EUROPEAN COLLABORATIVE ACTION

INDOOR AIR QUALITY & ITS IMPACT ON MAN

Environment and Quality of Life

Report No 20

Sensory Evaluation of Indoor Air Quality

EUROPEAN COMMISSION
JOINT RESEARCH CENTRE - ENVIRONMENT INSTITUTE

1999

EUR 18676 EN

EUROPEAN COLLABORATIVE ACTION

INDOOR AIR QUALITY & ITS IMPACT ON MAN

(ECA-IAQ)

Environment and Quality of Life

Report No 20

Sensory Evaluation of Indoor Air Quality

prepared by

WORKING GROUP 15

Birgitta BERGLUND, Department of Psychology, University of Stockholm, Stockholm, Sweden

Philomena BLUYSSEN, TNO, Department of Indoor Environment, Delft, The Netherlands

Geo CLAUSEN, Centre for Indoor Environment and Energy, Technical University of Denmark, Lyngby, Denmark

Ana GARRIGA-TRILLO, Univ. Nacional de Educación a Distancia, Madrid, Spain

Lars GUNNARSEN, Danish Building Research Institute, Hørsholm, Denmark

Helmut KNÖPPEL, European Commission, Joint Research Centre, Environment Insitute, Ispra, Italy

Thomas LINDVALL (chairman), Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden

Patrick MacLEOD, Laboratory of Sensory Neurobiology, Ecole Pratique des Hautes Etudes, Massy, France

Lars MØLHAVE, Institute of Environmental & Occupational Medicine, Aarhus University, Aarhus, Denmark

Gerhard WINNEKE, University of Düsseldorf, Düsseldorf, Germany

reviewed and approved by

The STEERING COMMITTEE

EUROPEAN COMMISSION

JOINT RESEARCH CENTRE - ENVIRONMENT INSTITUTE

1999

EUR 18676 EN

A great deal of additional information on the European Union
is avaiable on the Internet.
It can be accessed through the Europa server
(httpp://europa.eu.int).

Cataloguing data can be found at the end of this publication

Luxembourg: Office for Official Publications of the European Communities, 1999

ISBN 92-828-5699-2

Printed in Italy

Human subjects are indispensable in the measurement of perceived indoor air quality. Chemical and physical methods of characterisation often are insensitive to odorous and sensory irritating air pollutants, or do not take account of combinations of singular pollutants in a biologically meaningful way. Therefore, sensory methods many times are the only or the preferred tool for evaluation of perceived indoor air quality.

This report presents background to and advice on methodologies for sensory evaluation of perceived indoor air quality. It proposes methods which apply to source assessments as well as field investigations. The methods will assist in labelling of building materials, characterising air quality in indoor spaces, controlling ventilation performance, and measuring occupant responses in questionnaire field studies of the sick building syndrome.

The proposed methods will enable designers, manufacturers, chemical and ventilating engineers, consumers, building and health authorities, and other decision makers to compare and select appropriate building materials, furnishings etc. Thereby the design, supply and control for good perceived air quality in indoor spaces will be easified which will lower the costs and minimize waste of energy.

In this series the following reports have already been published

- Report No. 1: Radon in indoor air. (EUR 11917 EN) *
- Report No. 2: Formaldehyde emission from wood-based materials: guideline for the determination of steady state concentrations in test chambers. (EUR 12196 EN) *
- Report No. 3: Indoor pollution by NO_2 in European countries. (EUR 12219 EN)
- Report No. 4: Sick building syndrome - a practical guide. (EUR 12294 EN) *
- Report No. 5: Project inventory. (S.P.I. 89.33) *
- Report No. 6: Strategy for sampling chemical substances in indoor air. (EUR 12617 EN)
- Report No. 7: Indoor air pollution by formaldehyde in European countries. (EUR 13216 EN) *
- Report No. 8: Guideline for the characterization of volatile organic compounds emitted from indoor materials and products using small test chambers. (EUR 13593 EN)
- Report No. 9: Project inventory - 2nd updated edition. (EUR 13838 EN) *
- Report No. 10: Effects of indoor air pollution on human health. (EUR 14086 EN)
- Report No. 11: Guidelines for ventilation requirements in buildings. (EUR 14449 EN)
- Report No. 12: Biological particles in indoor environments. (EUR 14988 EN)
- Report No. 13: Determination of VOCs emitted from indoor materials and products. Interlaboratory comparison of small chamber measurements. (EUR 15054 EN)
- Report No. 14: Sampling strategies for volatile organic compounds (VOCs) in indoor air. (EUR 16051 EN)
- Report No. 15: Radon in indoor air. (EUR 16123 EN)
- Report No. 16: Determination of VOCs emitted from indoor materials and products; second interlaboratory comparison of small chamber measurements. (EUR 16284 EN)
- Report No. 17: Indoor Air Quality and the use of Energy in Buildings. (EUR 16367 EN)
- Report No. 18: Evaluation of VOC emissions from building products: solid flooring materials. (EUR 17334 EN)
- Report No. 19: Total volatile organic compounds (TVOC) in indoor air quality investigations. (EUR 17675 EN)

* out of print

Abstract

ECA-IAQ (European Collaborative Action 'Indoor Air Quality and Its Impact on Man'), 1999. Sensory evaluation of indoor air quality. Report No 20. EUR 18676 EN. Luxembourg: Office for Official Publications of the European Communities

This report presents background to and advice on methodologies for sensory evaluation of indoor air quality (IAQ). The report gives a short introduction to sensory mechanisms and responses and to the theory of measurement underlying sensory evaluations and discusses in detail available sensory evaluation techniques. After a critical methodological analysis of some recently published documents on IAQ, sensory methods best suited for the evaluation of material emissions and of IAQ and for population response studies are recommended. Also non-sensory techniques for the evaluation of odour and mucosal irritation are briefly discussed. However, it is concluded that, at present, human subjects are indispensable in the measurement of perceived indoor air quality.

The proposed methods will enable designers, manufacturers, chemical and ventilating engineers, consumers, building and health authorities, and other decision makers to compare and select appropriate building materials, furnishings etc. Thereby the design, supply and control for good perceived air quality in indoor spaces will be easified which will lower the costs and minimize waste of energy.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

I. Scope

This report is the outcome of an endeavour by a group of European experts with different disciplinary backgrounds to agree on methodologies for the sensory evaluation of indoor environmental quality and it reflects the difficulties of such an agreement. The report presents background to and advice on methodologies for sensory evaluation of indoor environmental quality, especially indoor air quality (IAQ). A comprehensive, critical overview is given of methods for measuring odour and sensory irritation, symptoms and annoyance/discomfort. The potential uses and misuses of the methods are indicated.

The proposed methods will enable designers, manufacturers, chemical and ventilating engineers, consumers, building and health authorities, and other decision makers to compare and select appropriate building materials, furnishings etc. Thereby the design, supply and control for good perceived air quality in indoor spaces will be easified which will lower the costs and minimize waste of energy. The practical application of the methods will require European standardisation. The investigator should have acquired basic laboratory skills in psychology, chemistry or laboratory engineering.

II. Principles and methods of sensory evaluations

Sensory mechanisms and responses (Section 2). Odour and sensory irritation are associated with two senses: olfaction and the common chemical sense. The underlying basic biological principles are partly well understood. Much more uncertain is how the basic processes relate to the complex, psychological responses to odorant/irritant stimulation, such as perceived odour quality, annoyance and symptom reporting. Population-based studies in the outdoor environment using questionnaires have revealed associations between e.g. odour exposure and annoyance as well as symptom reporting. However, such information cannot be directly translated to the indoor environment.

Theory of measurement underlying sensory evaluations (Section 3). Sensory analysis is based on the use of human subjects as measuring instruments. Therefore, it is essential that the sensory analyst knows the "basic rules" of subjective measurement. The main attributes of perception that can be measured by olfaction and the common chemical sense are: detection, discrimination and perceived intensity, and, by adding higher level processing, perceived quality and value judgements.

The mathematical level or the scale type (level of measurement) used for sensory measurements puts restrictions on the data analysis. Within the sensory field commonly four scale types are considered: nominal, ordinal, interval and ratio. A yes/no response in a detection task generates a nominal scale. Questionnaire data commonly are on nominal or ordinal scales. It does not make sense to add, subtract, multiply or divide the numbers of such scales.

Physical-chemical measurements produce commonly interval or ratio scales. Also the measurement of perceived odour/irritation attributes (like intensity or pleasantness) can

generate interval or ratio scales depending on the method used for the measurement. It makes sense to add and subtract the numbers from an interval scale (compare the Celsius scale). Within the ratio scale it makes sense to add, subtract, multiply and divide the numbers obtained (compare the Kelvin scale).

There is a great need for improving the level of measurement in sensory evaluations. For that purpose the discipline of psychophysics provides many techniques and testing procedures for obtaining sensory data from human subjects. In order to obtain comparability between sensory evaluations, they have to be calibrated or standardised. For calibration it is necessary to use references. A number of reference chemicals have been suggested for sensory investigations of air quality.

Sensory evaluation techniques (Section 4). Perceived quality is defined here as the perception of complex sensory stimulation, e.g., a pattern or a blend. There is little mental evaluation involved in this process. By value judgement is meant here a high order mental evaluation. It can be based upon perceived intensity and/or perceived quality of the sensory stimulation. The result can range from hedonic tone to acceptability.

Most sensory threshold methods are based on dilution of the air sample until half of a panel no longer detects odour or irritation. The thresholds vary widely with chemical substance, measurement procedure, quality of the equipment (olfactometer), purity of the chemical substance, and sample of subjects. This is reflected by the large spread in odour threshold values for single compounds as reported in the literature.

Threshold values do not give any measure of the perceived intensity above threshold levels. In materials testing, though, they can provide information on how much air dilution of the emission is needed to reach odorlessness.

Methods based on signal-detection theory do not require wide dilution series and, therefore, are suitable for detecting weak odours as they typically appear in indoor spaces. Both positive and negative false responses (false alarms and misses) can be estimated, given a sufficiently large number of stimulus presentations. Thus, the effects of the response criterion can be separated from the sensitivity measures. Classical threshold methods can only correct for false positive responses.

Measures for quantifying perceived intensity of odours and sensory irritation are obtained from subjects by various psychological matching and scaling methods. They include, i.a., equal-intensity matching and magnitude estimation. Direct scaling methods are the most common. Preferably, the perceived intensity measures shall be obtained from a calibrated scale so that measures can be compared between different materials, occasions and laboratories.

Non-sensory techniques (Section 6). Non-sensory biological and physical-chemical methods for "sensory" evaluations exist but they are not valid except for selected situations and selected compounds. Even in cases when they are valid they are not standardised to the level that they can be used for meaningful predictions of sensory responses to indoor environments. The non-sensory methods must always demonstrate a meaningful relation to human sensory responses before they can claim to indicate perceived indoor air quality.

Exposure conditions (Section 7). Many methods for sensory evaluation of perceived air

quality require advanced equipments for sampling, gas dilution and stimulus presentation to subjects. Dynamic systems for exposure generation should be preferred. In sensory assessments of relatively weak odours, as in non-industrial indoor spaces, grab-sampling is discouraged since the losses are expected to be significant from, i.a., adsorption and chemical reactions. Only exposures which do not impose an increased risk of health effects on the staff and subjects can be accepted. In practice, most exposures are whole body exposures in climate chambers under fully controlled conditions, field exposures of walk-in types, or partial exposures of nose and/or eyes. An air lock should be established between the test/waiting rooms and the surroundings. Adaptation should be minimised and air cleaning is often needed. Specified response modifying factors should be controlled.

Selection of subjects (Section 8). Some of the interlaboratory variations can be attributed to panel differences. Therefore, European standardised panel selection procedures should be established. Important factors are sensory sensitivity, representativeness to the target population and personality variables.

Panels may become more comparable by training to meet common performance criteria. Furthermore, panel members must be unbiased. Therefore, the use of occupants of problem buildings or polluted areas is problematic. Family or staff members cannot be used.

The documentation of the panel should include: gender, age, smoking habits, criteria for why some subjects may have been excluded, measures of sensory sensitivity and its representativeness as to a specified population.

Limits and quality requirements of sensory evaluations (Sections 9 and 11.1). Some harmful air pollutants are not sensed at all (e.g. carbon monoxide and radon) and the sensory effects of some other pollutants are not quantitatively linked with their toxicity. Therefore, perceived air quality cannot be used as the only measure of health effects.

Reliable graded measures of value judgements, such as comfort, pleasantness and acceptability, are not easily achieved since calibration is difficult and the outcome to a large extent depends on context factors. However, for design and control purposes, a reasonable assumption is that the perceived intensity of odours plays the major role in the generation of odour discomfort.

In all sensory evaluations where reliable data for comparison between buildings, laboratories and investigations are aimed at, calibration is a necessary part. That requires the use of references. Standardisation of methods helps but does not guarantee that meaningful comparisons can be made..

The perceived odour intensity scales should be calibrated by use of a reference scale of a standard odorant. In case another attribute than perceived odour intensity is being used, e.g. odour acceptability, its psychophysical exposure-response data should be compared to perceived odour intensity data. All perceived intensity measures for odour or sensory irritation should be reported also as concentration equivalents of the reference chemical being used.

Some methods require that a dilution series be made of the original air sample and therefore its odour/irritation must be fairly intense to start with. These methods include the

method of limit and the method of constant stimulus. Therefore, they are suitable only for laboratory testing.

The composition of the panel depends on the purpose of the test. This purpose should be displayed. Sometimes the test panel is recommended to comprise selected, sensitive persons, at other times naive subjects are preferred. In laboratory testing there should be at least 15 persons in the panel.

III. Critical analysis of some recently published IAQ assessments using psychological methods

Flooring materials (Section 10.1). For the purpose of a labelling of solid flooring materials, a simplified sensory assessment procedure has been proposed (ECA-IAQ, 1997a). The report accepts almost all possible endpoints of the sensory tests employed but leaves the choice to the authority or body establishing or granting a label: odour detectability, perceived odour intensity, percentage of test panel members dissatisfied, or equivalent quantities. Thus no specific methods are being proposed except for some basic functional requirements. There is no discussion of what are the minimum dimensions needed for sensory testing of materials. Furthermore, acceptance criteria are only given for sensory irritation, not for odour or other attributes.

Since the result of a test should enable consumers to compare the emissions from different materials, or to rank the materials, high demands are put on the sensory methods to be applied: a scientifically sound test design, a calibration procedure, identified inter-individual variation, reproducibility and validity, and quality assurance. These requirements seem reasonable. To meet them, calibration, or at least standardisation, will be needed on instrumental as well as inter-laboratory levels.

Required ventilation rate (Section 10.2). For the purpose of calculating the required ventilation rate in a space, a sensory assessment procedure has been proposed based on predicted dissatisfaction and load calculations in terms of person equivalents (ECA-IAQ, 1992). No sensory evaluation technique is being specified in the report although two examples are given in an appendix (binary classification of acceptability using untrained panel and magnitude estimation of perceived air quality (in decipol) with trained panel using several references).

The main advantage with the report is that it emphasises the importance of other contributions to the sensory load on the indoor air than just occupants. However, the guidelines only consider one endpoint, i.e., dissatisfaction. The claimed attribute is perceived air quality but no proof of the validity is presented.

In the methods described the test subjects are faced with the dual task to judge the acceptability of a presented stimulus and at the same time relate it to an imaginary context situation. However, judgements of acceptability are likely to have a higher inter-individual variability compared to judgements of perceived intensity. Therefore, "noise" is expected in the data which will diminish the resolution power of the methods described. On the other hand, little systematic bias is expected to occur.

Three categories of air quality are suggested by the ECA-IAQ (1992) but the criteria

fulfilment are not fully specified, e.g., as to accepted variance, although some statistical considerations are presented. It seems that there is too small a difference between the proposed three categories when considering the resolution power of the method. Higher percentages of dissatisfied occupants have been demonstrated in real life situations.

Questionnaire field studies of the sick building syndrome (Section 10.3). A thorough review and analysis of field studies of the sick building syndrome ("SBS") has been made by the collaborative action between CIB (International Council for Building Research Studies and Documentation) and ECA-IAQ (Berglund et al., 1996). In most non-experimental and experimental "SBS" field studies, structured self-administered questionnaires have been used to collect information on the experience of annoyance and symptoms. It is difficult to compare the results between different studies since there are considerable differences in the questions about the type, quantity, intensity and time reference of the symptoms. However, in most cases the questions have been consistent within each experimental study.

The common "SBS" measurements are on rank order scales and are not adequate for making precise economic calculations. Comparison between buildings on different occasions and with different, small groups of respondents would require procedures for calibration or standardisation of such scales. They do not exist at present.

A major threat to the validity of most interventions studies of "SBS" is information bias, which may take place if the intervention is not blinded to the participants. Another source of bias is the tendency of the participants entered in the study to be lost in the follow-up; the probability of loss may be related to the effect of treatment.

IV. Recommended methods for measurements of perceived air quality

Recommended methods for materials and compounds testing (Section 11.2). These methods should be used only for identifying those materials which are unacceptable to the test panel in the laboratory testing, and for calculating approximate indices of sensory effects, assuming that low material emissions are better than high.

For materials and compounds testing the sensory evaluation is recommended to include detectability and perceived intensity (or an attribute closely linked to perceived intensity) of odour and/or sensory irritation. In addition a yes-no classification should be made of whether the sample is perceived unpleasant or not, at conditions simulated to be typical for the intended use of the material. The following methods are recommended by the group (see also Table 1).

A. When high inter-laboratory comparability, resolution power and precision are required (as for the purpose of labelling materials and products):

- For the detection of odours
  1. Method of limit
  2. Method of constant stimulus

(A combination of the two methods may be preferred in some cases. Detection measures also can be obtained by jointly measuring detectability and perceived intensity.)

