± 0.26	0.15 ± 0.36
Lymphocyte	0.7 ± 0.1	**2.8 ± 1.8**	0.7 ± 0.5	1.0 ± 1.0
Epithelial	3.1 ± 2.4	2.2 ± 1.9	1.3 ± 1.6	1.9 ± 2.6
Unknown	8.7 ± 7.3	14.6 ± 17.7	7.1 ± 3.7	14.7 ± 15.5

Bold number indicates significant difference from Treatment B.

Histopathology Several microscopic lesions were observed in treatment groups that either were not diagnosed or occurred with a lower incidence or severity in the non-exposed, unrestrained control mice. The incidences of these histopathological lesions are reported in Table 14. Minimal to mild necrosis/inflammation of the heart and its arteries, focal liver necrosis, pituitary hemorrhage, thymic cortical necrosis, and hemorrhage/inflammation of the pinnae were common findings in all treatment groups.

Table 14. Incidence of Selected Histopathological Lesions in Treated and Non-Exposed Mice

Treatment	A	B	C	Non-Exposed
Total Number of Tissues Examined	8	8	8	12
Heart, Necrosis	1	4	2	0
Heart Artery Inflammation (minimal)	3	5	5	3
Heart Artery Medial Necrosis/Hemorrhage (minimal)	2	2	0	1
Heart Artery Medial Necrosis/Hemorrhage (slight/mild)	0	2	1	0
Heart Artery Medial Necrosis/Hemorrhage (moderate)	0	1	2	0
Liver, Focal Necrosis	3	2	1	0
Pituitary, Pars Distalis, Hemorrhage	2	4	4	0
Thymus, Cortical Necrosis (slight)	6	7	4	0

Minimal myocardial necrosis was observed in mice from all of the treatment groups, but was not observed in the non-exposed, unrestrained controls. The incidence of necrosis was greatest with Treatment B. This lesion was characterized by relatively small focal accumulations of cellular debris with an associated minimal infiltration of mononuclear cells and neutrophils replacing myocardial fibers.

Inflammation and medial necrosis/hemorrhage in the arteries of the heart were also observed in some mice from all groups. The incidence of inflammation, and to a lesser degree, the medial necrosis/hemorrhage was greatest in Treatments B and C. The severity of the medial necrosis/hemorrhage, which ranged from minimal to moderate, was also greatest in the affected mice in Treatments B and C. The inflammation and medial necrosis/hemorrhage involved the coronary arteries and their branches and rarely the large elastic arteries (aorta). The inflammatory changes were characterized by a minimal infiltration of mononuclear cells and lesser numbers of neutrophils in the perivascular area and occasionally, the walls of the arteries. Medial necrosis/hemorrhage consisted of the replacement of smooth muscle fibers in the tunica media with an eosinophilic fibrinoid material and occasionally erythrocytes. In most instances, the changes were accompanied by a small amount of a pale basophilic, somewhat granular substance in the perivascular area.

Focal hepatic necrosis was diagnosed in one to three mice of each treatment group, but in none of the non-exposed control mice. The incidence of this lesion was greatest in Treatment A. The focal necrosis was typically just beneath the hepatic capsule and was occasionally accompanied by a minimal infiltration of inflammatory cells.

Hemorrhage of the pars distalis of the pituitary gland was also diagnosed in some mice in each of the treatment groups, but in none of the non-exposed control mice. The incidence of pars distalis hemorrhage was greatest in Treatments B and C. This lesion was frequently bilaterally symmetrical in

distribution and was characterized by the focal accumulation of extravasated erythrocytes with secondary coagulative necrosis.

Thymic cortical necrosis was diagnosed in all mice. Most of the mice in the treatment groups (A, B, and C) had mild cortical necrosis compared to the minimal necrosis observed in the non-exposed mice. Minimal cortical necrosis was characterized by occasional pyknotic or karyorrhectic nuclei and occasional small foci of nuclear debris. This degree of cortical necrosis was considered to be due to normal cell turnover (necrobiosis). In contrast, mild cortical necrosis, as seen in the treatment groups, was characterized by more frequent small foci of nuclear debris.

The pinna of several mice in each of the treatment groups had either hemorrhage and/or subacute inflammation. Both lesions frequently occurred together in the pinna and consisted of extravasated erythrocytes, with or without an infiltration of neutrophils and mononuclear cells. The incidences of these lesions were greatest in Treatment groups B and C. Subacute inflammation was not diagnosed in any of the non-exposed control mice; however, minimal hemorrhage was present in the pinna of one non-exposed control mouse.

Discussion

Interpretation of the Results Overall, more than 70 different measurements were evaluated for each mouse and statistically significant differences were detected among the three treatment groups. However, many of these differences, although statistically significant, were not consistently observed in any carpet-exposed treatment group and those differences from Treatment B may be of little or no biological significance. For some parameters, it is not clear that the significant statistical or biological difference represents an adverse effect. Furthermore, significant differences were also observed between Treatment B (zero-air exposed group) and the non-exposed, unrestrained cage controls, indicating that the exposure procedure itself had significant consequences associated with it. A discussion of the two primary evaluations (measurements of irritancy and the neurobehavioral screen) and the postmortem assessment follows.

Sensory and Pulmonary Irritation Irritation was evaluated (1) by an adaptation of the ASTM method, (2) by examining the reflex changes in breathing frequency for individual animals, and (3) by subjectively inspecting and scoring the respiratory waveform morphology. The ASTM method of combining the four animal responses seemed to indicate that there was a greater decrease in frequency of breathing (a sign of irritation) with increased numbers of exposure that was somewhat more prominent in Treatments A and C than for Treatment B (the air-exposed group). This finding would provide minimal support that irritation was associated with carpet exposure. However, the rate decreases observed would only be scored slight, at best, according to the ASTM criteria (Table 7). To gain a further, more in-depth understanding of the rate decrease, individual animal data were used as input to a statistical model that allowed us to examine and separate treatment-related effects from exposure procedure effects. Visual inspection of Figure 4a emphasizes what the conclusions from the statistical analysis indicated, which was that there were no differences in response among the three groups to any of the four exposures. The most conspicuous difference between groups (Figure 4a) was observed during the first exposure when more mice

from Treatment B (air-exposed group) had greater decreases in frequency of breathing than carpet-exposed mice (Treatments A or C); however, this difference was not significant. Statistical differences in breathing frequency during the recovery period, which could be interpreted as a manifestation of toxicity, were also observed. However, the percent difference in frequency from the control period ranged from 0 to 3% for the three treatment groups (Figure 5a), a difference that would not be considered biologically significant.

Although the rate determinations discussed above were the primary determinants of irritancy, subjective analyses of respiratory waveforms were also evaluated to distinguish the type of irritant effect (i.e., sensory, pulmonary, or mixed). Using a very aggressive method to detect altered waveforms, evidence of sensory and pulmonary effects was observed for all treatment groups. During the exposure period, a greater incidence of sensory irritation was found for Treatment B (air-exposed) as compared to Treatments A or C (carpet-exposed), during exposures two and four. The rate analysis described above, however, indicated that the largest rate decrease for Treatment B occurred during the first exposure period. Thus, in contrast to the evidence of sensory irritation based on waveform morphology, the altered waveforms did not translate into a decrement in breathing frequency. This would indicate that biologically significant sensory irritation did not occur according to the ASTM procedure for which rate is the ultimate determinant of effect. Although this may seem discordant, it can be explained by the waveform analysis technique. For example, during the scoring of the waveforms, if a pattern indicative of sensory irritation appeared for several breaths (3 to 5), the entire minute rating period was scored as showing evidence of sensory irritation. However, if only five breaths had prolonged early expiratory pauses characteristic of sensory irritation, there would be minimal or no impact on the median rate calculated for the approximately 220 breaths. Additionally, it was observed that for many of the one-minute segments in which altered waveform morphology was observed, these periods were also accompanied by periods of rapid breathing. Juxtaposition of breaths with decreased rates and increased rates during the same minute would further dilute the impact of the expiratory pause to cause a decline in frequency of breathing.

Evidence of possible "pulmonary" irritation was also detected in all three treatment groups, ranging from a slight rounding at the end of expiration (score=1, slight) to a pause between breaths (score=2, moderate). Although there was one significant effect noted during the four exposure periods, it was attributed to a difference between Treatments A and C and not Treatment B. Visual inspection of the data collapsed across all exposure periods (Figure 7) would suggest that Treatment A had more pulmonary irritation than Treatment B, and that Treatment C had the least among the three groups; however, these differences were not significant. Again, evidence of pulmonary irritation based on waveform morphology did not correspond with observed decrements in frequency of breathing. However, in contrast to sensory irritation, pulmonary irritation is not solely diagnosed by a decrease in breathing frequency. Therefore, it is possible that Treatment A showed a slight pulmonary irritant effect.

Finally, in an attempt to quantify other abnormal breath shapes that would not be defined as sensory or pulmonary irritant-like, a disruption index was formulated. Most of these abnormal waveforms were movement artifacts related to the animals struggling in the plethysmograph, but it was of interest if such struggling would increase or decrease in association with an inhaled toxic exposure. The statistical analyses indicated mice in Treatment A were somewhat more disrupted than Treatment B during the control and recovery periods, but not during the exposure periods, whereas no differences in the index were observed between Treatment C compared to B. The interpretation of the relative adversity of this finding is uncertain.

Subjective Appearance and Functional Observational Battery After each exposure, both the behavior and the general appearance of the animals were observed. Formal testing using the functional observational battery was applied only after the second and fourth exposure. In general, many of the differences from the preexposure baseline performance were either similar across treatment groups or else were more prominent in the air-exposed group (Treatment B). There were also several significant differences between Treatments A and C, with Treatment C appearing more benign. Indeed, as these data

were being analyzed and before the code was broken, it appeared that Treatment C was the control.

Four prominent findings related to the animal's appearance were observed after each exposure. Virtually all of the animals had some lacrimation and dilated vessels of the pinna (outer ear), whereas many had facial swelling and evidence of hemorrhaged vessels within the pinna. There were statistical differences between treatment groups indicating that more mice in Treatment C (carpet-exposed) had dilated vessels of the pinna and more mice in Treatment B (air-exposed) had lacrimation. The occurrence of these four effects and the treatment group differences may, in part, be attributed to the high temperatures and low humidity in the animal exposure chamber (Tables 5 and 6). Higher temperatures occurred during both of the experiments involving Treatment B (range, 25.0 to 26.4 °C) and were the worst for the first experiment using Treatment C (range, 26.1 to 30.0 °C). Vessel dilation is one mechanism that rodents use to dissipate heat.

The significant treatment effects detected using the functional observational battery were not clustered in any one neurological domain. The only significant neuromuscular effect was the altered jar task performance observed in Treatments A and C. The increased number of falls observed in Treatment A could be interpreted as a clear, though not severe, effect. On the other hand, the improvement of performance in Treatment C, while also significant, would be difficult to interpret as an adverse effect. The other neuromuscular measures, which should also detect vestibular changes, weakness, or incoordination that might compromise the ability to climb on the jar, were not altered by carpet exposure. This lack of correlative findings reduce the certainty of a significant neuromuscular/vestibular effect of Treatment A.

The sensorimotor changes observed in Treatments A and C, although statistically significant, could not be considered clear evidence of carpet toxicity. Although it appeared that mice exposed to carpet (particularly Treatment A) had lower reactions to both the click and tail-pinch responses than those exposed

to zero-grade air, in reality the Treatment B mice responded more vigorously than they did before exposure. Moreover, in all groups, most of the responses were scored as "slight" or "clear", which are the responses expected for "normal" mice. Therefore, it is difficult to conclude that the differences in response between treatment groups was an adverse consequence of carpet exposure.

On the open field test, the only clearly significant effect was an apparent decrease in level of alertness in mice exposed to carpets. The same situation applies, however, as with the sensorimotor responses in that the effect was due not to lower scores in carpet-exposed mice, but an increase in the score of Treatment B mice, when compared to their preexposure data as well as what has historically been expected as "normal" behavior.

Postmortem Evaluation Although a crude indicator of toxicity, a treatment-related change in body weight is often a sensitive indicator of effect. In this study, body weights declined with exposure, but there were no treatment-related differences (Figure 2). Liver and thymus weights were found to be less than those of the non-exposed cage control group with a somewhat larger reduction in liver weight in the Treatment C group. Typically, liver damage produces an increase in liver weight, whereas the factors responsible for loss of body weight (physical, humidity, and temperature stress; food and water deprivation) may account for the reduction between treatment groups and the non-exposed animals. However, even after corrections for the differences in body weight, Treatment C mice were still significantly affected (10% decrease), suggesting a small, treatment-related effect.

Because the lung is the portal of entry for the carpet vapor exposure, it might be a primary target for toxicity. Thus, several screening measurements were conducted to evaluate the impact of exposure. Lactate dehydrogenase (LDH) in the BAL was unaffected, indicating that there was no direct cytotoxicity to pulmonary cells. Lavage fluid protein was increased in Treatments A and C compared to Treatment B.

However, the BAL protein from the non-exposed cage control group was higher than that observed with Treatment B. Thus, although an increase in BAL protein might be considered evidence of increased lung permeability (leakage of plasma protein from the blood to the lung surface), the BAL protein values of the cage control group would suggest that the effect was not biologically significant. The only other significant effect observed in the BAL was a small, but significant, increase in the number of lymphocytes for Treatment A, which might suggest an increased immune system activation from an antigenic protein or viral infection. However, no other physical (particle concentration), microbiological, or histopathological evidence would support these findings.

Histopathological examination revealed several differences between the cage control group and the three treatment groups (Table 14). Myocardial necrosis and necrosis of the small heart vessels were more frequently observed in Treatments B and C. Liver necrosis was seen in 3 of 8 mice exposed to Treatment A, with Treatments B and C showing less of an effect, respectively. Similarly, thymic cortical necrosis was observed in all mice, including the cage controls; however, the severity was much greater in the treated mice, especially in Treatments A and B. Hemorrhage of the pars distalis of the pituitary gland was observed predominately in Treatments B and C (50% incidence), but was also observed in Treatment A (25% incidence). The origin of these effects is uncertain, but is possibly related to the food and water deprivation during exposure, and physical restraint stress associated with the exposure procedure employed in this study. Overall, the histopathological effects were minimal such that the pathologist could not clearly distinguish which group was the air-exposed treatment group (i.e., Treatment A was incorrectly identified as the air-exposed control group).

Several measurements were obtained from the blood that focused on alterations of hemoglobin, white blood cell population profiles (differential), and serum chemistries. A marginally significant decrease in hemoglobin was observed for Treatments A and C compared to both Treatment B and the non-exposed

caged controls. A decrease in hemoglobin content could signify a decrease in oxygen-carrying capacity; however, only a 6% decrease was observed, a change that would not be considered clinically significant. No significant differences were seen in the peripheral blood white cell differential among treatment groups. When compared to the cage controls, all of the treatment groups had significantly elevated percentages of neutrophils (approximately threefold increases). Although this may be a stress-induced demargination of neutrophils, it could also signal an acute infection, chemical intoxication, acute hemorrhage or hemolysis, or acute tissue necrosis. However, correlative pathology (described above) and serum enzyme chemistries (described below) do not support these alternate interpretations. The fact that this condition was observed equally in all treatment groups, including the air-exposed group (Treatment B), suggests strongly that it is related to the exposure procedure (Table 10).

Significant alterations in serum chemistries were also observed (Table 11). Total protein, albumin and cholesterol were lower in Treatment C than in Treatment B; however, Treatment C was not different from the cage controls for these three measurements. On the other hand, these three serum chemistries were significantly greater in Treatments B (air-exposed) and A than in the cage controls. One common cause of increased serum proteins is dehydration, which certainly could have resulted from the exposure procedure. Increased cholesterol is nonspecifically associated with more chronic changes in the cardiovascular, hepatic, kidney, and pancreatic systems. Why Treatment C was less affected is uncertain; however, the Treatment C-related decrease in liver weight noted above may correspond to the decreased ability of this treatment group to express these particular proteins. What toxicity might be associated is unclear, especially in light of the similarity in these measurements to the cage control mice.

Five additional clinical serum chemistries were analyzed (AST, BUN, glucose, ICD, and triglycerides) and were found to be significantly different between the unrestrained, non-exposed cage controls and the restrained and air-exposed controls (Treatment B). Increases in AST, BUN, and ICD were observed,

whereas glucose and triglycerides were decreased. Although these differences were statistically significant, they are difficult to ascribe to a specific toxicity because the magnitude of most of the changes would not be considered of biological significance in human clinical medicine. Interestingly, Treatment C showed an even more pronounced difference from the cage controls than did Treatment B (Table 11), possibly reflecting the altered liver weights and serum chemistry results discussed above.

Increased AST, in humans, is typically associated with myocardial infarction, muscle and liver injury, acute liver necrosis, or acute pancreatitis, but also may increase with acute stress. Increased BUN occurs with impaired kidney function, salt and water depletion, hemorrhage, stress, and acute myocardial infarction. Glucose is decreased with hypoglycemia and liver disease, but is typically increased with stress. Liver cell injury is associated with an increase in ICD, for which this measurement is both sensitive and specific; however, malnutrition has been shown to sometimes increase ICD. A decrease in triglycerides is associated with stress and malnutrition. It is likely that all of the effects can be accounted for by food and fluid restriction and restraint-induced stress. However, it is possible that the treatments may also have some effect on liver and heart tissue. Table 15 shows the significant clinical chemistry and peripheral blood responses observed in Treatment B as related to potential causes of these changes from the human clinical literature (Wallach, 1978).

Table 15. Potential Causes of Clinical Chemistry and Blood Neutrophil Responses

Measured Parameter	Treatment B	Stress	Dehydration	Food Depri-vation	Liver Injury	Heart Injury
AST	↑	↑			↑	↑
BUN	↑	↑	↑			↑
Glucose	↓	↑		↓	↓	
Protein	↑		↑	↓	↓	
ICD	↑			↑	↑	
Albumin	↑	↑		↓	↓	
Triglycerides	↓	↓		↓		
Cholesterol	↑			↑	↑	↑
% Neutrophils	↑	↑				↑

Potential Confounds The carpet study protocol was founded on the premise that two different carpets previously shown by Anderson Laboratories to induce severe neurotoxicity or death would produce similar effects when retested under analogous conditions in the laboratories of both Anderson and EPA. In addition, it was anticipated that carpet-related responses in exposed mice would be distinguishable from sham, clean-air exposed control mice. The biological end points evaluated by EPA included those used in the standard Anderson Laboratories protocol and several additional biomarkers indicative of systemic organ function or toxicity. However, following carefully conducted experiments and after thorough analysis of the study data, no biologically significant effects attributable to carpet vapor exposure could be discerned. All treatment groups, A, B, and C, appeared essentially unaffected; only when the restrained clean-air controls were compared to unrestrained, non-exposed cohorts did there appear a consistent difference. The occasional small, but statistically significant changes in one or another biomarker in the actual carpet-exposed groups, although notable, could not be reasonably attributed to a toxic syndrome. Although this lack of response would appear to support the contention that the carpets tested simply did not possess inherent toxicity, it would seem reasonable that various factors, which could have influenced the response or its detection, be examined carefully as potential confounds.

Animal Model The use of Swiss-Webster (SW) mice in these studies must be considered a potential source of substantial variability. The carpet-effect in the SW, an outbred strain, is inherently variable from mouse to mouse (as reported by Anderson Laboratories), and with regard to the nonlethal effects described, appears to be quite transient and sporadic. The strain of mouse used in the EPA portion of the study (including gender, age and size, and even vendor) was identical to that used by Anderson. A complete microbiological evaluation of the mice, both upon receipt and in sentinels, verified the health status and lack of infection in the animals. The mice were housed in an AAALAC-certified vivarium upon receipt and care was taken to avoid an infectious outcome over the course of the study (about 10 days including the 7-day acclimatization). The Anderson mice, otherwise the same, were housed in a facility ostensibly clean and

healthy, but actually of unknown microbiological status (e.g., ducts, air conditioning, etc.). Whether some relatively passive organism residing in the Anderson mice could render them "susceptible" as compared to those of EPA seems unlikely, but is unknown. Moreover, by what mechanism a frequently rapid-onset, acute response might result from such an interaction is also unknown.

Exposure Factors The large variability in animal response described previously by Anderson, even within a given run of four animals, raises the question of exposure. The exposure system used in these studies by EPA was designed after that in use at Anderson Laboratories. The aquarium source chamber and plumbing were identical. However, there were some essential differences in the actual exposure method used in the EPA study compared to the standard Anderson exposure protocol. These differences arose from the desire to standardize and better control the quality of the exposure itself. The EPA tests used zero-grade bottled air that was certified "clean" by the air chemistry co-investigators working on this project. This bottled air was humidified through a distilled water trap and was provided to a manifold connected to the inlet port of the chamber. The chamber air was drawn from this manifold at a flow rate similar to that used in Anderson's tests (7 LPM). The excess air was exhausted to the room. Hence, the air passing through the source (carpet) chamber was as clean as possible (as certified in test samples). This procedure contrasts with the unfiltered "room air" drawn into the chamber system at Anderson Laboratories. The contaminant status of room air is probably quite variable, but previous spot samples that EPA chemists collected at Anderson Laboratories did not indicate unique components that would lead to the conclusion that this difference in source air is the origin of the biologic differences. Similarly, it would seem unlikely that streaming of air through the EPA chamber would somehow avoid entraining the diffusion-distributed vapors that would have emanated from the carpet during the one-hour incubation or during the test run. However, this possibility merits direct testing and validation because the inlet and exit outlets on the exposure chamber are on the same face and no provision is made to mix air within the chamber.

After the carpet test series was completed, quality assurance checks revealed that the humidity sensor used in the EPA test system had failed over the course of the study. This failure was not abrupt, but was a slow loss of accuracy; previous experience with these sensors had never encountered a failure and the effect on the dilution air adjustments was minor and not noticed. The result of the failure was that the humidity, rather than being standardized at 50%, ranged from 18.7% to 29.2%. The effect of this on the carpet emission is uncertain. Similarly, how this compares to the earlier runs with these carpets by Anderson Laboratories is uncertain because Anderson Laboratories does not control humidity in the laboratory and records room humidity with a simple "home-style" hygrometer. Given that Anderson Laboratories has tested many of the sample carpets during the winter season, it is likely that the humidity range of the EPA test was not unusually deviant from the typical conditions in Dedham, MA, though this difference needs further investigation.

Each animal was sealed and secured into individual plethysmographs with its head protruding into the glass exposure unit by a thin rubber dam with a sized hole. The standard head hole used in the EPA tests was slightly smaller than that used in the Anderson Laboratories system, and the gauge of the rubber dam was somewhat thinner. The hole size adopted was determined as a balance between the restraint provided by the dam and the comfort of the animal. Ancillary tests conducted by EPA suggest that the slightly smaller hole size may reduce breathing frequency slightly, but does not appear to affect the form of the tidal breath trace, indicating no exogenous bronchoconstrictive stress on the animal. Because tidal volume depth cannot be accurately measured in this plethysmographic arrangement, it is possible that there was a slight difference in the depth of each breath between the laboratories involved, but it would seem unlikely to be of such magnitude as to affect the dose to the animals and, hence, their response.

Carpet Factors The lack of response may reside in the heterogeneity of the carpets and the samples thereof. The CPSC samples were selected randomly from various locations of the test carpets without

regard for use areas or other potential variables. With "carpet" being the toxicant as best as can be currently defined, it seems quite likely that regions of the total carpet area would have less or none of the factors contributing to the observations of toxicity by Anderson Laboratories. Even Anderson Laboratories reports that some test results of carpet from the same source are not readily reproduced for reasons unknown. Such a possibility suggests that retests of these carpets or parallel tests as are being conducted at Anderson Laboratories as part of the overall test plan for the ORD Carpet Study could result in disparate findings. At this point in time, there is no way of establishing a probability that carpet heterogeneity and sample distribution events could occur. Additional potential confounding factors might arise from differences in storage and handling of the original test samples as compared to the samples used in the ORD study, but in light of the wide range of handling conditions prior to testing in Anderson Laboratories, this variable would appear to not be a major confounder if reasonable care in storage of the samples is followed.

Conclusion There appeared to be no severe toxic effects associated with exposure to the off-gassing of the two tested complaint carpets. Incidental findings, of statistical significance, but unlikely to be of biological significance, were observed. Of all the effects that were observed, only the number of falls in the functional observational battery would be suggestive of an "adverse" effect (Table 16).

Table 16. Summary of Toxicological Findings

	Treatment A	Treatment B	Treatment C
Adverse			
Postexposure	↑ Falls off Jar		
Possibly Adverse			
Irritation	↓ Frequency (ASTM)		↓ Frequency (ASTM)
Irritation	↑ Pulmonary Irritation	↑ Sensory Irritation	
Postexposure		↑ Lacrimation	
Postexposure		↑ Hyper Alertness	
Postmortem			↓ Liver Weight
Postmortem	↓ Hemoglobin		↓ Hemoglobin
Postmortem	↑ BAL Lymphocytes		
Postmortem		↑ Heart Necrosis	↑ Heart Necrosis
		↑ Heart Vessel Necrosis	↑ Heart Vessel Necrosis
Postmortem	↑ Liver Necrosis		
Postmortem	↑ Thymic Necrosis	↑ Thymic Necrosis	
Adversity Unknown			
Postexposure	↑ Disruption Index		
Postexposure			↑ Dilated Pinna
Postexposure		↑ Tail Pinch Response	
Postexposure		↑ Click Response	
Postmortem	↑ BAL Protein		↑ BAL Protein
Postmortem			↓ Serum Protein/Albumin
Postmortem			↓ Serum Cholesterol
Postmortem		↑ Pinna Hemorrhage	↑ Pinna Hemorrhage
Postmortem		↑ Pituitary Hemorrhage	↑ Pituitary Hemorrhage

Approximately the same number of "possibly adverse" effects occurred in the three treatment groups, with almost all treatment groups showing an effect in each of the three categories of measurements: irritation, postexposure (including general appearance and the functional observational battery), and postmortem evaluation (Table 16). Most likely, this indicates spurious findings relative to the number of unprotected multiple comparisons performed. It is interesting that Treatment A, which showed some evidence of pulmonary irritation (not significant) without a change in rate, also had an increase in BAL protein and lymphocytes, whereas Treatment B, which showed more sensory irritant patterns, also had more lacrimation. Treatment C had more postmortem signs of systemic toxicity (decreased liver weight and small changes in serum chemistry), but had less histopathological signs of liver compared to Treatments A and B. In conclusion, based on this assessment of irritation, neurobehavioral effects and a general toxicity screen, there is no indication that exposure to off-gassing from these two carpets poses a toxicological threat.

References

Alarie, Y. (1973) Sensory irritation by airborne chemicals. CRC Critical Reviews in Toxicology, pp. 299-363.

Alarie, Y. (1981) Bioassay for evaluating the potency of airborne sensory irritants and predicting acceptable levels of exposure in man. Fd. Cosmet. Toxicol. 19:623-626.

Beaton, J.M., J.D. Prejean, R.D. Irwin, J.K. Dunnick, D.S. Baal, H.D. Giles, and A.A. Bartolucci (1990) "The neurobehavioral assessment of the effects of 28-day administration of acrylamide on B6C3F1 mice." Paper presented at the 8th International Neurotoxicology Conference: Role of Toxicants in Neurological Disorders, Little Rock, AK, October.

Bos, P.M.J., A. Zwart, P.G.J. Reuzel, and P.C. Bragt (1992) Evaluation of the sensory irritation test for the assessment of occupational health risk. Critical Reviews in Toxicology 21(6):423-450.

Brown, L.J. (1980) A new instrument for the simultaneous measurement of total hemoglobin, % oxyhemoglobin, % carboxyhemoglobin, % methemoglobin and oxygen in whole blood. IEEE Transactions On Biomedical Engineering, V BME—E-27, 3, pp. 132-138

Creason, J.P. (1989) Data evaluation and statistical analysis of functional observational battery data using a linear models approach. J. Am. Coll. Toxicol. 8:157-169.

Feldman, D.B. and J.C. Seely (1988) Necropsy Guide: Rodents and the Rabbit CRC Press, Inc. Boca Raton, FL; pp 51-76

Ghio, A.J., T.P. Kennedy, G.E. Hatch, and J.S. Tepper (1991) Reduction of neutrophil influx diminishes lung injury and mortality following phosgene inhalation. J. Appl. Physiol. 71(2) 657-665

Irwin, S. (1968) Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiologic state of the mouse. Psychopharmacologia (Berlin) 13:222-257.

Moser, V.C., J.P. McCormick, J.P. Creason, and R.C. MacPhail (1988) Comparison of chlordimeform and carbaryl using a functional observational battery. Fund. Appl. Toxicol. 11: 189-206.

Moser, V.C. (1989) Screening approaches for neurotoxicity - a functional observational battery. J. Am. Coll. Toxicol. 8:85-93.

Moser, V.C. (1990) Approaches to assessing the validity of a functional observational battery. Neurotoxicol. Teratol. 12:483-488.

Moser, V.C. (1991) Applications of a neurobehavioral screening battery. J. Am. Coll. Toxicol. 10:661-669.

Moser, V.C. (1990) Application and assessment of neurobehavioral screening methods in rats. In Advances in Neurobehavioral Toxicology: Applications in Environmental and Occupational Health, edited by B.L. Johnson. Chelsea, MI: Lewis Publishers; pp. 419-432.

Operator's Manual IL282 CO-Oximeter, Instrumentation Laboratory, Inc. (1977)
Lexington, MA

Smialowicz, R.J., J.E. Simmons, R.W. Luebke, and J.W. Allis (1991) Immunotoxicological assessment of subacute exposure of rats to carbon tetrachloride with comparison to hepatotoxicity and nephrotoxicity. Fund. Appl. Toxicol. 17: 186-196.

SAS Institute Inc. (1990) SAS/STAT User's Guide, Version 6. Cary, NC/ SAS Institute. 1990.

Standard Test Method for Estimating Sensory Irritancy of Airborne Chemicals. American Society for Testing and Materials, ASTM Designation E-981-84 (also appendix A)

Tegeris, J.S. and R.L. Balster (submitted) Comparison of the acute behavioral effects of alkylbenzenes using a functional observational battery in mice. Fund. Appl. Toxicol.