Table 1. Features of recommended methods for sensory evaluation of indoor air quality

Features/ Methods	Method of Limit	Method of const.stimul.	Signal detection	Equal attribute matching	Category scaling	Magnitude estimation				Descriptor profiling	Classification (yes/no)
						Memory reference	One reference	Several references	"Master scaling"		
						The concept of magnitude estimation is well established					
Theory	Well established	Well established	Very well established	Less well established but similarity empirically verified	Not well established but postulated; empirically shown to work at times	Risk of high variances; empirically shown to work	Risk of systematic bias depending on location of standard	Well established for specific applications but not empirically verified for IAQ measurements	No additional theory but empirically shown to work	Less well IAQ established in the context of	Less well established in the context of IAQ
Equipment	Advanced equipment is required (olfactometer)	Advanced equipment is required (olfactometer)	Advanced equipment is required (olfactometer)	Some cross-modal matchings require little, e.g. visual analog scales; advanced equipm. in intra-modal matching.	Few requirements	Few requirements in testing but semi-advanced e quipment required for training	Few require-ments	Semi-Advanced equipment is required	Advanced equipment is required	Few requirements ments	Few require-
Subjects	Small-medium panel size. No training required, Selection as to sensitivity	Small-medium panel size. No training required, Selection as to sensitivity	Medium-large panel size. No training required.	Small-medium panel size. No training required; selection as to adherence to performance criteria; should belong to assumed target population	Small-large panel size. No training required. Selection as to representativ eness	Small-medium size panel. Training required. Selection as to adherence to performance criteria; should belong to assumed target population5	Small-large. panel size. No training required	Small-medium size panel. Sometimes training to anchors required, e.g., decipol method. Selection as to adherence to performance criteria. Should belong to assumed target population	Small-medium size panel. No training is allowed. Few requirements but panel must belong to target population	Small-large panel size. Typically no training. Selection as to representativen.	Small-large panel size. Typically no training. Selection as to representativen.
Existing for application to IAQ	For single compounds & a few building investigations	For single compounds only	For single compounds only	Visual analog scales in wide use. Few other studies done.	In wide use	In some use earlier but not now	Not common	In wide use in chamber studies	In limited use	Widely used in surveys & chamber studies. Few on odours	Widely used in surveys & in earlier sensory studies
Calibration	Postulated. Both instrumental & interlaboratory comparisons have been made	Postulated. Both instrumental & inter-laboratory comparisons have been made	None	Postulated. No inter-laboratory comparison made	Typically postulated. In a few investigations calibrated after Thurstonian scaling	Some response calibration have been made	Some response calibration have been made	In some cases postulated through training. For ratio estimation no calibration; possible to reference physical scale	Individual scale calibration possible by transforming to scale of reference stimuli	No calibration	Typically postulated
Level of measure-ment	Nominal	Nominal	Interval (d')	Interval-ratio	Ordinal - interval	Interval - ratio	Interval - ratio	Interval - ratio	Ratio	Nom. -ordinal - interval-ratio	Nominal
Recommend. application	Materials testing	Materials testing	Materials test. IAQ evaluation	Materials testing	Materials testing IAQ evaluat. Popul. response	IAQ evaluation	IAQ evaluation Popul. response	Materials testing IAQ evaluat. Popul. response	Materials test. IAQ evaluat.	IAQ evaluation Popul. response	Materials testing IAQ evaluation Popul. response

3. Signal detection index, d´

   (For checking the odour threshold concentration, as measured or arrived at by extrapolation.)

- For the determination of perceived intensity of odour and/or sensory irritation:
  1. Equal-attribute matching
  2. Magnitude estimation with several references
  3. Master scaling
- For the judgment of unpleasantness:

  Classification (yes/no)

  (For checking whether the perceived air of the sample is deemed unpleasant by a majority of the panel members, or not)

B. When high intra-laboratory comparability and resolution power are required (as for the purpose of guiding company product development):

1. All of the methods above
2. Category scaling
3. Methods for multidimensional analysis

Recommended methods for IAQ evaluations (Section 11.3).The sensory evaluation is recommended to include detectability and/or perceived intensity (or an attribute closely linked to perceived intensity) of odour and/or sensory irritation. The following methods are recommended by the group (Table 1).

A. When high inter-investigation comparability, resolution power and precision are required (as for the purpose of auditing buildings):

- For the detection of odours:

  Signal detection index, d´
- For the determination of perceived intensity of odour and/or sensory irritation
  1. Magnitude estimation with several references
  2. Master scaling

B. When moderate intra-investigation comparability and resolution power are required (as for the purpose of comparisons of sources and interventions within a building):

1. All of the methods above
2. Category scaling
3. Magnitude estimation with memory reference
4. Magnitude estimation with one reference
5. Descriptor profiling
6. Classification (yes-no)

Recommended methods for population response studies in buildings (Section 11.4). The choice of procedure is likely to depend on the size of the building being studied and the amount of time available for the study. The scope is greater for large buildings and long studies .

The evaluation can include measures of environment and body perceptions (symptoms) as well as of perceived quality and value judgements. When perceived intensities and value judgements are being measured calibration should be strived for. Thereby, physical as well as memory references may be used, within the same modality or with a different one. If calibration cannot be made, the scales at least should be standardised by transforming the empirical response distributions to a standard distribution of z-scores. The following methods are recommended by the group (Table 1).

A. When high inter-investigation comparability, resolution power and precision are required (as for the purpose of *comparing buildings or occupant groups*):

- For the determination of perceived intensity of odour and/or sensory irritation
  1. Magnitude estimation with several references
  2. Category scaling (calibration needed, e.g. after Thurstonian scaling)

B. When moderate intra-investigation comparability and resolution power are required (as for the purpose of *comparing systems and interventions within a building*):

1. All of the methods above
2. Category scaling
   (At least standardisation is needed)
3. Magnitude estimation with one reference
4. Descriptor profiling
   (A thoroughly developed measuring instrument is needed)
5. Classification (yes-no)

Magnitude estimation scales of any attribute should be calibrated by use of a reference scale. The reference does not have to belong to the same modality.

In case another attribute than perceived intensity is being used, e.g. perceived acceptability, its psychophysical exposure-response data should be compared to perceived intensity data achieved in an additional experiment. The purpose is to allow for a comparison between the attributes. All perceived odour intensity measures should be reported also as concentration equivalents of the reference odorant chemical being used.

The control group of subjects should be selected in the same building if possible, double blinding be made whenever possible and a short recall period be used in order to maximise recall accuracy.

1. BACKGROUND

The perceptual or comfort aspects are essential for a healthy building. These involve a complex of sensations. The most prominent are thermal sensations (e.g. thermal comfort), air quality sensations (e.g. odour and "freshness"), visual sensations (e.g. lighting and visual outlooks), and auditory sensations (e.g. sound and room acoustics). Other important complex psychological aspects are related to sensory effects: perceived personal space (crowding), perceived physical space (closeness or openness), and aesthetic quality of building and room design. In the following the focus is on perceived indoor air quality.

Human subjects are indispensable in the measurement of indoor air quality and other environmental factors, e.g., of odour and sensory mucosal irritation. Since chemical and physical methods of characterisation often are insensitive to odorous or sensory irritating air pollutants, or do not integrate in a biologically meaningful way, sensory methods many times are the only or preferred tools for evaluation.

Throughout this report, the term "sensory evaluation" refers to the assessment of perceived environmental quality, specifically indoor air quality, by human subjects and their senses of olfaction and somesthesia (the common chemical sense). The methods applied are generally psychological and taken from the subdiscipline of psychophysics.

Sensory evaluation of indoor air quality may be used, i.a., for:

- study of the impact of physical factors on perceived air quality and symptoms;
- investigation of exposure and response relationships;
- assessment of the effect of remedial actions regarding determinants as well as outcomes;
- evaluation of indoor air quality in new or refurbished buildings;
- operational control of buildings;
- identification and quantification of pollution sources in buildings;
- evaluation of emissions from building materials, furnishings etc.;
- development of a testing system for building materials.

The practical application of a sensory evaluation of environmental quality will require that the investigator has acquired basic laboratory skills in psychology, chemistry and/or laboratory engineering. The purpose of the evaluation should be clear to the investigator as well as to the interpreter of the study results. The purpose of the investigation determines which effect measure should be used (odour, symptom pattern, acute or late effects, etc.), the particular study design and the choice of measurement method. It also sets the level of sophistication required with respect to, i.a., type of scale, necessary resolution, calibration needs, view on inter-individual variation, reproducibility, validity, predictive ability and quality assurance means.

2. SENSORY MECHANISMS AND RESPONSES

The sensory systems in humans are largely designed along identical principles. The important difference lies in the interface between the external and the internal environment which, in the present context, refers to mucosae and skin. All human response to chemical stimulation results from an interaction between external molecules and sensory receptors, which are proteins. This implies that the response to chemical stimulation depends equally on physico-chemical stimulus and receptor properties as well as on personal factors.

Among the human senses, somesthesia ("the common chemical sense") and olfaction are directly involved in the perception of environmental quality, including indoor air quality. Sensory mucosal irritation is associated with somesthesia and odour with olfaction. Somesthetic sensations are mediated through, i.a., the trigeminal nerve, the free nerve endings of which are located all over the nasal, oral and ocular mucosae.

The nose is a paired organ. Behind each nostril, the olfactory epithelium is located at the top of the nasal cavity. Two types of nerve structures, the olfactory receptors and the free endings of the trigeminal nerve, are embedded in this tissue.

The raw sensory information delivered by the receptors is the first link of the sensory system (or "chain"). The total quantitative and qualitative information available for stimulus perception, is encoded into a robust impulse frequency code.

The remainder of the sensory chain can be divided in two parts. The first part which is more peripheral can be viewed as a serial, on-line, wind logic which filters. This part of the chain reshapes the raw information, ending with a stabilised, reduced and well discriminable sensory image, i.e., a perception.

The second, central part of the sensory chain, processes this sensory image in a parallel, programmed logic, including memory, learning, emotion and other incoming sensory messages. The result is a final integrated, perceptual information towards which consciousness and behavioural responses take place.

2.1 Somesthesia

Over the entire facial and forehead skin and the mucous membranes of the nasal and mouth cavities free endings of nerve fibres are distributed which together form the trigeminal nerve. This somesthetic or "common chemical" sense mediates sensations of warmth, cold, pain and pungency.

Trigeminal and olfactory sensations are typically activated together at different proportions, such that in most environmental situations involving chemical stimulation both the olfactory and the trigeminal systems produce blended sensations. However, the two types of sensation can be separated by adequate stimulation techniques as well as in certain disease states (e.g. Kallmann syndrome). Furthermore, the time-course of stimulation and of adaptation is probably different for the olfactory and the trigeminal systems (Cain, 1992).

In order to determine structural determinants of potency, measurements of odour thresholds in subjects with normal olfaction and nasal irritation have been compared with thresholds in anosmics for single compounds (Cain and Cometto-Muniz, 1993). A positive relationship has been shown between carbon chain-length and potency within classes of chemicals (primarily alcohols), although this structure-activity-association does not necessarily help in distinguishing between chemicals.

Interactions in terms of olfaction and irritation take place between components within mixtures as illustrated by observations made for three-component mixtures (Cometto-Muniz and Cain, 1995; Cometto-Muniz, Cain and Hudnell, 1997). Detection thresholds for odour, nasal pungency and eye irritation were measured for single and mixed VOCs, while nasal pungency was measured in anosmic subjects in order to avoid confounding by odour sensations. Various degrees of additivity were found, such that for the mixture to be perceived the concentrations of the individual components making up the mixture were below or markedly below their individual thresholds. Such additivity increased with increasing complexity of the mixture. This observation suggests that chemosensory detection, including irritation in complex chemical environments, is possible even if single VOCs are well below their individual thresholds.

2.2 Olfaction

Of the human special senses olfaction has the longest evolutionary history. The senses of olfaction and taste are viewed as chemical senses because chemical compounds interacting with receptor molecules are the stimulus prerequisites for eliciting the sensations (for a review, see e.g. Doty, 1995). Odorous properties of volatile molecules are linked to the formation of reversible, low energy bindings with protein receptors. The specificity of these receptor bindings depends on the actual topography of the receptor site which is still unknown. The binding energy is accounted for by van der Waals forces, hydrogen bonds and hydrophobic binding.

The peripheral processes involve stimulation of the olfactory cells. The olfactory epithelium, comprising a total area of about 4 cm^2, contains about 10 to 30 million receptor cells, which terminate in knobs with about ten cilia or microvilli. The microvilli form a network covering the mucosa. It is believed that the olfactory receptor sites are located in the ciliary surface membranes. The total ciliary membranes develop a hundred fold actual surface as compared to the neuroepithelium (i.e. about 400 cm^2 of receptive membrane).

There are around 1000 genes that encode altogether about 1000 odour receptors (Axel, 1995). Genes provide the template for proteins, the molecules that carry out the functions of the cell. Each type of receptor is expressed in thousands of neurons. Since mammals can detect at least 10 000 odours, each of the 1000 different receptors must be able to respond to several odour molecules, and each odour must bind to several receptors. The results shown by Axel (1995) suggest that each neurone features only one type of receptor. The problem of distinguishing which receptor was activated by a particular odour is then reduced to the problem of identifying which neurone fired.

The olfactory epithelium is roughly divided into four zones according to the types of receptors present. Within each zone receptors are randomly distributed, such that no precise spatial pattern of neurons in the epithelium can be found.

So far, no specific receptors have been identified. It is believed that about 100 to 300 receptor classes exist. Each sensor cell is more or less sensitive to specific odorants and, therefore, a large number of combinations are possible. It is not yet possible to predict odour sensations from the chemical structure of an odorant and to establish an objective classifications system for odorants.

As previously said, the elementary processes underlying olfactory transduction are not unlike those of other receptor-mediated neuronal communication processes. Odorant molecules bind to a protein receptor site in the membrane of the receptor cell, the receptor activated by the stimulus in turn activates other proteins which trigger an enzyme cascade (second messenger system). Finally this results in phosphorylation of channel proteins which in turn may open ion channels into the receptor cell. Through these channels positively charged kations (particularly sodium) will enter which produces a depolarisation of the receptor cell and an electric potential.

The axons of receptor cells form bundles. Between 30 and 50 such bundles carry olfactory information to the olfactory bulbs. Here the olfactory fibres synapse with the dendrites of mitral and tufted cells. Several hundreds of primary olfactory axons converge on a single mitral cell. It has been shown by antibody labelling that all axons converging in a given synapse have originated in receptor cells expressing the same receptor protein. Any given odour is therefore quality-coded by specific spatial pattern of synapse activation.

The olfactory bulb structure is characterised by the presence of numerous inhibitory interneurones providing both synapses and mitral cells with strong lateral inhibition. The overall effect is a powerful contrast enhancement of the spatial pattern of bulbar activation, resulting in an all-or-none "olfactory image" made of several well delineated patches. After a further contrast enhancing through the piriform cortex of the brain, the olfactory image is then ready for being processed by higher brain centres in the same way as all other sensory images.

Olfactory adaptation is similar to visual adaptation to light: within a few minutes of continuous exposure a dynamic equilibrium is reached between active and desensitised receptors. The result is a reduced, but not suppressed, response; odour thresholds increase with time of constant exposure and within a few minutes perceived odour intensity can drop to 1/3 of its initial value (Ekman et al., 1967; Berglund, Berglund and Lindvall, 1978); recovery, or readaptation, occurs within less than a minute after removal from odours (VDI, 1986). There is also a centrifugal inhibition mediated by the inhibitory interneurones, originating in higher brain centres, that may completely suppress the odour response. In contrast to adaptation also slight facilitation effects have been demonstrated for combinations of weak singular odours: for odour detectability (Corbit and Engen, 1971) as well as for perceived odour intensity (Berglund, Berglund and Lindvall, 1978).

For odours as for other sensory stimuli, perceived odour intensity increases as a power function of concentration (Stevens, 1957). The general equation of the psychophysical relationship for odour intensity is

$$R = c\,(S-S_0)^n$$

where R denotes perceived intensity, S stimulus concentration and S_0, c and n are constants. S_0 is related to the absolute threshold, c adjusts for differences in unit of measurement, and the exponent n is a descriptor of the input-output characteristics of the olfactory system.

The exposure-response power function still holds if stimulus concentration (S) is expressed as multiples of threshold (Svensson and Lindvall, 1974). For nearly all odours studied, the exponent n is less than one which means that the olfactory system would attenuate the stimulus information more at high concentrations (Patte, Etcheto and Laffort, 1975).

In accordance with Weber's law, the ratio (Weber fraction) between the just noticeable difference and the intensity of the stimulus is assumed to be a constant and be characteristic for each sense modality. Persons can reliably resolve differences in odorant concentrations in the range of about 10 to 50% (Wenzel, 1949; Stone and Bosley, 1965; Cain, 1977); the average Weber fraction being around 30%. Thus, the sense of smell is less keen in discrimination than any other sense. For vision, audition and touch the Weber fraction is less than 10%.

A majority of the indoor volatile air contaminants are odorous. Characteristic perceived odour patterns of indoor air can be identified and described with methods of pattern recognition such as principal component analysis (Berglund et al., 1982). However, the "odour print" of an air sample from a building is typically different from its "chemical print". The main reason is that the growth functions for perception are different among the air contaminants. Furthermore, the odour interaction for mixtures of constituent odorants is governed by a strongly attenuating function, hypoaddition (Berglund, Berglund and Lindvall, 1976). For example, the perceived odour strength of a mixture of five odorants, each of equal perceived odour intensity, does not exceed that of the single component odorant by more than 10% (Berglund, 1974). A vector model describes this process well (Berglund et al., 1973). In fact, the overall perceived odour intensity of an indoor air sample, which constituents typically are large in number and low in concentration, can be predicted simply from the number of components most frequently reported to have a "strong" odour (Berglund et al., 1982).

Air is usually perceived as an entity and humans cannot pinpoint the particular constituent odours in an air sample by the sense of smell alone. Results of pattern recognition analysis of indoor air samples (Berglund et al., 1982) point out the joint impact of the large number of chemical and sensory components for the qualitative character of indoor air quality. It is probable that the chemical senses perform a pattern analysis on exposure to complex air pollution (Berglund and Lindvall, 1986). The concentration of numerous compounds may be less important to the perceived indoor air quality than the addition or subtraction of a few specific compounds to the gas mixture (Berglund and Lindvall, 1990). Pattern recognition is particularly important in perception when the signal-to-noise concentration ratio for critical compounds is near to one. This is the predominant case for the majority of compounds which commonly appear at low concentrations in indoor air.

It is not fully understood to what extent odour perception is influenced by the sniffing behaviour of the subject. The odour area of the nasal cavity is poorly ventilated and most of the inhaled air passes through the lower part of the cavity. But the configuration of the nasal cavity creates turbulent air flows when sniffing and the air stream is directed upwards by increasing the flow rate of the inhaled air (Berglund and Lindvall, 1982). This is also believed to increase the number of molecules available to the olfactory receptors.

2.3 Systemic Responses

Since long annoyance has been studied as an adverse psychological response to community noise and odours. Annoyance-responses are typically viewed as endpoints in a chain of events starting with exposure-based sensations. But, they may alternatively be treated as

mediating the reporting of symptoms (or complaints) in such a manner, that sensation-based annoyance is a necessary (although not sufficient) determinant of symptom reporting. This simplified assumption is depicted in Figure 1 (from Cavalini et al., 1991).

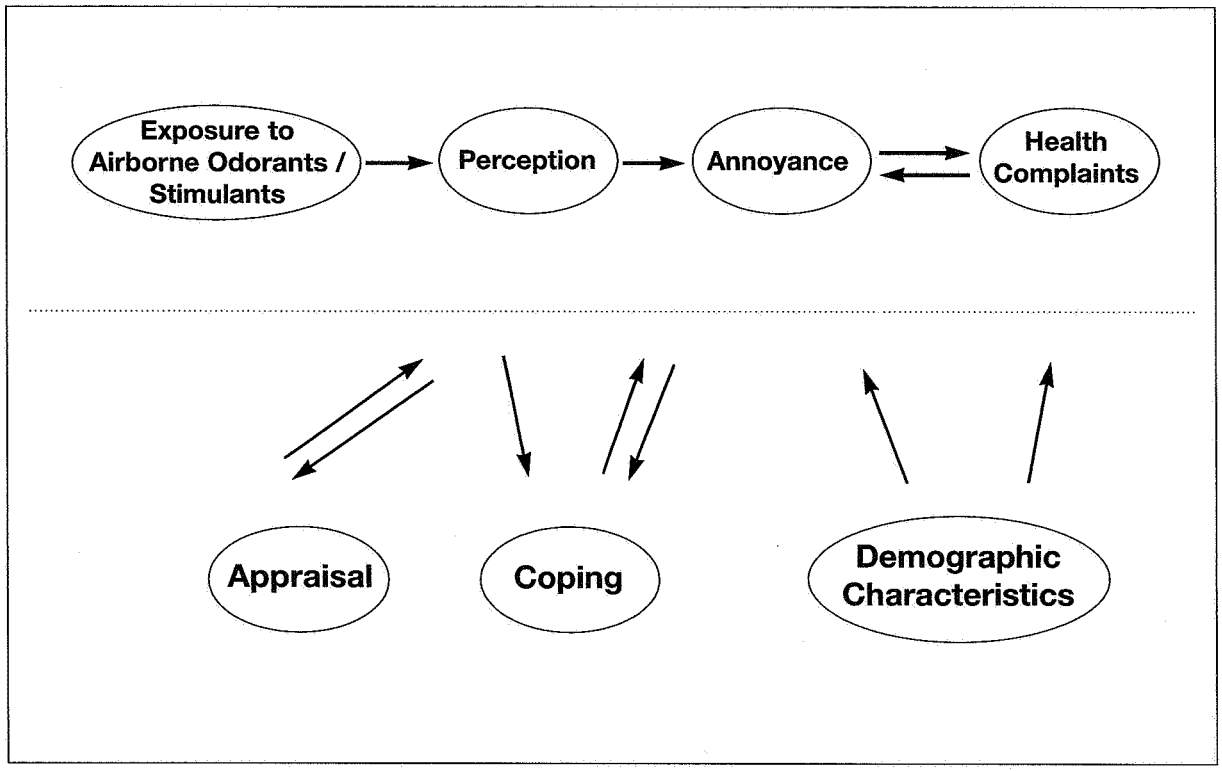

***Figure 1**: Schematic modelling of the relationship between odorant exposure and health complaints as mediated through annoyance, and the assumed role of appraisal- and coping processes (from Cavalini, 1992)*

Although in this figure "reverse causality" of symptoms to affect annoyance in principle is accepted as a possibility, to date research on the relation odour-annoyance only supports the directional sequence from annoyance to symptoms/complaints (Cavalini, 1992). Additional aspects related to appraisal and coping processes have been shown to be relevant for response modification: within a transactional stress-model (Lazarus and Launier, 1978) or based on socio-demographic characteristics (Steinheider and Winneke, 1993).

2.3.1 Annoyance

The annoyance response is the most frequently studied adverse outcome describing negative responses to sensory stimulation resulting from community noise or odour exposure. Although frequently studied in an environmental context and frequently used for regulatory purposes its place in a theoretical framework is still poorly defined. First, it is used for both acute and chronic exposure scenarios although, for regulatory purposes, it is typically related to chronic situations. Second, no systematic efforts have been made to compare annoyance responses to community noise or odour exposure in terms of

similarities and differences. Consequently definitions of annoyance vary (Koelega, 1987).

Clark (1984) describes annoyance in terms of three components: (1) an emotional component (e.g. a feeling of anger), (2) an interference component (e.g. disturbance of desired activities), and (3) a somatic component (e.g. headache, nausea). A similar three-dimensional concept of annoyance, covering both odour and community noise annoyance, was developed by Kastka (1976). The three dimensions of annoyance as identified by means of factor analysis include a perceptual-sensory, a socio-emotional and a somatic component. A questionnaire developed along these lines was found to exhibit satisfactory test-retest reliability (Kastka et al., 1986) and to differentiate between different sources (Winneke and Kastka, 1987). Since the three dimensions have been found to correlate well, simple one-dimensional annoyance scales typically have been used in most subsequent field studies.

> Annoyance has been described as a feeling of displeasure associated with agents or events believed to be detrimental to individuals or groups of individuals (Lindvall and Radford 1973). This definition emphasises the attribution of potentially damaging qualities as being essential for triggering annoyance responses; it is the cognitive evaluation which produces negative feelings.
>
> Another annoyance concept emphasises the emotional component as being dominant (Russell and Pratt 1980; Craik 1987). Russell and Pratt analysed verbal responses of individuals to describe environmental qualities and, by means of factor analysis, came up with a perceptual-cognitive and an affective component of annoyance. The affective component was found to be a two-dimensional, bipolar plane. The first dimension was one of activation with the two endpoints of "sleep" and "arousal", whereas the 2nd dimension was a pleasantness-unpleasantness-continuum emphasising the evaluation-component. According to Craik (1987) annoyance is characterised by an arousing, unpleasant quality, such that both these continua are sufficient for a comprehensive definition of annoyance; the perceptual-cognitive is neglected in defining the annoyance concept.
>
> According to Guski (1987), who has been dealing with noise-annoyance alone, annoyance may be defined as a negative evaluation following the conscious experience of interference with desired activities (reading, conversation, relaxation etc.).

2.3.2 Somatic symptoms and socio-behavioural complaints

The prevalence of general somatic symptoms (e.g. headache, fatigue, muscular pain, sleep disorders) in a normal population is relatively high (Kroenke and Price, 1993) and the vast majority of them remain unexplained (Kroenke and Mangelsdorf, 1989, cited after Wessely, 1997). The term "medically unexplained symptoms" is, therefore, receiving increased attention to describe non specific syndromes (Wessely, 1997). As yet much of the symptomatology of the so called "environmental syndromes" (as e.g. Sick Building Syndrome, Chronic Fatigue Syndrome, Multiple Chemical Sensitivity, amalgamism, electro-sensitivity) remains unexplained in biomedical or physico-chemical terms. Mechanisms of attribution as well as psychological factors may be needed as supplement to environmental stimulation in order to explain the occurrence of "environmental syndromes".

Both socio-behavioural complaints (e.g., "can't open the windows", "have to sleep with my windows shut", "can't relax in my backyard") and symptoms (e.g. nausea, vomiting, loss of

appetite, impaired breathing, sleep disorders) are reported by subjects living in odour problem areas (National Academy of Sciences, 1979; Winneke and Kastka, 1987; Neutra et al., 1991). Except for extreme situations of environmental exposure, symptom reporting is typically mediated through annoyance (Cavalini et al., 1991; Steinheider, 1997). If adjusted for degree of annoyance, exposure-response contingencies for symptoms vanish, whereas the reverse is not true.

2.3.3 Exposure-response relationships

Highly significant correlations between odour exposure and annoyance have been established in the vicinity of different sources. This holds for both predicted odour concentrations in terms of odour units (Cavalini et al., 1991; Cavalini 1992) and frequency of odour events as assessed by subject panels in systematic field inspections (Steinheider et al., 1993), However, the determination coefficients rarely exceed 0.16. For annoyance to community noise such associations typically have been found to be somewhat higher, around 0.25 (Guski, 1987). Much stronger exposure-response associations are established, if population or area means rather than individual responses are taken as the base of measurement. This is essentially due to large interindividual variability of exposure assessments and in responding as a result of person related influences.

> Age and perceived health (reduction) as well as problem-orienting stress-coping (aggravation) were found to modify odour annoyance (Cavalini et al., 1991; Steinheider et al., 1993). Females often exhibit stronger annoyance responses than males. Also annoyance exhibits both state and trait characteristics (Winneke and Neuf, 1992). Subjects preclassified as exhibiting either strong or weak annoyance towards odour or noise in their everyday living environment, show similarly different responses when exposed to either odour, community noise or even cigarette smoke in a controlled laboratory setting. Thus, environmental annoyance also exhibits a substantial degree of generalisation across different environmental situations characterised by chemosensory or physicosensory activation. This observation may have implications for some of the so called "environmental syndromes" (e.g., Multiple Chemical Sensitivity or Sick Building Syndrome).

Both complaints and somatic symptoms exhibit significant associations with outdoor odour exposure but less pronounced than for annoyance (Steinheider et al., 1993). If rank-ordered for strength of association, socio-behavioral complaints (e.g. "cannot open the windows, cannot stay outdoors, cannot invite friends") precede gastric symptoms (disgust, loss of appetite, nausea), sleep disorders (difficulties falling asleep, waking up) and general health-related symptoms (breathing problems, cough).

The main body of knowledge briefly summarised above is based on experience from chemosensory outdoor exposure settings. Information from indoor exposure settings in terms of annoyance and symptom reporting as related to markers of exposure is less developed. For example, low outdoor air flow rate in offices has been shown to be associated with an elevated presence of symptoms in the so-called "sick building syndrome" ("SBS") (Sundell 1994). The perception of "dry air" indoors has been shown to be associated with "SBS" symptoms but not with physical air humidity (Sundell and Lindvall 1993). So far no consistent associations have been established between measures of indoor air quality and reported well-being or symptoms.

2.4 Conclusions

The basic biological principles underlying odour and irritation are fairly well understood with respect to some phenomena: receptor activation, mid-central quality coding and, to a less extent, information processing at higher centres of the brain. Much more uncertain is how these basic processes relate to the more complex psychological responses to odorant/irritant stimulation, such as perceived odour quality, annoyance and symptom reporting. The latter two have typically been studied in population-based studies using questionnaire methods. Such studies have revealed consistent associations between e.g. odour exposure and annoyance as well as symptom-reporting in the outdoor environment. However, such information cannot be directly translated to the indoor environment.

There is no universally accepted definition of annoyance (Koelega, 1987). Sometimes it is seen as a perception, or a feeling, or an attitude, or as a mix of these. Most researchers have apparently studied the same phenomenon, namely a sensory and experience-based negative evaluation of a perceived environmental setting, which may be measured by attitude-scales in population-based studies. Annoyance-data collected in such studies have been shown to exhibit consistent associations with exposure in the outdoor environment.

3 THEORY OF MEASUREMENT UNDERLYING SENSORY EVALUATIONS

Sensory analysis is based on the use of human subjects as measuring instruments. Therefore, it is essential that the sensory analyst knows the "basic rules" of perceptual measurement. The main importance of subjective measurement is in the development of scales. Scales determine the type of statistical analysis that is adequate for the measurements obtained (Gaito, 1980). Scales derived from sensory evaluations can only be said to truly represent the test subjects'sensations if the scales follow the measurement theory that defines them.

3.1 Measurement Scales

Measurement is a procedure of applying a standard scale to a variable or to a set of values (Last, 1988). In a very direct way measurement is a process of assigning numbers to observations or empirical data in such a way that empirical relations and operations are preserved. However, a scale may not only be a set of numbers but the total "picture" of what has happened: observations, the relationships between them, a transformation of the observations to numerical data that maintains the original relationships and, finally, the numbers obtained under these conditions.

Measurement scales may be characterised with regard to the type of measurement approach (level of measurement, i.e., the mathematical level at which measurements are being made), which will put restrictions on potential data analysis (e.g., Torgerson, 1958). The most common data are in the form of a nominal scale only (presence/no presence of an occurrence or a set of qualitatively different classes such as different types of symptoms). Sometimes the responses are assessed on ordinal (rank order), interval (distance is constant, relative to a zero-point), or ratio scales (ratios are constant, absolute zero point). Physical-chemical measurements are usually interval or ratio scales in contrast to the questionnaire data with self reports that are commonly on nominal or ordinal scales. It seems that there is a great need for improving the level of measurement in sensory evaluations. This is particularly true for applications in intervention studies where it is desirable to draw more refined conclusions. The allowable operations in each scale and the permissible transformations are summarised in Table 2.

Table 2. *Allowable operations and permissible transformations for various scale types.*

Scale Type	Permissible Transformations	Allowable Operations
Ratio	$h(x) = ax$, $a>0$	Equality, Order, Distances and Ratios
Interval	$h(x) = ax + b$, $a>0$	Equality, Order and Distances
Ordinal	$x \leq y$ if $h(x) \leq h(y)$	Equality and Order
Nominal	Any one to one correspondence	Equality

A nominal scale is one in which you can only speak about two elements being equal or unequal. If they are equal they belong to the same set, if they are unequal they belong to different sets. There is only one allowable operation within a nominal scale: each number

can be transformed to only one other number and vice versa. For example, a yes/no response in a detection task generates a nominal scale. It does not make sense to add, subtract, multiply or divide the numbers assigned to these categories.

The scaling of air quality can be an ordinal scale if it can be unique up to order. For example, it can assign a one to "unhealthy" air, a two to "unsatisfactory", a three to "acceptable", a four to "good" air and a five to "excellent " air (Roberts, 1976). But other numbers could be assigned to each category of air quality, as well, as long as you preserve the order. Another five numbers that could be assigned to the five categories of air quality could be e.g.: -1, 0, 5, 14, 20. It does not make sense to add, subtract, multiply or divide the numbers assigned to these orders.

The scaling of the degree of perceived air quality of a sample can be an interval or a ratio scale depending on the method used to generate the results. It makes sense to add and subtract the numbers from an interval scale (compare the Celsius scale, or an equal intensity visual analog scale for sensory evaluation). Within the ratio scale it makes sense to add, subtract, multiply and divide the numbers obtained (compare the Kelvin scale, or the Master Scale for odour evaluations).The grading of scales may have different resolutions for different types of measuring instruments. For perceptual measurements the resolution is related to the discrimination capability along the scale. Older people have lost some of their discriminatory power for sensory stimulation and therefore the resolution of response scales obtained from old persons may be less than that for young persons.

The procedure for calibrating scales is tied to the level of measurement. Calibration of scales is performed in order to rectify the grading of quantitative measurements obtained. Then the act of measuring often is performed by comparison with one or several references. Ordinal scales are dubious to calibrate. Interval scales may be calibrated with regard to a distance between two scale values and ratio scales may be calibrated with regard to one scale value only. Calibration procedures are imperative in chemical and physical measurements but are relatively new to scales in sensory evaluations. Further information on calibration is given in chapter 5.

3.2 Measurement Theory

The meaning of measurement scales is based on measurement theory. The theory offers a frame of reference and its application guarantees the validity of the subjective data, i.e., if one follows the "rules", the test subjects´responses will respresent their sensations. The theory also guides in the choice of the adequate adequate mathematical analyses of the data.

Measurement is a process of assigning numbers to observations or empirical data in such a way that empirical relations and operations are preserved. More formally, the starting point is an empirical relational system (ERS) that includes a set of empirical data and the relations defined on the set. There exists a mathematical application (f) that assigns a number to each element of the empirical set in such a way that all operations defined in the empirical set can be "reproduced" in the numerical set. In this way a numerical relational system (NRS) can be obtained, whose components are numbers corresponding to data and the "reflected" relationships between them. So, fundamental measurement is the assignment of a mathematical application from an observed or empirical relational system to some specified numerical relational system. For example, a subject says that the air in room A is more polluted than the one in room B. The measure of this observation

can be quantified by whatever numbers the subject chooses provided the condition is fulfilled that the number assigned to the air pollution in room A is larger than the number assigned to the air pollution in room B. That is, the numbers that represent the empirical fact have to maintain the order relationship of the empirical set.

The above mentioned measurement definition leads to the representation theorem. This theorem provides for the assignment of numbers to the empirical observations. Under some conditions, or axioms, empirical phenomena can be represented by numbers in the positive real set, $\Re$. The mathematical application gives the representation.

In sensory evaluation the importance of building a good scale is evident. It should minimise the subject's response bias and the empirical setting should consider previous knowledge of physical and chemical factors that govern the sensory variable to be scaled.

Once the sensory variable to be measured is determined, what components does a scale have? A scale may be defined as the set {ERS, f, NRS}. So, underlining what has already been said, a scale is not only a set of numbers but the total "picture" of what has happened: observations, the relationships between them, a transformation of the observations to numerical data that maintains the original relationships and, finally, the numbers obtained under these conditions. Sometimes researchers consider the scale as the measuring instrument per se (a questionnaire, a test, etc.), or a function that associates the elements in the ERS to the ones in the NRS (for example, the power function in psychophysical scaling) or only the numbers obtained (for example, magnitude estimates of odour intensities). All of those considerations are incomplete. Numbers or measurements, functions or rules lie within a scale but the scale is defined as a bigger set. A sensory scale is then defined as the set {ERS, f, NRS} where:

$$ERS = \{S, S(1), S(2),..., S(n)\}$$

$$NRS = \{QR, QR(1), QR(2),..., QR(n)\}$$

f is a mathematical application from ERS to NRS

ERS = An empirical relational system

NRS = A numerical relational system

S = Sensations

S(i) = Relations between sensations, i = 1,...,n

QR = Quantified responses

QR(i) = Relations between quantified responses.

In sensory evaluation, previous to the analysis of sensations, there is a physical relational system (PRS) whose elements are stimuli (ST), and ST(1), ST(2),...,ST(n) are the relationships between stimuli. Another mathematical application, g, goes from PRS to ERS. A complete diagram is shown in Figure 2.