Tegeris, J.S. (1991) "Acute Behavioral Effects of Alkylbenzenes Evaluated Utilizing a Functional Observational Battery and Schedule-Controlled Operant Behavior." Ph.D. dissertation, Medical College of Virginia (Richmond, VA).

Tepper, J.S. and D.L. Costa (1992) Will the mouse bioassay for estimating sensory irritation (ASTM E-981-84) be useful for evaluation of indoor air contaminants. Indoor Environ. (1992)

Tilson, H.A. and V.C. Moser (1992) Comparison of screening approaches. NeuroToxicology 13:1-14.

U.S. Environmental Protection Agency (EPA) (1991) Neurotoxicity Test Guidelines Addendum 10, Pesticides Assessment Guidelines Subdivision F. Publication Number PB91-154617, Springfield Va. National Technical Information Services.

Wallach, J. (1978) Interpretation of Diagnostic Test: A Handbook Synopsis of Laboratory Medicine. Little, Brown and Company, Boston, MA.

Figure Legends

Figure 1. Diagram of the exposure system for evaluating carpet emissions.

Figure 2. Mean body weight of mice in each treatment group, recorded before each of the four exposures.

Figure 3. Mean frequency of breathing of mice during the control period: (a) collapsed over all four exposures for each treatment, and (b) collapsed across treatments for each of the four exposures.

Figure 4. Percent decrease in frequency of breathing of mice for each of the four exposures during the exposure period: (a) individual mouse data for each treatment for each exposure, and (b) mean percent decrease collapsed across treatments for each of the four exposures.

Figure 5. Mean percent change in frequency of breathing of mice for each of the four exposures during the recovery period: (a) collapsed over all four exposures for each treatment, and (b) collapsed across treatments for each of the four exposures.

Figure 6. Percent incidence of mice in each treatment group receiving a rating of none, slight, moderate, or severe sensory irritation during the exposure periods. Data are collapsed across all four exposures.

Figure 7. Percent incidence of mice in each treatment group receiving a rating of none, slight, moderate, or severe pulmonary irritation during the exposure periods. Data are collapsed across all four exposures.

Figure 8. Percent incidence of mice in each treatment group receiving a rating of none, slight, moderate, or severe disruption index. Data are collapsed across all periods of all four exposures.

Figure 9. Percent incidence of mice in each treatment group showing lacrimation following each of the four exposures.

Figure 10. Percent incidence of mice in each treatment group showing dilated pinna vessels following each of the four exposures.

Figure 11. Incidence of mice in each treatment group falling off the jar 0, 1, 2, or 3 times, after the second exposure (day 1) and the fourth exposure (day 2).

Figure 12. Incidence of mice in each treatment group showing decreased, the same, or increased scores in response to the tail-pinch stimulus. The score for each mouse was calculated as the difference from that mouse's preexposure value, and ranged from -1 ("<day 0"; 1 rank less than the preexposure score) to +2 (">>day 0"; 2 ranks higher than the preexposure score). Data are collapsed across the two postexposure tests (days 1 and 2).

Figure 13. Incidence of mice in each treatment group receiving a ranking of none, slight, clear, or extreme in response to the click stimulus. Data are collapsed across the two postexposure tests (days 1 and 2).

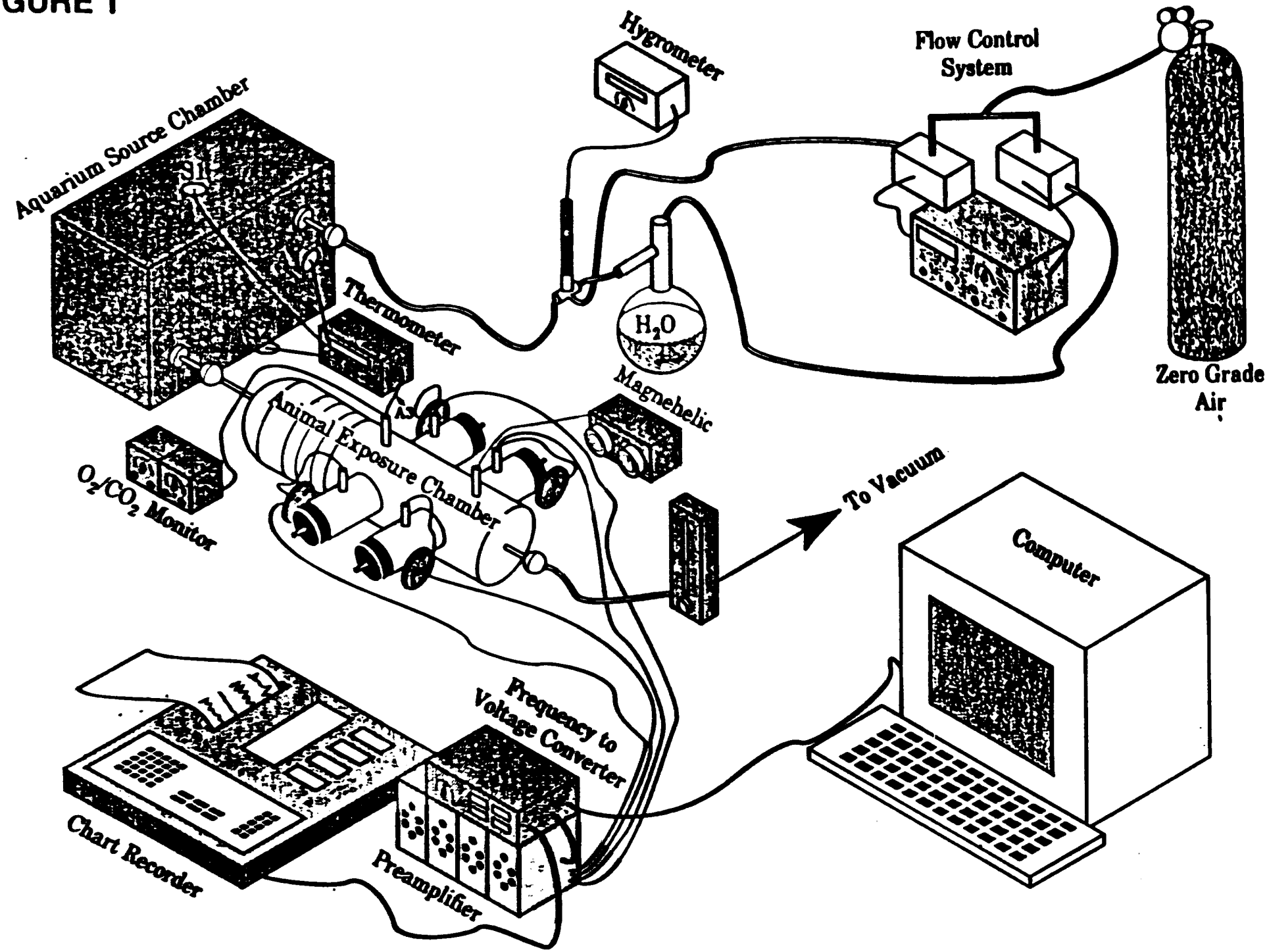
Figure 14. Incidence of mice in each treatment group receiving a score of stupor, low, alert, or excited during the open-field observation period. Data are collapsed across the two postexposure tests (days 1 and 2).

Figure 15. Mean hemoglobin values for mice in each treatment group and the non-exposed cage-control mice.

Figure 16. Mean serum chemistry values: (a) cholesterol, (b) protein, and (c) albumin. Data are presented for mice in each treatment group and the non-exposed cage-control mice.

Figure 17. Mean lavage values: (a) glutathione in nasal lavage fluid, and (b) protein in lung lavage fluid. Data are presented for mice in each treatment group and the non-exposed cage-control mice.

FIGURE 1



EPA - 87

FIG. 2 Mean Postexposure Body Weight

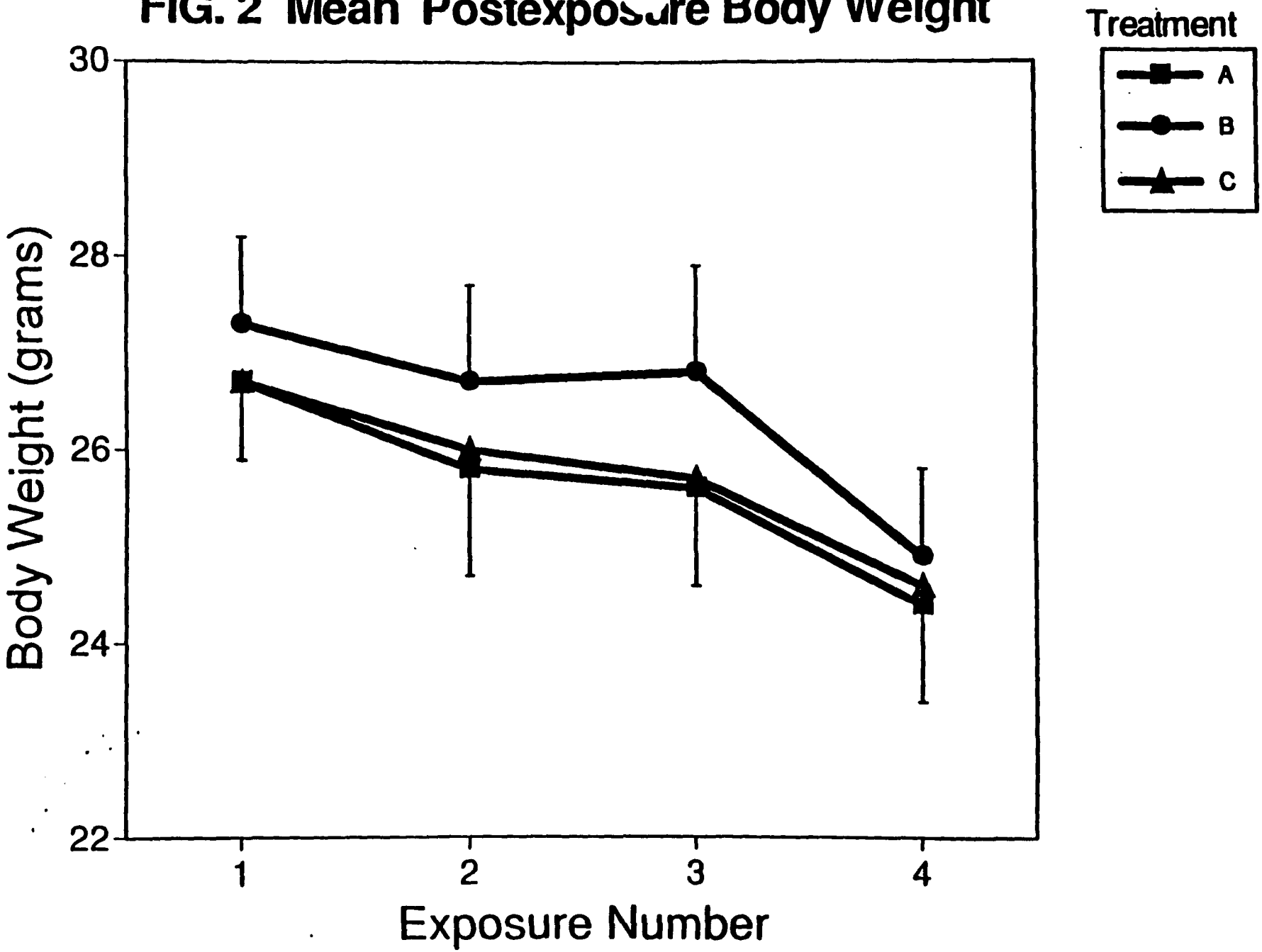
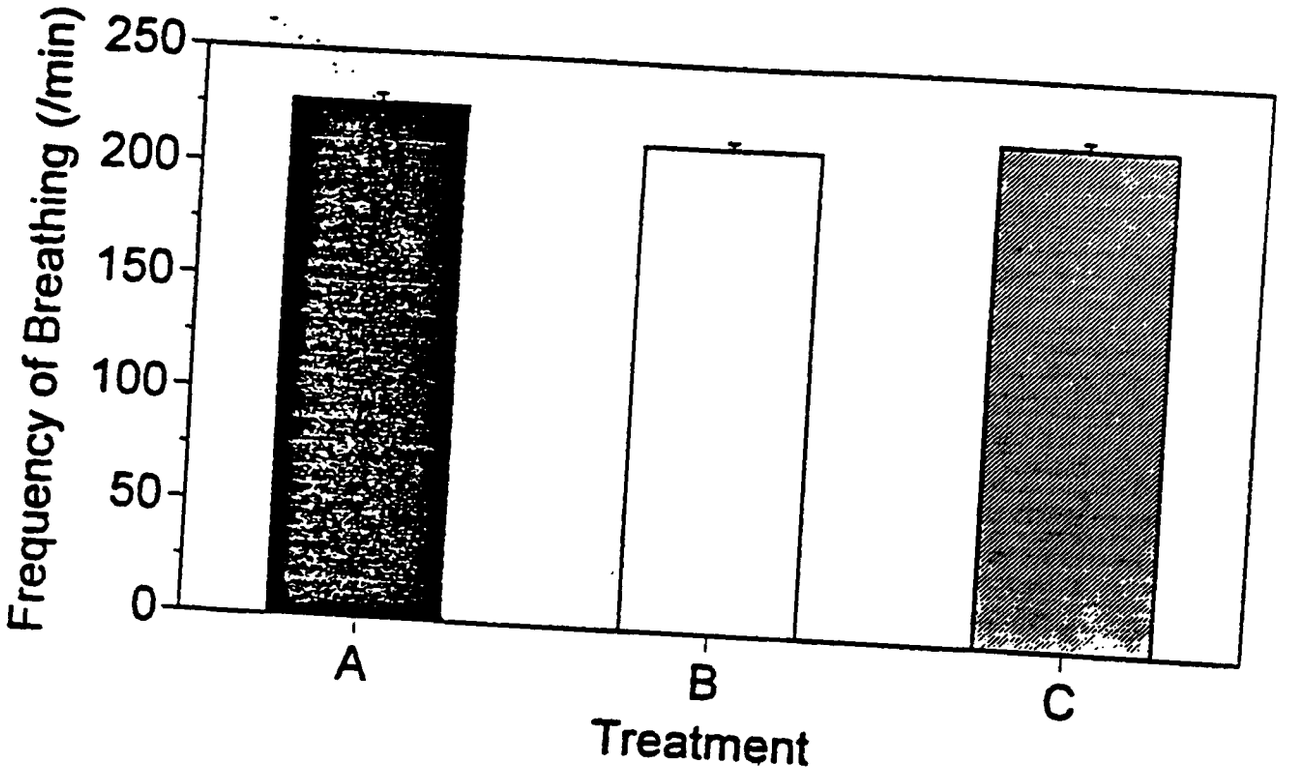


Fig. 3 Control Period



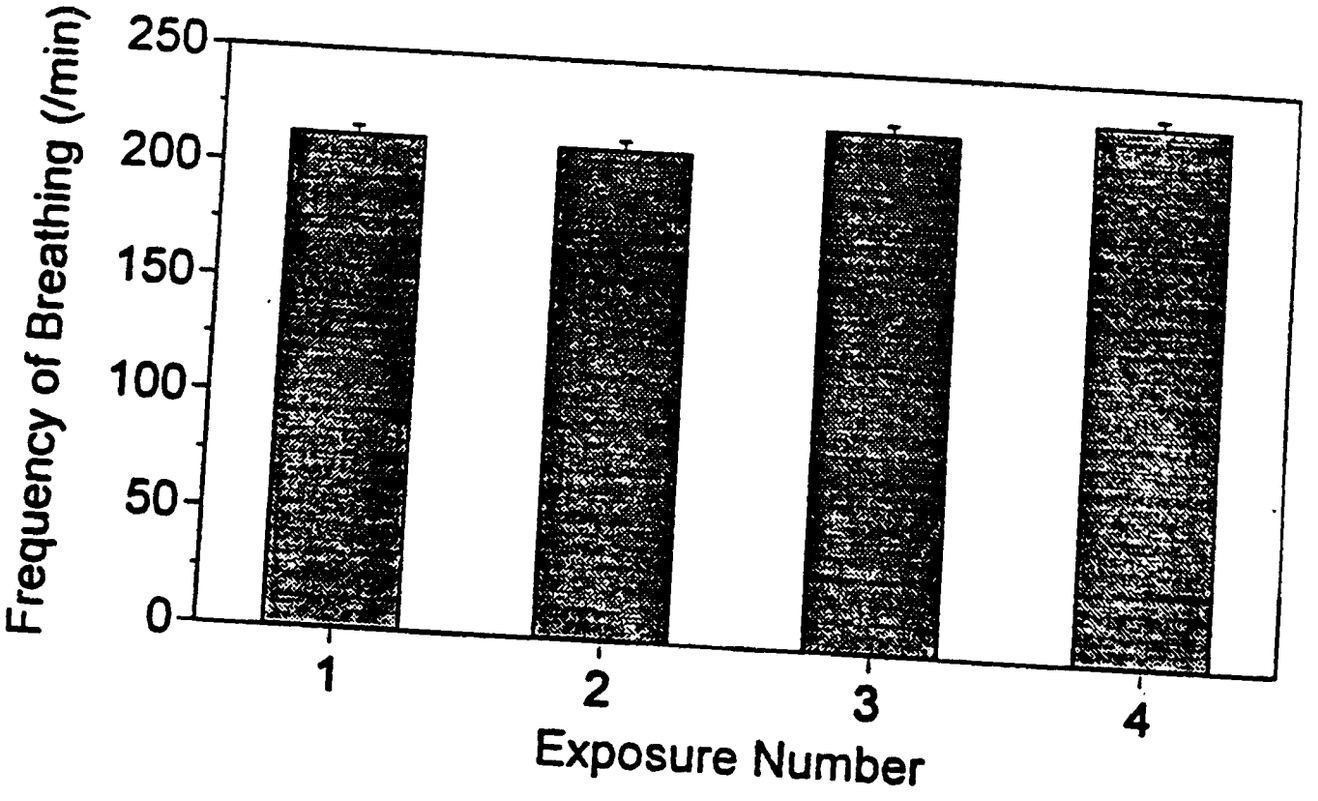
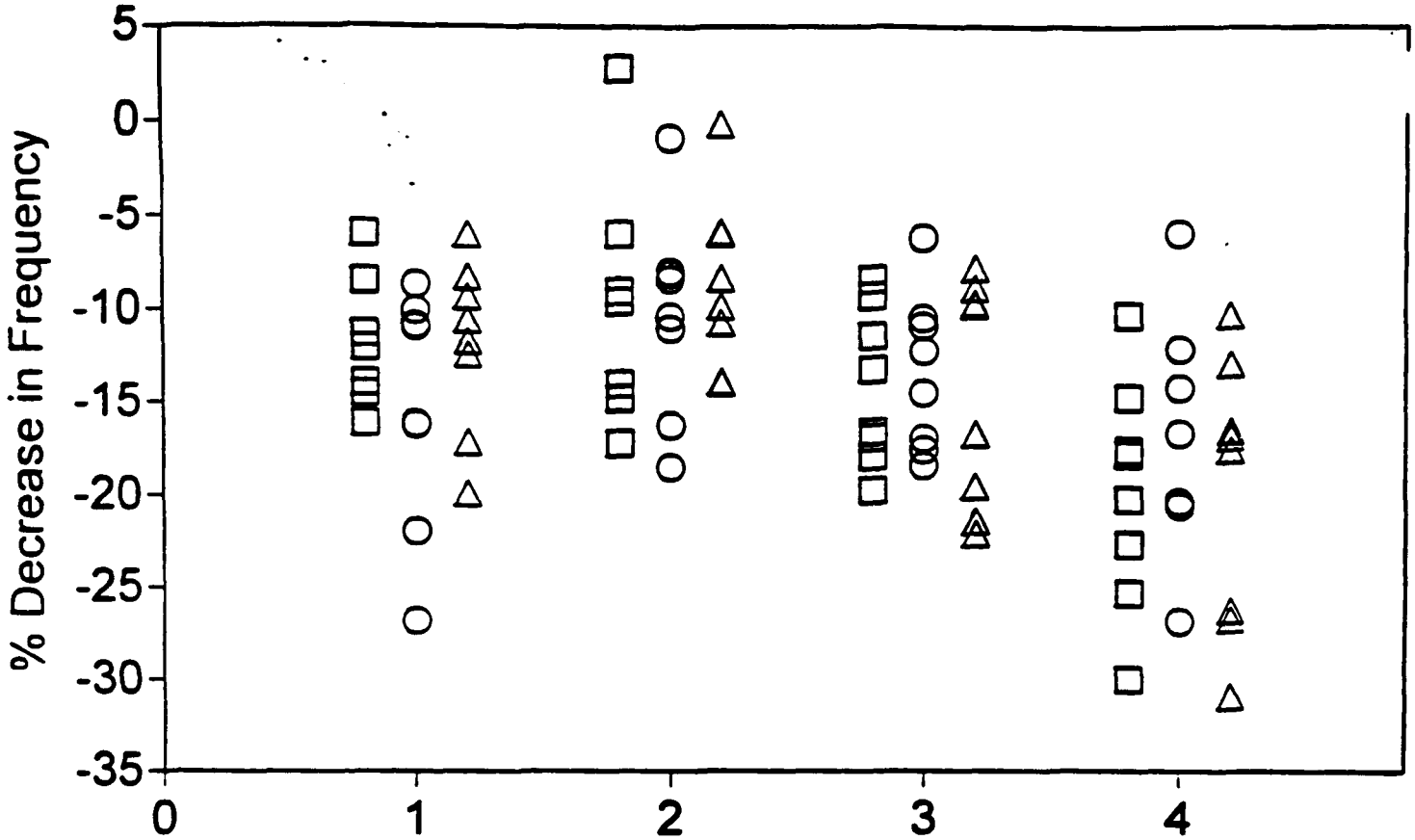


Fig. 4 Exposure Period



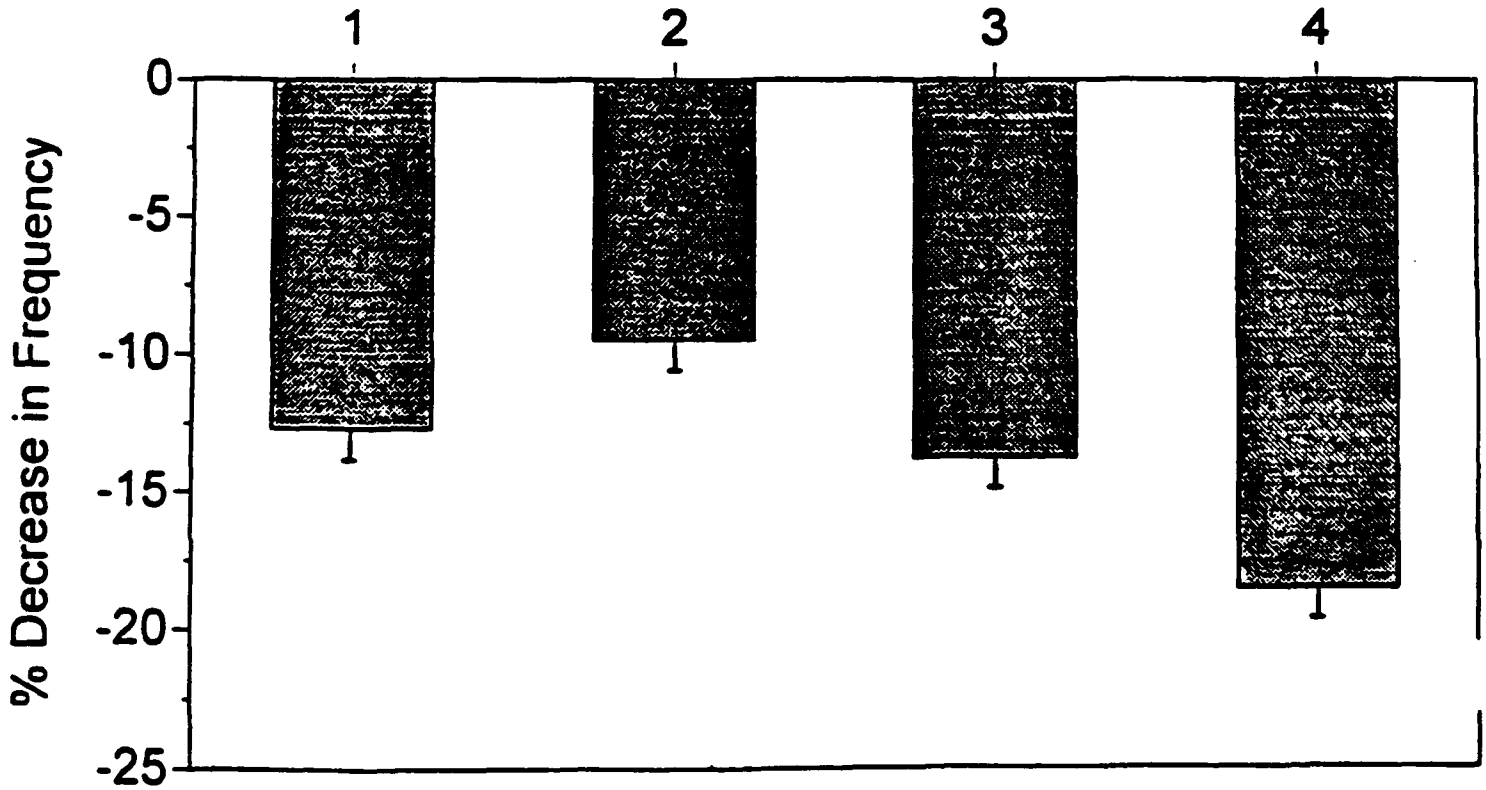
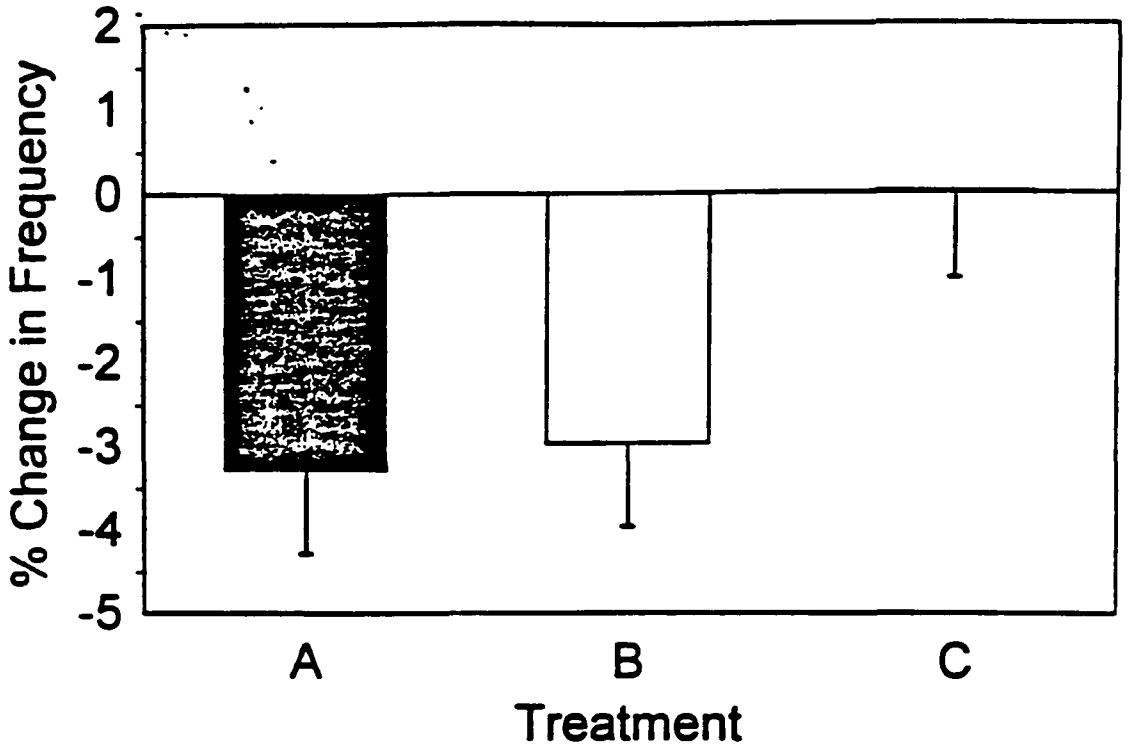