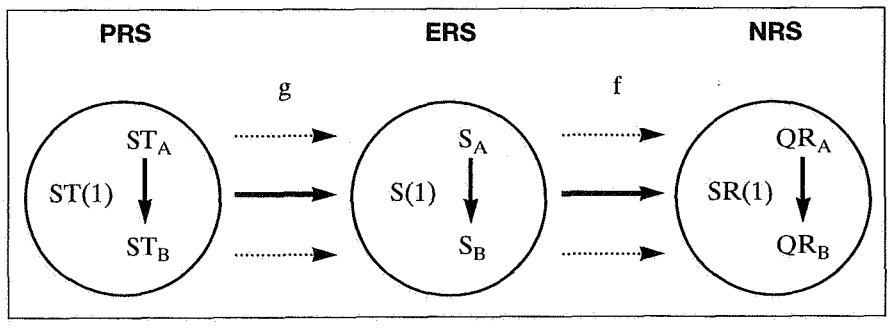

Figure 2. *Total representation of a sensory scale.*

Let us interpret the diagram of Figure 2 in an odour detection task considering the absolute threshold ST_A of a pyridine solution of 1.5 log ppb. ST_A can be related to other concentrations of pyridine, let us say to $ST_B = 2$ log ppb. So, $ST_A < ST_B$. ST_A, ST_B and their relationship ST(1) belong to PRS. ST_A produces a sensation S_A and ST_B produces another one, S_B. Both are interrelated by S(1), a "similar" relationship to less than (<). These elements belong to the ERS. S_A and S_B produce a quantified response in NRS, QR_A and QR_B. These responses are such that if the answer to the question "Do you smell something?" (presenting substance B) is yes, that would imply that $QR_B > QR_A$ or $QR_A < QR_B$. These responses and their relation, SR(1), belong to NRS. Since the relation between responses is an image of the relation between sensations we will have an olfactory sensory scale for the detection task proposed. Considering the first set, PSR, we can also say that responses are a re-scaling of the physical domain.

If we consider just the representation theorem, many numbers can be assigned to the same empirical observation still maintaining or "mirroring" the relationship between the elements of ERS and NRS. If there are many representations, what is the use of these measures since they are not unique? What is the use of giving numbers to sensations if an infinite number of them could fit the requirement of "keeping the relations of the ERS"? How unique is the mathematical application f?

If we could find a function that "groups" all the representations we might then say that all representations are the same. The theory then would define for each measurement scale a permissible transformation. This transformation is a function h from $\Re$ to $\Re+$ such that its composition with f also results in a mathematical application. Then, since there is a specific function that relates two or more numerical representations, you can say that all these representations are the same since h[f(ERS)] is a mathematical application. Once you have this function, the representation of the ERS is unique. This is what the uniqueness theorem states.

3.3 Conclusions

Sensory analysis is based on the use of human subjects as measuring instruments. Therefore, it is essential that the sensory analyst knows the "basic rules" of subjective measurement. The main attributes of perception that can be measured by olfaction and the common chemical sense are: detection, discrimination and perceived intensity, and, by adding higher level processing, perceived quality and value judgements.

The level of measurement puts restrictions on the data analysis. Within the sensory field commonly four scale types are considered: nominal, ordinal, interval and ratio. A yes/no response in a detection task generates a nominal scale. Questionnaire data commonly are on nominal or ordinal scales. It does not make sense to add, subtract, multiply or divide the numbers of such scales.

Physical-chemical measurements are commonly interval or ratio scales. Also the scaling of perceived odour/irritation attributes (like intensity or pleasantness) can be an interval or a ratio scale depending on the method used to generate the results. It makes sense to add and subtract the numbers from an interval scale (compare with the Celsius scale). Within the ratio scale it makes sense to add, subtract, multiply and divide the numbers obtained (compare with the Kelvin scale).

There is a great need for improving the level of measurement in sensory evaluations. For that purpose the discipline of psychophysics provides many techniques and testing procedures for obtaining sensory data from human subjects. In order to obtain comparability between sensory evaluations, they have to be calibrated or standardised. For calibration it is necessary to use references. A number of reference chemicals have been suggested for sensory investigations of air quality.

4 SENSORY EVALUATION TECHNIQUES

The attributes of perception that can be measured by the olfactory and somesthetic senses are the same as for all other sensory modalities:

- limit values for absolute detection (sensory threshold),
- limit values for discrimination (minimum difference between two perceptions that makes them distinguishable),
- ordered discrimination between two or more perceptions (e.g., which of two perceptions is more intense?),
- intensity of a defined perception (e.g., odour intensity, sensory irritation intensity),
- quality of a defined perception (e.g., what kind of odour or what sensory modalities mediate the perception of inhaled air).

By perceived quality is here meant the perception of the complex sensory stimulation which as a result gives, for example, a pattern (like a form) or a blend (like a colour). There is little mental evaluation involved in this process. This definition is not the same as the expression "perceived air quality" as being used when assessing decipol with a trained panel.

In addition to the perceptual aspects mentioned above, various value judgements may be assessed. By value judgement is here meant the high order mental evaluation based upon, i.a., perceived intensity and/or perceived quality of the sensory stimulation. The result of the mental evaluation can range from hedonic tone (like fragrance, pleasantness/unpleasantness) to acceptability. Value judgements can be made by using most methods for matching and direct scaling which are otherwise used for perceived intensity measurements.

The discipline of psychophysics provides many techniques for obtaining human sensory data on IAQ. Odour emitting building materials which make occupants complain, may be counteracted by reducing the emission load on the air below the level where odour is no longer detectable. An odour threshold evaluation of the source strength would provide the building manager with the necessary basis for decision. If the manager wants to just lower the odour load to an acceptable level but not further, then a discrimination evaluation would offer the basic information. If, for cost-effectiveness reasons, the building manager is interested in knowing how the perceived intensity of the odour emissions will change with alternative mitigation solutions, a magnitude estimation evaluation would be appropriate. Below the main ways of sensory evaluation of IAQ are described briefly. (For general overviews on the topics, see e.g., ASTM, 1968, 1984; Engen, 1972; Baird and Noma, 1978; Gescheider, 1985; Berglund, Berglund and Lindvall, 1986).

4.1 Detection

In the usual sense there is no fixed odour or irritation threshold of absolute detection for a particular individual or a particular pollutant but rather a gradual transition from total

absence to definitely confirmed sensory detection. Thus, theoretically thresholds do not represent a fixed point but a value on a continuum (Garriga-Trillo, 1985).

Classical threshold theory assumes the existence of a momentary absolute sensory threshold. Methods based on this concept have been used for assessments of outdoor air pollution for some time (CEN 1995; VDI 1986, Guideline 3881; Lindvall, 1970). Some measurements have been made on indoor air, ventilation systems and emmissions from materials (Berglund and Lindvall, 1979; Bluyssen and Walpot, 1993).

Most threshold methods are based on dilution of the air sample until 50% of a panel no longer detects odour. The odour threshold concentration of a gas sample can be given as a mass concentration number (as in the case of a single substance; e.g., Devos et al., 1990) or as a number of dilution factors, or "odour units" (as in the case of a gas mixture with a known or unknown composition; Berglund and Lindvall, 1979; Meilgaard, Civille and Carr, 1991). In neither case do multiples of threshold values express the degree of perceived odour intensity above threshold.

The absolute detection threshold varies widely with chemical substance. This is reflected by the large spread in odour threshold values for single compounds as reported in the literature (Devos et al., 1990). Recognition thresholds (i.e, requiring identification of quality) are generally higher than detection thresholds (detection of the mere presence of a stimulus). Odour threshold values reported in the literature often vary considerably for the same odour substance. The reason is that threshold values are to a large extent defined by threshold measurement procedure, quality of the olfactometer, purity of the chemical substance, sample of subjects, etc. (Lindvall, 1970). For example, if the samples are presented in ascending order the estimated detection limits are generally lower than when the samples are presented in randomised order. If the method of constant stimulus is used, detection limits will be generally higher than for the method of limits.

Threshold values do not give any measure of the perceived intensity above threshold levels, but threshold determinations can provide source strength information for controlling singular odorant/irritant producing materials and the distribution of detectable pollutants in a space.

Whereas detection methods are used for assessing the presence of a subtle perception (e.g., odour), identification methods are used for identifying (or recognising) a particular perception (e.g., odour of formaldehyde). These methods are the same as those for detection, i.e., method of limits and method of constant stimuli. In addition, the choice theory (Luce, Bush and Galanter, 1963) or the signal detection theory (Green and Swets, 1966) may be applied to identification (or recognition) data using the same type of assumption as for the detection data mentioned above.

4.1.1 Method of limits

For odours the method of limits is the most direct method for establishing an absolute sensory threshold and it is based on a well established theory. In its classical form the stimuli are presented in alternating ascending and descending series starting at different

points to avoid having the subject fall into a routine. The subject is required to report whether the sample can be detected or not. A response criterion is established at the start so that the experimenter will know when the presentation series is to be interrupted. The threshold value for each separate test series is defined as a point in-between the last undetected and the first detected points in the stimulus continuum.

Applied for IAQ studies, the method of limit requires an advanced dilution and presentation equipment (olfactometer) (Lindvall, 1970; Berglund and Lindvall, 1979). The method is in wide use for determining odour thresholds for single compounds (Devos et al., 1990) and has been applied in a few building investigations (Lindvall and Berglund, 1979). A small-medium sized panel is required with the subjects mainly selected as to their sensitivity. There is little need for training the subjects. Typically no calibration is made but is postulated. The resulting data appear on nominal scales and, therefore, it is not meaningful to add, subtract, divide or multiply the threshold values.

> In determining odour thresholds there is a risk that adaptation (fatigue) may develop when long diminishing stimulus series are used. To minimise these effects an alternative is to use only an ascending series of stimuli (Lindvall, 1970). Thus Pangborn et al. (1964) in studying the influence of methodology on olfactory response obtained the lowest thresholds with a sequential-up presentation order and the highest thresholds with a sequential-down order.
>
> A further variant of the method of limits is the "up and down" method (Cornsweet, 1962). In this procedure the order of stimulus presentation is changed as soon as the test subject changes his manner of answering (from yes to no, and viceversa). In this way adaptation can be minimised, and a large amount of information can be obtained in a short period of time. However, this method presents difficulties in estimating the magnitude of response error.
>
> All modifications of the method of limits have the common disadvantage that the presentation of stimuli in a monotonic sequence may introduce anticipation and response perseveration; for example, the test subject due to earlier experience in the experiment may persist in a response pattern although the experience itself may have changed. These phenomena can be counteracted, or at least an estimate of their extent be made, by selecting a principle of response indication based on forced choice (Jones, 1956). An additional disadvantage of the method of limits lies in the fact that the determination of the threshold value it is based on observations from only a limited part of the stimulus range used.

4.1.2 Method of constant stimulus

Also this method has a well established theory. The method of constant stimulus (method of frequency) is based on the assumption that the momentary individual threshold value varies from time to time and that this variation has a normal distribution. The stimulus range is selected in discrete intervals so that the expected frequency of positive answers is distributed over this range between 1–99 %. In general the frequency of positive responses is cumulatively normally distributed over a geometric intensity continuum and may be graphically presented by expressing them as z-scores or probits thus transforming the dose-response curve to an approximately straight line. The absolute odour threshold can then be defined as the effect dose corresponding to an arbitrarily selected frequency of positive response, typically 50 %, "effective dose-50, ED-50".

The frequency index can be corrected in various ways for the possibility of correct answers by guessing. The stimulus is usually presented in a random selection of intensities which minimises response perseveration, among other things. In its classical form the method of constant stimulus requires careful planning, and preliminary experiments are frequently necessary. In IAQ studies the method requires an advanced equipment (olfactometer) and takes long time to conduct. It is in limited use and only for single compounds. As with the methods of limit, the test panel can be small- to medium-sized. The subjects may be mainly selected as to their sensitivity and no training is required.

Typically no calibration is made but is postulated. Both instrumental and interlaboratory comparisons have been made. The level of measurement is nominal and it follows that addition, subtraction, division and multiplication of the thresholds is discouraged.

A practical variant of an ascending method of limit with paired comparison and frequency analysis similar to that in the method of constant stimulus has been used extensively in outdoor and indoor investigations of environmental odours (Lindvall, 1970; Berglund, Berglund and Lindvall, 1987).

4.1.3 Signal detection

The theory of signal detectability is very well established (Tanner and Swets, 1954; Green and Swets, 1966; Engen, 1972). It claims that no absolute threshold exist. While the classical view stresses the relationship between stimulus and response, signal-detection theory stresses the relationship between correct and incorrect positive responses (hits and false alarms). In signal-detection theory, sensory excitation from one and the same repeated signal is assumed to have a defined distribution. Furthermore, it is assumed that excitation from another origin than the stimulus, for example from spontaneous nervous activity, appears as an integrated part of the sensory response. This "noise" is always present and is, thus, confused with the effect of the stimulus on a hypothetical excitation continuum. Every sensory excitation is evaluated by the subject on a probability basis, and the response criterion used depends on her expectations in the experimental situation as well as the consequences of her decisions.

Various detectability indices are used to estimate the overlap of the "noise" and "signal+noise" distributions. Such indices are less influenced by guessing and lack of co-operation on the part of the subject than is the classical threshold concept. A frequently used detectability index is d´. In calculating this index, it is assumed that "noise" and "signal+noise" are both normally distributed along the same excitation continuum, and that these distributions have the same variance. The index d´ is calculated from the proportions of hits and false alarms obtained from subjects when they are exposed to stimuli and "blanks" in irregular order, assuming that the response criterion is constant over the conditions studied. The index d´ is the difference between z-scores for p (hit) and p (false alarm).

With signal detection methods the task of the subject is simple. No training of the subjects is required. However, an advanced presentation equipment is needed (olfactometer) and

the procedure is very time consuming. The resulting data are on interval scales and additions/subtractions are meaningful operations. No calibration is being made.

The main advantage of applying the signal detection theory is that both positive and negative false responses (false alarms and misses) can be estimated. Thus, the effects of the response criterion can be separated from the sensitivity measures. Classical threshold methods typically only correct for false positive responses.

Signal-detection theory also provides methods with very short sampling time suitable for measurements of weak sensory signals (Berglund, Berglund and Lindvall, 1974). However, as a relatively large number of observations are required, the signal level has to be constant for some length of time, or be repeated a large number of times with good reproducibility.

The main drawback of the signal detection approach is the small range of physical quantities that can be investigated if the number of observations in the studies is to be held within reasonable limits. On the other hand the method does not require wide dilution series and, therefore, potentially is suitable for measuring weak odours as they typically appear in indoor spaces (Berglund, Berglund and Lindvall 1987). The signal-detection approach is technically easy to adapt to sensory environment control; for example, it has been used for investigating street odours in trafficked and less-trafficked areas (Lindvall 1973). Its application to IAQ has been limited to testing of a few single compounds and to a few building investigations (Berglund and Lindvall, 1979).

> The proper evaluation of detection data using signal-detection theory depends on the validity of the assumptions underlying the theory. A basic assumption concerns distribution forms for signal+"noise" and "noise" along the hypothetical excitation continuum. In field studies, the distributions often have been assumed to be Gaussian with differing means but equal variance since data are based on groups of subjects. Even if the results obtained from individual subjects have skewed distributions, the calculations of group means cause the resulting distributions to approach the Gaussian. However, when d´ is calculated as an integrated value over an observation period during which the stimulus (e.g. an odour) may have varied, the result could be a greater variance for the signal+noise than for the noise distribution. If the variances are unequal, the d´ value will be encumbered with systematic error and its size become affected.
>
> In laboratory situations, the distributions for noise and signal+noise can be determined from the separate individuals´ so-called ROC curves (receiving operating characteristics) (Green and Swets, 1966). A ROC curve is obtained by experimentally altering the response criteria for noise and signal+noise responses. Such experiments require that the signal strength be under experimental control. The process is time consuming, and it is uncertain whether the ROC curve obtained in the laboratory can be generalised to field conditions. In field situations ROC curves cannot be determined, because the signal varies in an uncontrolled manner.

4.2 Discrimination

In a discrimination evaluation the task demanded from the subject can be summarised in the questions "Is stimulus A greater, smaller or equal to stimulus B?" or, when pleasantness is the attribute, "Is presentation A more, less or equally pleasant than presentation B?". The rank order between the two substances can be obtained from either question and an ordinal scale can be constructed. An example is found in Clausen et al. (1993). In addition to generating rank order scales, the results obtained from discrimination tasks can be used

for determining difference thresholds and, by using indirect scaling methods (Thurstonian scaling, e.g., using Thurstone's law of comparative judgement; Thurstone 1927, reedited 1994; Torgerson 1958), for generating annoyance/pleasantness scales (Berglund et al. 1977).

> There are many discrimination techniques that generate order responses: method of adjustment, method of limits using a standard for making comparisons, method of constant stimuli, pair comparisons, duo-trio and triangle test (Baird and Noma, 1978; Meilgaard, Civille and Carr 1991).
>
> In the first three techniques a standard is presented and subjects match its perceived magnitude by adjusting a comparison stimulus by himself or by the answers he gives to the experimenter (less than, greater than or equal to). One direct application is in the determination of difference thresholds. Even though thresholds vary a lot from person to person and from group to group they can be used: (1) to tell about recipe tolerance in manufacturing and quality control, and (2) to determine scale steps for sensory intensity evaluations (Weber fraction).
>
> In pair comparisons the subjects compare stimuli and the experimenter notes the number of times (in percentage form) one stimulus is perceived to be greater than another. The number of times a stimulus is selected over the others is converted to a random variable having a specific distribution. The quantitative model leads to an unidimensional scale for all stimuli. For the attribute "pleasantness" (A is more pleasant than B) a Thurstonian scaling technique can also be applied directly. The resulting scale tells the order of pleasantness considering all stimuli along the same axis.
>
> In the duo-trio technique, also a standard is given and subjects are being asked which of two other stimuli is identical to the standard. Chance answers are 50%. The duo-trio test can be used: (1) to know if differences in general (or in pleasantness) are due to specific changes in ingredients, processing, packaging or storage, and (2) to determine an overall difference.
>
> The triangle test adds a third stimulus and considers two of the three stimuli as duplicates when asking subjects to select the different stimulus. Chance answers are reduced to 33%. The triangle test can be used to: (1) detect differences between stimuli (or in their pleasantness), (2) detect the factors that build up the differences in sample products, and (3) to select those subjects that have discriminated more efficiently as panellists.

4.3 Equal-Attribute Matching

Equal-attribute matching is based on a less well established theory but the reliability of similarity matchings have been empirically verified. The practical usefulness of cross-modal equal-attribute matching in developing "objective" yardsticks of perceived intensity has been demonstrated for several pairs of sensory attributes . For example, visual analogue scales are in wide use in IAQ investigations, force of handgrip has been used as a matching continuum for equating the perceived intensity of noise, light, warmth and cold (Stevens, Mack, and Stevens, 1960; Stevens and Stevens, 1960) and finger span has been used as the matching continuum for odours (Ekman et al., 1967; Berglund, Berglund and Lindvall, 1971).

By "intra modality matching" matching scales are developed within one and the same sense modality. In this technique the subject matches the subjective intensity or some other attribute of, for example, two different odorants (Svensson and Lindvall 1974). Thus, in cross-modality matching the matching is performed between continua, while in intra-modality matching the matching is between two qualities within one modality.

Some cross-modal matchings require little equipment, such as visual analogue scales (e.g. line length). Advanced equipment is required in intra-modal matching (olfactometer in the case of odours). The panel can be small- to medium-sized and no training of subjects is required. However, the panel members should be selected as to how well they adhere to the specific performance criteria and they should belong to the assumed target population. Typically no calibration is made but is postulated. No inter-laboratory comparisons have been made. The resulting data are on interval-ratio scales and arithmetical operations are meaningful.

Intra-modality intensity matching has been shown to be a reliable and valid method for olfaction (Beck, Kruger, and Calabresi, 1954; Kruger, Feldzamen, and Miles, 1955; Cain, 1969; Lindvall and Svensson 1974). The consistency of intramodal intensity matching in olfaction has been demonstrated in that matchings of three different chemical compounds were found transitive and symmetric with respect to odour (Svensson and Lindvall 1974). For applied studies a time-saving method of successive approximations has been developed (Svensson and Szczygiel 1974).

Equal-magnitude matching also has been used in acceptability/preference testing. By simultaneous exposure to two different stimuli and allowing the subjects to adjust the stimuli, they may trade one off with the other (Santos and Gunnarsen, 1997). Less uncertainty is introduced since adjustments may be continuous without categories.

Equal-magnitude cross-modal matching is common in questionnaire evaluations of IAQ parameters, such as air temperature and humidity, or for symptom ratings, such as headache. Typically, the ratings are made using a visual analogue rating scale consisting of a drawn fixed line with end points marked, e.g. "No symptoms" and "Unbearable". The subject indicates his present perceived magnitude by placing a mark on the line (Mølhave, Jensen and Larsen, 1991).

A continuous visual scale has been used for rating of acceptability (Gunnarsen and Fanger 1992). The middle of the scale is indicated as the transition between just acceptable and just not acceptable. Votes may therefore be interpreted both as binary votes and as votes on a continuous visual scale. This allows for a conversion of the votes on the continuous scale to an estimate of the percentage of subjects voting not acceptable with reduced standard deviation compared to direct binary votes.

4.4 Category Scaling

Category scaling is not based on a well established theory but empirically it has been shown to work at times. For example, the method has been used to scale indoor air qualities based on semantic scales. The five point intensity scale introduced by Yaglou was initially used as a category scale and later modified to be continuous (Yaglou, Riley and Coggings, 1936). Based on everyday experience subjects were asked to assign a point on the visual scale ranging from "No odour" to "Overpowering odour" to an air sample. The semantics of the

category names give some information of absolute levels. Comparison of results rely however on the representativeness of the subjects and their consistent interpretation of the descriptors.

Category scaling is in wide use in IAQ investigations, such as materials and compounds testing, IAQ evaluations and population response surveys.

Category scaling puts little requirement on equipment and is fast to conduct. The panel size can range from small to large. No training of subjects is required. The panel usually is selected as to representativeness. Typically no calibration is made but is postulated. In a few investigations data have been calibrated after Thurstonian scaling. The data from category scaling will appear on ordinal-interval scales depending on the application.

A common approach in surveys of discomfort/annoyance is to use verbal category scales. In the data treatment, typically only one of the response categories are being considered or several neighbouring categories are being combined into one. These procedures are commonly adopted when comparisons are made between areas or buildings. However, the distribution of responses over categories within the different areas or buildings will determine the outcome.

The comparisons of proportions in a specific response category involve certain assumptions about the response criteria. The problem is especially important for comparisons between areas or buildings where different criteria may be used by the different groups of respondents. The exposure conditions are likely to influence not only the perception of the environmental factors, but also the response criteria of the exposed populations. Thus, the assumption of stability in the conception of the response category is crucial. Although reported surveys have been based on this assumption only a few have discussed the matter.

Calibration procedures have been used to improve the comparability between annoyance data from different populations (Berglund et al. 1977; Berglund, Berglund and Lindvall 1987). A possible calibration procedure is to introduce a defined psychological distance into the data. Although such methodological precautions only rarely have been used in environmental studies, in many cases validation of traditional annoyance data against physical-chemical measures of odour or noise pollution have yielded useful results as long as group data are concerned.

> The Danish indoor climate labelling uses category scaling to distinguish between accepted and rejected construction products (Larsen, A., Nielsen, P. and Wolkoff, P., 1996). The period required for detected and emitted chemicals to be below known health and sensory detection limits is estimated, based on literature data. A sensory assessment is performed at that estimated age of the product. A naive sensory panel uses both a continuous category scale for intensity and a continuous value scale for acceptability. The voting scales are shown in the Figure 3.

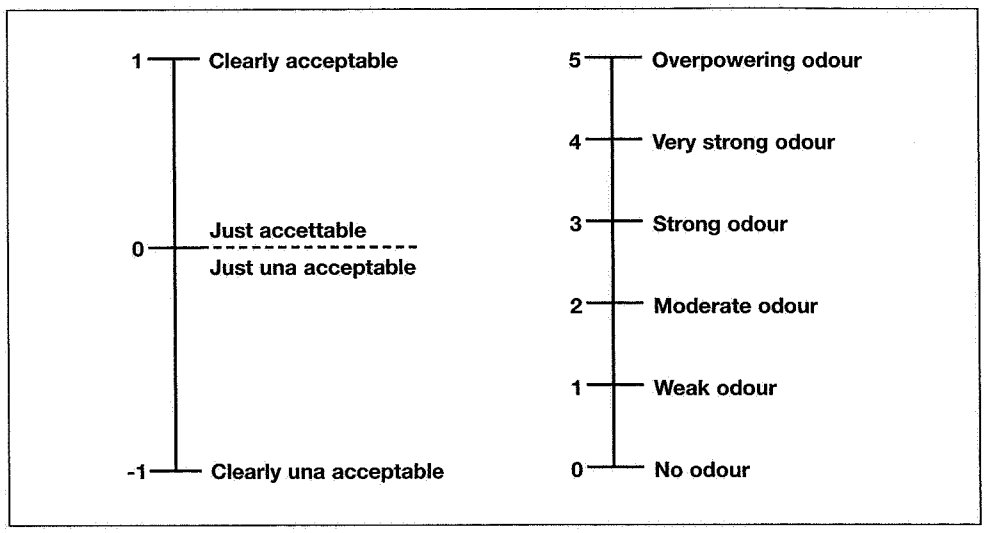

Figure 3. *Voting scales for acceptability (left) and intensity (right) used in the Danish indoor climate labelling scheme.*

If the sensory assessment results in a mean intensity vote below 2 on the intensity scale and above 0 on the acceptability scale, the time value based on chemical analyses is reported as final. If not, repeated sensory tests must be performed after further ageing until the acceptance criteria for sensory testing is met. The required period to finally meet the sensory criteria will be reported in the product standard. If the time value exceeds 160 days (for some products less) the product will be rejected.

Data can be collected by asking the respondents to judge their experiences of annoyance, or discomfort, about each environmental factor in a set (e.g., perceived odour, eye irritation, warmth, loudness, draught) in two ways: on a verbal category scale and through paired comparisons of annoyance for all possible pairs of factors. At least five stimuli are necessary in both category scaling and paired comparisons to ensure a wide frame of reference and a satisfactorily high number of confusions in the judgements. In category scaling also at least five response categories are needed (e.g., (1) not at all disturbing, (2) hardly disturbing or hardly noticeable, (3) annoying, (4) fairly irritating, and (5) nearly unbearable.

The data from verbal category scaling may be treated by a so-called indirect scaling methods (e.g., Thurstone's law of category judgement; Thurstone 1927, reedited 1994; Torgerson, 1958). In order to apply this law to the matrices of frequencies for the different areas of investigation, the frequency values have to be transformed into cumulative proportions over categories. Then the resulting matrices are transformed into matrices of unit deviates. The solution requires a few specific assumptions about the variance of response categories or of stimuli and about the correlations between these variables.

Gulliksen (see Torgerson, 1958) has devised a least squares solution obtaining scale values in such matrices. The solution gives scale values of the category boundaries as well as of the stimuli included in the investigation. The measurements are obtained on interval scales.

Another method of data treatment is to use information specifically gathered for calibration purposes. Some stimuli (or questions) may be included in the study solely for

that purpose, e.g., a reference odour presented to the respondents and their perception of another nuisance in the neighbourhood that is common to all the respondents. The calibration requires that on the average the differences between the annoyance to the reference odour and the other nuisance be constant. Also the responses to these two stimuli may be assumed to be independent of the responses to the target stimulus (e.g., the indoor air) to which exposure may differ in different buildings.

The traditional treatment of data in terms of proportions is inferior to the Thurstonian solution in that the latter with the same assumptions gives a single measure of annoyance for an environmental factor. Further, since the measurements are made on an interval scale, they allow a grading of stimuli with respect to mean degree of annoyance.

Paired comparisons may be a better procedure than verbal category scaling. First, the judgement tasks of the respondents are characterised by simplicity, and, second, data treatment is based on comparisons of homogeneous items, i.e., stimulus compared to stimulus, which probably gives more reliable results. The paired comparison approach still has the disadvantage of a possible variability in response criteria between populations but presumably to a lesser degree than category scaling. In paired comparisons the decision is made with reference to individuals´ conception of a noticeable and relevant difference, which probably does not vary as much as the conception of different category labels.

4.5 Magnitude Estimation

The concept of magnitude estimation is theoretically well established. In its classical form subjects are required to make direct numerical estimations of the sensory magnitudes (like perceived intensity or pleasantness) produced by different stimuli. The typical task can be "How intense/pleasant do the following series of stimuli seem to you? Assign numbers to your subjective impression.". From the numbers obtained an interval or ratio scale can be constructed. The slope (or curvature) of the psychophysical exposure-response function tells how fast responses change with stimulus intensity.

There are many psychophysical techniques that generate magnitude estimates of intensity, pleasantness or acceptability. Some of them are classical magnitude estimation, magnitude estimation without a standard, magnitude estimation with a standard but without a modulus, absolute magnitude estimation (Gescheider, 1985), converging limits (Guirao, 1987) and the extended triangular technique.

Magnitude estimations of perceived intensity can be used for modelling relationships to consumer acceptance and the results can be applied to product R&D. The psychophysical exposure-response function provides a standard for reformulating products, such as ingredient reductions to create new environment-friendly compositions (Lawless and Klein, 1991).

> *Classical magnitude estimation.* Subjects are required to make direct numerical estimations of the sensory magnitudes produced by different stimuli. A standard is presented and the subjects are being told that the sensation it produces has a given numerical value called its modulus.
>
> *Magnitude estimation without a standard.* The subjects are required to do the same but no standard stimulus is given.

Magnitude estimation without a modulus. The subjects are required to do the same, a standard is presented but the experimenter does not give any specific value to it, i.e., it is not numerically defined by the experimenter.

Absolute magnitude estimation (or Free number estimation). The subjects are asked to match their subjective impressions of size of numbers to their subjective impressions of sensory stimuli. These instructions may be experienced unclear by the subjects.

Converging limits. This procedure is a modified version of the classical magnitude estimation technique. The experimenter informs the subjects about the size of the continuum to be measured by presenting stimuli located at both ends of the physical range. The subjects are then asked to set their own range of numbers and use an adjustable scale that can expand or contract to accommodate their judgements.

The extended triangular technique. Three stimuli are presented, two of them being equal in intensity and one different. The experimenter asks the subjects which stimulus is the different one. When the subject answers, the experimenter assigns a number to that stimulus and asks for magnitude estimation of the two equal stimuli.

Besides being used for measuring perceived intensity of sensory stimuli, magnitude estimates can be used as a dependent variable in pre-post experimental designs to detect changes (Gescheider, 1985) and allow subjects to use numbers to describe the intensity of value judgements such as pleasantness and strain.

Magnitude estimates permit the calculation of various sensitivity measures (like Stevens exponent, Pearson's product moment correlation between the natural logarithm of the responses and the natural logarithm of the stimulus intensities P, and the M, MR and CI measures; Garriga-Trillo, 1985, 1987, 1996, 1997). From these sensitivity measures one can choose the subjects that have had the highest indexes as panellists and interpret the index as a proportion of correct responses. With this last interpretation in mind and asking for the confidence subjects have in the judgements they make, magnitude estimation responses can be calibrated. Subjects can be overconfident, underconfident or well calibrated.

4.5.1 Magnitude estimation with memory references

Magnitude estimation using memory references implies that panel members are basing their judgements on the recollection of how they perceived the references and moduli at presentations before the test situation. The technique increases the risk of high variances but empirically it has been shown to work.

There are few requirements on equipment in the test situation but semi-advanced equipment is required for the training sessions. The test procedure is fast but the training takes long time.

A small- to medium-sized panel can be used. Training of the subjects is required and they need to be selected as to how well they adhere to specific performance criteria. Anyhow, they should belong to the assumed target population. Some response calibration have been attempted. The data appear on interval-ratio scales.

The method earlier has been in some use in IAQ investigations (e.g., assessment of decipol levels using trained panels (see below) as applied in the European Audit project; Clausen, Pejtersen, and Bluyssen, 1993) but it is no longer in wide use.

In the European audit project (Bluyssen et al., 1996), in each of nine European countries a panel was trained to evaluate the perceived air quality in decipol of a total of 56 buildings. The panels where trained in the laboratory and then transported to the building to be investigated where the panel evaluated air qualities of outdoor air, room air, corridor air and supply air of ventilation system. The reference gas concentrations used for training was not presented at the building site, so the panel members made their judgement on a memory basis.

4.5.2 Magnitude estimation with one reference

With the one reference method there is a risk of systematic bias depending on where the reference is located on the stimulus continuum. There a few requirements on equipment and the procedure is fast. No training is required of the subjects. Some response calibration have been attempted. The data appear on interval-ratio scales but may be distorted pending on where the reference has been located. One-reference magnitude estimation is not a common method in IAQ investigations.

4.5.3 Magnitude estimation with several references

Magnitude estimation using several references is a well established method for specific applications. For IAQ measurements a semi-advanced equipment is required as well as a small- to medium-sized panel. Sometimes training to anchors is required (e.g., when decipols are being assessed; see below). Subjects may be selected as to their adherence to performance criteria but should belong to an assumed target population.

In some cases calibration is postulated through the training procedure. For ratio estimation there are calibration procedures. With the several standards method it is always possible to calibrate to a reference physical scale. The data appear on interval-ratio scales. The method is in wide use in chamber studies of IAQ and compounds testing.

Assessment of decipol levels. An example of magnitude estimation using several references is assessing decipol levels with a trained panel. In this method, a panel of subjects is selected and trained to evaluate the perceived air quality in decipol, with the use of a reference gas, acetone (2-propanone), supplied by an olfactometer (decipolmeter) (Bluyssen, 1990).

The theory of olf and decipol (Fanger, 1988) is based on the assumption that the pollutants in buildings all have the same relation between exposure and response after one factor normalisation based on human bioeffluents. Emission rates are measured in olf where one olf is defined as the emission rate causing the same level of dissatisfaction as bioeffluents from one seated person at any airflow. Concentration or perceived air quality is measured in decipol. One decipol is defined as the concentration of pollution causing the same level of dissatisfaction as emissions from a standard person diluted by a clean airflow of 10 l/s.

The relation between the percentage of dissatisfied persons and seven acetone (2-propanone) concentrations in air has been determined by 265 test persons using a yes-no classification scale of acceptability (yes the air is acceptable, or no the air is unacceptable) (Bluyssen 1990). From this relation the relation between perceived air quality in decipol and the acetone concentrations (see Figure 4, Bluyssen 1990) has been derived using the exposure-response relationship between perceived air quality in decipol and the percentage of dissatisfied persons shown in Figure 5 (Fanger, 1988; Knudsen, Clausen

and Fanger, 1997). The relation in Figure 5 was established based on experiments involving more than one thousand persons (Fanger et al., 1988).