Fig. 5 Recovery Period



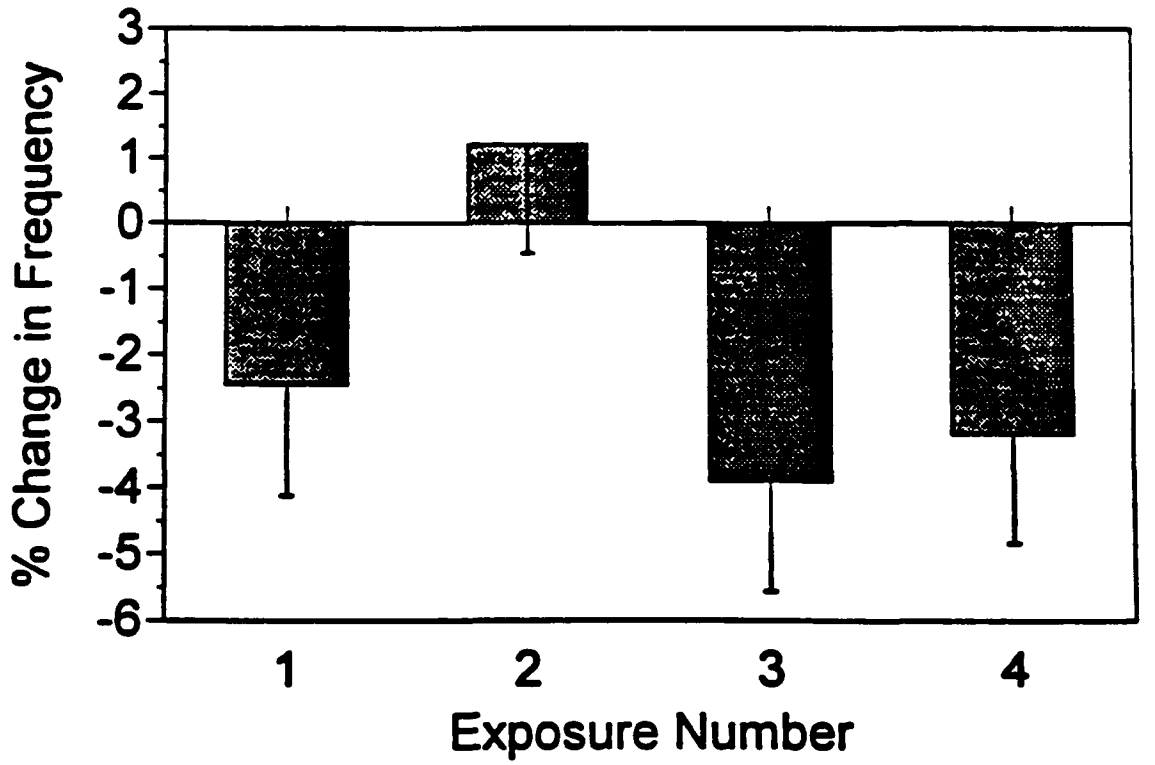
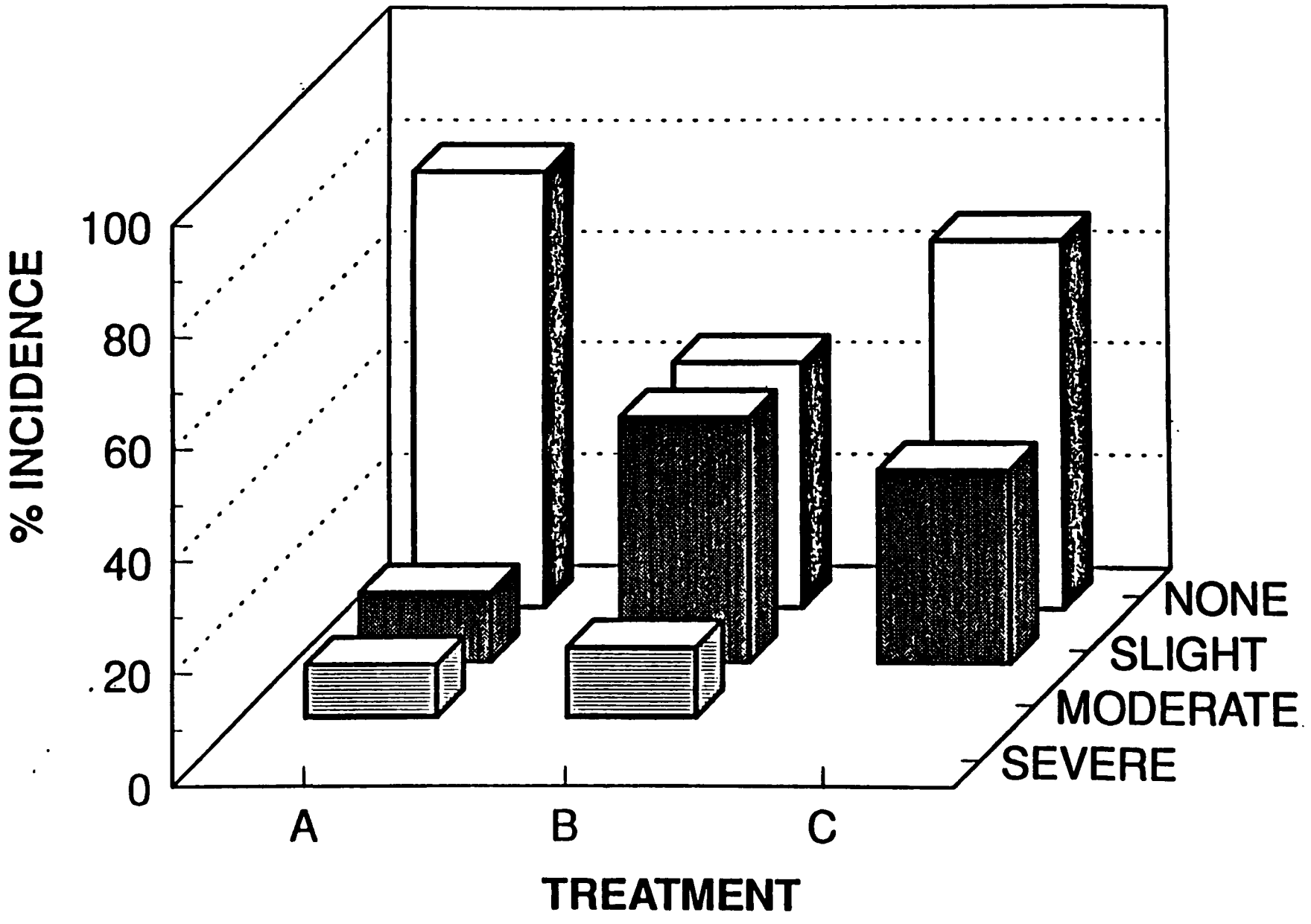
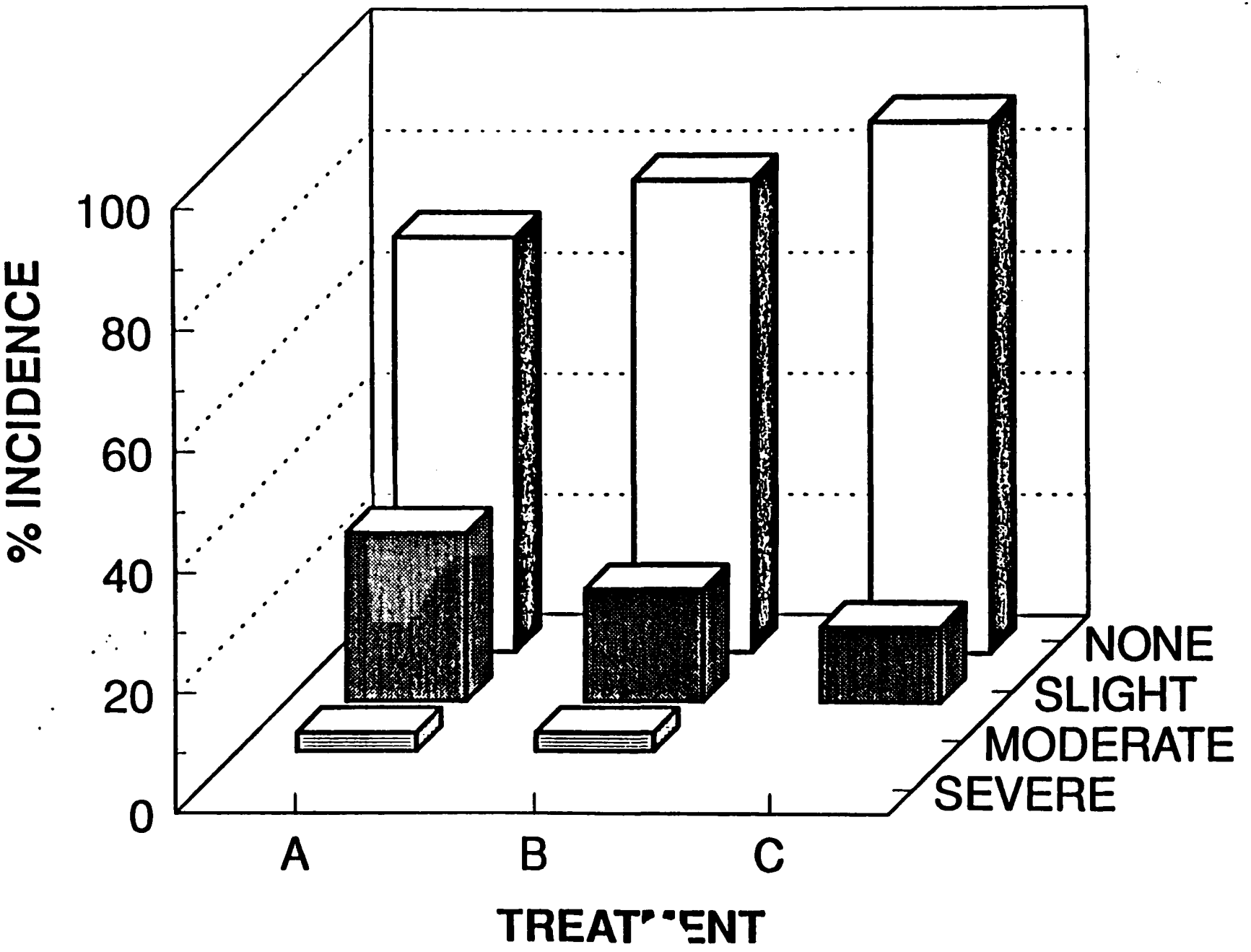


FIG. 6 SENSORY IRRITATION - EXPOSURE



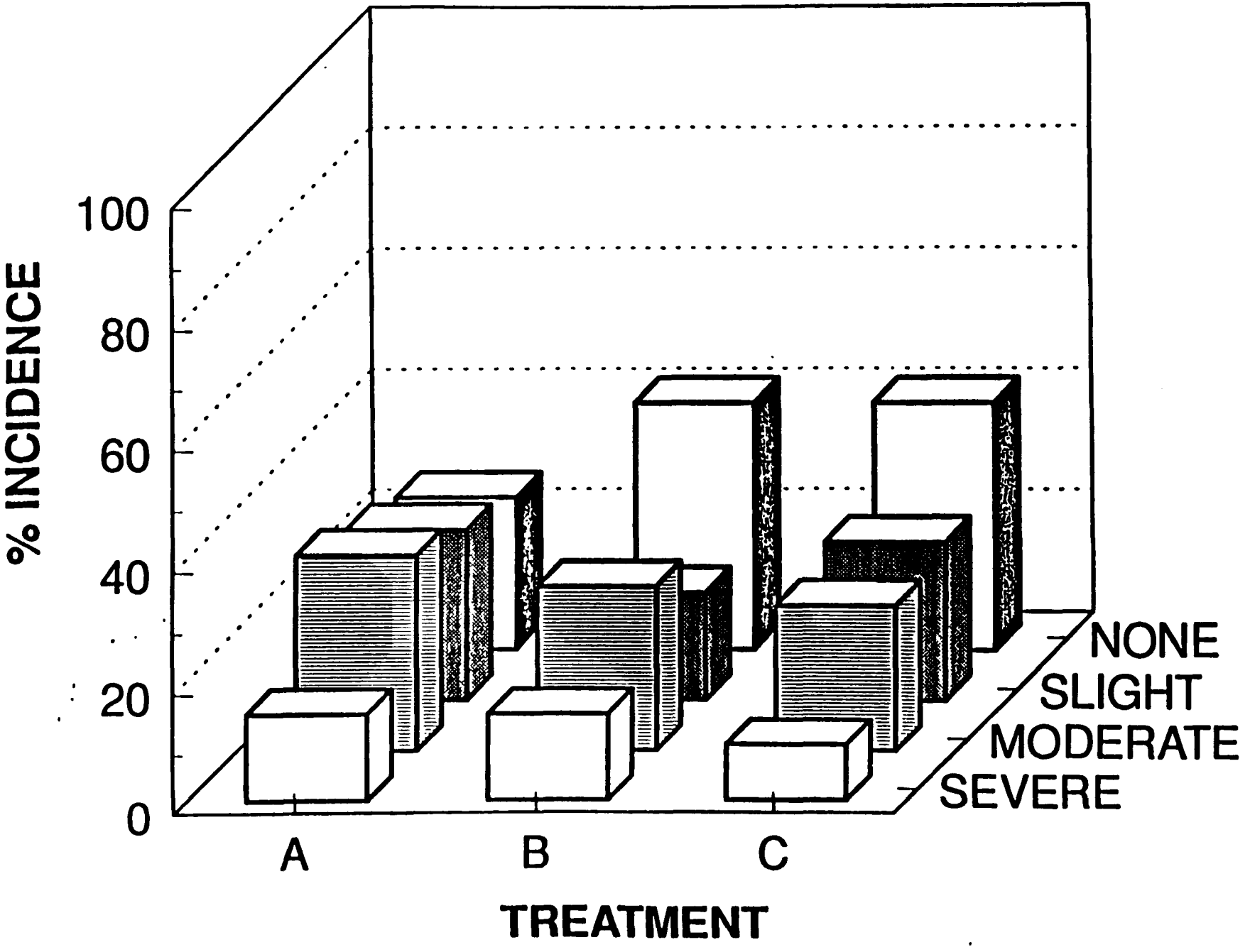
EPA - 92

FIG. 7 PULMONARY IRRITATION - EXPOSURE



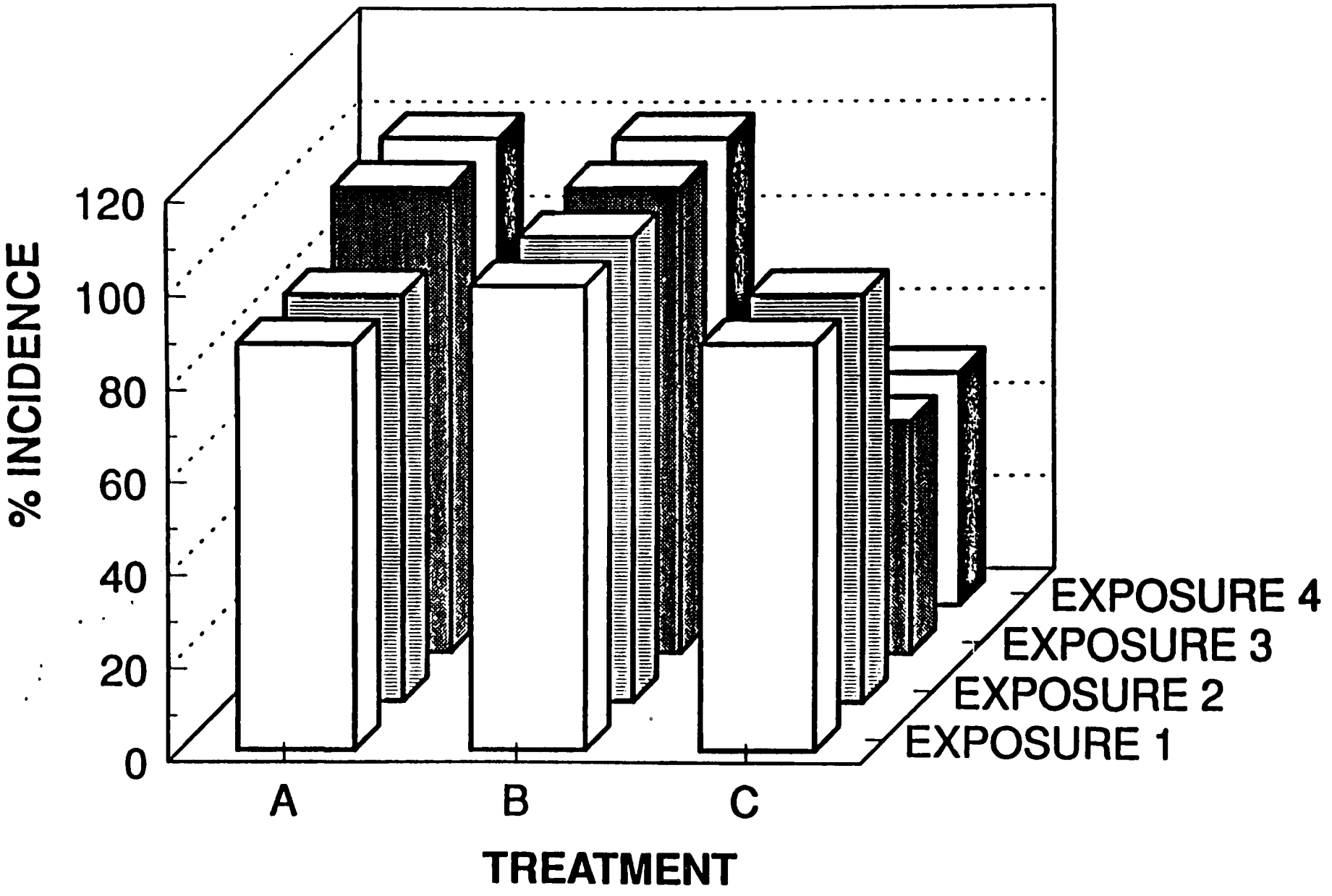
EPA - 93

FIG. 8 DISRUPTION INDEX - COLLAPSED



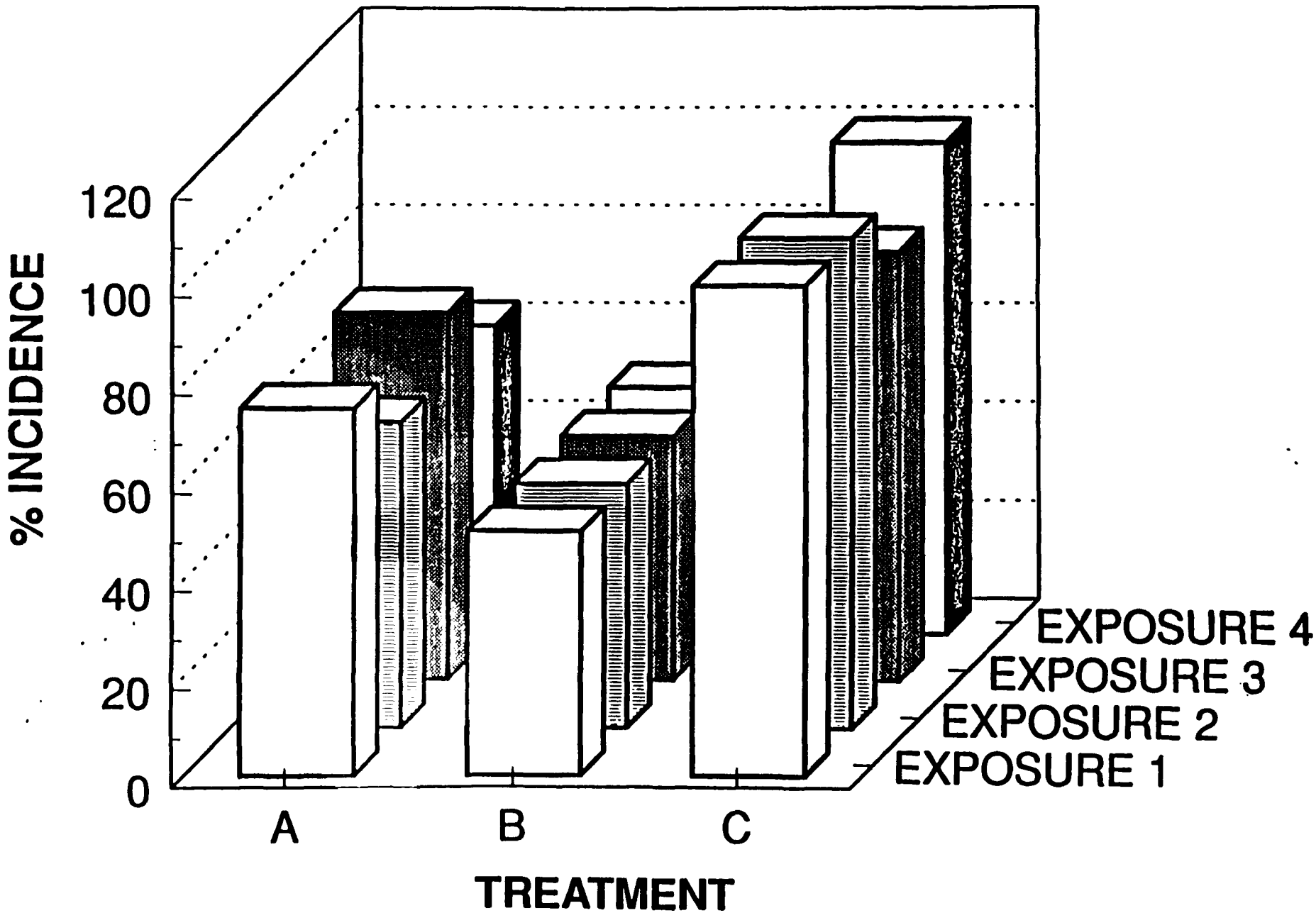
EPA - 94

FIG. 9 LACRIMATION



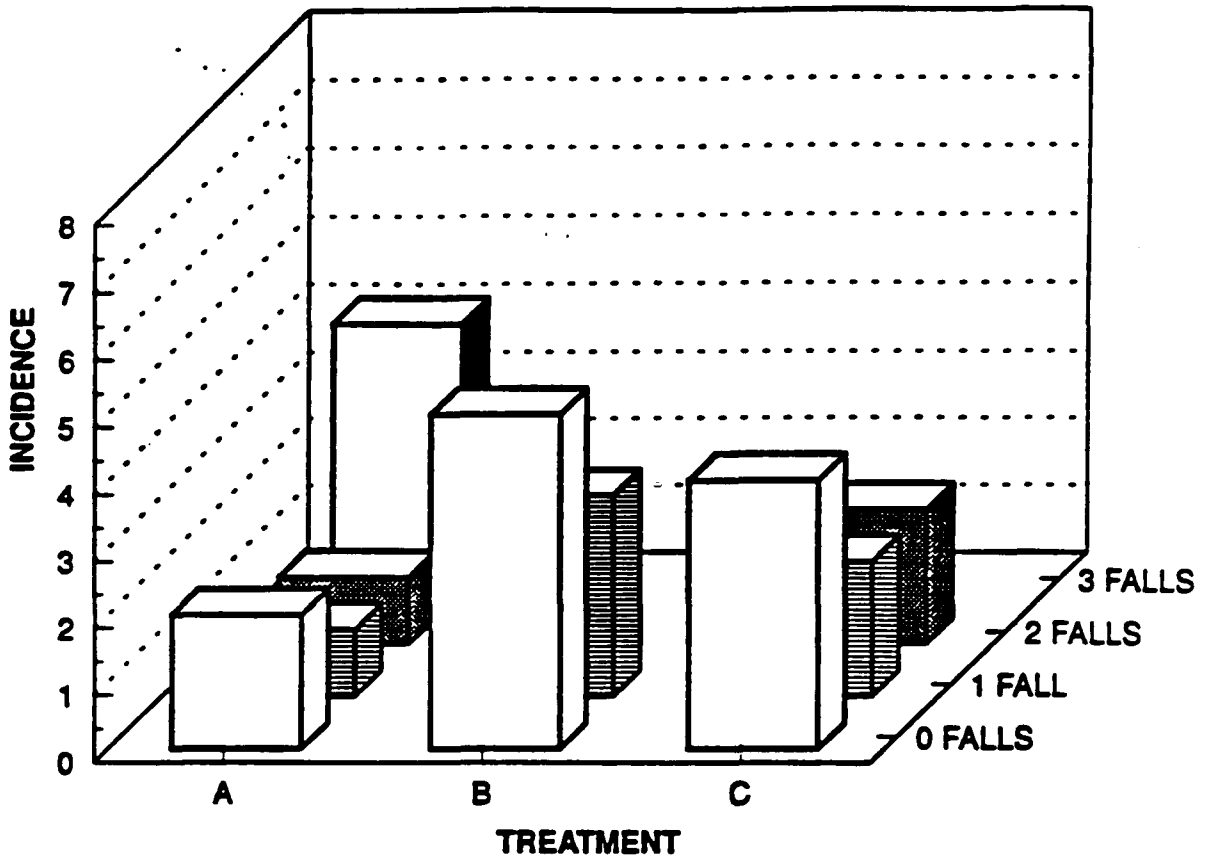
EPA - 95

FIG. 10 DILATED PINNA VESSELS



EPA - 96

FIG. 11 JAR TASK PERFORMANCE - DAY 1



JAR TASK PERFORMANCE - DAY 2

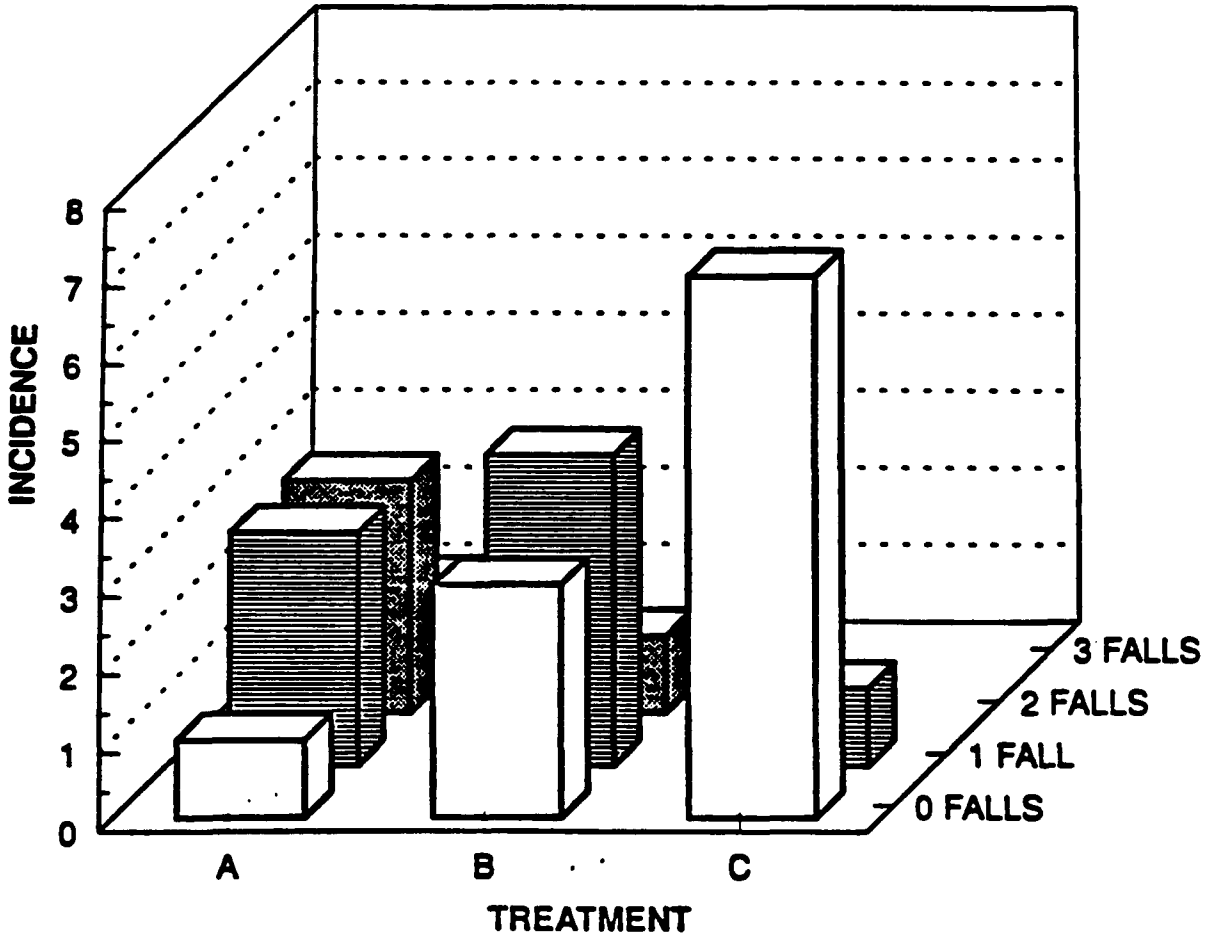
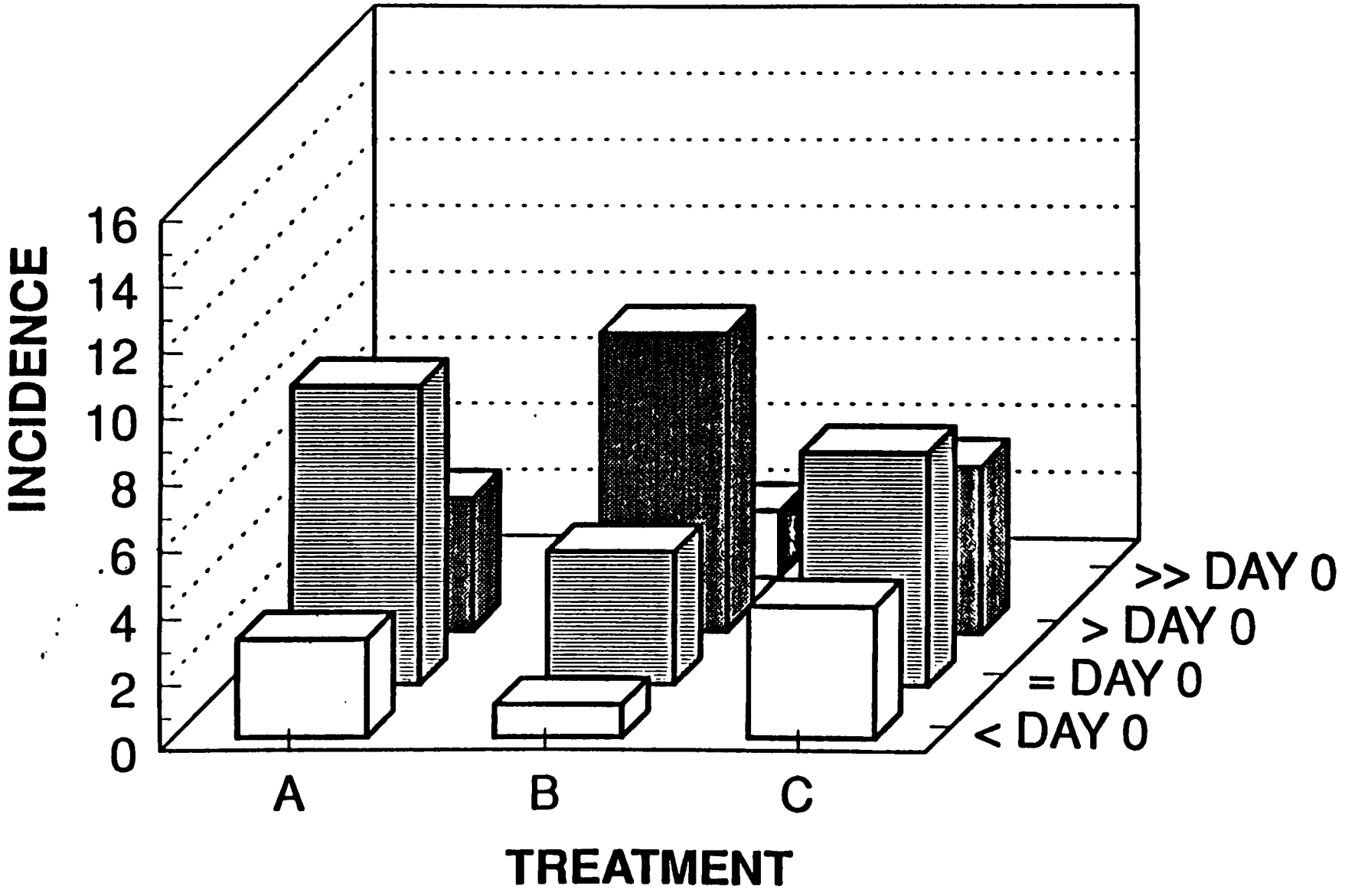