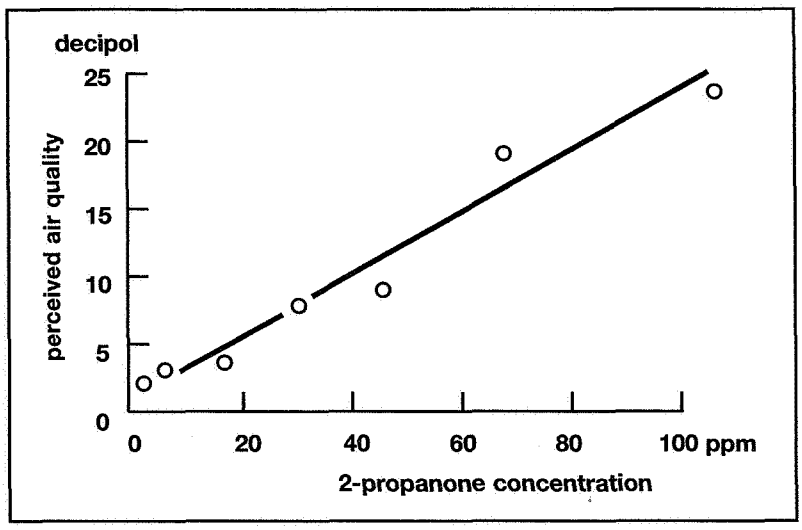

Figure 4. *Relation between perceived air quality expressed in decipol and 2-propanone (acetone) concentration in air, as determined by 265 persons (Bluyssen 1990).*

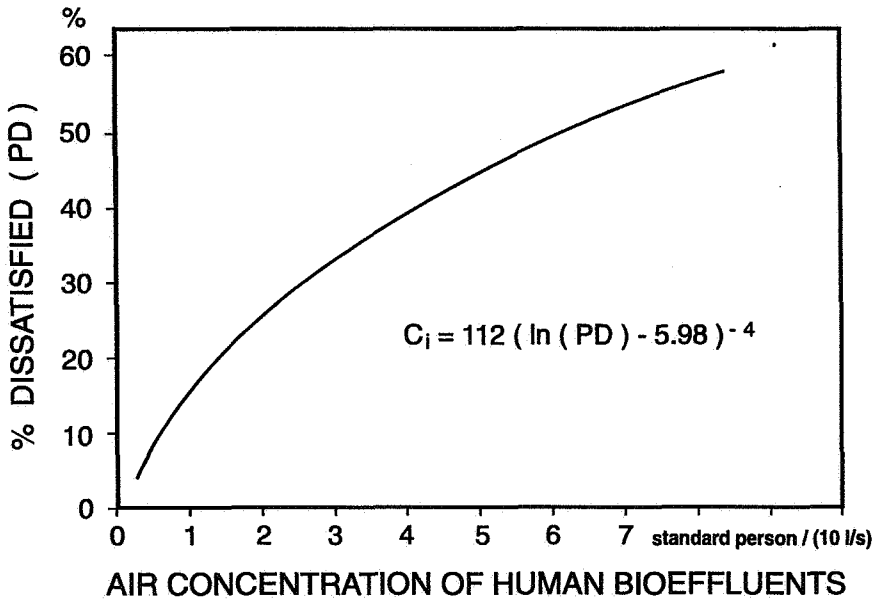

Figure 5. *The exposure-response relationship between the air concentration of human bioeffluents expressed in decipol (standard person/(10 l/s)) and percentage of dissatisfied (Fanger 1988; Knudsen, Clausen and Fanger, 1997).*

Magnitude estimation of odour intensity. Magnitude estimation with several references (1-butanol) has been used to measure perceived odour intensity of flooring materials (Ramalho, 1998; Karpe et al, 1995a, Karpe, 1995). Subjects were selected and trained to assess unknown butanol samples against the 8 reference concentrations scale increasing in a geometric manner (ASTM, 1981). Flooring material samples were either placed in large environmental chambers where panel members successively evaluate the perceived odour intensity of the samples (Karpe et al., 1995a) or introduced in small individual bags supplied to each panel member (Ramalho, 1998).

Trained Panels. In the European IAQ-Audit project exams were used to select a panel member as well as a whole panel for the assessment of decipol levels (Clausen et al. 1993). Four acetone (2-propanone) concentrations (equivalent to 2, 5, 10 and 20 decipol)

generated by four decipolmeters, called the "milestones", served as the reference for the panel members. During the training several unknown decipol levels (acetone (2-propanone) concentrations) were evaluated several times using the four milestones as a reference. After each evaluation the correct answer was given to the subject. During the training the panel members were also exposed to other pollution sources than acetone (2-propanone), comprising several common materials from buildings.

After the training it was assumed that any source can be evaluated by the trained panel with the help of the milestones and can give any value between 0 and 30 decipol to the pollution emitted by that source. During the training and the evaluation of sources after training, the panel is taught that it is not the intensity that should be evaluated but the acceptability of the perceived air quality. When an evaluation is made the panel members are asked the following question: "Imagine that you have to work for eight hours in this air, how would you then feel?"

Addition of sources. Simple addition of pollution loads would imply hypo-addition in terms of sensory impression because of the non-linearity and shape of the underlying exposure-response relationship. Hypo-addition would also be in agreement with laboratory findings on the interaction of perceived odour intensities and the vector model of odour interaction (Berglund, Berglund and Lindvall, 1976).

Several experiments have been made to investigate addition of source strengths. In one study, simple addition of the source strengths (in olf) of two single sources was found to overestimate the actual measured source strength of the combined sources (Bluyssen and Cornelissen, 1997). However other studies would imply that addition of source strengths (in olfs) may be meaningful in practice for some groups of sources such as bioeffluents from human beings, tobacco smoke and emissions of chemicals from building materials (Lauridsen et al., 1988; Iwashita and Kimura, 1995; Wargocki et al., 1996). But it is stressed that the addition of source strengths in person equivalents (olfs) should not be confused with the addition of the perceived air quality magnitudes. The reason for this is that the percentage dissatisfied and the perceived intensity of odour are not proportional to the pollutant concentration or to source strengths of pollution in person equivalents (olf). Furthermore, exposure-response relationships seem to differ between different materials (Knudsen et al., 1997).

4.5.4 Magnitude estimation with master scaling

Master scaling is based on free number magnitude estimation with several references but without moduli. The method gives to the subjects a large freedom of responding similar to that in free-number estimation. However, there is no risk of distorted scales and the instructions are clear to the subjects. The master scale is defined by five or more concentrations of a reference chemical that are jointly scaled with the target air samples during the whole investigation. The whole set of sensory data is transformed with factors necessary for transforming the individual psychophysical functions of these references to the group function for the references.

An advanced equipment (olfactometer) is required. The panel can be small- to medium-sized. No training is allowed. There are few requirements on the test subjects but the panel must belong to the target population. Individual scale calibration is possible by transforming to the scale of reference stimuli. Master scaling produces data on ratio scales.

Master scaling has been in limited use in IAQ investigations but it has been shown empirically to work (Berglund and Lindvall, 1979).

A master scale with references will account for the subject differences. It serves as a common reference of judgements of sensory stimuli independent of the judgement peculiarities of individual subjects participating in the particular experiment. The scale provides a defined unit of measurement of the attribute and the procedure will retain the information about the subjects' individual differences in scaling performance. When applied to a psychophysical problem the target can be expressed in terms of the perceptual or physical units of the master function (Berglund, Berglund and Lindvall, 1978, 1984; Berglund and Lindvall, 1979). On one and the same master scale of perceived intensity, it has been possible to compare, e.g., the class room air in a school building, filtered class room air, ventilating inlet air to the class room, and the perceived strength of kitchen odours as they appear after 20 min of cabbage cooking in a standard kitchen situation (Berglund, Berglund and Lindvall 1987).

The usefulness of the master scaling is that the objects of investigation can be compared in absolute terms by their position on the same master scale. Further advantages are that one can compare present results with earlier studies and it is not necessary to know the target stimuli in physical units (like in mixtures of emissions from building materials).

The introduction of reference stimuli to the psychophysical experiment requires that they do not distort the perception of the target. The psychophysical scaling of the master agent and the target agent jointly should not affect either of the two scales of perceived intensity. Furthermore, the reference master should be technically easy to produce, reproduce and measure as well as be perceptually distinct and easy to discriminate in small steps over a wide range of the perceptual dimension.

The use of master scaling will

- serve as an indicator of the subject's scaling behaviour, e.g., panel performance,
- be used for calibrating perceptual scales,
- make standardisation possible by the aid of the psychophysical function of the master,
- control for the range effect of perceptual scales.

4.6 Descriptor Profiling

Descriptor profiling as applied in IAQ investigations is based on a less well established theory. The method puts few requirements on equipment. The testing is fast but it takes a long time to develop the descriptor profiling instrument. Typically there is no training of subjects. It is important that they are selected as to representativeness. No calibration is possible. The level of measurement varies with the application; the data can be on nominal, ordinal, interval or ratio scales. Descriptor profiling is widely used in surveys and chamber studies of IAQ.

Ideally, many aspects of perception should be taken into account in the sensory evaluation of building materials' emissions. By measuring a set of relevant attributes of "goodness" of a building material it is possible to construct a score which may qualitatively be evaluated for each material, including e.g., detectability, perceived air quality, perceived intensity, modality-specific sensations (e.g., cooling, warmth), freshness/staleness, appeal, preference, intrusiveness, irritation and its target organ, adequacy for intended use, etc. However, a measurement model for such a combined score has not yet been established.

The questionnaire is the common tool to assess the occurrence of symptoms in large populations. However, there are several dimensions, which should be taken into account in the assessment of the symptoms, i.a., presence, quality, intensity and severity of a given symptom (Berglund and Lindvall, 1992).

In so-called SBS investigations symptom prevalences are measured by counting the number of persons reporting each particular symptom from a restricted list of symptoms. Typically, SBS is operationally measured (and thus defined) by adding the prevalences of the various symptoms for a group of occupants. The building is then classified as a sick building if this SBS measure is "high". This approach involves both a quantification by postulating rules of correspondence between the SBS and the counting of people (the prevalences of symptoms assumed to be relevant) and a classification by postulating that a certain measured value of SBS divides buildings into sick and healthy ones.

> The measurement power of so-called "SBS" questionnaires has been accepted uncritically. Is it at all possible to quantify symptoms with a questionnaire? Is the questionnaire sensitive enough for the measurement task? These questions relate to reliability (i.e. the precision of the measurement) as well as to validity.
>
> A common model of measurement is that every respondent in "SBS" questionnaire surveys uses identical response criteria in reporting the presence or absence of a symptom or assigning the frequency of symptoms into categories. An alternative model is that the individual response criteria are distributed, perhaps normally, along an underlying sensitivity continuum such that the criterion adopted by an individual is influenced by personal experience and abilities to cope with the so-called "SBS" (Berglund, Berglund and Lindvall, 1975). This could be represented, for example, by Thurstonian scaling (Thurstone 1927, reedited 1994; Torgerson, 1958); or signal detection (Green and Swets, 1966). Still more complex models have proven useful in indoor air quality research, like pattern recognition (Berglund et al., 1982; Noma et al., 1988; Baird et al., 1990).
>
> There is little agreement on diagnostic tools by which the investigator may identify people afflicted with so-called "SBS". However, various not yet perfected means for classification of people exist; one of them being simple self reports regarding the respondents' own opinions as to whether they are suffering from "SBS".

4.7 Classification (yes/no)

Methods of classification of IAQ are based on a less well established theory. The typical example is counting persons reporting on a binary yes-no response task. There are few requirements on equipment, the procedure is fast, it can be used by small or large panels and typically no training of subjects is required. However, the subjects must be selected as to representativeness. No calibration is possible and is typically postulated. The results are on nominal scales and, therefore, it does not make sense to add, subtract, multiply or divide such values.

Classification is widely used in surveys and in earlier panel studies of IAQ. The most recognised document describing classification of IAQ is probably ASHRAE 62-1989 (1996) which is now under revision. In the proposed revised version it states: "Among a panel of 20 untrained impartial persons more than 80% should find the air acceptable." However, there are limitations in using crude frequency measures, such as counting persons. This method is easy to describe but little attention is paid to response criteria, precision and efficient use of subjects.

Problem buildings often are being characterised by their high frequency (prevalence rates) of adverse symptoms among their occupants. However, counting persons is not sufficient,

because various determinants other than indoor air quality may have an effect on the occurrence of symptoms. Persons also have different criteria for reporting a specific symptom or its severity.

4.8 Multidimensional Analyses

When dealing with subjects' responses and their measurement, the best overall picture of results would be multidimensional. Unidimensional analyses never disclose the relationships among different attributes and they seldom include more than one response variable.

In sensory evaluations of materials and IAQ, multidimensional techniques can be used to:

- identify variables that interact with others to maximise product characteristics,
- screen variables that have large effects on a product,
- identify variables that affect some product characteristics,
- summarise large numbers of variables to fewer ones or more meaningful ones,
- identify relationships among responses, and
- group subjects with similar patterns of behaviour.

> Examples of unidimensional measures or functions are: absolute thresholds, proportion of correct responses, corrected proportion of correct responses, d', differential thresholds, Weber's law, Thurstonian scaling, psychophysical exposure-response functions, magnitude estimates, sensitivity measures, confidence ratings and calibration.
>
> Multidimensional techniques are useful in assessing interdependencies of variables (as in principal components analysis techniques and exploratory and confirmatory factor analysis; Stevens, 1992; Bernstein, Garbin and Teng, 1988), functional interrelationships (as in multiple regression analysis; Bernstein, Garbin and Teng, 1988), data from multiple sources (cluster analysis and multidimensional scaling; Baird and Noma, 1978; discriminant analysis, Stevens, 1992), and differences or similarities between samples of data (MANOVA and ANOVA, Stevens, 1992; procrustes analysis, Lawless and Klein, 1991; the Kruskal-Wallis, Jonkheere, Friedman and Cochran statistics).

4.9 Conclusions

Most sensory methods to determine thresholds are based on dilution of the air sample until half of a panel no longer detects odour or irritation. The thresholds vary widely with chemical substance, measurement procedure, quality of the equipment (olfactometer), purity of the chemical substance, and sample of subjects. This is reflected by the large spread in odour threshold values for single compounds as reported in the literature.

Threshold values do not give any measure of the perceived intensity above threshold levels. In materials testing, though, they can provide information on how much air dilution of the emission is needed to odorlessness.

Methods based on signal-detection theory do not require wide dilution series and,

therefore, are suitable for detecting weak odours as they typically appear in indoor spaces. Both positive and negative false responses (false alarms and misses) can be estimated, given a sufficiently large number of presentations. Thus, the effects of the response criterion can be separated from the sensitivity measures. Classical threshold methods can only correct for false positive responses.

Measures for quantifying perceived intensity of odours and sensory irritation are obtained from subjects by various psychological matching and scaling methods. They include, i.a., equal-intensity matching and magnitude estimation. Direct scaling methods are the most common. Preferably, the perceived intensity measures shall be obtained from a calibrated scale so that measures can be compared between different materials, occasions and laboratories.

Perceived quality is defined here as the perception of complex sensory stimulation, e.g., a pattern or a blend. There is little cognitive evaluation involved in this process. By value judgement is meant here a high order mental evaluation. It can be based upon perceived intensity and/or perceived quality of the sensory stimulation. The result can range from hedonic tone to acceptability.

In the field of indoor air quality, the most recognised document describing the assessment of value judgement of indoor air quality is probably ASHRAE 62-1989 which is now under revision. The scale used there has the apparent advantage of dealing with acceptability. However, it is difficult for the subjects to distinguish between intensity, quality and acceptability. Little attention is paid to precision and efficient use of the subjects. Furthermore, the procedure of using frequency of responses for measurement puts requirements on sampling statistics.

Investigations performed to quantify building-related discomfort, annoyance and symptoms rely on verbal reports or other kinds of judgements. The typical approach is the sociological survey. One difficulty is to select an attribute that adequately reflects peoples conception of environmental factors. A further difficulty is that scales of discomfort/ annoyance/symptoms obtained may be based on different units of measurement. Such scales cannot be compared easily unless they are calibrated or, at least, standardised to distributions of standard scoress.

5 CALIBRATION

It is not quite obvious how one should proceed in calibrating perceptual scales as compared with scales in physics. In measuring perceived odour intensity of indoor air, or any inhaled gas, by human subjects, the individual momentary scale values given by the subject will not yet belong to a common scale for the group. The different subjects can be viewed as different measuring instruments which are measuring their own perceived odour intensity, which will differ perceptually between individuals although the stimulus may be the same. In order to obtain comparability between scales, the individual scales have to be calibrated. For that references are needed. The calibration procedure should preferably leave the subject free in this judgements by calibrating the individual scale units to a common scale unit for the group.

Typically, magnitude estimation is based on successive numerical intensity scaling of references and unknown air samples. The subjects are being asked to assign numbers to the presented air samples trying to get the relative magnitudes right. Assuming a common zero the scale values may be multiplied with a constant to make up for the individual uses of the number scale. The method provides differences between air samples. Absolute levels may only be derived from the references.

The use of references vary. Some investigators use only one reference and present it openly at the beginning of the sessions together with its assigned value. Others use several references randomly mixed with the unknown air qualities and do not give information on the supposed scale values of the references.

The reference stimuli must not distort the perception of the target stimulus. Furthermore, the reference should be technically easy to produce, reproduce and measure in an olfactometer. It should be perceptually distinct and easy to discriminate in small steps over a wide range of the perceptual dimension. Not many chemicals satisfy these requirements (Johansson, 1990). A relatively high number of assessments are required by each subject to justify the estimation of individual calibration factors.

In order to accomplish comparability between scales in behavioural science (e.g. in questionnaire surveys) the practice have been to standardise scales by transforming the empirical response distributions to a standard distribution of z scores (e.g., Lord and Novick, 1968) rather than to calibrate the scale as such.

Scale calibration to equal intercepts. Problems related to acclimatisation, anticipation and habituation put special requirements on the methodology, e.g., in chamber studies. For example, the measurement of discomfort under different environmental conditions in a semi-or-full-scale laboratory setting causes difficult scaling problems. Changing the environment of a chamber may take hours, and the subjects´ adaptation to the environment, as well. A subject can never directly compare the comfort one day with that of another. The frame of reference, that is the subject´s concept of discomfort, may change with the climatic conditions. This change must be controlled for by an experimental design in which one of the elements is a calibration procedure. A simple approach is scale calibration to equal intercepts (Berglund, Berglund and Lindvall, 1984).

Based on the fundamental assumption that exposure condition and discomfort covary in various ways depending on time of exposure and level of stimulation, a transform of the data can be made. The intercepts for the different psychophysical exposure-response functions obtained for different environmental conditions are then set equal and adjusted for differences in initial exposure and discomfort. The transform compensates for the subject's anticipation of discomfort in terms of unit of measurement as well as differences in initial base values for exposure and discomfort for the different environmental conditions.

Scale calibration to a reference scale. The design of a field study of indoor environments must permit comparability between the annoyance/discomfort responses from different observers. The exposure conditions are only partly under experimental control and typically only one or a few subjects occupy the same space in the building. This situation requires a scale calibration of perceptual responses to a reference scale (Lindvall, 1974).

The odour of the indoor air may be assessed by using magnitude estimation relative to one reference odorant in a sniff bottle. In an additional experiment, the psychophysical function for the reference odorant is determined for each subject. With the aid of the psychophysical functions the odour responses obtained from each subject is calibrated to a reference scale based on the mean psychophysical relationship for the group.

An example is the goggles method used for assessing the eye irritating potency of air pollutants (Hempel-Jørgensen et al. 1997). A set of concentrations of a reference eye irritant (CO_2) are used as reference exposures and each pollutant concentration is compared to those references with respect to perceived eye irritation.