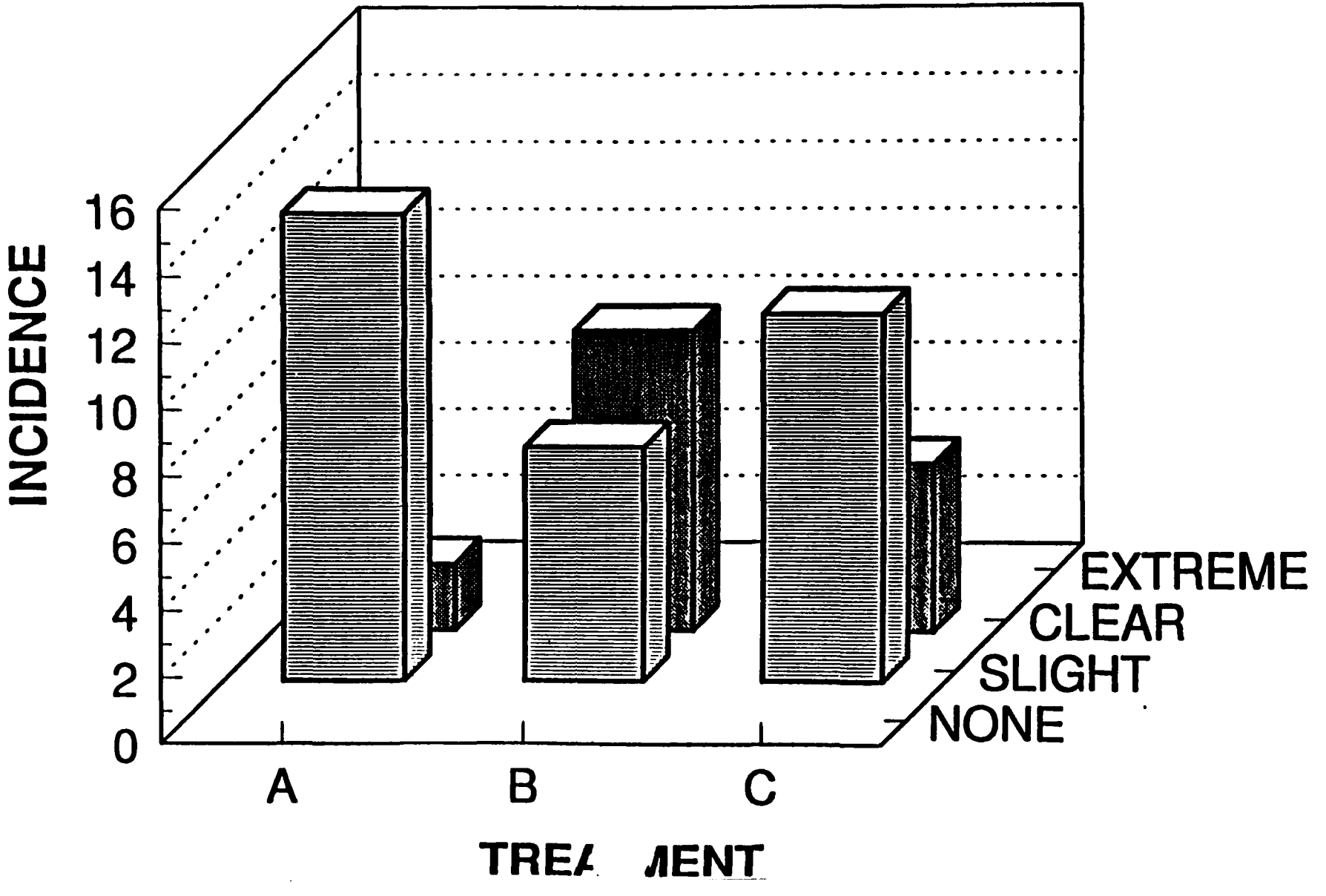
FIG. 12 TAIL PINCH RESPONSE

DAY 1 & 2 COMBINED



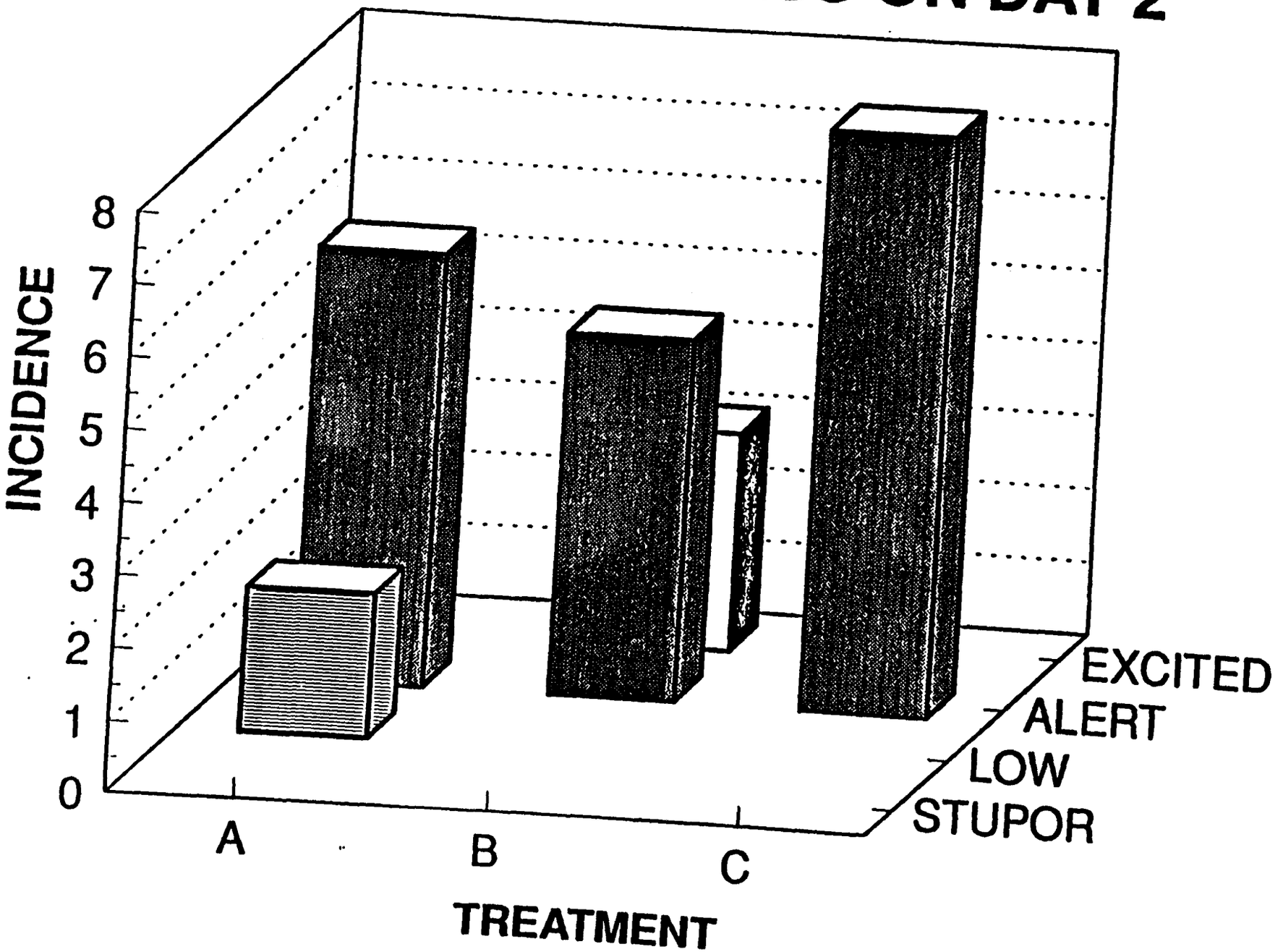
EPA - 98

FIG. 13 CLICK RESPONSE
DAYS 1 & 2 COMBINED



EPA - 99

FIG. 14 ALERTNESS ON DAY 2



EPA - 100

20. Gulumian, M.; Sardianos, F.; Kilroe-Smith, T.; Ockerse, G. Lipid peroxidation in microsomes induced by crocidolite fibers. Chem. Biol. Interact. 44:111-118; 1983.

21. Bonneau, L.; Pezerat, H. Etudes des sites donneurs et acceptures d'an election en surface des amiantes. J. Chem. Phys. 80:275-280; 1983.

22. Kamp, D. W.; Graceffa, P.; Pryor, W.A.; Weitzman, S.A. The role of free radicals in asbestos induced diseases. Free Rad. Biol. Med.12:293-315;1992.

23. Fontecave, M.; Mansuy, D.; Jaquen, M.; Pezerat, H. The stimulatory effects of asbestos as NADPH-dependent lipid peroxidation in rat liver microsomes. Biochem. J. 241:561-565; 1987.

24. Morehouse, L.A.; Thomas, C.E.; Aust, S.D. Superoxide generation by NADPH-cytochrome P-450 reductase: The effect of iron chelators and the role of superoxide in microsomal lipid peroxidation. Arch. Biochem. Biophys. 232: 366-377; 1984.

25. Aust, S.D.; Morehouse, L.A.; Thomas, C.E. Role of metals in oxygen radical reactions. Free Rad. Biol. Med. 1:3-25; 1985.

26. Diplock, A.T. Vitamin E, In: Fat soluble vitamins, their biochemistry and their applications; Diplock, A.T., ed, Technonic Publishing Co. Inc., Lancaster-Basal; 1985:154-224.

27. Kornbrust, D.J.; Mavis, R.D. Relative susceptibility of microsomes from lung, heart, liver, kidney, brain and testes to lipid peroxidation: Correlation with vitamin E content. Lipid 15:315-322; 1980.

28. Burk, R.F. Glutathione-dependent protection by rat liver microsomal protein against lipid peroxidation. Biochim. Biophys. Acta, 757:21-28; 1983.

Table:1

Chrysotile-mediated enzymatic lipid peroxidation in rat lung microsomes supplemented with vitamin E in presence of GSH and other quenchers of reactive oxygen species.

Incubation system	Control group	+ Vitamin E group	% Protection by vitamin E
Microsomes	0	0	-
+ NADPH (0.4 mM)	0.56±0.04	0.19±0.06	-
+ Chrysotile (500 ug)	0.48±0.01	0.16±0.04	-
Complete	2.50±0.35	1.23±0.17	51
+ GSH (1 mM)	1.27±0.17^{c}	0.30±0.06$^{a,(a)}$	76
+SOD (100 units)	1.83±0.21^{NS}	0.67±0.04$^{a,(a)}$	63
+ Mannitol (1 mM)	1.77±0.23^{NS}	0.70±0.05$^{c,(b)}$	55
+Catalase (150 units)	1.80±0.22^{NS}	0.74±0.04$^{c,(a)}$	59
+ B-Carotene (0.5 mM)	1.98±0.31^{NS}	0.72±0.08$^{c,(b)}$	64

Values are mean ± SE (n=3) and expressedas nmol MDA formed/mg protein. Complete = Microsomes + NADPH + Chrysotile

Letters without parentheses = Compared with respective complete systems. Letters in parentheses = Compared with respective control groups. $^{a}p< 0.01$; $^{b}p< 0.02$; $^{c}p< 0.05$; NS = Not significant.

Table:2

Effect of chrysotile on the activity of vitamin E regeneration factor in rat lung microsomes supplemented with vitamin E in presence of GSH and other quenchers of reactive oxygen species.

Incubation system	Control group		+ Vitamin E group	
	20 min.	40 min.	20 min.	40 min.
Microsomes + Ascorbate	0.096±0.006	0.108±0.004	0.080±0.006	0.089±0.007
Microsomes + Ascorbate + GSH (1mM)	0.032±0.002	0.036±0.003	0.015±0.001	0.025±0.002
Complete	0.164±0.009	0.194±0.016	0.120±0.009	0.146±0.008
+ GSH (1mM)	0.100±0.005[b]	0.115±0.005[a]	0.046±0.002[a]	0.058±0.006[a]
+ SOD (100 units)	0.117±0.004[a]	0.130±0.006[a]	0.079±0.006[b]	0.085±0.009[a]
+ Mannitol (1 mM)	0.120±0.005[b]	0.129±0.003[a]	0.078±0.010[c]	0.088±0.008[a]
+ Catalase (150 units)	0.115±0.009[a]	0.120±0.015[b]	0.081±0.006[b]	0.090±0.011[b]
+ B-Carotene (0.5 mM)	0.110±0.010[b]	0.124±0.001[a]	0.076±0.005[b]	0.085±0.012[b]

Values are mean ± SE (n=3) and expressed as $\Delta A_{535-600}$. Complete = Microsomes + Ascorbate + Chrysotile.

Letters = Compared with respective complete systems. [a] $p<0.01$; [b] $p<0.02$; [c] $p<0.05$.

Table:3

Effect of chrysotile on the vitamin E content in rat lung microsomes supplemented with vitamin E in the presence of GSH and other quenchers of reactive oxygen species.

Incubation system	Control group	+ Vitamin E group
Microsomes	0.70±0.08	1.92±0.06
+ NADPH (0.4 mM)	0.38±0.08$^{(d)}$	1.61±0.06$^{(d)}$
+ Chrysotile (500 ug)	0.40±0.07$^{(d)}$	1.50±0.08$^{(c)}$
Complete	0.26±0.05$^{(b)}$	1.06±0.08$^{(a)}$
+ GSH (1 mM)	0.64±0.11^{d}	1.99±0.09^{b}
+ SOD (100 units)	0.31±0.10^{NS}	1.49±0.04^{b}
+ Mannitol (1 mM)	0.34±0.08^{NS}	1.45±0.02^{b}
+ Catalase (150 units)	0.36±0.07^{NS}	1.52±0.04^{b}
+ B-Carotene (0.5 mM)	0.33±0.08^{NS}	1.55±0.07^{b}

Values are mean ± SE (n=3) and expressed as n moles/mg protein. Complete = Microsomes + NADPH + Chrysotile. Letters in parenthesis = Compared with microsomes only. Letters without parenthesis = Compared with respective complete systems. $^{a}p<0.001$; $^{b}p<0.01$; $^{c}p<0.02$; $^{d}p<0.05$; NS = Not significant.

Table:4

Effect of chrysotile on the glutathione-S-transferase activity profiles in rat lung microsomes supplemented with vitamin E in presence of GSH.

Incubation system	Control group	+ Vitamin E group	% Protection by vitamin E
Microsomes	50.42±4.79	49.33±3.88	-
+ NADPH (0.4 mM)	150.89±10.11	121.01±9.25	-
+ Chrysotile (500 ug)	149.71±11.75	115.60±10.40	-
Complete	181.35±10.21	143.34±9.75$^{(b)}$	21
+ GSH (1 mM)	79.43±9.24^{a}	54.30±8.29$^{a,(NS)}$	32

Enzyme activity is mean±SE (n = 3) and expressed as nmol CDNB conjugated/mg protein/min.

Complete = Microsomes + NADPH + Chrysotile. Letters without parentheses = Compared with respective complete system. Letters in parentheses = Compared with respective control group. $^{a}p<0.02$; ^{b} $p<0.05$; NS = Not significant.

FIG. 15 Hemoglobin

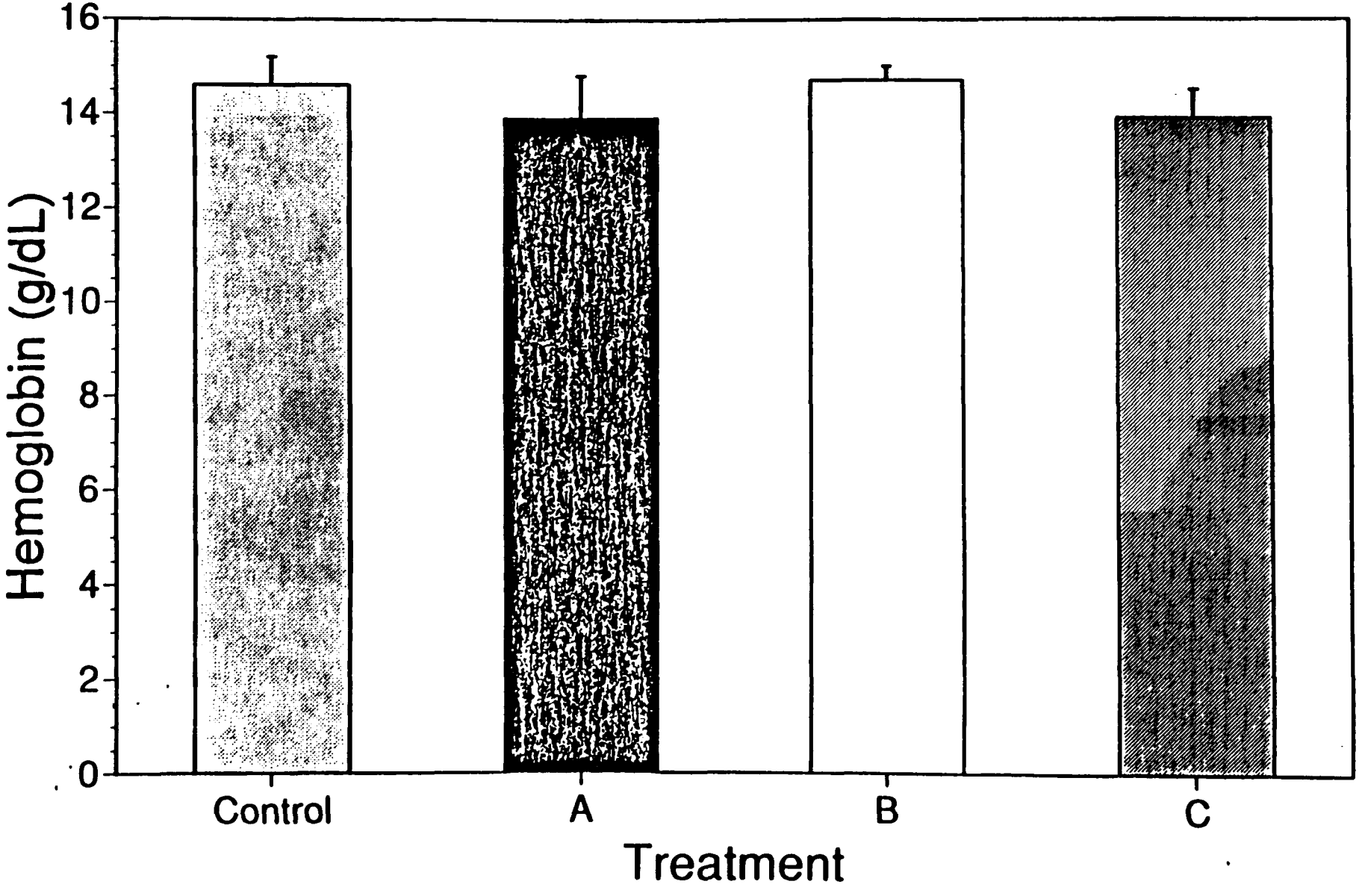


Fig. 16 Serum Chemistries

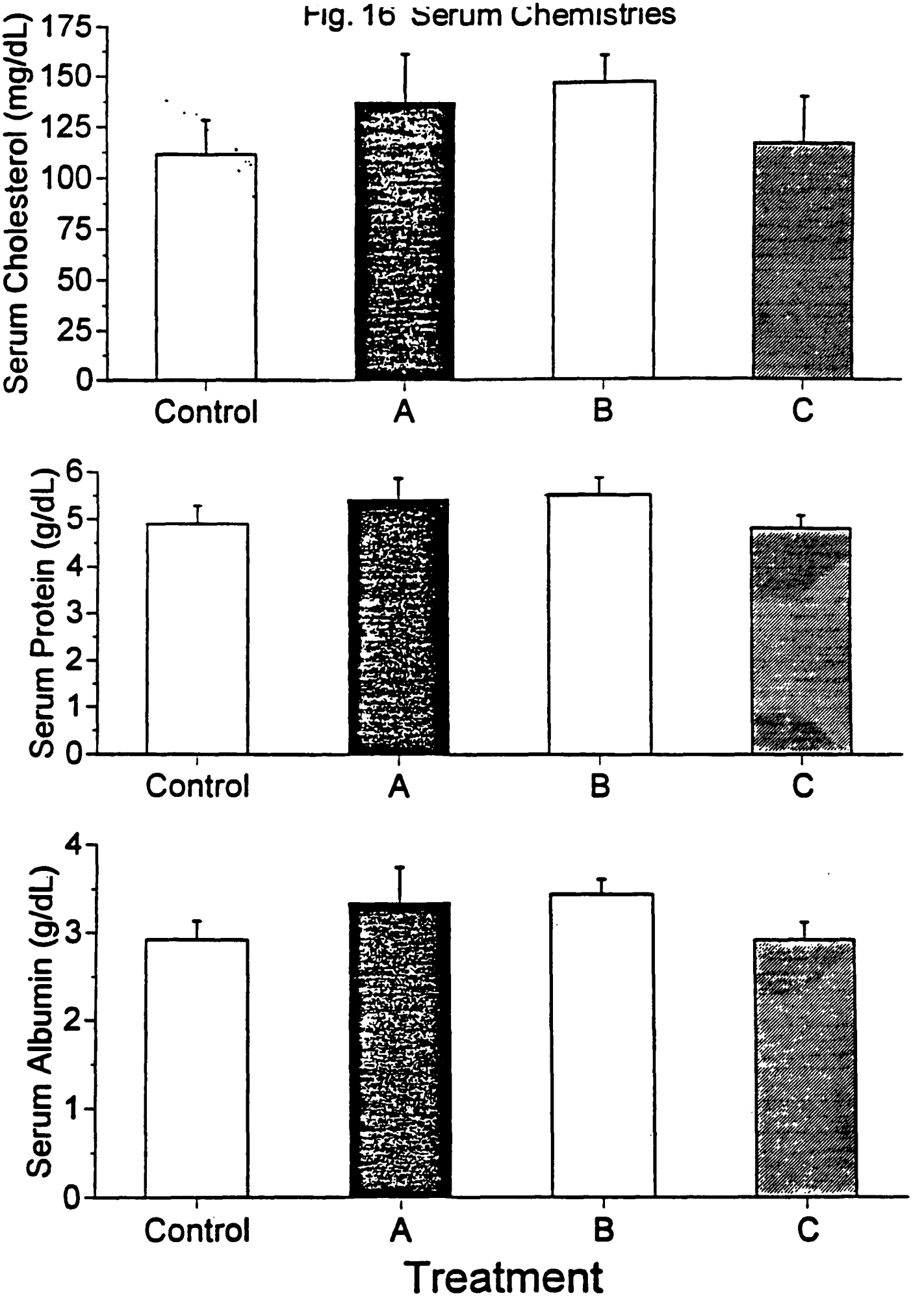
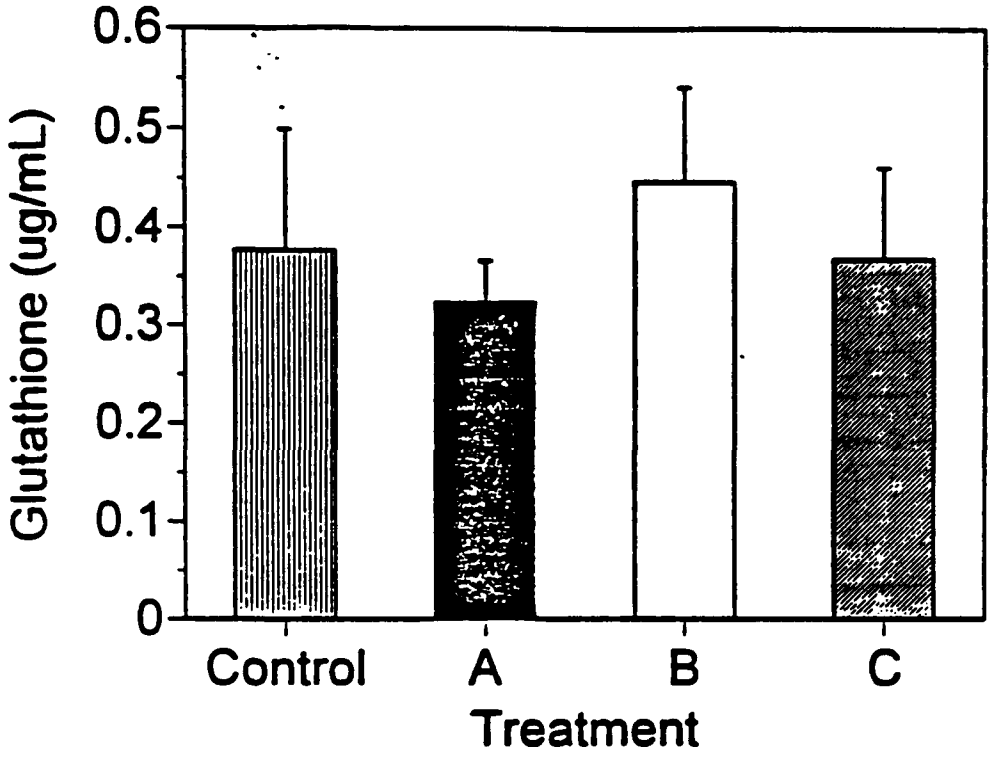
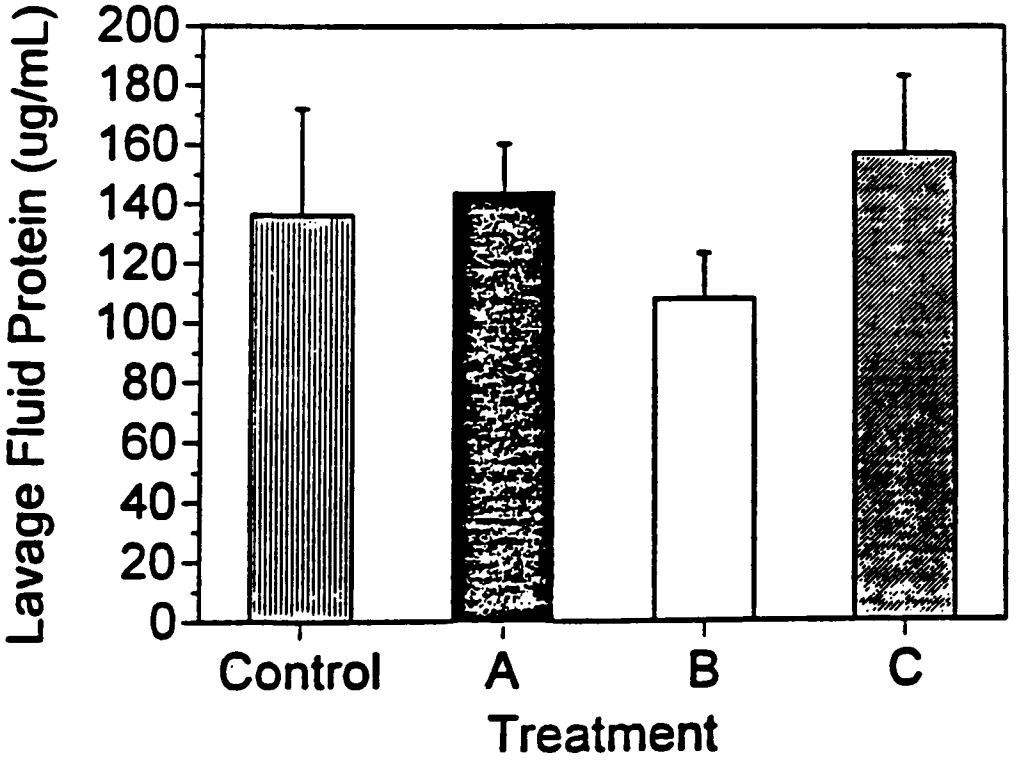


Fig . 17

Nasal Lavage



Lung Lavage



APPENDIX A

SUMMARY OF FOB DATA FOR EPA FORMAL REPLICATION STUDY

SUMMARY OF FOB DATA FOR EPA FORMAL REPLICATION STUDY

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Neuromuscular Function:									
Jar Task									
fell 0x	4	2	1	6	5	3	4	4	7
1x	2	1	3	1	3	4	2	2	1
2x	1	1	3	1	0	1	2	2	0
3x	1	4	1	0	0	0	0	0	0
Grip Strength									
dropped 0x	8	8	8	8	7	7	8	8	8
1x	0	0	0	0	0	0	0	0	0
2x	0	0	0	0	1	1	0	0	0
3x	0	0	0	0	0	0	0	0	0
Body Tone									
1 hypotonia	0	0	0	0	0	0	0	0	0
2 slightly flaccid	1	2	4	0	2	2	0	1	1
3 normal	7	6	4	8	6	6	8	7	7
4 hypertonia	0	0	0	0	0	0	0	0	0

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Righting Reflex									
1 normal	8	8	8	8	7	7	8	7	8
2 slightly slow	0	0	0	0	1	1	0	1	0
3 difficult	0	0	0	0	0	0	0	0	0
4 not present	0	0	0	0	0	0	0	0	0
Body Posture									
1 lying on side	0	0	0	0	0	0	0	0	0
2 pelvis flat	0	0	0	0	0	0	0	0	0
3 upright	8	8	8	8	8	7	8	8	8
4 hunched	0	0	0	0	0	1	0	0	0
Body Tilt									
1 normal	8	8	8	8	8	8	8	8	8
2 head tilts	0	0	0	0	0	0	0	0	0
3 shoulder leans	0	0	0	0	0	0	0	0	0
4 body lists	0	0	0	0	0	0	0	0	0