Master scaling. It is clear that individual differences exist in people´s perception of odour and irritation and in their response behaviour. To fully account for the subject differences a master scale with references may be used. Such a scale serves as a common reference of judgements of sensory stimuli independent of the judgement peculiarities of individual subjects participating in the particular experiment, and provides a defined unit of measurement of the attribute. The procedure is described in section 4.5.4.

6 SELECTION OF SUBJECTS

Human panels have been used to assess IAQ at least since the work by Yaglou and Witheridge (1937). Individuals, panels and populations do differ in sensory sensitivity, response behaviour and value judgements. Some of these differences are environmentally induced, some are linked to person and personality characteristics. Much of the interlaboratory variations which occur even when the same technical equipment is applied, can be attributed to panel differences. To decrease the effects of such differences standardised panel selection procedures should be established. These procedures should include such factors as sensory sensitivity, representativeness to the target population and personality variables. For sensory irritation it is important to consider if persons with allergy or hypersensitivity should be included or not.

All panel members are subject to individual differences in sensory scaling behaviour. This is reflected in the fact that not all the variance of quantitative responses is accounted for by the physical characteristics of the stimulus. Personality variables may be important (e.g., Costa and McCrae, 1992). Low sensory sensitivity is claimed to be associated with neuroticism, high neurotic scores have been associated with the use of "large" numbers when quantifying sensations, and introverts have been shown to be more sensitive than extroverts (Garriga-Trillo et al., 1994).

Furthermore, a number of biological variables influence olfactory sensitivity. Most important is the decrease in sensitivity with age.

> The odour sensitivity (absolute detection thresholds to dimethyl monosulfide) decreases to about 3/4 in the age group 25-50 yrs and to about 1/5 in the age group 5l yrs or above, all compared to the age group 18-24 yrs (Lindvall, 1970). Only small differences in odour sensitivity are reported between sexes, although women are claimed to be more sensitive than men to certain odours (Koelega and Köster, 1974). Some individuals are claimed to have a genetically controlled inability to smell certain substances, e.g., potassium cyanide or musk (Kirk and Stenhouse, 1953; Whissel-Buechy and Amoore, 1973). Mostly, however, the sensitivity to one odorous compound is accompanied by a proportional sensitivity to another (Brown, MacLean and Robinette, 1968).

A critical question is whether a panel should be "naive" (plain consumers) or "trained" (experienced). In using naive panels, an underlying assumption would be that the experience or training effects are of minor importance. A certain level of training is required even for the most simple evaluation and this training is enhanced if the subject participates in several test sessions. Thus, a truly naive untrained panel does not exist. Several guidelines for the selection of subjects have been given (e.g., ASTM,1981; ISO 6658, 1985; Meilgaard, Civille and Carr, 1991).

The selected and trained panel members are the most suitable for many sensory quality control tasks. Panel members must be unbiased. Therefore, the use of occupants of problem buildings or polluted areas may become problematic. Family or staff members cannot be used.

The panel's success depends on the selection and training process. The panellists are not representative of a large population since they have been chosen and /or trained for specific sensory capabilities. For sensory detection tasks the selection criteria commonly are: low

absolute and differential odour or irritation thresholds, a high signal detection index (d'), and a high proportion of correct responses. If pair comparisons are involved, the congruence of Thurstonian scaling for subjective judgements of physical intensity over time could also be used as a criterion. If magnitude estimates are to be obtained, the calculation of sensitivity measures could help in the selection of subjects.

The sensitivity measures, M, MR and CI, have been defined by Garriga-Trillo (1987, 1996). M is the proportion of correct ordered responses. Its mathematical formula is:

$$M = NI/Cn,2$$

where
NI = number of non inversions in the subject's response to the quantification of the stimulus S, considering the real order of stimuli

Cn,2 = combinations of the n stimuli taken two at a time.

In defining MR, we consider a weighted proportion of correct ordered responses (R) taking into account the distance between the stimuli in the chosen comparison pair. The largest weight would be assigned to the smallest distance and the smallest weight to the largest distance. In this way it results that:

$$MR = \frac{\sum W_{NI}}{\sum W_T}$$

where
W_{NI} = weights for non inversions
W_T = total weights.

This measure considers an ordinal level of measurement but based on distances between stimuli and not on the stimulus magnitude itself as contemplated for M. We start with n stimuli, n(n-1)/2 distances, di, between them and n quantitative responses to the stimuli. The steps for calculating MR are:

(1) Calculate: di, i = 1,..., n(n-1)/2.

(2) Determine how many different distances are there, let us say r.

(3) Assign the number one to the largest distances, two to the second largest ones, etc., until you assign r to the smallest distance between stimuli. These numbers will be the weights, w, assigned to each comparison between pairs of stimulus-response.

(4) Identify which pairs of (Si, Ri) show a non inversion (the order shown between the two stimuli should be in the same direction as the order between their corresponding responses).

(5) Add the weights for the non inversions. This will be the numerator of the above defined formula.

(6) Add all the weights assigned to each stimulus-response pair comparison. This will be the denominator of the formula.

The next measure, CI, is designed to solve the problem of having to group the S-R data within one experimental block, as we had to do to calculate M. CI will let us consider each response to a stimulus as correct or incorrect. This measure is based on the notion of a confidence interval for the population's mean response (U) for each stimulus. A response

would be correct if it belonged to the confidence interval and incorrect if it lied outside this interval. Its definition is:

$$CI = RCI/TR$$

where
RCI = Number of responses that belong to the confidence interval
TR = Total number of responses.

The definition of the confidence interval ($\Re$) for the population's mean response considers an $\alpha = .05$ and standard scores. It follows that:

$$\bar{R} - 1.96\sigma_{\bar{R}} < \Re < \bar{R} + 1.96\sigma_{\bar{R}}$$

where:

$\bar{R}$ = Mean response for each stimulus in standard scores

$\sigma_{\bar{R}}$ = Standard error for the mean.

Once we have developed these sensitivity measures we have studied their probability distributions (Garriga-Trillo, 1997) and have built up a categorisation in five intervals. Scores indicate then who has very high, high, average, low or very low sensitivity. So, like thresholds, these measures may also be used as a means of selecting and testing panellists adding the advantage of knowing their probability distribution. Also the mathematical function for the population distribution can be found and differences between values and their associated probabilities can be statistically tested for differences. High measures imply that response biases have been small or non-existent. Using a multiple regression model one can find which variables influence the sensitivity measures.

For panel members that are intended to be just *plain consumers*, their representativeness has to be considered. The selected group of test subjects has to be a sample from the same population as to the one the sensory analyst will make inferences. Thus, if the study goal is to promote a specific product for a specific segment of the population, subjects should be carefully chosen from that segment. The panel will become still more adequate if the subjects are further selected to include only the ones who are highly sensitive in discrimination tasks. If the resulting panellists do not discriminate between alternative products the target population is unlikely to note a difference.

The main variables to consider when selecting a panel of plain consumers (Meilgaard, Civille and Carr, 1991) are:

1. split the users of the product to be studied in categories (heavy users, moderate users or light users) and sample from the different groups,
2. age is relevant in most studies, so the selected subjects should be chosen according to the percentage of the population that belongs to each age group,
3. sex is another relevant variable depending on the product to be studied,
4. geographic location is relevant because there are regional differences in preferences or pleasantness, and
5. nationality, race, religion, education and employment should also be considered in sample selection.

The description of the panel should be reported together with the results and include: gender, age, smoking habits, criteria for why some subjects may have been excluded, measures of sensory sensitivity and representativeness as to a specified population.

In so-called *"SBS" investigations*, various kinds of panels may be used, e.g.: (a) representative samples of the population, (b) persons in the general population who likely have "SBS", (c) persons who are sensitive in reporting symptoms, e.g. a sub-sample of allergic people, and (d) persons who have been shown to get "SBS"-like symptoms in provocation tests. The selection of a control group poses a large additional problem mainly due to the few occupants per building and the uniqueness of a building "exposure". Finally, panel designs of surveys commonly suffer from large no-response and drop-out which require repeated revisits to non-responders and an analysis of the potential biases to the results introduced by the drop-outs.

In the ideal intervention study both the subjects and the investigators assessing the effect should be "blinded" to the assignment, i.e. unaware whether the subject belongs to the treatment or control group. In some experiments the intervention or treatment can not be blinded, but sometimes it is possible to apply alternative treatments between comparisons can be made, or apply a placebo treatment, i.e. to carry out procedures which simulate the actual intervention but do not have the true effect.

Some of the symptoms of the so-called "SBS" have seasonal variation and thus an uncontrolled experimental change of any condition might appear to increase or decrease the symptoms over a given time period. The main purpose of the control (reference) group is to represent the change in the outcome of interest over time due to the effects of extraneous factors. Also repeated administration of questionnaires might affect the reporting of symptoms. While the change in reporting is expected to be similar between two randomised groups, the reporting of a control group represents the outcome without the effect of the treatment.

In some cases, designing a survey as a panel study is the only way to obtain the desired information. If we wish to know something about an occupant for a special weekday throughout the year, then it is necessary to use diaries and usually a panel is set up. By using appropriate panel designs one may also reduce the memory effects that are inherent in personal diaries (van de Pol, 1989).

7 EXPOSURE CONDITIONS

A controlled exposure is "exposure with controlled variations". All relevant response modifying factors are supposed to be identified and their effect nullified either through controlling their variation, by randomisation of the variation in the measuring protocol or by including the measured variation in the statistical analyses.

Establishing of a protocol starts with identification of the relevant variables and discussion of how to nullify their effects on the effect measurements. In this process introduction of clean air exposures, reference exposures, spiked exposures or duplicate exposures to subjects are essential to document that the exposure is unaffected by response modifiers.

The most representative exposure context is obtained in the rooms where these exposures normally occur. However, more accurate and reproducible measurements may be obtained by exposing subjects in a strictly controlled environment, e.g., in a climate chamber. In practice, most exposures belong to one of three types:

- whole body exposures in climate chambers under fully controlled conditions
- field exposures of walk-in types of panels (e.g., in mobile vans or actual rooms) and
- partial exposures of nose only, eyes only, or nose and eyes (e.g., through funnels, hoods, face masks or goggles) in an equipment placed under controlled conditions in a climate chamber or in a van at the test site.

> The exposure may be generated by an exposure equipment or by sources in the environment itself. The exposure may be delivered to the subjects either directly by the room air or through equipments controlling the concentrations, timing and type of exposures given to the subjects. The exposure may reach the whole body or only a part of the subjects body. The main components of an exposure instrumentation therefore are:
>
> 1. The test room, which may be a real life room (e.g. an office or living room) or an artificial but neutral room for evaluation of air samples presented through auxiliary equipments.
> 2. Auxiliary equipments in the test room including the exposure applicator (e.g. nose funnel or pollutant injection into room air), exposure timer, air sampler, environmental monitors, etc.
> 3. Auxiliary equipment outside the test room including ventilation, air conditioning, air filters, air samplers and analytical equipment, separate laboratory/hood for odorant generation/handling, pressurised clean air, etc.
> 4. Acclimatisation or waiting room for subjects.

The relevant exposure measures for investigations of dose-response relations are the concentrations of pollutants at or around the receptors in the sensory tissues.

The air concentration of pollutants vary around the exposed subject. The concentration in the inlet to an exposure funnel may not be an acceptable exposure measure if exhalation or adsorption affects the air in or around the nose. Instead, the exposure should be measured close to the sensory tissue.

Some sensory effects such as irritation may depend on accumulation of pollutants, on

release of mediators in the tissues, or of the activation of a sequence of effects with a delay. Other sensory effects such as odour show an acute increase within few seconds and are followed by a decrease caused by adaptation. Therefore, an average air concentration may not be a sufficient exposure indicator. The exposure may have to be characterised both through distributions of peak concentrations measured with a few seconds averaging period and through the average of the entire exposure period (Berglund, Berglund and Lindvall, 1974).

The exposure and recording of effects should be strictly timed and the timing be better than a few seconds if reproducible results are to be obtained. In some cases this is not possible and the subjects have to use their memory perception. This increases the risk of bias.

The environmental exposure includes both the presence of the deliberately added pollutants and the presence of unintended compounds from other sources. The presence of these other compounds act as a disturbing background and must be measured, or at least their effects must be shown to be negligible (e.g., through sensory tests of the background).

To avoid the effects of the unintended compounds all surfaces of the exposure equipment, furniture, room etc. must be selected to reduce the effects of background sources, sinks, adsorption, chemical reactions etc. and the surface should be easy to clean. Regular cleaning procedures and background tests of contamination must be established.

Further, the ventilation must be sufficient to prevent background pollution from bio-effluents emitted from the subjects or emissions from equipments etc. Subjects also must be instructed to maintain a sufficient level of personal hygiene, not to use perfumes, or eat strongly spiced food before the tests. In addition, air movements, local ventilation rates around the subject, air temperature, etc., must be controlled.

The air humidity and temperature of the surroundings may affect the sensitivity of the exposed tissues and receptors to the chemical exposure. The effects of such non-chemical exposure should be kept at a minimum, the relevant factors be identified and their variations measured or documented to be negligible.

When test chambers are being used the exposure facilities should include a waiting room with a controlled IAQ and climate in which the subjects are brought to a sensory and psychological equilibrium for about half an hour before the exposures. Test rooms should be pleasant, not crowded and as little as possible look like a laboratory. To avoid contamination of the air in the chamber an air lock should be established between the test and waiting rooms and the surroundings.

The ethical rules of using experimental human exposures in science has been defined in the Helsinki declaration (Mølhave, 1997). The ethical aspects of all intended experimental exposures of humans should be evaluated and found acceptable using the best available toxicological principles and data. In all sensory evaluations the requirements of the Helsinki declaration must be followed. This, among other things, requires that the investigator prior to any exposure of panel members:

- evaluates the potential consequences of the exposure; only exposures which do not impose an increased risk of health effects on the staff and subjects can be accepted,

- identifies potential risk groups and excludes them from exposures,
- informs panel members about the risks and sees to that they sign an accept of the risk,
- provides the panel members with both verbal and written instructions,
- uses an experimental design that includes clean air exposures before and after exposures to avoid adaptation and habituation, and blinded exposures to avoid biases.

> The acceptability of chemical exposures of human subjects in indoor air quality research should be based on the following conditions:
>
> 1. Only reversible non-adverse health effects can be accepted. The risk of adverse effects associated with the exposure must be documented to be acceptably low.
> 2. Exposures to well-known chemicals or emissions from commercially available sources can be accepted if these sources have a history of many years of problem-free use on the market and if they are used in ways during the experimental exposures which correspond to normally encountered exposures. Studies including exceptional exposures or sensitive subjects should be registered and evaluated by the local ethical committee.
> 3. Exposures to emissions from new types of sources can be accepted if the exposures are chemically identified and are below official guidelines for exposures. These exposures should be registered and evaluated by the local ethical committee.
> 4. The selection of subjects must include special pre-tests and defined criteria to exclude risk groups. Potential risk groups must be excluded and be defined in the selection protocol.

8 NON-SENSORY TECHNIQUES

Till now non-sensory methods have not been successful as general measuring devices in the evaluation of odour and mucosal irritation. The main reason is lack of validity since, in essential parts, the mechanisms of the olfactory and somesthetic senses are not known. Hence no analogues to the sensory systems have been possible to develop. It is true that the perceived intensity of any given odorant or irritant increases as the concentration of the chemical increases but different chemicals with the same concentration will not have the same perceived intensity. Therefore the outcome is difficult to predict. In addition, other non-chemical factors often contribute to the final sensory effect of chemical exposure. However, biological and physical-chemical methods as well as "artificial noses" may provide supplementary or even surrogate information in selected situations and for selected groups of compounds.

Indoor air quality may be studied by measurement of several physical-chemical properties (e.g. air temperature, air humidity, VOC, particles). For this, a measuring instrument is needed. Ideally, the instrument should be based on a physical mechanism (e.g., thermal expansion of mercury) allowing the quantity to be expressed in terms of a magnitude (e.g. volume increase of mercury). Thus measurement requires a reference magnitude and an accepted measurement rule (e.g. the Celsius thermometer). The measured magnitude should be expressed in a defined and calibrated unit of measurement.

Indicators are substitutes for measures. Usually indicators are used when suitable measures cannot be obtained. The basic requirements of an indicator are that it bears a relation to the underlying effects and mechanisms and is convenient to use (see Stevens, 1951). An indicator can be a simple measure of one marker pollutant (e.g. CO2 for indoor air quality). It can also be a combination measure such as the arithmetic sum of several marker pollutants, a complex equation of a defined set of pollutants, or a model for combining these (e.g. total hydrocarbons in air). The distinction between measures and indicators disappears when the quantitative relation between the indicator and the effect of interest is known. The indicator can then be calibrated and used as a substitute measure with adequate validity and convenience. Candidate variables for substituting sensory evaluations do not per se have to have a causal connection to the biological process of interest. However, if the indicator is backed by a reasonable explanation, it becomes a much stronger predictive tool for assessing the sensory effects.

A wide range of parameters may be used to indicate exposure levels and their potential sensory consequences. Combinations of VOCs and TVOC have been suggested as such an indicator of chemical indoor air exposures of sensory relevance (see ECA-IAQ, 1997b). The quality of the indicators depends on the degree of association between the indicator and either the exposure or the target response variable. This association may only be of an acceptable quality within a limited range of variations and under certain conditions, depending on the confounding variables affecting the underlying physical or biological mechanisms. The acceptable quality relates primarily to the extent to which the relationship can be approximated with a monotonic function between the indicator and the variable or construct for which it is an indicator. Other important factors are the level of measurement and resolution obtained.

8.1 Biological Techniques

There are a number of techniques to evaluate the potency of odorants and irritants other than by direct sensory evaluation. These techniques include, i.a., recording of negative mucosal potentials from the olfactory epithelium, recording of cortical evoked potentials (olfactory as well as chemosomatosensory evoked potentials; Kobal and Hummel 1991), measurement of neural mediators or modulators, or, in the case of irritants, measurement of inflammatory products and reflexes (Alarie, 1984; Cain and Cometto-Muniz, 1993). Assays of the irritating potency of chemicals have not yet been validated with symptoms.