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Ataxic Gait									
1 none	8	6	7	8	8	6	8	8	8
2 slight	0	2	1	0	0	2	0	0	0
3 somewhat	0	0	0	0	0	0	0	0	0
4 marked	0	0	0	0	0	0	0	0	0
5 severe	0	0	0	0	0	0	0	0	0
Abnormal Gait Score 1 none									
2 slight	5	1	1	5	0	1	3	1	1
3 somewhat	3	3	4	3	2	0	5	3	2
4 marked	0	4	3	0	5	6	0	3	5
5 severe	0	0	0	0	1	1	0	1	0
	0	0	0	0	0	0	0	0	0
Inverted Screen									
fast <30 sec		5	5		4	5		3	5
slow ≥30 sec		2	2		2	0		4	3
hanging		1	1		1	2		1	0
dropped		0	0		1	1		0	0

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Missteps									
1 none	7	5	3	5	3	4	4	5	6
2 few	1	3	5	3	4	2	4	3	2
3 some	0	0	0	0	1	2	0	0	0
4 legs hanging	0	0	0	0	0	0	0	0	0
Sensorimotor Measures:									
Forelimb Placing									
present	8	8	8	8	8	8	8	8	8
absent	0	0	0	0	0	0	0	0	0
Air Puff Response									
1 none	0	0	0	0	0	0	0	0	0
2 slight	2	0	2	2	0	2	0	1	1
3 clear	6	8	5	5	5	4	6	5	4
4 extreme	0	0	1	1	3	2	2	2	3

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Click Response									
1 none	0	0	0	0	0	0	0	0	0
2 slight	4	7	7	4	3	4	6	4	7
3 clear	4	1	1	3	5	4	2	4	1
4 extreme	0	0	0	1	0	0	0	0	0
Tail Pinch Response									
1 none	0	0	0	1	0	0	0	0	0
2 slight	2	2	2	3	1	1	0	1	1
3 clear	6	6	5	4	5	5	7	5	4
4 extreme	0	0	1	0	2	2	1	2	3

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Activity and Excitability Measures:									
Alertness									
1 stupor	0	0	0	0	0	0	0	0	0
2 low	0	0	2	0	1	0	0	0	0
3 alert	8	8	6	7	6	5	8	8	8
4 hyperalert	0	0	0	1	1	3	0	0	0
Handling Reactivity									
1 low	1	0	0	1	0	1	1	2	1
2 slight	5	4	5	3	5	2	4	3	1
3 active	2	4	3	4	2	4	2	2	6
4 moderate	0	0	0	0	1	1	1	1	0
5 high	0	0	0	0	0	0	0	0	0

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Activity Level									
1 none	0	0	1	0	0	0	0	0	0
2 low	0	1	1	1	0	0	0	0	0
3 somewhat low	2	3	3	2	2	2	1	2	1
4 active	3	2	1	2	3	1	3	3	1
5 clear	3	0	1	3	2	2	4	3	6
6 hyperactive	0	2	1	0	1	3	0	0	0
Rears									
X	10.0	11.5	8.6	7.4	7.6	13.3	9.5	9.1	7.3
SEM	2.4	3.0	2.7	2.4	2.7	3.2	2.4	2.5	1.8
Clonic Movements									
1 none	8	8	7	8	8	7	8	8	8
2 smacking	0	0	0	0	0	0	0	0	0
3 quivers	0	0	1	0	0	1	0	0	0
4 mild tremors	0	0	0	0	0	0	0	0	0
5 severe tremors	0	0	0	0	0	0	0	0	0
6 myoclonus	0	0	0	0	0	0	0	0	0
7 convulsions	0	0	0	0	0	0	0	0	0

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Tonic Movements									
1 none	8	8	8	8	8	8	8	8	8
2 extension	0	0	0	0	0	0	0	0	0
3 opisthotonus	0	0	0	0	0	0	0	0	0
4 emprosthotonus	0	0	0	0	0	0	0	0	0
5 popcorn	0	0	0	0	0	0	0	0	0
6 convulsion	0	0	0	0	0	0	0	0	0

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Other Measures:									
Tilted Screen									
up>2x	4	1	1	6	4	4	5	2	2
even>2x	1	2	2	0	1	2	1	1	0
down>2x	3	3	3	2	3	2	2	3	6
1,1,1	0	2	2	0	0	0	0	2	0
Vocalizations									
present	0	0	0	0	0	0	0	0	0
none	8	8	8	8	8	8	8	8	8
Bizarre Behaviors									
present	0	0	0	0	0	0	0	0	0
none	8	8	8	8	8	8	8	8	8
Stereotypies									
present	0	0	0	0	0	0	0	0	0
none	8	8	8	8	8	8	8	8	8
Diarrhea									
present	0	0	0	0	0	0	0	0	0
none	8	8	8	8	8	8	8	8	8

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
General Appearance Measures:									
Facial Swelling									
present	0	8	3	0	8	7	0	8	4
absent	8	0	5	8	0	1	8	0	4
Chromodacryorrhea									
present	0	2	1	0	4	3	0	3	0
absent	8	6	7	8	4	5	8	5	8
Lacrimation									
1 none	8	0	1	8	0	0	8	0	4
2 mild	0	7	4	0	6	6	0	8	4
3 severe	0	1	3	0	2	2	0	0	0
Gasping									
present	0	0	0	0	0	0	0	0	0
absent	8	8	8	8	8	8	8	8	8
Cyanosis									
present	0	0	0	0	0	0	0	0	0
absent	8	8	8	8	8	8	8	8	8

	CARPET A samples 1,6			CARPET B samples 2,3			CARPET C samples 4,5		
	Test Day			Test Day			Test Day		
	0	1	2	0	1	2	0	1	2
Exophthalmus									
present	0	0	0	0	0	0	0	0	0
absent	8	8	8	8	8	8	8	8	8
Palpebral Closure									
1 eyes open	8	8	8	8	8	8	8	7	8
2 slight droop	0	0	0	0	0	0	0	0	0
3 ptosis	0	0	0	0	0	0	0	1	0
4 eyes shut	0	0	0	0	0	0	0	0	0
Piloerection									
present	0	1	0	0	2	0	0	0	0
absent	8	7	8	8	6	8	8	8	8
Salivation									
1 none	8	8	8	8	8	8	8	8	8
2 mild	0	0	0	0	0	0	0	0	0
3 severe	0	0	0	0	0	0	0	0	0
Splotchy ears									
present	0	6	1	0	7	5	0	5	3
absent	8	2	7	8	1	3	8	3	5

DRAFT REPORT
Double Blind Study
Biological Effects of Carpets

Anderson Laboratories, Inc.
Dedham, Massachusetts

CARPET STUDY REPORT

I. EXECUTIVE SUMMARY

This study is a double blind evaluation of biological potency of emissions from two carpets (Experiment A, samples 1, 6, Experiment C, samples 4, 5) and one control (Experiment B, samples 2, 7). Each is tested two times in the randomized order determined by Consumer Product Safety Commission, (CPSC).

The questions being addressed ask:

1. Is there a measurable biological effect in animals following exposure to the test samples?
2. Do the test data reveal differences between the effects of the carpet samples and the control sample?
3. How do the findings from Anderson Laboratories, Inc. relate to the findings of the EPA team?

Blinding Precautions

The sample custodian receives the sample, transcribes identification and places the sample into a chamber which he then seals. The contents, if any are, not visible. He positions the sample chamber in the test laboratory. At the completion of the test the sample custodian repackages the sample to be returned to the CPSC.

Sample Preparation

Three one square foot pieces of a test sample formed into cylindrical shape with face fiber outside are placed in the chamber. Some heating pads are used to warm the air temperature to 37°C.

Animals

For the test, male Swiss Webster mice are positioned in the glass animal chamber. The head extends into the central cylinder. A flexible seal around the neck of the animal allows the animals to breathe the test atmosphere. The body is enclosed within the side arm which serves as a whole body plethysmograph. Room air (7 liters/minute) is used for ventilation of the chamber. During baseline and recovery periods the air is taken directly into the animal exposure chamber. During the animal exposure, the air is

drawn through the sample chamber and then to the animal exposure chamber.

Respiratory activity is detected and recorded by microphones attached to the plethysmograph and connected to a high speed recorder and computer.

Respiratory Effects

These data provide information concerning irritant chemical presence and potency in the test atmosphere. The upper and lower airway reflexes are studied by ASTM E 981 "Estimating Potency of Air Borne Irritant Chemicals." The graphic indication of sensory irritation [(SI), upper airways] is a diagnostic pause or slowing at the completion of inspiration as seen on tracings of individual animal respiratory movements. The respiratory rate decreases proportionally.

Pulmonary irritation [(PI), lower airways] is detected as a change in the slope of the respiratory tracing at the end of expiration which may not be accompanied by a rate change. The moderate or severe effects are reported.

Animal Evaluation

Acute toxic effects of the nervous system are studied by an EPA protocol (Functional Observational Battery (FOB) by Virginia Moser) designed to detect neurological changes through systematic observation of animals. One day prior to study and following each exposure, animals are observed and the status of each is recorded using this FOB.

Test Chamber Monitoring

Temperature and humidity are measured in specified locations throughout the testing. Total volatile organic compounds (TVOC) is measured following each animal exposure.

Results and Discussion

The summary of results is presented on Table 1.

The study showed both sensory and pulmonary irritation in test animals exposed to carpet emissions (Experiments A and C). The sensory effect was slight or moderate never severe. The pulmonary effect was severe for two animals in each experiment. Neither sensory nor moderate/severe pulmonary irritation is seen on records of Experiment B (control samples 2 and 7). FOB data

Summary of Findings
Double Blind Study,
Anderson Laboratories, Inc.

	CARPET Sample 1,6	CARPET Sample 4,5	AIR Control 2,7
end point			
Sensory Irritation	slight	slight/ moderate	no
Pulmonary Irritation	severe, two animals	severe, two animals	no
FOB Extreme changes*			
appearance	9/9	8/9	2/9
activity excitability	7/9	9/9	1/9
neuromuscular	12/13	13/13	0/13
Death from Toxicity			
during exposure	0/8	0/8	0/8
after exposure	2/8	3/8	0/8

* # test categories showing extreme change
unblinded 5/12/93 and 5/13/93
Study in conjunction with EPA, CPSC

Table 1

shows that the emissions studied in Experiments A and C caused clear changes of animal behavior and appearance.

There were no deaths due to toxicity during the exposure period for any experiment. A total of 5 animals died following the fourth exposure (day 2) on days 3, 5, 6 and 7. Two of these animals were from Experiment A and three from Experiment C.

The results of this double blind study demonstrate health effects including death resulting from exposure of male Swiss Webster mice to emissions from these carpet examples. The control sample did not have similar consequences.

A similar study is being conducted by EPA staff and contractors at Research Triangle Park which differs in the post exposure treatment of animals. This will introduce variation into the data collected. Due to an error, sample 3 was incorrect. These data were discarded and the sample was replaced by sample 7.

CARPET STUDY REPORT
DATE OF REPORT: ^

II. INTRODUCTION

This study is a double blind evaluation of biological potency of emissions from three tested samples. Two of the samples are carpet, one is a control. Each sample is tested two times in a randomized order.

The questions being addressed ask:

1. Is there a measurable biological effect in animals following exposure to the test samples?
2. Do the test data reveal differences between the effects of the carpet samples and the control sample?
3. How do the findings from Anderson Laboratories, Inc. relate to the findings of the EPA team?

A similar study is being conducted by EPA staff and contractors at Research Triangle Park. The protocols of the Anderson Laboratories and EPA (RTP) differ in certain details. In particular, the post exposure treatment of animals will introduce variation into the data collected.

III. METHODS

A. Exposure System

1. Air and Humidity

a. Air Source

The source of the air for all studies is the laboratory. During baseline and recovery periods, the air is pulled directly from the room into the glass animal exposure chamber. During animal exposure, the air is drawn from the room through the sample chamber and through a 4 inch glass connection (3/4" ID) to the animal exposure chamber.

b. Air Flow Control - Calibration

Air flow is 7 liters/minute, moved by a peristaltic pump (Cole Parmer) downstream of the animal exposure system. The air flow is controlled by a flow meter (Gilmont) which is calibrated by comparison to a soap bubble meter (SKC) each day before animal exposures. The dry gas meter is used to check air flow at the

inlet of the sample chamber before and after each animal exposure. This information is recorded in laboratory record books.

c. Humidity

The humidity of the ambient air is recorded at 15 minute intervals throughout each animal exposure interval. A sling psychrometer (Bacharach, Inc., PGH, PA) is used to determine room air humidity according to the prescribed schedule.

2. Exposure System/Aquarium

a. Source and Preparation

Two new glass aquaria were obtained from a local pet supply store. The sealant adhesive in the interior of the chamber was removed as far as possible by means of a sharp blade. A plastic reinforcing band and lip around the open edge of the chamber is not removed. After removal of the sealant, the chamber is washed with water and Alconox and wiped dry with 95% alcohol. At the completion of each set of four exposures with a single sample, the chamber was again cleaned with water, Alconox and alcohol.

b. Chamber Face and Fittings

The open side of the chamber is covered by a three panel assembly. Two side panels (10.25 by 2.5 inches) are fashioned from plastic faced on the inside by Teflon film. The central panel (10.25 by 14.75 inches) is glass. Duct tape is used to hold the composite panel together. Nylon fittings are inserted into the side panels. During use, the front panel is attached to the aquarium with duct tape.

c. Heating

Home heating pads are used to warm the system. One pad measures 11.5 by 15 inches and is placed under the bottom of the chamber, the second measures 20 by 30 inches and is positioned on the top, back and sides. The air temperature is heated to 37°C by turning both devices to high for 90 to 120 minutes before the exposure begins. The temperature is adjusted by

changing settings of each switch. During exposures, the settings are typically on high.

d. Insulation

A double layer of metal faced fiberglass insulation is used to cover all of the sample holding chamber during the experiment.

e. Test Sample Source and Placement

Each test sample is delivered to the laboratory on the day preceeding its use. The samples are to have been provided as three one square foot pieces of a test material. The diagonal corners of each piece are fastened together to form a cylinder. The face fiber is the external surface of the cylinder. Two of the cylinders are placed directly on the floor of the chamber parallel to the long dimension of the aquarium. The third cylinder is positioned on top of the first two samples. The chamber is then sealed until study the following day.

f. Blinding Precautions

The samples are received in our laboratory in a Federal Express delivery box. After working hours of the day of receipt, the sample custodian opens the box, transcribes all identification into a sample book (which is kept in his posession) and places the sample pieces into a chamber, which he then seals. The chamber is entirely covered by a layer of duct tape so that the contents, if any, are not visible. The sample custodian places the sample chamber in the test laboratory. At the completion of the test, the sample custodian opens the chamber, removes and repackages the sample in the original delivery box. The sample is returned to the sender.

3. Exposure Chamber with Plethysmograph

a. Source and Preparation

The animal exposure chamber is custom made by a glass blower according to the drawings of the ASTM E 981 publication (Appendix I). Between samples, it is washed with Alconox, dried with 95% alcohol and then

air dried. Between trials with a single sample the glass chamber is wiped out with 95% alcohol.

b. The Plethysmograph

Animals are positioned in the sidearms of the glass animal chamber. The head extends through a hole into the central cylinder. A flexible seal around the neck of the animal allows the test animals to breathe the test atmosphere while the body is enclosed within the sidearm (closed with a hard rubber stopper). The neck seal is made of a latex dental dam (medium gage) reinforced and held in place by duct tape. A central hole in the latex dam is of a size appropriate for the animal weights and is made with a cork borer. Typically, the hole through the reinforcing duct tape is size 8, through the latex, size 6. The reinforced dam is carefully positioned in the chamber and the duct tape seal checked for smooth fit.

c. Transducers and Attachment

The four transducers are miniature microphones manufactured by Fakuda Denshi (Japan). Each microphone is attached by a short piece of flexible PFC tubing to the sample port on a sidearm. The cable is directly connected to the Gould RS 3400 recorder.

4. Exposure Measurement System

a. Temperature System and Calibration

The temperature measurements are made by means of a Cole Parmer Thermister/Thermometer. The thermisters, range -40 to 100°C, are manufactured by YSI, Inc. and are NIST traceable. The device has been calibrated by comparison of digital output to thermometer readings using hot and cold water as convenient test conditions. The thermometer will be provided to EPA for reference measurements at a later time.

During testing, three temperature probes are used. They are positioned at sites identified by the EPA in the protocol, one is outside the sample chamber between the heating pad and the bottom glass surface, two is outside the sample chamber between the top glass surface and the heating pad, and the third is inside the sample chamber in air.

b. Other Monitoring

Oxygen, carbon dioxide and chamber pressure are not measured. Total volatile organic chemicals (TVOC) are measured by flame ionization detector (Rosemont) calibrated against a methane standard before each use. This measurement is made following each animal exposure session by attaching the sample chamber outlet to the inlet of the FID device.

B. Animals

1. Procurement Specification

Swiss Webster outbred male mice are obtained from Taconic Farms, Germantown, New York. They are delivered by truck and arrive on Tuesday of each week. The purchase weight specified is 13 to 15 grams. Ten percent of the animals are weighted at delivery to confirm the animal weights.

2. Housing Conditions

For one week before testing, animals are housed in an animal room with controlled heat, ventilation and light (12/12). No other species is kept in this facility. The food (Formulab Chow 5008, Purina Mills) is available ad lib and is used within its marked shelf life. Dedham tap water is available except during actual experimental procedures. The bedding is Bed O Cobs (The Andersons) changed 2 times/week. Animals are housed in groups of 10 for a minimum of 7 days prior to test.

After testing has begun, animals are housed in groups of four. The exposed animals are kept in a separately ventilated compartment of the same animal room.

3. Health Evaluation

Animals are evaluated by brief visual examination and by body weight measurements at receipt and before selection for testing. The animals are between 25.5 and 28 grams at the initiation of the study.

C. Respiratory Frequency Measurements

1. Data Acquisition System

The data acquisition system is described in the ASTM E 981 protocol. The computer hardware was selected by Dr. Yves Alarie, author of the protocol because of its suitability for the intended uses.

The software was developed for this use at the University of Pittsburgh where the validation was conducted.

In addition, at intervals, the program output is directly compared against results of a hand counted recorder tracing.

2. Subjective Evaluation of Waveforms

Scoring Criteria

a. Sensory Irritation

The graphic indication of sensory irritation is a pause or slowing at the completion of inspiration as seen on tracings of individual animal respiratory movements. As the pause becomes longer with increased potency of the irritant exposure, the respiratory rate decreases proportionally. Copies of tracings showing the examples of SI and PI are included on Appendix II. Example 1 is a classic example of a response to a chemical having strong sensory irritant properties. Example 2 is a slight SI response but a recognizable change from normal. In the presence of a detectable pattern change (example 2) the calculated mean group rate change of baseline is used to quantify the response. Rate change is summarized as maximum & deviation (three minute duration) from normal baseline value for each exposure. If no pattern change is found a rate change is not an indication of sensory irritation.

b. Pulmonary Irritation

This is detected as a change in the slope of the respiratory tracing at the end of expiration or initiation of the next inspiration. The pattern change may not be accompanied by a rate change. The severity of the effect is categorized as slight, moderate or severe as defined in the ASTM E 981 protocol. The

effect is described as slight if the pattern tracing shows only a rounding of the expiratory peak. A tracing resulting from exposure to a pulmonary irritant of moderate potency is described as a flattening of the expiratory pattern with indications of a hook or tail. During expiration the severe effect shows a more pronounced change of slope with an accentuated tail. These patterns are printed in the ASTM E 981 protocol with examples in Appendix II.

In our laboratory only moderate or severe pattern changes are reported.

D. Functional Observational Battery (FOB)

The description of this evaluation procedure and the scoring form are included as Appendix III.

Test Procedure

One day prior to study, animals are observed and scored using the EPA FOB.

At the conclusion of each animal exposure period, the animals are removed from the exposure chamber and placed in a prepared cage with food water and bedding. After 15 minutes, the animals are observed and scored according to the written description of the EPA FOB.

As an exception to the EPA procedure, animals are placed on a three sided elevated shelf (as contrasted to a low laboratory cart having 4 sides) for observation of gait, activity and posture. The shelf is cleaned after each animal is observed.

E. Study Design

1. Exposure Protocol - Sequence of Events

- Check and adjust air flow.

- Measure room humidity. Start to heat chamber containing sample.

- Position thermisters to determine temperatures, record findings at t_0 and at 15 minute intervals throughout the study.

- Weigh and record animal weights and mark tails for ID.
- Position each mouse in plethysmograph with head extending into the central core of the exposure chamber.
- Close open-end of each animal chamber sidearm with a hard rubber stopper.
- Ventilate animal exposure chamber using room air at a rate of 7 liters/minute.
- Fix microphone to the sampling port on each sidearm.
- Insert thermometer into central sampling port of animal chamber.
- Allow 15 to 20 minutes for animals to adjust to the chamber.
- Collect 15 minute baseline respiratory rate data on recorder and computer.
- Observe record for signs of pulmonary and sensory irritation.
- Connect sample chamber to the animal exposure chamber.
- Collect respiratory rate data on recorder and computer for 60 minutes.
- Observe record for signs of pulmonary and sensory irritation.
- Disconnect sample chamber from animal exposure chamber. Mark chart. Continue ventilation of animal chamber with room air.
- Determine TVOC in test atmosphere.
- Collect respiratory rate data on recorder and computer during 15 minute recovery period.
- Observe record for signs of pulmonary and sensory irritation.
- Remove animals from exposure chamber, place in container with food and water.

- Evaluate each animal using FOB after a minimum of 15 minutes in container.

- Return animals to animal room for desired internal.

2. Experimental Protocol

For each experiment (A,B,C) there is a sequence of four animal exposures. Each group of animals is exposed two times on day one followed by two times on day two. There is a minimum of two hours between the end of the first exposure of the day and the start of the second. Body weights are determined and recorded before each exposure.

3. Study Protocol

A total of six experiments were planned for the study, two experimental samples and one control sample. Two examples of each were to have been provided to each test lab in a random order by the CPSC. Because of an error, one sample was incorrectly provided to Anderson Laboratories, Inc. but not to EPA. The data from the incorrect sample (number 3) have been discarded. A new sample has been provided by CPSC and has been tested at Anderson Laboratories as sample 7 (Appendix 4).

F. Data Analysis and Statistics

The data are summarized (as we have been advised by M. Mason of EPA) by combining samples 1 and 6, samples 4 and 5 and samples 2 and 7. No statistical analysis is offered.

IV. RESULTS

A. Experiment A (Samples 1 and 6)

1. Exposure System

The temperature and humidity are reported on Tables 1 through 8, Appendix V.

TVOC's measured by flame ionization detector (calibrated against methane standard) ranged from 4.4 ppm, high to 2.4 ppm, low.

2. Body Weight Data

Day one, the mean starting weight of the 8 animals was 26.3 grams. Day 2, the mean starting weight was 23.7 grams. This represents an approximate 10% body weight loss over the course of one day. The weights following the exposures are not used for this calculation due to the possibility of dehydration effects.

3. Irritation Measurement

a. Sensory Irritation (SI)

Examination of the record showed a change in the respiratory pattern which is defined as SI. No irritation pattern was noted in two exposures and in a third exposure the numerical change was at the no effect level (<12%). For the remaining 5 exposures, the three minute values were 14, 16, 17, 20 and 23%. This is at the borderline between slight and moderate effect.

b. Pulmonary Irritation (PI)

The tracing of each animal is examined and scored during five intervals in each exposure. This results in a maximum of 160 observations. Examination of the record shows a change in the respiratory pattern defined as PI. In the first two exposures (day 1, samples 1 and 6), the PI was classed as moderate in only 3 of 80 observations. During day two (exposures 3 and 4) of 80 observations the PI was scored as moderate once and severe 20 times. Two animals reacted strongly, the others did not.

4. Functional Observational Battery (FOB)

The observations have been grouped as neuromuscular function, activity and excitability and general appearance. For each observation, each individual score is compared against the prescore value for the group of animals being studied. The number of observations outside the prescore range for the group have been presented in Appendix VI. In addition, the observations showing a severe effect (maximal deviation from normal) are listed on Table 2.

Deaths

CARPET STUDY REPORT
DATE OF REPORT: ^

There were no toxicity related deaths during the exposures. There were, however, 2 deaths which occurred on days 3 and 6 of the experiment.

B. Experiment B (Examples 2 and 7)

1. Exposure System

The temperature and humidity are reported on Tables 6 through 16, Appendix V.

TVOC measurements range from 2.1 ppm, high to 1.6 ppm, low.

2. Body Weight Data

Day one, the mean starting weight of the 8 animals was 25.5 grams. Day two, the mean starting weight was 25.0 grams. This represents an approximate 2% body weight change over the course of one day. The weights following the exposures are not used for this calculation due to the possibility of dehydration effects.

3. Irritation Measurement

a. Sensory Irritation

Examination of the records from samples 2 and 7 showed no change in the respiratory pattern indicative of SI.

b. Pulmonary Irritation

The tracing of each animal was examined and scored during five intervals in each exposure. This results in a maximum of 160 observations. Examination of the record showed that there were several instances where a detectable change in the respiratory pattern was defined as slight PI. However, in no case was the change classed as moderate or severe.

4. Functional Observational Battery

The observations have been grouped as neuromuscular, activity/excitability and general appearance. For each observation, the score of each individual animal is compared against the prescore value for the group of animals being studied. The number of observations

which are outside the prescore range for the group have been presented in Appendix VI. The only observations showing a change distinct from normal are vocalization, piloerection, and ear petechiae. These are presented in Appendix on Tables 3.

Deaths

There were no toxicity related deaths.

C. Experiment C (Samples 4 and 5)

1. Exposure System

The temperature and humidity are reported on Tables 17 through 24, Appendix V.

TVOC measurements range from 5.4 ppm, high to 3.7 ppm, low.

2. Body Weight Data

Day one, the mean starting weight of the 8 animals was 26.0 grams. Day two, the mean starting weight was 23.8 grams. This represents an approximately 9% body weight loss over the course of one day. The weights following the exposures are not used for this calculation due to the possibility of dehydration effects.

3. Irritation Measurement

a. Sensory Irritation

Examination of the record showed a change in the respiratory pattern which is defined as SI. An irritation pattern is noted in all exposure records. The maximum mean group % decrease from baseline rate (three minute duration) are 12, 12, 12, 13, 14, 14, 15 and 18%. This describes a consistent effect reflecting sensory irritants having slight potency.

b. Pulmonary Irritation

The tracing of each animal was examined and scored during five intervals in each exposure. This results in a maximum of 151 observations (two animals died from accidental causes during exposure four). Examination of the record showed no detectable change in the

respiratory pattern defined as PI during exposures 1 and 2. During day two (exposures 3 and 4) of 71 observations the PI was scored as moderate 2 times and severe 24 times. The effect was strong in three animals and absent in others.

4. Functional Observational Battery

The observations have been grouped as neuromuscular function, activity and excitability and general appearance. For each observation, the individual animal score is compared against the prescore value for the group of animals being studied. The number of observations which are outside the prescore range for the group have been presented in Appendix VI. The observations showing a severe effect (a maximal deviation from normal) are listed on Table 4.

Deaths

There were no toxicity related deaths during the exposures. There were, however, 3 deaths which occurred on days 5 (two deaths) and 7 of the experiment.

DISCUSSION

Information was received by phone from Mark Mason that the samples were grouped as 1/6, 2/3 and 4/5. Because the records of our sample custodian did not confirm the grouping of similar elements, Mark Mason was contacted for discussion and instructions. A replacement sample was received from CPSC on May 5th. All usual blinding procedures were followed. The sample was tested on May 10th and 11th. For this report, therefore, Experiment B includes samples 2 and 7. On May 12th, Anderson Laboratories, Inc. was informed again by phone that samples 2 and 7 are the control materials. May 13th a letter was received from CPSC which identifies 1, 6 (Experiment A) and 4, 5 (Experiment C) as carpet samples. Experiment A and Experiment C were carpets.

The data from Experiment A indicates the presence of sensory [illegible] chemicals in the test atmosphere. The effect was somewhat variable in its occurrence. This could reflect either changes of the sample as a result of heating, instability of the [illegible] or some complex mixture of interactions of the test sample with its absorbent layers resulting in emission and remission of the chemicals biologically potent.

The pulmonary irritation response was primarily shown by two individual animals (one with sample 1 and one with sample 6) and was pronounced on the second day. This increased response on the second day is fairly typical of the effects noted with previous carpet samples tested in this laboratory. The effects which were identified by way of the FOB showed several changes compared to prescore which were not mirrored in the sample results. In the observations grouped as alertness/excitability the greatest change from normal included alertness, handling and reactivity, activity and rearing. Both increased and decreased activity scores were recorded. Although the vocalization scores were rather different from prescore the comparison with the control indicated that the effect might be related to the procedure as well as to the test exposure.

In the group "general appearance", facial swelling, lacrimation, gasping, palpegbral closure and ear petechiae are the findings which were not consistent with prescore or control. The general appearance group is primarily scored as present or absent and thus represent a clear difference from the starting condition.

The severe neuromuscular changes in Experiment A include the jar task, body tone righting, reflex body posture, gait score, impaired gait, inverted screen, mis-steps reach reflex, tilted screen and body tilt.

The data from Experiment B (control) indicates the absence of sensory and pulmonary irritants. The slight, somewhat random changes which appear to represent PI are not considered to be clear indicators of the effect and are not reported as positive findings by the test developer (Y. Alaire, personal communication). The major findings of the FOB evaluation are limited to vocalization, piloerection and ear petechiae. No toxicity related deaths were recorded.

Experiment C reveals clear sensory and pulmonary effects as well as a sequence of FOB finding of considerable severity. Three deaths were recorded following the exposure. It appears after unblinding that sample B (the control) is easily differentiated from samples A and C (carpets) by its lack of activity. In addition, samples A and C, tested double blind, resulted in acute biological effects, excitability and appearance changes and post exposure deaths.

Table 2A

Appearance Data Summary

Extreme deviation from pre test values

test	% observations showing extreme deviation from pre test		
	sample 1,6	sample 4,5	sample 2,7
facial swelling	34	25	0
bleeding	3	13	0
lacrimation	13	13	0
gasping	52	37	0
cyanosis	6	17	0
exophthalmus	6	13	0
palpebral closure	13	7	0
piloerection	9	0	3
ear petechiae	63	50	32

AL-20

Activity/Excitability Data Summary

Extreme deviation from pre test values

test	% observations showing extreme deviation from pre test: sample 1,6	sample 4,5	sample 2,7
alertness	25	17	0
handling reactivity	28	37	0
activity	22	10	0
vocalization	40	33	19
air puff	3	3	0
click	19	3	0
tail pinch	3	3	0
clonic movements	0	3	0
tonic movements	0	3	0

Table 2C

Neuromuscular Data Summary

Extreme Deviation from Pre test values

test	% observations showing extreme deviation from pre study sample 1,6	sample 4,5	sample 2,7
jar	19	27	0
grip strength	9	27	0
body tone	22	20	0
righting reflex	13	13	0
body posture	56	50	0
ataxic gait	6	7	0
gait score	13	17	0
impaired gait	25	27	0
inverted screen	19	23	0
mis steps	16	3	0
reach reflex	31	50	0
tilted screen	19	20	0
body tilt	0	17	0

783 P23

MAY 20 '93 15:09 ANDERSON

AL-22

Anderson Laboratories, Inc.

APPENDIX I
ASTM E 981

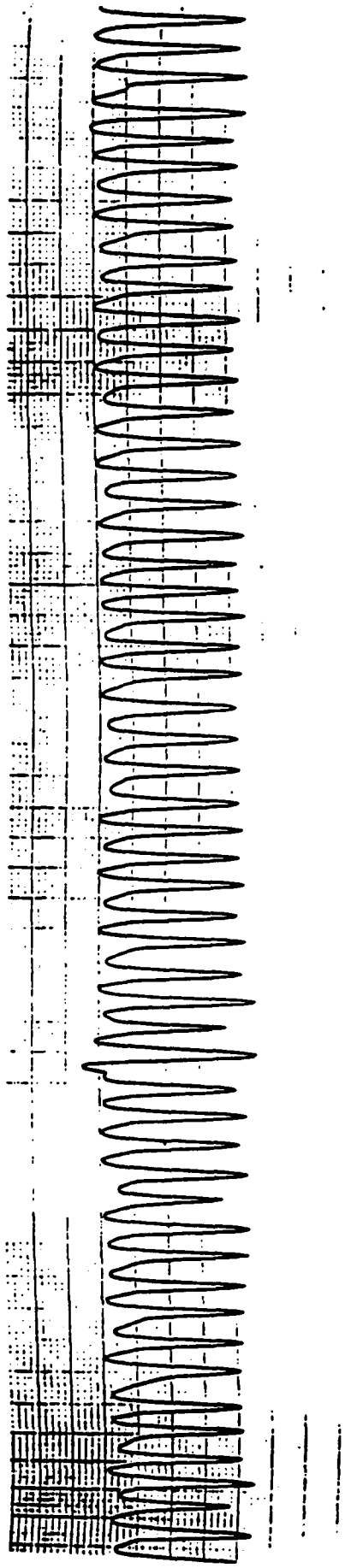
Anderson Laboratories, Inc.

Date_____ Study_____ Test period_____ Observers _____

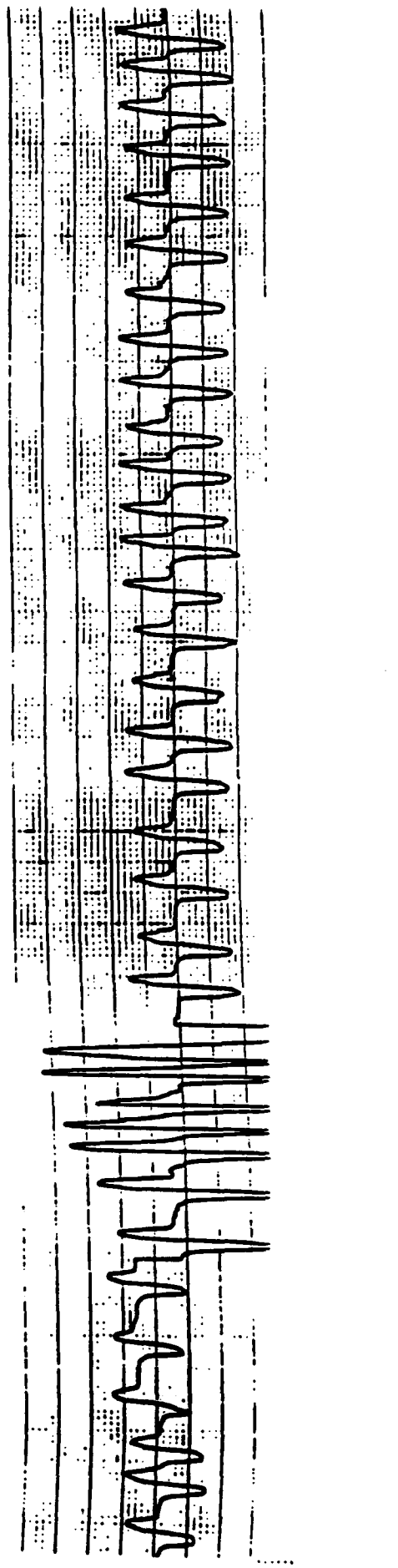
Mouse number						
Jar task count: # times fell (3 tries)						
Tilted screen test count: # times UP.EVEN.DOWN (3 tries)						
Mis-steps rank: 1-4						
Forelimb placing quantal: 1,2						
Grip strength count: # times dropped (3 tries)						
Handling reactivity rank: 1-5						
Lacrimation rank: 1-3						
Palpebral closure rank: 1-4						
Salivation rank: 1-3						
Piloerection quantal: 1,2						
Exophthalmus quantal: 1,2						
Cyanosis quantal: 1,2 (indicate where)						
Gasping quantal: 1,2						
Facial swelling quantal: 1,2						
Bleeding (eye,ear,nose) quantal: 1,2 (indicate which)						
Body tone rank: 1-4						
Righting reflex rank: 1-4						
Rears count: # rears (2 min)						
Body posture descriptive: 1-4						
Body tilt rank: 1-4						
Clonic movements descriptive: 1-7						
Tonic movements descriptive: 1-6						
Gait score rank: 1-5						
Impaired gait descriptive: 1-4						
Ataxic gait rank: 1-5						
Alertness rank: 1-4						
Activity rank: 1-6						
Diarrhea quantal: 1,2						
Air puff response rank: 1-4						
Click response rank: 1-4						
Tail pinch response rank: 1-4						
Inverted screen test # sec to climb.HANG.DROP						
Vocalizations quantal: 1,2 (if ≥3)						
Stereotypy descriptive: repetitive behaviors						
Bizarre behaviors descriptive: unusual behaviors						
Other comments						

Appendix II
EXAMPLES OF RESPIRATORY TRACINGS

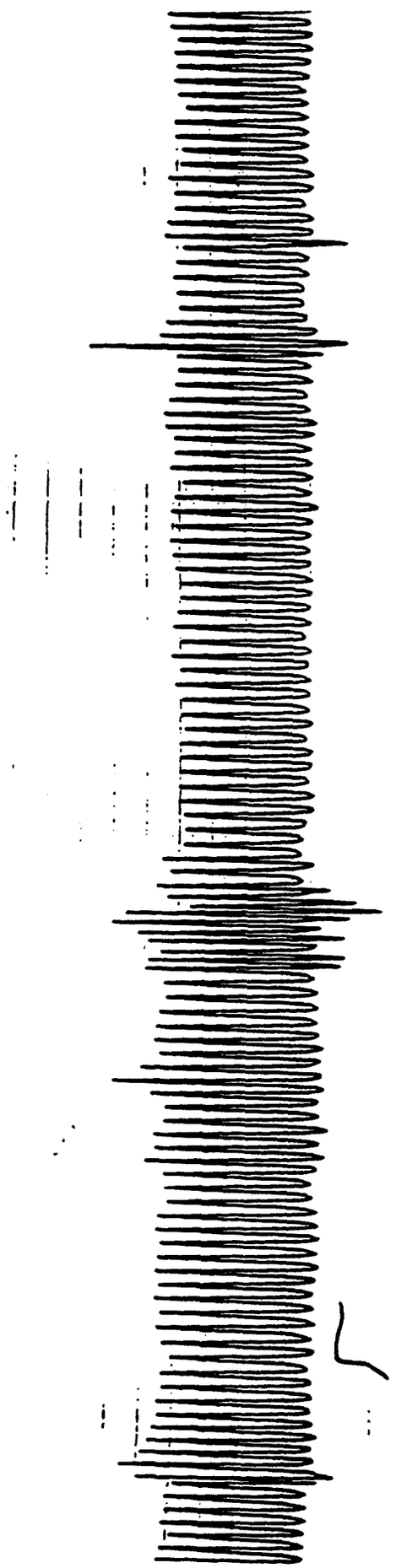
Anderson Laboratories, Inc.



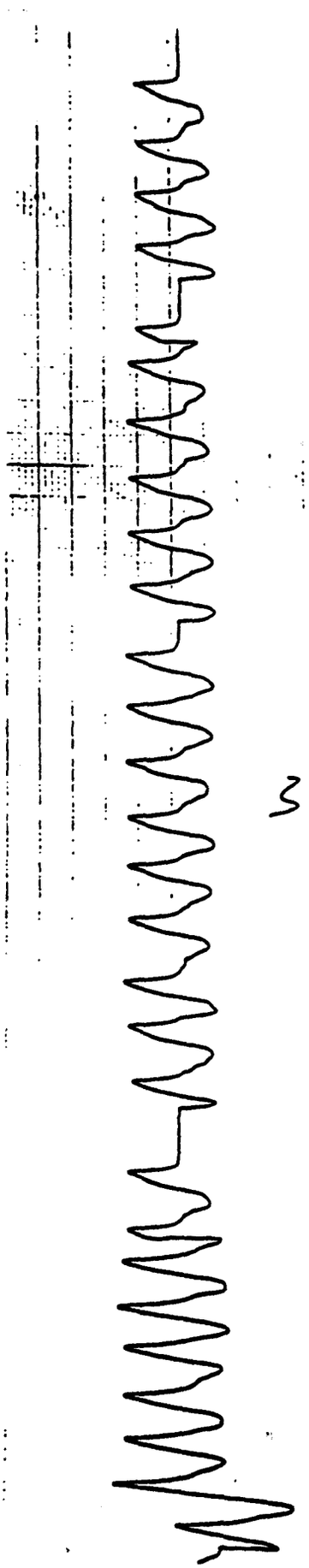
Sensory Irritation
(Detectable)



Sensory Irritation
(Severe)



Pulmonary Score of 2 (moderate)



Pulmonary Score of 3 (Severe)

Appendix III
FUNCTIONAL OBSERVATIONAL BATTERY

Anderson Laboratories, Inc.

HYBRID FUNCTIONAL OBSERVATIONAL BATTERY
FOR USE WITH MICE EXPOSED TO CARPET SAMPLES

V.C. Moser
ManTech Environ. Tech.
revised 1/14/93

Mice will be individually tested. A pre-exposure, or baseline, test will be conducted (preferably the day before exposures begin, or else immediately before the first exposure). Post-exposure tests will be initiated 10-15 minutes after removal from the exposure chamber. At a minimum, tests will be conducted after the second and fourth exposures, and more often as necessary to document behavioral changes. Tests are performed in the order listed below.

Type of data:

D=descriptive data. Note-more than one description is possible.
R=rank order data
Q=quantal, or yes/no data
I=interval or continuous data, including count data

Jar task (I) - Place mouse on 1-gallon glass jar, hanging from the rim of the jar by only one hind leg. Count how many times, out of three tries, the mouse falls off the jar (usually the mouse pulls up to stand upright on the jar rim).

Tilted screen test (I) - Holding the screen at a 45° angle, place the mouse sideways (horizontal). Count how many times the mouse either rotates and walks up the screen (UP), walks straight across the screen (EVEN), or rotates downward and walks down (DOWN). This is repeated three times, with a maximum of 20 sec on each try.

Mis-steps (R) - How often the mouse slips a paw down between the wire mesh while on the tilted screen.

1) none, paws are placed correctly on wire mesh
2) legs slip down only few times
3) legs slip down about half of the steps
4) legs are hanging between mesh with no attempt to pull them up

Hold mouse by tail and lower it towards edge of wire mesh screen. Watch for forelimb placing response as mouse reaches for screen, then allow mouse to grab edge of screen and test grip strength.

Forelimb placing (Q) - Note extension and reaching with forelimbs as mouse reaches towards screen. Allow up to 5 seconds for response to take place.

1) placing response is present
2) no response

Grip strength (I) - Raise up the mouse until the screen is about perpendicular. Count the number of times, out of three tries, the mouse drops the screen before it reaches the 90° point.

Holding the mouse in hand, assess the following measures.

Reactivity to being handled (R) - The extent of activity and resistance to being held.

1) low, no resistance, mouse does not move
2) moderately low, slight resistance, some movement
3) active, constantly moving around in hand
4) moderately high, frantic movement, may be tense or rigid in hand
5) high, squirming, or twisting, or attempting to bite

Lacrimation (R) - evidenced by wetness around eyes

1) none
2) slight
3) severe

Palpebral closure (R)

1) eyelids wide open
2) eyelids slightly drooping
3) eyelids drooping approximately half-way
4) eyes completely shut

Involuntary motor movements (D)

Clonic - Repetitive contractions and relaxations of muscles

1) none
2) repetitive movements of mouth and jaws, smacking
3) fine quivers of limbs, ears, head, or skin
4) mild tremors, moderately coarse
5) severe or whole body tremors, extremely coarse
6) myoclonic jerks
7) clonic convulsions

Tonic - Prolonged contractions of muscles

1) none
2) contraction of extensors such that limbs are rigid and extended
3) opisthotonus - head and body rigidly arched backwards
4) emprosthotonus - head and body rigidly extended forward
5) explosive jumps into the air with all feet leaving the surface
6) severe clonic and/or tonic convulsions resulting in dyspnea, postictal depression, or death

Gait score (R) - Degree of any abnormality of gait excluding ataxia (see below). If only ataxia present, then gait score = 1, ataxic score > 1

1) none
2) slightly abnormal
3) somewhat abnormal
4) markedly abnormal
5) severely abnormal

Impaired gait (D) - Note, if mouse did not move during the 2-min observation period, it may be gently prodded in order to observe the gait. If gait score is 2 or higher, the type of impairment should be indicated here.

1) none
2) hindlimbs show uncoordinated placement, exaggerated or overcompensated movements, or are splayed
3) walks on tiptoes, hindlegs perpendicular to surface
4) flat foot walk, leg(s) flat on surface, crawling

Ataxic gait (R) - Swaying, lurching, rocking, stumbling

1) none
2) slight but definite
3) somewhat, can locomote without falling
4) marked, falls over occasionally
5) severe, cannot locomote without falling

Alertness (R) - level of arousal in the open field

1) very low, stupor, coma, slight or no vibrissae movement
2) low, some head or body movement, occasionally attends to surroundings
3) alert, interested in surroundings, exploration, sniffing
4) high, tense, excited, sudden darting or freezing

Activity (R) - amount of activity in the open field

1) lethargy, no body movement
2) low, somewhat sluggish, little movement
3) somewhat low, some exploratory movements
4) low but active, exploratory movements but mostly walking with very little or no running
5) clearly active, exploratory movements, includes walking and running
6) high, very active, darting or running

[redacted]iarrhea (Q)

1) none
2) present

Salivation (R) - evidenced by wetness around mouth and chin
1) none
2) slight
3) severe

Piloerection (Q)
1) no piloerection
2) indicates presence of piloerection, coat does not lie down after stroking

Exophthalmus (Q)
1) none
2) indicates presence of bulging eyes

Cyanosis (Q) - evidenced by blueness of skin, paws, ears, or tail
1) none
2) present, indicate where

Gasping (Q) - evidenced by laborious and/or convulsive breathing
1) none
2) present

Facial swelling (Q)
1) none
2) present

Bleeding of eyes, ears, or nose (Q)
1) none
2) present, describe where

Body tone (R) - Assess the resistance of abdominal muscle.
1) hypotonia, completely flaccid, limp
2) slightly flaccid, abdomen shows slow return when compressed
3) body tone present, abdomen has resistance when compressed
4) hypertonia, body stiff, abdomen displays great resistance

Righting reflex (R) - Holding mouse between palms of the hands, flip hands over so that mouse is on its back. Do this several times, rating overall ease with which it turns over and regains normal posture
1) flips over immediately
2) slightly uncoordinated, but flips over within 1-2 sec
3) difficulty turning over, takes several seconds
4) cannot turn over within 10 sec

Place mouse in open-field (cart-top with 3-inch lip around the perimeter) for exactly 2 minutes. During this time, the number of rears is counted and other observations are made.

Rearing (I) - Defined as each time the front legs of the mouse come completely off the surface, although the mouse does not necessarily have to raise itself up. Includes when the mouse uses the side or lip of a cart top as support, also includes instances when the mouse lifts front paws for grooming.

Body posture (D)
1) lying on side
2) completely flattened, pelvis flat on surface
3) sitting or walking upright, pelvis off surface
4) hunched, back raised up

Body tilt (R)
1) typical posture, with head straight forward
2) head tilts to one side
3) hand and shoulders lean or tilt to one side
4) whole body leans or tilts to one side, including falling over

Stimulus reactivity - performed while mouse is sitting on open field

Air puff response (R) - Blow into face of mouse.
1) no reaction or response
2) slight or sluggish reaction, e.g., blink or flinch
3) clear reaction, visible startle response
4) exaggerated reaction, e.g., jumps or flips into air

Click response (R) - position clicker approximately 3 cm above the back of the mouse and make sudden sound.
1) no reaction or response
2) slight or sluggish reaction, some evidence that noise was heard , e.g., ear flick
3) clear reaction, visible startle response, quick darting
4) exaggerated reaction, e.g., jumps, bites, or attacks

Tail pinch response (R) - metal tweezers are used to squeeze the tail approximately 1-2 cm from the tip
1) no reaction or response
2) slight or sluggish reaction, some evidence that pinch was felt, e.g., tail flick
3) clear reaction, visible response or jump, quick darting
4) exaggerated reaction, e.g., jumps, bites, or attacks

Inverted screen tests (I) - Place mouse on screen and invert. Measure seconds necessary for mouse to climb to the top. Note "D" if mouse drops off, or "H" if mouse remains hanging, to a maximum of 60 sec. During the time-0 test, mice will be given a total of 3 chances in order to "train" the mouse to climb to the top within 20 sec. A notation will be made as to how many trials were required or if a mouse is unable to invert after 3 trials.

Vocalizations (Q) - Note number of vocalizations, whether or not provoked by handling, testing, etc.
1) none, or only 1 or 2 instances
2) 3 or more instances

Stereotypy - record any behaviors that are excessive or highly repetitive such as circling, stereotypic grooming, pacing, repetitive sniffing, or head weaving.

Bizarre behavior - record any unusual behaviors such as self-mutilation, Straub tail (tail straight up), retropulsion, writhing, flopping, spinning, or attacking.

Other - includes soiled fur, fur discoloration, crustiness around face or eyes, emaciation, or any other findings of interest.

APPENDIX IV
COMMUNICATION OF UNBLINDING AND SAMPLE REPLACEMENT

Anderson Laboratories, Inc.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
OFFICE OF RESEARCH AND DEVELOPMENT
AIR AND ENERGY ENGINEERING RESEARCH LABORATORY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

May 4, 1993

Dr. Rosiland Anderson
Anderson Laboratories Inc.
30 River Street
Dedham, MA 02026
Fax: 617-364-6709

Dear Dr. Anderson,

After our conversation last Friday, I spoke with Dr. Schaeffer at CPSC and with Mr. Ezra Chan, your sample custodian for the split sample test phase. It is clear to me that an inadvertent mixup occurred in preparation of the third sample that was sent to Anderson Labs. I cannot reveal the details of the mixup without foiling the blinding procedure so please bear with me until after we have swapped results, at which time I'll explain what happened. Please repeat the exposure sequence with the sample that CPSC is sending out today. You should receive the package by 10 AM Wednesday, May 5th. Please follow the established blinded procedure for sample loading and testing. Pair the new data with exposure series number two and disregard the original data, it is not relevant to the test plan.

Please accept my apologies for the delay and extra effort this mistake has caused. I realize that this will make it necessary to move the date for exchanging preliminary reports to Monday, May 10th. Please call me later this week and we will discuss details regarding exchange of preliminary reports.

Sincerely,

Mark Mason

Mark Mason
Project Coordinator

cc: Dan Costa
Robert Dyer

MAY 13 1993

U.S. CONSUMER PRODUCT SAFETY COMMISSION

WASHINGTON, D.C. 20207

Rosalind Anderson, Ph.D.
Anderson Laboratories Inc.
30 River Street
Dedham, MA 02026

Dear Dr. Anderson:

This letter serves to officially reveal the order in which carpet subsamples and their sham controls were shipped to an Environmental Protection Agency (EPA) contractor, Acurex Environmental, and Anderson Laboratories during the phase I evaluation of the carpet toxicity bioassay. Six shipments were sent by the Consumer Product Safety Commission (CPSC) staff to the sample custodians at each testing laboratory at staggered intervals during March, 1993. Four shipments were Federal Express packages that contained two subsamples (1 and 2) each of two different carpets (A and B) sealed in labeled Tedlar bags. Two packages containing labeled bags without carpet were sent to indicate that control tests should be run with no carpet in the source chamber. The testing was intended to be "blinded" so that the individuals conducting the animal experiments did not know the identity of the samples being tested. The order in which the six packages were shipped are tabulated below along with the dates the shipments were made, the source identification, and the contents of the labeled bags.

Shipment	Date	Source ID	Contents
1	March 4, 1993	Carpet A1	Pink Carpet
2	March 8, 1993	Control 1	Empty
3	March 18, 1993	Control 2	Paper
4	March 22, 1993	Carpet B1	Blue Carpet
5	March 25, 1993	Carpet B2	Blue Carpet
6	March 29, 1993	Carpet A2	Pink Carpet

Some explanation is needed in regard to the control shipments 2 and 3. Prior to the initiation of the study, it was agreed that a package solely containing an empty bag would serve

AL-35

to signal that a control experiment was to be performed at the testing laboratories. As a result, shipment 2 contained an empty bag. However, just prior to the third shipment the EPA Project Coordinator instructed CPSC staff to weight the next control shipment so that it would be indistinguishable in weight from a package containing three square feet of carpet. This was instructed because of a concern that testing personnel at Anderson Laboratories would have access to the unopened shipments prior to the test run, which could potentially compromise the blinding procedure. CPSC staff weighted the third shipment by including an unsealed Tedlar bag containing a stack of computer paper with the same approximate weight as three square feet of carpet.

On April 30, 1993, the EPA Project Coordinator, after consulting with the sample custodian at Anderson Laboratories, informed the CPSC staff that a problem occurred in loading their source chamber with the third shipment. The EPA Project Coordinator instructed CPSC staff to send another control shipment to Anderson Laboratories but the paper used to weight the package was not to be inside the Tedlar bag. This shipment was sent on May 4. A May 4 letter from the EPA Project Coordinator to Anderson Laboratories stipulates that the test results from this additional shipment will replace the data obtained from Control 2 for the purposes of the phase 1 experiments. If you have any further questions, please do not hesitate to contact me at (301)504-0994.

Sincerely,

Val Schaeffer

Val Schaeffer, Ph.D
Health Sciences Directorate

APPENDIX V
TEMPERATURE AND HUMIDITY VALUES

TEST NUMBER: #921
SAMPLE SOURCE: CPSC SAMPLE 1

EXPERIMENT A

TABLE 1: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	72.6	42.3	35.3	23.0	38.0	71
TIME 0 (START EXPOSURE)	73.7	43.4	36.4	23.0	38.0	71
15 MIN	74.6	45.0	32.8	24.0	37.0	71
30 MIN	75.4	47.1	33.0	24.0	38.0	71
45 MIN	76.0	47.4	33.4	24.0	38.5	71
60 MIN	75.9	47.8	33.3	24.0	38.5	72
75 MIN	75.7	47.9	33.2	24.2	38.5	71

TEST NUMBER: #922
SAMPLE SOURCE: CPSC SAMPLE 1

EXPERIMENT A

TABLE 2: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUT-SIDE S2 (°C)	TOP OUT-SIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	75.4	48.0	37.9	23	38.0	71
TIME 0 (START EXPOSURE)	75.7	48.1	38.2	23.5	38.0	71
15 MIN	76.0	47.8	38.0	24.5	37.5	71
30 MIN	76.7	47.5	37.1	24.5	38.0	71
45 MIN	76.8	47.1	36.6	24.0	36.5	71
60 MIN	77.2	47.0	36.7	24.0	36.5	71
75 MIN	76.3	47.1	37.6	24.0	36.5	71

AL-39

TEST NUMBER: #923
SAMPLE SOURCE: CPSC SAMPLE 1

EXPERIMENT A

TABLE 3: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	73.9	39.2	35.7	23.5	37.0	70
TIME 0 (START EXPOSURE)	74.4	40.4	35.6	24.0	36.5	70
15 MIN	75.6	41.3	34.0	24.0	36.1	70
30 MIN	76.1	42.3	33.5	24.0	36.0	70
45 MIN	76.0	42.7	33.7	24.5	36.0	70
60 MIN	76.5	42.3	33.6	24.5	36.0	70
75 MIN	76.4	43.8	33.6	24.0	36.0	70

TEST NUMBER: #924
SAMPLE SOURCE: CPSC SAMPLE 1

EXPERIMENT A

TABLE 4: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	76.3	46.8	35.2	23.0	36.0	70
TIME 0 (START EXPOSURE)	76.9	47.0	35.3	23.0	36.0	70
15 MIN	76.6	47.0	34.0	24.0	35.5	70
30 MIN	77.0	46.8	33.9	24.5	35.5	70
45 MIN	77.4	46.6	33.8	25.0	36.0	70
60 MIN	77.8	46.6	33.8	25.0	36.5	70
75 MIN	78.0	46.6	34.0	25.0	36.5	70

TEST NUMBER: #962
SAMPLE SOURCE: CPSC SAMPLE 6

EXPERIMENT A

TABLE 5: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	74.3	38.0	34.2	23.0	45	68
TIME 0 (START EXPOSURE)	75.2	37.8	35.2	23.0	45	69
15 MIN	75.0	39.0	33.5	23.5	46	70
30 MIN	75.6	39.0	34.1	24.0	46	70
45 MIN	76.2	39.9	34.4	24.0	46	70
60 MIN	75.9	40.5	34.2	24.0	45	70
75 MIN	76.4	40.7	34.9	24.0	46	70

TEST NUMBER: #963
SAMPLE SOURCE: CPSC SAMPLE 6

EXPERIMENT A

TABLE 6: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	72.6	45.6	36.6	23.5	45	70
TIME 0 (START EXPOSURE)	74.8	45.6	36.9	23.5	46	70
15 MIN	77.5	44.0	36.1	24.0	46	70
30 MIN	76.5	43.3	37.0	24.0	46	70
45 MIN	76.0	43.2	37.4	24.0	46	70
60 MIN	76.0	43.1	36.4	24.0	46	70
75 MIN	75.8	43.1	37.5	24.0	46	70

TEST NUMBER: #964
SAMPLE SOURCE: CPSC SAMPLE 6

EXPERIMENT A

TABLE 7: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	78.4	36.2	32.0	23.0	47	70
TIME 0 (START EXPOSURE)	78.8	37.7	34.6	23.0	46	70
15 MIN	78.9	38.9	33.8	24.0	45	70
30 MIN	78.9	39.9	34.1	24.0	45	70
45 MIN	78.8	40.3	34.2	24.0	45	70
60 MIN	78.7	41.1	34.0	24.0	45	70
75 MIN	79.1	41.5	35.6	24.0	45	70

TEST NUMBER: #965
SAMPLE SOURCE: CPSC SAMPLE 6

EXPERIMENT A

TABLE 8: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	70.2	43.5	34.0	23.5	44	70
TIME 0 (START EXPOSURE)	73.4	42.8	35.3	24.0	44	70
15 MIN	73.6	41.4	34.8	24.0	44	70
30 MIN	73.3	39.8	34.0	24.0	44	71
45 MIN	72.9	39.3	33.7	24.0	44	71
60 MIN	72.7	39.1	33.6	24.0	44	71
75 MIN	72.0	39.0	34.3	24.0	44	70

TEST NUMBER: #925
SAMPLE SOURCE: CPSC SAMPLE 2

EXPERIMENT B

TABLE 9: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	65.4	43.5	37.1	23.5	36.0	71
TIME 0 (START EXPOSURE)	67.0	44.8	38.7	24.0	36.0	71
15 MIN	66.8	45.7	37.5	24.5	36.0	71
30 MIN	66.4	46.0	37.5	24.5	36.0	71
45 MIN	66.1	46.4	37.3	25.0	36.0	71
60 MIN	66.6	46.5	37.3	25.0	36.0	71
75 MIN	67.5	47.1	39.1	24.0	36.0	71

TEST NUMBER: #926
SAMPLE SOURCE: CPSC SAMPLE 2

EXPERIMENT B

TABLE 10: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	66.5	42.0	37.2	23.5	36.0	71
TIME 0 (START EXPOSURE)	65.8	41.9	37.2	23.5	36.0	71
15 MIN	65.4	41.6	35.1	24.0	35.0	70
30 MIN	65.7	41.2	34.6	24.5	35.0	70
45 MIN	65.9	40.9	34.6	25.0	35.0	70
60 MIN	65.6	40.8	34.6	25.0	35.0	70
75 MIN	65.7	41.2	36.0	24.5	35.0	70

TEST NUMBER: #927
SAMPLE SOURCE: CPSC SAMPLE 2

EXPERIMENT B

TABLE 11: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	66.0	37.5	34.2	23.0	29.0	70
TIME 0 (START EXPOSURE)	66.0	39.0	35.5	23.5	29.0	71
15 MIN	65.0	39.4	33.8	24.0	29.0	71
30 MIN	64.4	39.7	33.7	24.5	29.0	71
45 MIN	64.9	39.9	33.9	25.0	27.0	71
60 MIN	64.7	40.0	33.8	25.0	27.0	71
75 MIN	65.6	40.5	35.6	24.5	27.0	71

AL-48

TEST NUMBER: #928
SAMPLE SOURCE: CPSC SAMPLE 2

EXPERIMENT B

TABLE 12: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	67.2	43.0	38.8	23.5	28.0	70
TIME 0 (START EXPOSURE)	67.9	43.1	39.1	23.5	28.0	70
15 MIN	66.8	42.6	36.8	24.0	27.0	70
30 MIN	66.4	42.2	35.5	25.0	27.0	70
45 MIN	66.6	41.9	35.2	25.0	27.0	70
60 MIN	66.2	42.0	35.8	25.0	27.0	70
75 MIN	66.1	42.2	37.2	24.5	27.0	70

TEST NUMBER: #1034
SAMPLE SOURCE: 7

EXPERIMENT B

TABLE 13: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	66.6	39.8	35.1	22.0	55	70
TIME 0 (START EXPOSURE)	66.8	39.9	35.4	22.0	55	70
15 MIN	65.0	41.8	35.8	23.0	55	70
30 MIN	65.0	43.5	36.8	23.5	55	71
45 MIN	65.4	44.7	37.5	23.5	55	71
60 MIN	65.8	44.9	38.0	23.5	55	72
75 MIN	65.9	45.5	38.2	23.0	54	72

TEST NUMBER: #1035
SAMPLE SOURCE: 7

EXPERIMENT B

TABLE 14: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	68.0	43.3	39.1	23.0	54	72
TIME 0 (START EXPOSURE)	67.7	43.4	39.4	23.0	54	72
15 MIN	67.1	44.2	39.1	23.0	54	72
30 MIN	67.1	44.3	39.2	23.5	54	73
45 MIN	67.0	45.1	39.3	24.0	54	74
60 MIN	67.0	45.0	39.2	24.0	54	74
75 MIN	67.1	45.0	39.3	24.0	55	74

AL-51

TEST NUMBER: #1036
SAMPLE SOURCE: 7

EXPERIMENT B

TABLE 15: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	65.3	41.7	38.6	21.5	55	71
TIME 0 (START EXPOSURE)	65.6	41.4	35.9	22.0	55	71
15 MIN	65.5	40.9	35.4	22.0	55	71
30 MIN	65.0	40.4	35.5	23.0	55	71
45 MIN	65.2	40.2	35.5	23.5	54	72
60 MIN	65.5	40.2	35.1	23.5	54	72
75 MIN	65.3	40.3	35.6	23.5	54	72

TEST NUMBER: #1037
SAMPLE SOURCE: 7

EXPERIMENT B

TABLE 16: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	70.2	43.1	39.7	23.5	56	73
TIME 0 (START EXPOSURE)	70.1	43.3	39.6	24.0	56	73
15 MIN	70.1	43.7	38.8	24.0	56	72
30 MIN	69.9	43.2	38.2	24.5	56	72
45 MIN	69.9	43.0	38.5	24.5	58	72
60 MIN	69.6	43.1	38.5	24.5	58	72
75 MIN	70.1	43.8	39.6	24.0	58	71

AL-53

TEST NUMBER: #944
SAMPLE SOURCE: CPSC SAMPLE 4

EXPERIMENT C

TABLE 17 TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	68.1	44.2	36.2	23.0	36	70
TIME 0 (START EXPOSURE)	68.4	44.8	36.4	23.5	36	70
15 MIN	69.8	45.1	35.1	23.5	36	70
30 MIN	71.0	45.1	35.2	23.5	37	70
45 MIN	71.3	45.3	35.3	24.0	37	70
60 MIN	71.4	45.6	35.0	24.0	37	70
75 MIN	71.5	45.7	35.0	24.0	37	70

TEST NUMBER: #945
SAMPLE SOURCE: CPSC SAMPLE 4

EXPERIMENT C

TABLE 18: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	65.3	48.0	34.6	23.0	37	71
TIME 0 (START EXPOSURE)	65.9	47.8	35.0	23.5	37	71
15 MIN	67.6	47.6	34.8	24.0	37	70
30 MIN	68.0	47.0	34.5	24.0	37	70
45 MIN	68.2	46.8	34.5	24.0	37	70
60 MIN	69.4	46.6	36.0	24.0	37	71
75 MIN	69.9	46.8	36.0	24.0	37	71

TEST NUMBER: #946
SAMPLE SOURCE: CPSC SAMPLE 4

EXPERIMENT C

TABLE 19: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	68.5	43.1	37.0	23.0	39	70
TIME 0 (START EXPOSURE)	69.6	45.3	37.7	23.0	39	70
15 MIN	71.3	46.1	33.0	23.5	39	70
30 MIN	72.3	46.7	33.1	23.5	39	70
45 MIN	73.6	46.4	33.1	23.5	39	70
60 MIN	73.8	46.3	33.4	23.5	38	60
75 MIN	72.5	46.5	35.0	23.0	39	70

TEST NUMBER: #947
SAMPLE SOURCE: CPSC SAMPLE 4

EXPERIMENT C

TABLE 20: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	72.9	48.1	36.5	23.0	39	70
TIME 0 (START EXPOSURE)	72.8	48.3	35.8	23.0	40	70
15 MIN	73.8	48.0	37.6	23.5	40	70
30 MIN	74.2	47.9	38.4	24.0	39	69
45 MIN	74.7	47.8	39.0	24.0	39	70
60 MIN	74.0	47.6	39.7	24.0	41	70
75 MIN	74.1	47.7	39.9	24.0	41	70

TEST NUMBER: #958
SAMPLE SOURCE: CPSC SAMPLE 5

EXPERIMENT C

TABLE 21: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	69.7	44.6	35.8	23.0	48	70
TIME 0 (START EXPOSURE)	70.3	45.9	36.4	23.0	48	70
15 MIN	71.5	46.5	36.1	23.5	47	70
30 MIN	71.5	47.2	36.5	23.5	47	70
45 MIN	71.8	47.6	36.7	24.0	43	70
60 MIN	71.9	48.1	37.1	24.0	48	70
75 MIN	71.2	48.5	38.0	23.5	48	70

TEST NUMBER: #959
SAMPLE SOURCE: CPSC SAMPLE 5

EXPERIMENT C

TABLE 22: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	72.2	42.6	39.5	23.0	48	71
TIME 0 (START EXPOSURE)	72.6	44.4	39.6	23.0	47	71
15 MIN	72.5	43.5	38.5	23.5	47	70
30 MIN	72.4	42.9	38.2	24.0	48	70
45 MIN	72.3	41.5	37.6	24.5	48	70
60 MIN	72.7	41.4	37.4	25.0	48	71
75 MIN	71.6	41.5	37.8	24.5	48	71

TEST NUMBER: #960
SAMPLE SOURCE: CPSC SAMPLE 5

EXPERIMENT C

TABLE 23: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	72.8	38.0	37.1	23.0	47	70
TIME 0 (START EXPOSURE)	72.1	43.8	38.1	23.0	47	70
15 MIN	74.5	45.6	38.0	24.0	48	70
30 MIN	74.4	45.9	38.6	24.5	48	70
45 MIN	74.3	46.7	39.2	24.0	48	70
60 MIN	74.4	47.1	39.2	24.0	48	70
75 MIN	73.8	47.6	39.4	24.0	48	70

TEST NUMBER: #961
SAMPLE SOURCE: CPSC SAMPLE 5

EXPERIMENT C

TABLE 24: TEMPERATURE, HUMIDITY RECORD

TIME	SAMPLE CHAMBER			MOUSE CHAMBER SAMPLING PORT (°C)	ROOM HUMIDITY (%)	ROOM TEMPERATURE (°F)
	BOTTOM OUTSIDE S2 (°C)	TOP OUTSIDE S7 (°C)	AIR A1 (°C)			
START BASELINE	71.2	44.6	38.5	23.5	46	70
TIME 0 (START EXPOSURE)	71.3	44.9	38.9	23.0	46	70
15 MIN	70.8	45.5	38.2	23.5	46	70
30 MIN	71.4	45.5	38.7	24.0	46	70
45 MIN	68.6	45.6	38.8	24.0	45	70
60 MIN	68.3	45.6	38.9	24.0	45	70
75 MIN	68.5	45.8	40.0	24.0	45	70

AL-60

APPENDIX VI
FOB, EXPERIMENT A

Anderson Laboratories, Inc.

ACTIVITY / EXCITABILITY
Sample 1, 6
DAYS 1, 2

OBSERVATION	PRE SCORE #OBS X SCORE	# OBS OUTSIDE PRE SCORE RANGE	# OBS	% OUT OF RANGE
alert	7X3, 1X2	15	32	47
handling reactive	7X3, 1X2	11	32	34
activity	5X4, 3X5	17	32	53
vocalization	8XA	10	32	31
air puff	8X2	5	32	16
click	7X2, 1X3	7	32	22
tail pinch	8X2	8	32	25
clonic	8X1	31	32	97
tonic	8X1	1	32	3
rears	3-11	20	32	63

AL-62

P/A = present/absent
ep17
ANDERSON LABORATORIES, INC.

ACTIVITY / EXCITABILITY

EXTREME CHANGES, Samples 1,6

DAYS 1, 2

OBSERVATION	PRE SCORE #OBS X SCORE	OBSERVATIONS SHOWING EXTREME SCORE	TOTAL # OBS
alert	7X3, 1X2	6X STUPOR, 2X HYPER	32
handling reactive	7X3, 1X2	9X LOW,	32
activity	5X4, 3X5	6X NONE, 1X HYPER	32
vocalization	8XA	13X PRESENT	32
air puff	8X2	1X EXTREME	32
click	7X2, 1X3	6X NONE	32
tail pinch	8X2	1X EXTREME	32
REARS	45/8	84/32	
	5.6	2.6	
		11X NO REARS	

AL-63

P/A = present/absent

epactiv16

Neuromuscular Function

Samples 1, 6
DAYS 1, 2

OBSERVATION	PRE SCORE # OBS X SCORE	# OBS OUTSIDE PRE SCORE RANGE	TOTAL # OBS	% OUT OF RANGE
jar	7X0,1X1	13	32	41
grip strength	7X0,1X1	8	32	25
body tone	8X3	30	32	94
righting reflex	8X1	23	32	72
body posture	8X3	20	32	63
body tilt	8X1	11	32	34
ataxic gate	8X1	4	30	13
gait score	8X1	31	31	100
impaired gait	8X1	29	29	100
inverted screen	7XH, 1XFAST	14	32	44
mis-steps	8X1	15	31	48
reach reflex	8X1 (PRESENT)	10	32	31
tilt screen	8X3 up	17	32	53

AL-64

EPN16A

Neuromuscular Function
EXTREME CHANGES Samples 1,6
DAYS 1, 2

OBSERVATION	PRE SCORE # OBS X SCORE	OBSERVATIONS EXTREME SCORE	# OBS	% EXTREME
jar	7X0,1X1	6X 3 FALLS	32	19
grip strength	7X0,1X1	3X 3 DROPS	32	9
body tone	8X3	7X HYPOTONIA	32	22
righting reflex	8X1	4X NOT PRESENT	32	13
body posture	8X3	19X HUNCHED	32	59
ataxic gate	8X1	2X MARKED	30	7
gait score	8X1	4X SEVERE	31	13
impaired gait	8X1	8X SCORE 4	29	28
inverted screen	7XH, 1XFAST	6X DROP	32	19
mis-steps	8X1	5X LEGS HANGING	31	16
reach reflex	8X1 (PRESENT)	10X ABSENT	32	31
tilt screen	8X3 up	6X 3 DOWN	32	19

EP9
ANDERSON LABORATORIES, INC.

GENERAL APPEARANCE

Samples 1, 6
DAYS 1, 2

OBSERVATION	PRE SCORE	# OBS OUTSIDE PRE SCORE RANGE	Total # OBS	% OUT OF RANGE
face swelling	8XA	11	32	34
bleeding	8XA	1	32	3
lacrimation	8X1	20	32	63
salivation	8X1	0	32	0
diarrhea	8XA	0	32	0
gasping	8XA	16	31	52
cyanosis	8XA	2	32	6
exophthalmus	8XA	2	32	6
eye closure	8X1	14	32	41
piloerection	8XA	3	32	9
ear petechiae	8XA	20	32	63

AL-66

P/A = present/absent
EPAPEAR3

GENERAL APPEARANCE
EXTREME CHANGES, Samples 1, 6
DAYS 1, 2

OBSERVATION	PRE SCORE	OBSERVATIONS EXTREME SCORE	# OBS	% EXTREME
face swelling	8XA	11X PRESENT	32	34
bleeding	8XA	1X PRESENT	32	3
lacrimation	8X1	4X SEVERE	32	13
gasping	8XA	16X PRESENT	31	52
cyanosis	8XA	2X PRESENT	32	6
exophthalmus	8XA	2X PRESENT	32	6
eye closure	8X1	4X EYES SHUT	32	13
piloerection	8XA	3X PRESENT	32	9
ear petechiae	8XA	20X PRESENT	32	62

EPA16 MAX
P/A =PRESENT/ ABSENT
ANDERSON LABORATORIES, INC.

OTHER OBSERVATIONS
SAMPLES 1, 6

OBSERVATION	FREQUENCY
POPCORN	2
FALL OVER	1
NO ACTIVITY	12
BLEEDING EYE	1
REPETITIVE MOTION	
FOOT	2
CIRCLE	2
NOSE BUMP	4
BUMP INTO WALL	1

APPENDIX VII
FOB, EXPERIMENT B

Anderson Laboratories, Inc.

ACTIVITY / EXCITABILITY
Samples 2,7
DAYS 1, 2

OBSERVATION	PRE SCORE #OBS X SCORE	# OBS OUTSIDE PRE SCORE RANGE	TOTAL # OBS	% OUT OF RANGE
alert	8X3	0	31	0
handling reactive	7X3, 1X2	0	31	0
activity	4X4, 4X5	0	31	0
vocalization	8XA	6	31	19
air puff	7X2, 1X3	2	31	6
click	8X2	4	31	13
tail pinch	4X1, 4X2	9	31	29
clonic	8X1	8	31	26
tonic	8X1	0	31	0
rears	1-10	5	31	16

AL-20

P/A = present/absent
ep11
ANDERSON LABORATORIES, INC.

ACTIVITY / EXCITABILITY
EXTREME CHANGES, Sample 2,7
DAYS 1, 2

AL-71

OBSERVATION	PRE SCORE #OBS X SCORE	# OBSERVATIONS SHOWING EXTREME SCORE	# OBS	% EXTREME
vocalization	8XA	6X PRESENT	31	19

P/A = present/absent
ep13
ANDERSON LABORATORIES, INC.

Neuromuscular Function
Samples 2,7
DAYS 1, 2

OBSERVATION	PRE SCORE # OBS X SCORE	# OBS OUTSIDE PRE SCORE RANGE	# OBS	% OUT OF RANGE
jar	3X0, 1X1	0	31	0
grip strength	4X0	0	31	0
body tone	4X3	2	31	6
righting reflex	4X1	1	31	3
body posture	4X3	5	31	16
body tilt	4X1	0	31	0
ataxic gate	4X1	0	31	0
abnormal gait score	4X1	9	31	29
impaired gait	4X1	9	31	29
inverted screen	4XH	8	31	26
mis-steps	4X1	0	31	0
reach reflex	4X1 (PRESENT)	0	31	0
tilted screen	4X0 (DOWN)	3	31	10

EPAN2A
ANDERSON IABORATORIES. INC.

Neuromuscular Function
EXTREME CHANGES, Samples 2,7
DAYS 1, 2

OBSERVATION	PRE SCORE # OBS X SCORE	OBSERVATIONS SHOWING EXTREME SCORE	TOTAL # OBS
		NONE	

EPAN2A MAX
ANDERSON LABORATORIES, INC.

GENERAL APPEARANCE
Samples 2,7
DAYS 1, 2

OBSERVATION	PRE SCORE	# OBS OUTSIDE PRE SCORE RANGE	Total # OBS	% OUT OF RANGE
face swelling	8XA	0	31	0
bleeding	8XA	0	31	0
lacrimation	8X1	2	31	6
salivation	8X1	0	31	0
diarrhea	8XA	0	31	0
gasping	8XA	0	31	0
cyanosis	8XA	0	31	0
exophthalmus	8XA	0	31	0
eye closure	8X1	0	31	0
piloerection	8XA	1	31	3
ear petechiae	8XA	10	31	32

P/A = present/absent
EP15
ANDERSON LABORATORIES, INC.

GENERAL APPEARANCE
EXTREME CHANGES, Samples 2,7
DAYS 1, 2

OBSERVATION	PRE SCORE	# OBSERVATIONS EXTREME SCORE	Total # OBS	% EXTREME
ear petechiae	4XA	10X PRESENT	31	32

AL-75

EPAR 2,7 max
P/A =PRESENT/ ABSENT
ANDERSON LABORATORIES, INC.

OTHER OBSERVATIONS
SAMPLES 2, 7

OBSERVATION	FREQUENCY
REPETITIVE MOTIONS	
FOOT	1
CIRCLE	4

APPENDIX VIII
FOB, EXPERIMENT C

Anderson Laboratories, Inc.

ACTIVITY / EXCITABILITY
Sample 4, 5
DAYS 1, 2

OBSERVATION	PRE SCORE #OBS X SCORE	# OBS OUTSIDE PRE SCORE RANGE	# OBS	% OUT OF RANGE
alert	8X3	15	30	50
handling reactive	6X3, 2X2	11	30	37
activity	7X4, 1X3	17	30	57
vocalization	8XA	10	30	33
air puff	8X2	5	30	20
click	8X2	7	30	23
tail pinch	7X2, 1X1	8	30	27
clonic	8X1	30	30	100
tonic	8X1	1	30	3
rears	1-8	16 (3 UP, 13 DOWN)	30	53

AL-78

P/A = present/absent
ep14
ANDERSON LABORATORIES, INC.

ACTIVITY / EXCITABILITY
EXTREME CHANGES, Sample 4, 5
DAYS 1, 2

OBSERVATION	PRE SCORE #OBS X SCORE	OBSERVATIONS SHOWING EXTREME SCORE	# OBS	% EXTREME
alert	8X3	5X STUPOR	30	17
handling reactive	6X3, 2X2	10X LOW, 1X HIGH	30	37
activity	7X4, 1X3	2X NONE, 1X HYPER	30	10
vocalization	8XA	10X PRESENT	30	37
air puff	8X2	1X EXTREME	30	3
click	8X2	1X EXTREME	30	3
tail pinch	7X2, 1X1	1X EXTREME	30	3
clonic	8X1	1X MYOCLONUS	30	3
tonic	8X1	1X EMPROSTHOTONUS	30	3

AL_76

P/A present/absent
ep10
ANDERSON LABORATORIES, INC.

Neuromuscular Function

Samples 4, 5
DAYS 1, 2

OBSERVATION	PRE SCORE # OBS X SCORE	# OBS OUTSIDE PRE SCORE RANGE	TOTAL # OBS	% OUT OF RANGE
jar	6X0, 2X1	9	30	30
grip strength	7X0, 1X1	11	29	38
body tone	8X3	29	30	97
righting reflex	8X1	20	30	67
body posture	8X3	17	29	59
body tilt	8X1	13	30	43
ataxic gate	8X1	2	28	7
gait score	8X1	21	30	70
impaired gait	8X1	21	30	70
inverted screen	7XH, 1XFAST	10	30	33
mis-steps	8X1	16	30	53
reach reflex	8X1 (PRESENT)	15	30	50
tilt screen	8Xup	15	30	50

EPAN4,5

Neuromuscular Function
EXTREME CHANGES, SAMPLES 4,5
DAYS 1, 2

OBSERVATION	PRE SCORE # OBS X SCORE	OBSERVATIONS EXTREME CHANGE	TOTAL # OBS	% EXTREME
jar	6X0, 2X1	8X3 FALLS	30	27
grip strength	7X0, 1X1	8X3 DROPS	29	28
body tone	8X3	6X HYPOTINIA	30	20
righting reflex	8X1	4X NOT PRESENT	30	13
body posture	8X3	1X ON SIDE, 14X HUNCHED	29	3
body tilt	8X1	5X SHOULDER OR BODY LIST	30	17
ataxic gate	8X1	2X SEVERE	28	7
gait score	8X1	5X SEVERE	30	17
impaired gait	8X1	8X4	30	27
inverted screen	7XH, 1XFAST	7X DROP	30	23
mis-steps	8X1	1X LEGS HANGING	30	3
reach reflex	8X1 (PRESENT)	15X ABSENT	30	50
tilt screen	8X 3UP	6X 3DOWN	30	20

EPAN4,5 max

ANDERSON LABORATORIES, INC.

GENERAL APPEARANCE
Samples 4, 5
DAYS 1, 2

OBSERVATION	PRE SCORE	# OBS OUTSIDE PRE SCORE RANGE	Total # OBS	% OUT OF RANGE
face swelling	8XA	7	30	23
bleeding	8XA	4	29	14
lacrimation	8X1	16	30	53
salivation	8X1	2	30	7
diarrhea	8XA	0	30	0
gasping	8XA	11	30	37
cyanosis	8XA	5	30	17
exophthalmus	8XA	4	30	13
eye closure	8X1	9	30	30
piloerection	8XA	0	30	0
ear petechiae	8XA	15	30	50

AL-23

P/A = present / absent
EP5
ANDERSON LABORATORIES, INC.

GENERAL APPEARANCE

EXTREME CHANGES, SAMPLES 4,5

DAYS 1, 2

OBSERVATION	PRE SCORE	OBSERVATIONS EXTREME SCORE	# OBS	% EXTREME
face swelling	8XA	7X PRESENT	30	23
bleeding	8XA	4X PRESENT	29	14
lacrimation	8X1	4X SEVERE	30	13
gasping	8XA	11X PRESENT	30	37
cyanosis	8XA	5X PRESENT	30	17
exophthalmus	8XA	4X PRESENT	30	13
eye closure	8X1	2X EYES SHUT	30	7
ear petechiae	8XA	15X PRESENT	30	50

AL-83

ep6
P/A =PRESENT/ABSENT
ANDERSON LABORATORIES, INC.

OTHER OBSERVATIONS
SAMPLES 4, 5

OBSERVATION	FREQUENCY
POPCORN	3
NO ACTIVITY	5
BLEEDING NOSE	2
BLEEDING EYE	1
REACHING WITH ONE ARM	1
SPASMODIC CHIRPING	1
REPETITIVE MOTION	
CIRCLE	3
WALK OFF EDGE	3

REVIEW OF PHASE I OF EPA/ORD CARPET STUDY

Prepared for the U.S.E.P.A.

by

Richard B. Schlesinger, Ph.D.

Department of Environmental Medicine
New York University Medical Center
New York, NY

June 1, 1993

PR-1

This report provides my views of the Phase I component of the EPA/ORD carpet study as described in the written reports and in oral presentations at EPA/HERL on May 26 and 27th, 1993.

GENERAL CONCLUSION

The goal of Phase I was to perform, to the extent possible, replicate studies at EPA/HERL and at Anderson Laboratories. The study design was originally provided by EPA and was subsequently modified by a peer review panel. The experiments at both sites were conducted within the constraints imposed by the protocol. With this in mind, the study as performed by EPA clearly conformed to GLP, and was conducted with appropriate and adequate QA in all components and for both the bioassays and the chemical/physical characterization of carpet emissions. Some concern about data analysis for the FOB is expressed, but this is not a major problem. Accordingly, it is my opinion that the EPA did perform a scientifically valid study of the effect of carpet emissions on mice, and was not able to find any toxic effects which could be ascribed to the carpets tested under the conditions imposed by the experimental design. Furthermore, the levels of emissions of any potentially toxic chemicals from these carpets was very low, and unless there was some exotic interaction between these chemicals, the lack of biological response is not surprising.

On the other hand, the study as performed by Anderson Laboratories did not conform to GLP, inasmuch as the double blind condition was not maintained, the animal housing was substandard (possibly resulting in compromised mice), and the data presentation and analysis for the FOB and analysis of the respiratory waveforms are of major concern. The lack of double blind conditions at Anderson Laboratories could have resulted in exaggeration of the severity of response for the FOB and the respiratory waveform analyses, which tend to be somewhat subjective in nature under ideal conditions. Futhermore, the inadequate housing of the mice at Anderson Laboratories could have resulted in some change in their health status, perhaps increasing susceptibility to even low levels of chemicals emitted by the heated carpets or, more likely, to some combination of these emissions and stress/restraint. However, this conclusion cannot be supported with hard evidence, since no necropsy was performed after exposure nor was

there any provision for sentinel (colony control) animals in the Anderson facility protocol.

SPECIFIC CONCERNS

Double Blind Nature of the Study

One of the most important conditions of the study that must be maintained is its double blind nature, especially since some of the bioassays tend to be subjective. To this end, and in both laboratories, a sample custodian received the carpet samples and loaded them into the source chamber. The chamber was then taped so the carpet was not visible, and this source chamber was transported to another area for connection to the exposure chamber. Thus, theoretically, the technician actually perfoming the exposures should not be aware of whether the exposure was to any carpet emissions or was a sham control.

At EPA, the individual performing the exposures was not the same individual who performed the bioassays. Furthermore, the technician who attached the source chamber to the exposure chamber was not the same individual who performed the actual exposures. Thus, the individual analyzing the data was unaware of the exposure being performed, except for any suspicion based upon response. However, in a number of cases, the sham exposure resulted in greater responses than those noted with carpet samples in the source chamber, which would serve to suggest a lack of bias towards exaggeration of responses to carpet samples.

At Anderson Laboratories, as at EPA, samples were received and loaded by the sample custodian. However, only one other technician was responsible for performing pre-exposure screening, for connecting the source chamber to the exposure chamber, for conducting the exposures, for obtaining the biological data, and generally also for rating the FOB and scoring the waveform respiratory data. There was, therefore, opportunity for this individual to know whether the source chamber was loaded with carpet or was a sham exposure. This had the potential to bias the data analysis. Furthemore, a TVOC monitor was used after the exposures by the same technican that conducted the exposure, which provided this individual with additional opportunity to know whether the exposure was sham or carpet, since the carpet exposures were characterized by higher emissions than the

sham exposures. No on-line monitors were used by EPA during or after the exposures; all such measures were performed separately. Based upon the above, it is my conclusion that the study as performed at Anderson Laboratories was clearly not double blinded.

Measurements of Sensory and Pulmonary Irritation

Changes in respiratory rate were monitored by computerized systems in both laboratories. However, the waveform analysis is somewhat subjective. From discussions, it was noted that the technician at Anderson Laboratories would occasionally adjust the gain or sensitivity of the recording device, which would serve to change the apparent scale of the waveform tracing. This change in gain was not always noted on the recordings, and had the potential to result in conclusions of exaggerated responses. Futhermore, the technician would stop and start the recorder during the measurement period; any shift in baseline would not be known with such a procedure.

Ideally, all respiratory tracings obtained at both EPA and Anderson Laboratories should have been read by readers at both sites, so as to provide some inter-laboratory comparison. From discussions, it was concluded that the reader from EPA did look at some tracings obtained at Anderson Laboratories and found reasonable agreement when tracing scales were not changed during recording. However, it was also noted that when both readers did not agree, that the Anderson Laboratory reader tended to rank the effect as more severe than did the EPA reader. As indicated above, the lack of double blind conditions may have affected conclusion as to the degree of severity of response.

At EPA, respiratory waveforms and rates were recorded during the entire period of the exposure, as well as in pre- and post-exposure periods. At Anderson Laboratories, tracings were recorded only during discrete intervals during these times. There were, for example, three such periods during each animal exposure period. Furthermore at EPA, all breaths were analyzed. On the other hand, at Anderson Laboratories, the tracings were scanned and only series of breaths that met some criterion, which is not clear to me, for moderate to severe changes were analyzed. Thus, the database at EPA was larger, and provided a better basis for determination of any change due to toxicant exposure.

Presentations by Anderson Laboratories indicated that sham control exposed animals consistently had scores of zero for all parameters, indicating no effect compared to pre-exposure levels. On the other hand, control animals at EPA did show changes from pre-exposure levels, which in some cases were greater than those following carpet emission exposures. The low control levels for Anderson would further serve to exaggerate any response seen following carpet expousre. Furthermore, in my experience, there are always some control animals that show some response. If such animals are removed from the study, it would further serve to exaggerate effects in toxicant exposed animals.

Animal Housing Facility/Animal Health Status

The animal housing facility at EPA met appropriate standards. The facility at Anderson Laboratories did not. This latter consisted of a separate smaller room (about 12 ft x 12 ft) within a larger laboratory room in which exposures were performed. The relative humidity and temperature within the housing room was not rigidly controlled.; a through the wall air conditioner was apparently used for some ventilation. Airflow was from the laboratory into the animal room, and then out an exhaust. This is opposite from what should be maintained, namely positive pressure within the animal housing room driving air from the animal quarters into some other space. Under the conditions at Anderson Laboratories, any contaminants, chemical or biological, present in the larger laboratory room would be drawn into the animal quarters. Since the mice were housed for at least one week in this facility prior to their use, it is quite possible that their condition was not the same as that of the mice properly housed at EPA. But since no sentinel animals were necropsied at Anderson Laboratories, any adverse effect of housing in their facility is not known. However, it is possible that the mice may have become hypersensitive to any low level carpet emissions, or more likely to a combination of stress plus these emissions, due to some pathogen or other airborne contaminant entering the housing quarters.

All animals exposed at EPA underwent a necropsy. However, those exposed at Anderson Laboratories were not subjected to a valid necropsy. Thus, in the latter case, the cause of death following exposure to carpet emissions is unknown, and the health status of the animals compared to those at EPA is unknown. This lack of information is a major problem in

attempting to understand the differences in responses obtained at the two study sites.

Room Air v. Zero Air/Carpet Water Content

Air entering the source chamber, and eventually the exposure chamber, was room air at Anderson Laboratories and humidified zero air at EPA. From discussions at EPA, it was determined that the air at Anderson Laboratories had an RH of 30-50%. This is comparable to that used at EPA, which was actually lower than expected. Thus, differences in humidity of air entering the experimental system is likely not a factor accounting for differences in biological response.

During discussions at EPA, a suggested possibility for such differences was differences in water content of the carpet samples. In order to account for biological response differences, all of the samples used at Anderson Laboratories would have had to have a greater water content than the samples used at EPA. Based upon the sampling scheme used to obtain carpets for testing, the probability of this occuring is essentially nil. Thus, it is unlikely that differences in moisture content account for differences in biological responses between the two laboratories.

RECOMMENDATIONS

There are various routes which may be taken to attempt to resolve the apparant discrepancy between results in the two laboratories. One approach would be to have the study repeated by another, independent laboratory, but this would be quite difficult, since it would also have to be repeated at EPA to obtain some baseline for comparison. It may be better to attempt to determine the reasons for differences in response. This could be done by having EPA personnel together with Anderson personnel repeat the study at Anderson Laboratories with EPA exposure equiptment, EPA double blind study procedure and Anderson mice, which would under complete necropsy following each exposure series. This would go a long way to help resolve the apparent "disprepancy" in results of the two laboratories.

Review of EPA and Anderson carpet experiments

Harriet A. Burge

Based strictly on the quality of the reports (including description of tests, results, discussion, and quality control plans) and on the post exposure pathology data from EPA, one would have to place more confidence in the EPA results than in those of Dr. Anderson's lab. However, the negative case is clearly difficult to prove, and positive results tend to be more persuasive even when the study designs are less rigorous (or less rigorously described).

The data presented by the EPA appear to make a good case for little or no effect in spite of the fact that total VOCs were elevated in the presence of carpet, and some recognized toxins were measured (although in low levels). It should be noted that we already knew that carpeting emits VOCs. It would have been useful for the EPA report to contain examples of actual tracings that demonstrated slight, moderate or severe respiratory effects (these were included in Anderson's report). Without these, it is not possible to determine whether or not both groups interpreted the tracings in the same way.

Since Anderson's results differ from those of the EPA, and no obvious reason is apparent, the case continues to be unresolved (i.e., we don't know whether or not the carpet causes mouse toxicity). It is still possible that some difference between the systems that controls the nature of exposure to the carpet is responsible for the disparate results. It is also possible that some inadvertant action on the part of one (or both) of the investigative groups is controlling the outcome of these experiments independent of the carpet exposures. I suggest that during the next phase, at least 1 EPA observer attends Anderson's experiments, and that Dr. Anderson and/or a person she designates attend the EPA experiments. I would also suggest that an outside observer attend both sets of experiments (a panelist would be a likely choice, especially if one of the panel feels especially capable of evaluating the respiratory function tests). I would also suggest that the person (at both sites) responsible for observing the mice and recording observations be blinded to the concurrently collected ambient measurements, since the temperatures at the bottom of the tank appear to provide a clear indication of absence of carpeting in the chamber. An additional control should be a "no complaint" carpet that did not produce effects in Dr. Anderson's lab.

If, following the next series of experiments, a clear indication should exist that exposure to the carpet is causing toxicity in mice, I would suggest that the experiments be run under more normal temperature conditions. It is true that offgassing can be stimulated by heat. However, in the real world of carpeting, it is probably quite unusual for carpet materials to reach these temperatures in occupied residential or office interiors.

With respect to public release of this data, my first choice would be not to release information until the discrepancy problem has been solved. However, since some release will probably be necessary, I

would suggest a carefully worded statement stressing that the EPA experiments did not demonstrate an effect using methods that the expert panel considered appropriate, but that the EPA feels that the issue is important enough to continue testing.

Specific Comments:

Dr. Anderson stresses that the EPA post exposure treatment is different and that this introduced variation into the data collected. She doesn't discuss this further. Killing the mice for pathology (assuming this is the difference she means) obviously would prevent observation of late deaths due to treatment, but would not affect differences observed during the exposure experiments.

It is difficult (without actually seeing the procedure) to imagine why a "low laboratory cart having 4 sides" would produce different results from a three-sided elevated shelf in the FOB studies

Microbiological data:

As expected, these results are unremarkable and provide no insight as to possible mediators of either the "major" effects seen by Dr. Anderson or the minimal effects of carpet exposure seen by the EPA. As stated in the report, microbial analysis on samples collected long after alleged exposure/complaint episodes may not (and, in fact, are probably not) representative of that exposure. At the time of testing, unless the carpeting was supporting active microbial growth that would release VOC's, it is unlikely that, in a static system of this sort, particulate exposure to biological agents would occur.

Specific comments on the microbiology report:

Pg 5 (also on Pg 8): The use of the word "significantly" implies that some statistical test was used. It is entirely possible for very small consistent differences to be statistically significant, and multiple tests on these samples could well have produced such significance. Actually, all you can say is that the levels appear to be similar.

Pg 7: Other reasons for the variability between culture media include the overgrowth by members of the mucorales on MEA and 2% MEA.

I disagree that the culture results from R-800-3516 did not change after storage. Gliocladium positive plates went from 12 to 1 and Trichoderma positive plates went from 2 to 16. This indicates a clear change that is not reflected in numbers and points up the fact that numbers alone are of little use in this kind of study. Note that Aspergillus glaucus is no longer a valid species name, and that either the actual species should be determined, or the taxon listed as Aspergillus glaucus group.

The slightly higher levels and larger diversity of taxa recovered from R-800-3516 may also be due to normal variance in fungal populations in carpet, or to variance in the methods used. Actually, I would consider R-800-3516 to be considerably different than the other two samples. Only Cladosporium is consistently frequent across all three carpet types. R-800-3516 has higher frequencies (by a factor of at least two) for Penicillium, Aspergillus, yeast, Alternaria, Gliocladium, and Trichoderma. I developed the following chart from your data to make this evaluation.

	R872-72	R-800-1205		R-800-3516	
		Before	After	Before	After (Storage)
Penicillium	24	32	22	72	74
Cladosporium	33	26	32	37	44
Aspergillus	18	15	9	44	37
yeast	4	4	0	11	7
Rhizopus	7	0	4	4	0
Neurospora	0	2	2	2	4
Alternaria	4	0	0	6	9
Gliocladium				22	2
Trichoderma				4	30

Comparing fungal counts per gram of carpet and per gram of dust is irrelevant, and gives the (possibly) erroneous impression that the current samples contained fewer colony forming units than those in the listed literature citations. In fact, the frequency of Penicillium, Aspergillus, Gliocladium, and Trichoderma is an indication that R-800-3516 is probably "contaminated."

Pg 9: There is at least one paper in the literature that related dust recoveries to airborne recoveries (Gravesen 1978*). A bigger problem

here is the fact that you measured CFU/gram of carpet rather than per gram of dust. Carpet is unlikely to become airborne, and you have no idea of the fraction of recoveries that were from dust in the carpet or were from fungi and bacteria adhering tightly to the carpet fibers that are unlikely ever to become airborne. This is why we use dust rather than carpet.

---

*Gravesen S. 1978. Identification and prevalence of culturable mesophilic microfungi in house dust from 100 Danish homes. Comparison between airborne and dust-bound fungi. Allergy 33(5):268-72

The endotoxin part of this report is weak. For example, how was endotoxin extracted from the carpet? Extraction method controls how much endotoxin you get even from less complex substrates than carpet. Has the assay used been checked for interference by other carpet/dust components? Note that the endotoxin assay is a bioassay that is easily perturbed by extraneous materials in environmental samples. Also, as for bacteria and fungi, it is irrelevant to compare levels recovered from carpet to levels recovered from dust. Endotoxin is well-known to stick quite tightly to surfaces in ways that would be unlikely to allow aerosolization, but would allow release into washing solutions.

There is considerably more literature on the microbiology of carpet dust than is listed in the references. Considering the "volatility" and visibility of this project, the EPA would be wise to conduct an in-depth and critical literature review on carpet research.

Tables: What is the minimum level of sensitivity? i.e., what does "below detection limit mean? If this is different for each table or entry, then it would be better to use <xxx so that the sensitivity is known.

Lawrence Berkeley Laboratory
1 Cyclotron Road Berkeley, California 94720
Energy & Environment Division (510) 486-4000 • FTS 451-4000

June 1, 1993

Dr. Robert Dyer
Associate Director (MD-51)
Health Effects Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

RE: Review of Phase I of the EPA/ORD Carpet Study

Dear Dr. Dyer:

This letter report contains my review comments on the results of Phase I of the EPA/ORD carpet study. These comments are based primarily on the reports provided by the EPA study team and Anderson Laboratories, Inc. (AL) and on the oral presentations made at HERL on May 26, 1993. In addition, the group of reviewers separately interviewed the EPA study team and Dr. Anderson on May 27.

It is my opinion that the EPA team adequately reproduced in their laboratory the apparatus and procedures developed at AL to measure the respiratory and neurotoxic effects of carpets on mice. This was accomplished through numerous discussions with Dr. Anderson and by reciprocal visits, including one week at AL in which the EPA team worked side-by side with the AL technician. The comparability of the apparatus is, in part, demonstrated by the temperature data. The temperature range in the source chamber (location A_1) at AL was 32-40° C; at EPA the range was 36-39° C. The temperature range in the mouse exposure chamber at AL was 22-25° C; at EPA the range was 22-26° C with one value at 30° C. EPA made one significant modification to the apparatus which was to use humidified zero-grade air as the inlet air for the system. This is a justified improvement resulting in a more controlled test environment. As a result, the humidities of the inlet air were somewhat different. At AL the range was 27-58 % RH; at EPA the range was 19-29 % RH. Other differences between the test systems were the use of Swiss-Webster mice from different suppliers; behavioral examinations after every exposure at AL and only after the second and fourth exposures at EPA; and the use of thinner mouse collars with somewhat smaller holes at EPA. None of these differences would be expected to have an obvious impact on the ability of the systems to detect the toxic effects of emissions from carpets.

The focus of Phase I of the study was to attempt to replicate the previously reported AL results by conducting simultaneous experiments at both laboratories using carpets that had been shown by AL to have severe adverse effects on mice. The experiments were to be replicated to provide a larger sample size and the possibility of

statistical analyses. Blinding procedures were to be used at both laboratories to ensure that the individuals collecting the data did not know the contents of the source chambers.

In my opinion, the research conducted by EPA was of very high quality. Experts from various laboratories within EPA, as well as outside contractors, were brought together to participate in the study. The team consisted of nationally and internationally recognized experts in emissions testing, chemical characterization, microbiology, and toxicology. The animals were properly housed and cared for. Sentinel animals were used to assess the colonies for sub-chronic disease. The blinding procedures built into the study were effective and strictly adhered to. These blinding procedures included receipt of the samples by an offsite contractor, loading of the source chambers by this contractor, use of separate observers for both the irritation tests and the Functional Observational Battery (FOB) that were not present when the mice were loaded into the chambers or when the mouse exposure chambers were connected and disconnected from the source chambers, and chemical analyses conducted by another team using a separate but equivalent apparatus with no mice. The data for the pulmonary irritancy (PI) and FOB tests were rigorously analyzed in a manner that would tend to increase the chance of finding false positive results (*i.e.*, the analyses were conservative). Extensive quality assurance procedures were used throughout all aspects of the study, including the use of another offsite contractor to review procedures.

The same cannot be said about the quality of the research conducted at AL. The most serious problem, which could have had a significant impact on the results of the PI and FOB tests, was the likely breach of the blinding procedures at AL. This breach may have started from the time the samples first arrived at AL by Federal Express. The packages were delivered in the afternoon at approximately 3 pm and sat on the secretary's desk until the sample custodian arrived at 7 pm. The package for Test 2, the first blank, was marked "empty" on the outside by CPSC and not weighted to compensate for the weight of carpet. This package would have been visible and accessible to all personnel since the entire facility is quite small. For each sample, the sample custodian loaded the contents of the package into the source chamber, taped the chamber to conceal the contents, placed the chamber on a cart, wheeled the cart into the testing laboratory, and left. The next morning the laboratory technician started the test. This technician performed all aspects of the test. His functions included making the baseline sensory irritancy (SI) and PI measurements and then hooking the mouse exposure chamber up to the source chamber. At this point, it is quite possible that the contents of the source chamber could be viewed through the connecting port on the source chamber. After the one hour exposure period, the technician removed the exposure chamber from the source chamber and measured the concentration of TVOC in the atmosphere of the source chamber using a simple hydrocarbon analyzer which presumably produces a result within a few minutes and provides a readout. The results presented by AL, show a distinct difference between the TVOC concentrations of empty and carpet-loaded source chambers. The TVOC concentrations for Experiment B, the blank, are reported to be 1.6-2.1 ppm as methane. The TVOC concentrations for Experiment A were 2.4-4.4. For Experiment C, the concentrations were 3.7-5.4 ppm. With carpets, the highest concentrations would occur at the end of the first exposure. Based on all of their previous data using the same system (*i.e.*, the analysis of hundreds

of samples), the technician would probably know at that point whether the source chamber was empty or contained carpet. Next, the technician performed the FOB. Another possible clue that a blank was being tested would be the clean SI and PI tracing reported for the blanks. Test 3, another blank, was an inadvertent test of computer paper. The results of this test have not been reported by AL. When AL was informed of the sample pairings at the end of the study, AL indicated that their records showed that Tests 2 and 3 were not pairs. Another sample was sent, this time an empty bag with the box weighted with paper. However, since AL had been informed of the pairings and undoubtedly knew that two pairs of carpet samples had already been tested, it would be easy for them to assume that this final sample was indeed another blank.

It is surprising and of some concern that Dr. Anderson did not recognize the relatively easy ways in which the blind could be broken at AL and did not take any steps to eliminate the potential problems.

The extent to which breaking the blind might have had on the results produced by AL is impossible to assess. However, experience shows that the anticipation of results can introduce significant biases into subjective measurements, such as the PI and FOB tests. This is why blinding procedures are used in scientific investigations.

Such biases cannot explain the extreme behavior effects or the deaths reported by AL for Experiments A and C with carpets. Some other factor may be involved. One of the other reviewers (R. Schelsinger) raised the possibility of sub-chronic disease making the animals more susceptible to stress and toxic exposure. Another possibility is that stress induced by the collars and restraint in the plethsmographs might have made the mice more susceptible. J. Tepper told the reviewers that mice can die as the result of having their head shoved through the hole in the collar or from forcing their way out of a collar. Interestingly, Dr. Anderson told the reviewers that unrestrained mice exposed to emissions from carpets exhibit severe behavioral effects but not death. AL has not attempted to determine the cause of death of the mice from their tests.

It is possible that there is some subtle phenomena occurring that is not directly related to exposures to emissions of toxic chemicals from carpets. Other than factors involving the source and health of the mice, the major difference between the apparatus at the two laboratories is EPA's use of humidified zero air. The humidities of inlet air at AL were higher than those at EPA, but the typical difference is only about 10 % RH. It is difficult to postulate a mechanism, either biological or chemical, whereby this small difference could have such a dramatic effect.

I reviewed all of the chemical emissions data provided in the reports from Acurex Environmental Consultants and Research Triangle Institute (RTI). The multi-sorbent method used by both laboratories for the analysis of volatile organic compounds (VOCs) is appropriate for this application as it measures a broad spectrum of compounds with high sensitivity. The DNPH method used for the carbonyl compounds is the best available method for this class of compounds. An intercalibration was performed between Acurex and RTI to ensure comparability of the data. In addition, numerous calibrations and other quality assurance procedures were used by both laboratories for the chemical analyses.

The TVOC emissions from both carpet samples were elevated compared to what would be expected at room temperature, but still quite low compared to emissions of TVOC from "wet" sources, such as adhesives, paints, and other coatings. During the four exposure periods (4 hours total) the total emissions of TVOC were only 4-5 mg. This is equivalent to 14-17 mg m^{-2}, which compares to total emissions of TVOC from new styrene-butadiene rubber latex carpets over 24 hours of 8-26 mg m^{-2} (Hodgson, A.T., J.D. Wooley, and J.M. Daisey, 1993, Emissions of volatile organic compounds from new carpets measured in a large-scale environmental chamber, *J. Air Waste Manage. Assoc.* 43: 316-324). The composition of the VOC samples was also unremarkable. Both carpet samples were dominated by C_{11} - C_{12} alkene hydrocarbons and siloxane compounds. The siloxanes were also emitted by the aquarium. Carpet A had higher emissions of butylated hydroxy toluene and formaldehyde. The alkene hydrocarbons would be expected to have relatively low toxicity. Formaldehyde might play a role in sensory irritation, but would not be expected to cause behavioral effects or death in mice. In summary, the chemical emissions data do not seem to provide any clues about the cause of the severe effects reported by AL.

Additional work is required to resolve the differences between the two laboratories regarding the PI and FOB results and the deaths. Firstly, I recommend that the raw data be exchanged between the two laboratories. The investigators may interpret the other laboratory's respiratory tracings differently and, thus, account for some of the differences in the PI results that were observed. The FOB results presented by AL could benefit from the statistical analyses available at EPA. It is impossible to fully evaluate these results as they are now presented because of their highly condensed summary form. For example, differences among exposures for a test or between replicates for a treatment are obscured. Secondly, additional data obtained by AL should be examined. The issue of reproducibility could be assessed by comparing the results obtained by AL during this study to their previous results for the same carpets. The results of the experiment with paper would provide another control against which to evaluate the carpet results. Thirdly, an attempt should be made to determine the cause of death of mice at AL. Complete autopsies are needed. Relatively simple experiments could be conducted to evaluate the potential roll of the hole size and thickness of the mouse collars in causing death. Other experiments using known mixtures of VOCs could be conducted. Finally, additional tests with carpets and controls could be performed in which the blinding procedures were strictly adhered to by both parties. This might involve both AL and EPA teams making simultaneous observations at AL.

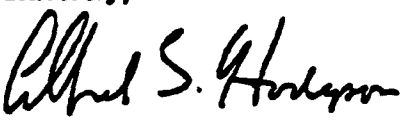
I would focus on efforts to understand the disparity in the Phase I results before proceeding with additional experiments. If additional experiments still seem warranted, they could be conducted with a limited number of carpet samples in the new all-glass chamber as suggested by D. Costa. It would be of particular value to establish dose-response curves for any observed effects. The SI and PI tests should definitely be separated from the FOB test. The SI and PI tests should follow ASTM E 981 in which a single exposure is used since the EPA data suggest that the test procedure of multiple restrained exposures produces measurable stress in the mice. The FOB test should be conducted with mice not restrained by collars to eliminate this confounding factor.

I recommend that the summary report, "Evaluation of Carpet Toxicity - Phase I," by R. Dyer and D. Costa be amended to specifically address the comparison of the results obtained by EPA and those obtained by AL. The high quality and the thoroughness of the research conducted at EPA should be emphasized. It is justified to conclude that even when all reasonable efforts were made to replicate the apparatus and procedures used by AL, it was not possible to reproduce any of the results of AL showing the study carpets to be toxic to mice.

I further recommend that EPA sponsor a scientific forum on the issue of carpet toxicity to mice to include other researchers and organizations who have either tested the same carpets using similar methods or have experience with these types of tests. Such a meeting might generate additional insight into the potential causes of the toxic effects reported by AL.

Please contact me if you have any questions regarding these review comments. Thank you again for the opportunity to participate in an interesting and important research effort.

Sincerely,



Alfred T. Hodgson
MS 70-193A
510-486-5301
FAX 486-6658

REPORT TO EPA ON EPA/ORD CARPET STUDY

D. E. McMillan

University of Arkansas for Medical Sciences

This report is an independent report written by the author listed above. The report will review the data provided to him in written form as well as the data presented at the meeting held at EPA on May 26-27, 1993. The focus of this report will be on the Functional Observational Battery (FOB) data provided by Anderson Laboratories (AL) and the EPA; however, some general issues regarding experimental design and conduct of the experiments will be included.

I. Quality of the research that was conducted

A. Carpet Samples

The decision was made to test samples from two different carpets about which complaints had been received. Although it makes sense to begin testing with carpets presumed to be toxic, the study of used carpets is a complicating factor. Even if both AL and EPA had confirmed a toxic effect from carpet emissions, it would not have been possible to establish whether the toxicity resulted from the manufactured carpet, or from contamination during use and subsequent handling of the carpet. Further testing would be required to answer this question. Since the EPA was not

able to replicate the AL findings this has become a moot point; however, in the future it might be appropriate to conduct some tests with new samples of the carpets listed as suspect to avoid this problem.

B. AL study

AL reported large effects in almost all treatment groups including the death of 5 exposed animals. Sham exposed animals showed remarkably few effects, thereby establishing a large treatment effect. There are a number of problems with the experiments conducted at AL that somewhat limit the interpretation of their data. The first problem concerns the double blind design of the study. Samples were received at AL in the late afternoon and were placed by the courier on the desk of a secretary. It is not clear who signed for the delivery, but apparently the sample container sat on the desk with no security until approximately 7:00 PM at which time the sample custodian arrived, removed the sample, placed it in the source chamber and sealed the chamber. The sample custodian then moved the chamber to the test room where it remained until the next morning. The lack of security of the sample to this point allowed a number of opportunities to break the blind. However, assuming that there were no deliberate attempts to break the blind, the testing procedure almost certainly did so.

The next morning the technician who conducted all subsequent testing initiated heating of the carpet samples and placed the animals in the exposure chamber. During exposure the technician measured the rate and patterns of respiratory responding of the animals. During this period the technician constantly monitored the animals. Dr. Anderson stated on several occasions how clean (lacking in effect) the data were from the control animals, which certainly is consistent with the data in the written report. Since the same technician conducted the FOB, after completing the respiratory measurements, it would be almost impossible for this technician not to know at the time of FOB testing whether an animal had been exposed to carpet emissions or sham exposure. If there was any doubt about group identity at this point, the measurement of total volatile organic chemicals by flame ionization detection by the same technician, would eliminate that doubt. Thus, it seems unlikely that the AL observations for the FOB were done according to a blind procedure as specified.

Another problem with the experiments done at AL is the data analysis. As I understand their procedure for the FOB, on the day prior to exposure to carpet emissions the 8 test animals are observed and scored on the FOB. On the basis of these pre-exposure observations, a range is established for each group. Subsequently, the FOB is repeated four times after sham or carpet-emission exposure. Data are reported as number of animals falling outside the range. Based on

the relatively small number of control observations relative to the number of treatment observations, it would be expected that many experimental values would fall outside the control range. This was clearly the case for the treatment groups, but not for the sham exposure groups.

Reasons that might account for the difference between sham and exposure groups include an effect of the carpet emission exposure, bias due to loss of the blind, or other yet to be determined factors. The manner in which the AL data are presented makes further detailed analysis of the data very difficult. It is unclear, for example, whether or not the effects reported by AL are seen repeatedly in only a few animals, or are seen sporadically in other animals. All of the data are heavily derived so that the absolute comparability of groups before, during and after exposure becomes very difficult. Dr. Anderson indicated that she would be willing to supply less derived data for further examination of this issue.

C. EPA Study

The EPA study appears to have been carefully conducted. The chambers at EPA were loaded by outside personnel. The FOB was conducted by individuals very well trained in conducting this battery and they were not the same personnel that conducted the other tests. There does not appear to have been much opportunity for those conducting the testing at EPA to have broken the blind (barring the possibility of

fraudulent conduct of the scientists, which seems most unlikely).

Relative to the rather simple analysis conducted by AL, the statistical analysis applied to the data developed by the EPA was sophisticated. The statistical analysis of the EPA data gives this reviewer greater confidence in the EPA findings than is possible from the AL data analysis. However, some of the EPA data also were presented in a derived form (percentage incidence, percent change from control, etc.) that made it very difficult to determine the comparability of control and treatment groups before exposure and other details of what actually happened. I also found the 3-dimensional bar graphs that they used for the presentation of some data difficult to follow, but this may be a personal reaction against this form of graphics. Clearly, the EPA data could be presented in a form that would provide more information than the present version does.

In summary, the general design and conduct of the EPA studies appears to be more rigorous than that done by AL and from a scientific viewpoint I am inclined to place much greater confidence in their data than that from AL. Nevertheless, there other possible explanations for the differences between EPA and AL findings and since the welfare of the public exposed to carpets is at issue it seems clear that further studies should be done to determine the reasons why AL finds effects that EPA cannot replicate

and to further determine whether or not carpet emissions constitute a public health problem.

II. INTERPRETATION OF THE DATA

Using their criterion of treatment values falling outside the range of control values, AL finds evidence of sensory irritation and changes in the FOB including changes in their groupings of altertness/excitability, neuromuscular function and appearance. It should be noted that the data analysis done by AL (proportion of animals that fall outside the control range established by pre-exposure observations in the same animals) gives equal weight to both increases and decreases in the magnitude of the dependent variables. In some instances the changes that they measured consistently moved in the same direction. In other instances both increases and decreases occurred, so that their treatment effect was to observe an increase in variability of the treatment group relative to the control group. It is possible that some of these differences would not be significant when analyzed by more rigorous statistics. Although the interpretation of the AL data are limited by the presumption of the broken blind and the minimal statistical evaluation, it seems unlikely that the AL effects would disappear with more rigorous attention to these problems. This is particularly true with respect to the deaths that they observed in the treatment groups.

The EPA investigators found few significant differences. In some instances the significant differences they found were due to changes in the control group over time, or were found in only one of the two treatment groups. On the basis of this analysis the EPA's conclusion that they found no differences seems appropriate. In fact, based on the analysis of the data presented, both AL's conclusion that there clear differences between sham exposure and actual exposure to carpet emissions and EPA's conclusion that there were no differences, are appropriate. At the review, Dr. Virginia Moser presented a preliminary analysis of the EPA data using the AL method of analysis. This preliminary analysis did not account for the differences between laboratories. Obviously, differences in statistical rigor cannot account for the significant death rate at AL which was not observed at EPA.

Although I have little expertise in the analysis of source materials, the chemistry data appear to support the EPA findings. There is no toxic chemical identified in the carpet emissions in an amount sufficient to produce the observed toxicity. Although the subject of toxic effects of mixtures has not been well studied by toxicologists, the very low levels of toxic chemicals make any powerful synergistic effects unlikely, although impossible to eliminate entirely.

III. FOCUS OF THE REPORTS

Both the draft report suppled by AL and that suppled by EPA are appropriately focused on the fundamental issues. The AL includes an executive summary, an adequate description of the methods and procedures, and a rather brief results section. Most of the relevant data a presented in accompanying appendices. After brief mention of the problem that occurred with sample submission, the AL report concludes with the obvious, namely that exposure to carpet emissions produced large effects in the exposed animals relative to sham exposure.

The draft report from the EPA is broader, but this is not inappropriate. The EPA study included a great deal of additional testing including tests on the identification of source emissions, postmortem evaluations, quality assurance, etc. The additional data required the expansion of the results section of the report as well. The discussion section of the EPA draft also includes a section that speculates on some of the reasons that might relate to the reasons why EPA was not able to replicate the results obtained by AL.

In summary, both draft reports are appropriately focused, despite some dissatisfaction of this reviewer with the details of the data presentation in both draft reports. The bottom line is that AL has replicated their previous results. EPA, despite an intensive effort, has not been able to replicate the AL findings. The studies conducted by EPA are superior scientifically to those conducted by AL;

however, it does not seem likely that any deficiencies in the AL procedures could account for all of the large differences that they obtained.

Because of the publicity surrounding putative problems with carpet emissions and the economic issues associated with these studies the issue of possible fraud on the part of scientists at AL, or EPA must be raised. Fortunately, there is no evidence that scientists at either institution have in any way acted inappropriately. Visits between laboratories have occurred. Both laboratories appear quite willing to allow others to examine their data in detail. The presumed loss of the blind at AL relates to the small size of their operation and lack of sufficient personnel to permit the isolation of experiments from each other and not to any intentional attempt to violate the blind conditions of experimentation.

IV. APPROPRIATE NEXT STEPS

It is the opinion of this reviewer that EPA cannot allow this matter to drop at this time. Although their attempt to replicate the AL results has failed, the issue of whether or not carpet emissions constitute a public health problem is not resolved and further testing is required. The following recommendations and the reasons for making them are offered:

A. Replication

An independent laboratory should conduct an independent systematic replication of the data. Because of the political sensitivity of the issue this laboratory should have no association with EPA, AL, the carpet industry, or anyone else already associated with the controversy. Alternatively, EPA and AL might conduct a joint study. For example, a replication might be done at EPA and supervised by EPA personnel using AL equipment, staff and animals, or vice versa.

B. Systematic replication

Any replication should not be an exact replication in the opinion of this reviewer. A better exposure system should be used to study carpet emissions. New carpets of the type previously implicated as problem carpets should be investigated. At least one increased emission exposure level should be studied. If EPA failed to replicate the AL result because their methods lacked sensitivity, effects may be seen at a higher exposure level. Great care should be taken to separate respiratory testing from FOB testing so that the blind cannot be accidentally broken. Perhaps consideration should be given to the use of different animals for respiratory and FOB testing with the FOB animals receiving whole body exposure, since the EPA's data seem to indicate that the restraint is very stressful and could contribute to the FOB effects. The statistical analysis to be applied to the data should be specified as part of the

replication attempt. Consideration should also be given to study of emissions from unheated carpet, since the heated carpet may not model most real world exposure. An completely objective behavioral measure should be added to the test procedure. Since experimental bias is an issue in these experiments, the addition of a behavioral procedure that does not require observer judgments would be a welcome addition. Measurements of spontaneous locomotor activity or schedule-controlled responding are suggested.

C. Laboratory Differences

Attempts to determine reasons for the differences in laboratory findings should continue to be pursued. Current methods of risk assessment often place emphasis on the extreme result, even when conflicting studies are scientifically stronger. Once a report of toxicity enters the literature, several negative reports may be required to decrease the perception of toxicity. Therefore, it remains important to continue to explore reasons for the differences between labs. If it could be established that something is causing effects at AL that is only remotely related to exposure to carpet emissions, or that something is happening at EPA to prevent them from obtaining what are real effects of emissions, the entire issue would be resolved. Some of the hypotheses to explain reasons why laboratory results differed evolved from discussions at the review. These hypotheses include a role for water vapor in the carpet samples affecting the toxicity of the emissions,

a role for conditioned stress during multiple carpet exposures, differences between laboratories in animals in the outbred strain that was used, possible pathogens in the laboratory, variability in the samples submitted to the testing laboratories and others. These and other hypotheses should be tested.

D. Cost-Effectiveness Considerations

The reviewer realizes that the suggestions for additional action may not be very helpful to EPA because they are broad, expensive and time consuming. Unfortunately, for the reasons outlined, I see few alternatives. If forced to prioritize among these choices I would suggest that EPA simultaneously pursue as extensive a systematic replication as funding and time permit with the replication to be conducted at an independent laboratory. At the same time investigations into the reasons for the different results should continue, beginning with a complete data exchange. EPA and AL staff should meet to discuss possible reasons for the differences, prioritize the order in which the hypotheses about testing are scheduled for testing and begin to do the appropriate tests. It is in the best interests of both EPA and AL, as well as the health of the public and the carpet industry, that these issues be resolved as soon as possible.

Donald E. McMillan
6/1/93