8.2 Physical-Chemical Techniques

Methods to measure volatile organic compounds (VOCs) and particles are sometimes important as a supplement to sensory evaluations, since these compounds are main pollutants that can reach the human nose and the majority are potentially odorous and/or sensory irritating. In fact, the overall odour strength of an indoor air sample was shown to be predicted simply from the number of components most frequently reported to have a strong odour (Berglund et al., 1982). If the "strong odour" compounds in a particular space are identified they would form the base for a surveillance by physical-chemical methods in that specific environment. However, no consistent relationship has been demonstrated between concentrations of different VOC mixtures in general and odour intensity.

Sampling and analysis of indoor air compounds are mainly based on two principles:

- one that first requires an enrichment step before making a physical or chemical measurement; knowledge of sample volume or sampling rate is required; e.g., chromatography;
- one that makes a direct physical measurement of some property of the sample in situ; no knowledge of sample volume or sampling rate is required; e.g., non-dispersive infrared spectrometry.

> *Chromatography.* Chromatography is a separation technique. In chromatography the mobile phase is a gas or liquid which is inert with respect to the sample. It flows at a constant rate in one direction through the stationary phase, which may be a solid with a large surface-to-volume ratio, or a high-boiling liquid on a solid support. The sample may be a gas or liquid; however, it must be soluble in the mobile phase. Typically, gas chromatography is used for separation of volatile, relatively non-polar materials or members of homologous series; liquid chromatography for separation of particularly those materials with low volatility and labile or unstable compounds; and thin layer and column chromatography for separation of inorganic or organic materials, and low-molecular weight species up to high chain-length polymers. A wide variety of chromatographic methods to separate organic compounds in air are available (Butcher and Charlson, 1972; Lodge, 1989).
>
> *Spectrophotometry.* Spectrophotometric methods are generally based on the measurement of the transmittance or absorbance of a solution of an absorbing salt, compound, or reaction product of the substance to be determined (Butcher and Charlson, 1972; Lodge, 1989). Spectrophotometric methods include absorption spectroscopy, emission spectroscopy, laser spectroscopy, photo-acoustic techniques and X-ray analysis.
>
> In spectrophotometry it is necessary to decide upon the spectral region to be used in the

determination. In general it is desirable to use a filter or monochromator setting such that the isolated spectral portion is in the region of the absorption maximum.

Ionisation. Mass spectrometry and flame ionisation can be placed under the category ionisation method (Butcher and Charlson, 1972; Lodge, 1989). In mass spectrometry a substance is analysed by forming ions and then sorting the ions by mass in an electric or magnetic field. Positive ions are produced in the ion source by electron bombardment or an electric discharge.

A flame ionisation detector (FID) makes use of the principle that very few ions are present in the flame produced by burning pure hydrogen, or hydrogen diluted with an inert gas. The introduction of mere traces of organic matter into such a flame produces a large amount of ionisation. The response of the detector is roughly proportional to the carbon content of the solute. The response to most organic compounds on a molar basis increases with molecular weight.

8.3 Multicoupling Techniques : GC/FID/MS/Sniffing

The multicoupling analytical method enclosed physico-chemical techniques and sensory techniques. The volatile organic compounds are identified and quantified using gas chromatography (GC) with flame ionisation (FID) or mass spectrometry (MS) detector. In the same time, sensory evaluation of the individual compounds are made by the mean of a sniffer (a device designed to inhale the volatile compounds at the capillary column outlet). This system allows to identify and quantify volatile organic compounds and odorous compounds resulting from air pollution.

Few studies on the determination of odorous compounds linked to pollution sources have been conducted so far. The few analytical methods described cover the use of a sniffer connected up at the capillary column outlet. The identification of odorous compounds is usually carried out by implementing two or three separate analytical systems (Gas Chromatograph (GC)/sniffer, Gas Chromatograph/Flame Ionization Detector (FID) and Gas Chromatograph/Mass Spectrometer (MS)), thus making the determination (identification, quantification and evaluation of odorous compounds) sometimes very difficult (Jensen et al., 1993 ; Sävenhed et al., 1985 ; Khiari et al., 1992). Another analytical system using the combination of thermal desorption, gas chromatography, mass spectrometry and sniffing (TD/GC/MS/sniffing) was used for the identification of odorous compounds (Heustache et al, 1988) but quantification is more difficult than with the FID. A further analytical combination of headspace/GC/FID/sniffing plus headspace/GC/MS (Taylor and Mottram, 1990) presented the drawback to double the analysis making the identification of odorous compounds sometimes difficult too.

A multicoupling method involving TD/GC/MS/FID/Sniffing has been used in the frame of indoor pollution in order to identify odorous compounds emitted from floor covering materials (Karpe et al., 1995b ; Karpe, 1995)

8.4 Electronic or Artificial Noses

Chemical sensors for gas molecules may in principle monitor physisorption, chemisorption, surface defect, grain boundary, or bulk defect reactions (Gardner and Bartlett, 1992). Several chemical sensors are available: mass-sensitive sensors, conducting polymers and semi-conductors.

Mass sensitive sensors include quartz resonators, piezoelectric sensors or SAW (Surface Acoustic Wave) sensors (Ema et al., 1989; Nakamoto, Fukunishi and Moriizumi, 1990; Bruckman et al., 1994). The basis is a quartz resonator coated with a sensing membrane which works as a chemical sensor.

With conducting polymers, a wide range of aromatic and heteroaromatic monomers undergo electrochemical oxidation to yield adherent films of conducting polymer under suitable conditions (Gardner et al., 1990). The conductivity of the polymer film is altered on exposure to different gases.

The principle with semi-conductor sensors is based on the change of the electrical characteristics of the semi-conductor when the gas to be measured is adsorbed. The change of the number of free load carriers or the change of polarisation of the bounded load carriers is then measured (Bruckman et al., 1994).

Several commercial instruments are available. Some comprise conducting polymers, others tin oxide gas sensors (thick or film devices) or metal oxide semiconductors, and combinations. A recent study exhibits good discrimination of six dry paint film samples using a 32 conducting polymer array sensor (Ramalho et al 1997).

Instruments with conducting polymer based sensors to distinguish between different beers and to detect low levels of specific unwanted vapours (Neotronics Scientific, 1994). The sensors may be based on several polymer material types (e.g., polypyrrole and polyanaline).

Tin oxide gas sensors (Gardner and Bartlett, 1992) are thick film devices made by coating a film of tin oxide onto a tubular ceramic former through the middle of which runs a heater. Electrode contacts are made to either end of the cylinder of tin oxide. In order to control the selectivity of the material for different classes of vapour, small quantities of catalytic metals are incorporated within the sintered SnO2. Thin film integrated devices (Gardner et al., 1990; Gardner and Bartlett, 1992) use a much thinner film and the power consumption is much lower.

Devices combining metal oxide semiconductors and conducting polymers can detect polar, sulphur, ammonia and aromatic compounds. The odorant is injected into a glass container; the responses of the sensors are recorded; "fingerprints" for the characteristic odour are generated; and these "fingerprints" are compared with known standards.

8.5 Conclusions

Much work has been done to use instrumental analysis for "sensory" evaluations. However, unequivocal detection and quantification of odour is generally not possible in this way, except for selected situations and compounds. The relation between the results of non-sensory methods and human responses needs to be established.

9 QUALITY ASSURANCE

In order to obtain high-quality data in sensory investigations, two facts have to be remembered:

1. No sensory measurement can be better than the limits set by the technical equipment and the psychophysical method.

2. Quality assurance in sensory studies is possible, in the field as well as in the laboratory, at least with regard to internal consistency. It is therefore mandatory.

Quality assurance refers to all steps which may be taken to ensure that data are reliable. Quality assurance covers the utilisation of scientifically and technically sound practices for the collection, transport and storage of samples, the laboratory analysis, as well as the recording, reporting and interpretation of results. Quality assurance of any measurement series has two main elements: quality control and quality assurance.

Quality control has two components. One is external quality control, which is a system for objective checking of laboratory performance by an external group. The other is internal quality control which is a set of procedures used by the staff of a laboratory for continuously assessing results as they are produced in order to decide whether they are reliable enough to be released. It is not sufficient that laboratory analyses are subject to internal quality control procedures alone. To ensure accuracy of results from an individual laboratory and to guarantee comparability among different laboratories, external quality assurance must also be practised.

Quality assurance consists in checks which guarantee a regular quality control and good quality results. The quality assurance in perceived odour intensity measurements may be referred to (1) actions with regard to the olfactometric procedure, and (2) actions with regard to the judgmental procedure. The actions involved are, for example, the introduction and control of judgements of blanks.

The purpose of the evaluation of perceived air quality should be clear to the investigator and the interpreter of the study results. The purpose of the investigation determines which effect measure should be used (odour, symptom pattern, acute or late effects etc.), the study design and the choice of measurement method. It also sets the level of sophistication required with respect to, i.a., calibration needs, view on inter-individual variation, reproducibility, validity, forecasting ability and quality assurance means.

In practical terms quality assurance is made in three separate steps:

1. The individual measurements are calibrated
2. The performance or credibility is checked
3. The documentation or the credibility of the final measuring result is communicated.

The performance and credibility part of quality assurance is focused on an overall management of the quality assurance procedures including:

1. Calibration of auxiliary systems such as scales, flow meters, volume meters.

2. Intercalibrations, duplicates analyses etc. and establishing quality indicators such as reproducibility, accuracy, sensitivity, and reproducibility.
3. Time series analyses of the variations of the quality with time.
4. Supervising the calibration, setting target values for the calibrations, follow up on deviations.
5. External audits.
6. Keeping track of reference samples used and external calibrations made.

The documentation or communication phase includes ensuring that the measuring results are reported together with adequate quality assurance information and documentation, and administering the filing system.

The experience gained from using different exposure equipment for sensory testing may be summarised as the following requirements on *exposure control:*

- To control subject performance as well as psychophysical method the dosage system should deliver references (e.g., reference odorous chemicals).
- The dosage system should be able to handle test and reference gases with minimal change in the composition of the samples and permit a controlled dilution of the samples.
- The sensory measurement equipment should be free from non-target pollutants since they may create adaptation effects in the test subjects, increase the background "noise" resulting in less resolution power of the method, and distort the subjects´ judgements.

More recent odour investigations have used reference substances such as acetone (2-propanone), dimethyl monosulfide, hydrogen sulphide, n-butanol and pyridine. The introduction of references in the psychophysical experiment requires that the reference does not distort the detection and perception of the target. The psychophysical scaling of the reference and the target jointly should not affect either of the two scales of the perceived attribute, i.e., no specific interaction should take place. Furthermore, the reference should be technically easy to produce and measure, and perceptually distinct and easy to discriminate in small steps over a wide range of perceived intensities.

The goal with sensory threshold testing is to arrive at absolute threshold values for detection and the need for quality control of exposure is evident. A comparison of odour thresholds for one single compound can be made in the following way.

First, one must introduce a reference odorous chemical to be determined by the laboratory in the same context as the test sample. This gives a practical check on such intervening factors as a possible biased selection of subjects and differences in adaptive state.

Second, when the investigator has decided upon a certain absolute odour threshold of the test substance, this specific threshold concentration may be studied by a signal detection approach. Methods based on signal detection theory permit a check not only of the existence of false positive responses on the part of the subject but also the existence of false negative responses. This means that an subject who does not perform satisfactorily, for example, due to lack of motivation, will not distort the results. A high detectability index, as measured by the signal detection approach, of the threshold concentration of the test substance would indicate that the threshold determination was not properly done.

One of the central problems in assessing perceptions of environmental factors is that different persons make judgements widely separated in time, context and compounds. This makes it difficult to compare judgements between experiments because it is clear that individual differences exist in people's perception and in their response behaviour. Therefore, a *response control* is necessary. Calibration is the preferred mean.

In sensory testing the introduction of well-defined reference substances is a primary requirement, both in threshold studies and perceived intensity studies. Securing results with a signal detection method furnishes supplementary internal checks on the reliability of the detection measures obtained by threshold methods.

The use of panels commonly aims at reducing inter-individual variation. This may be accomplished by performance training of panel members to conform with a set of references, and by calibration of individual scales for references through a transformation to a common unit of measurement. Principles and details of quality assurance for odour threshold determinations have been proposed by CEN (1995).

Range effects are important in sensory measurements. A common finding is that the exponent of the psychophysical exposure-response function is increased when the concentration range is decreased. For example, in calibrating sensory scales it is not enough to set the scale value for one stimulus to an equal level and then transform data by multiplication. A better approach is to match for equal-magnitude with a matching continuum or to use a master scale for scale calibration.

In studies of perceived air quality a high level of measurement scale (ratio or interval scale) is not always needed. Even rank order information may be sufficient but there are limitations in the comparability of such scales. However, calibration requires a minimum level of measurement (minimum an interval scale).

Basic requirements in *evaluation control* are that

- the inter-individual variance is displayed for the attribute being studied (perceived intensity, acceptability, etc.).
- the response distribution is displayed for the group in detection studies - not only the single threshold value of the target,
- the perceived intensity scales of the targets are calibrated by use of a reference scale of a reference chemical, and
- the reliability of threshold concentrations is checked by evaluating its index of detectability.

The perceived intensity scales may be checked for reproducibility by a split-half procedure or by repeated measurements. In the latter case it is important that the time between the two measurement occasions is long enough for ruling out memory effects in the subject.

For the perceived odour intensity of indoor air, it is quite uncertain what may constitute an external validation variable. For model prediction, the calibration by applying a master scaling procedure should result in a psychophysical power function. This should hold also when single stimulus judgements of perceived intensity are being produced by the subjects but made in the context of the reference stimulus series.

10 CRITICAL METHODOLOGICAL ANALYSIS OF SOME RECENTLY PUBLISHED DOCUMENTS ON IAQ

During the last 5-10 years a number of attempts have been made to prescribe sensory evaluation methods for indoor air quality. They have ranged from sensory testing of flooring materials and determination of required ventilation rates to questionnaire field studies of buildings with presumed indoor air quality problems. Here, three examples are taken and critically analysed mainly from the sensory measurement methodology point of view. Please note that in these earlier reports the terminology is not always as stringent as the terminology which has been developed in the present report.

10.1 Flooring Materials

For the purpose of a labelling procedure, the ECA-IAQ report "Evaluation of VOC emissions from building products: Solid flooring materials" proposes a provisional simplified sensory assessment procedure (ECA-IAQ, 1997a).

> In the report, the starting point is that indoor air pollution sources should not cause more than a maximum of 10 % of building occupants to perceive sensory irritation. Accordingly, a material will only receive a label if not more than 10 % of the test panel members assessing the material emission perceive sensory irritation.
>
> Furthermore, the evaluation procedure of material emissions requires that a sensory test of odour or perceived air quality is included. The result of this test should enable consumers to compare the emissions from different materials, or to rank the materials, with respect to odour detectability, perceived odour intensity, percentage of test panel members dissatisfied with the perceived air quality, or equivalent quantities.
>
> The report leaves to the authority or body establishing and/or granting a label to select and prescribe an appropriate test method among those described in the literature and used in practice. The authority or body is kept responsible for selecting a scientifically sound test design, choosing the measurement method(s) and the calibration procedure, identifying inter-individual variation, reproducibility and validity and assuring the quality of the methodology and the results. The evaluation method and results have to be displayed as well as the false positive rate for background air conditions, the traceability of the method and the results of quality assurance assessments.

The report accepts almost all possible endpoints of the sensory tests employed but leaves the choice to the authority or body establishing or granting a label: odour detectability, perceived odour intensity, percentage of test panel members dissatisfied, or equivalent quantities. Thus no specific sensory methods are being proposed. There is no discussion of what are the minimum dimensions needed for sensory testing of materials. Furthermore, acceptance criteria are only given for sensory irritation, not for odour. This hampers the recommend-ation since odours are always present.

Since the result of a test should enable consumers to compare the emissions from different materials, or to rank the materials, high demands are put on the sensory methods to be applied. In the report, the authorities or bodies establishing or granting a label are requested to select a scientifically sound test design, choose the measurement methods and the calibration procedure, identify inter-individual variation, reproducibility and validity, and assure the quality of the methodology and the results. These requirements seem

reasonable. To reach comparability between methods, calibration, or at least standardisation of the measurement procedure, will be needed. No guidance is given in the report.

10.2 Required Ventilation Rate

For the purpose of a calculation procedure for ventilation a sensory assessment procedure has been proposed based on predicted dissatisfaction and load calculations in terms of person equivalents (ECA-IAQ, 1992). No sensory evaluation technique is being specified in the report although two examples are given in an appendix (binary classification of acceptability using untrained panel, and magnitude estimation of perceived air quality (in decipol) with a trained panel using several references).

> In the report, the starting point is that perceived air quality may be expressed as the percentage of dissatisfied, i.e. those persons who perceive the air to be unacceptable just after entering a space. For air polluted by human bioeffluents the percentage of dissatisfied is a function of the ventilation rate per standard person. The pollution generated by such a standard person is called one olf.
>
> The report claims that the strength of most pollution sources indoors may be expressed as person equivalents, i.e. the number of standard persons (olfs) required to make the air as annoying (causing equally many dissatisfied) as the actual pollution source. Furthermore, one decipol is defined as the perceived air quality in a space with a pollution source strength of one olf, ventilated by 10 l/s of clean air, i.e. 1 decipol = 0.1 olf/l,s. The air quality in a space as expressed by the percentage of dissatisfied visitors is believed to be reflected in the decipol unit.
>
> The calculation of the required ventilation begins with a decision on the desired indoor air quality in the ventilated space, e.g. one of three categories of air quality corresponding to 10, 20 or 30% dissatisfied. The perceived outdoor air quality is also estimated.
>
> The next step is to estimate the sensory pollution load. The pollution load per occupant depends on physical activity and the tobacco-smoking behaviour of the occupants. Examples of occupancy, i.e. the number of people per m2 floor are given for different spaces. The pollution caused by the building including furnishing, carpets and ventilation system are being estimated as person equivalents. The total sensory pollution load is found by adding the loads (source emission rates) from the occupants and the building.

The main merit of the report is that it emphasises the importance of other contributions to the sensory load on the indoor air than just occupants. The guidelines only consider one endpoint, i.e., dissatisfaction with indoor air quality and the confidence limits of the underlying exposure-response relationship (Fanger, 1988) is not known (Oseland and Raw, 1993. The claimed attribute measured is perceived air quality. However, the term "perceived air quality" involves many other dimensions than dissatisfaction and acceptability, for example, odour intensity, stuffiness, perceived dryness, degree of unpleasantness. The report focuses on the first impression as a visitor and not on the occupants´ situation. In fact, later field studies failed to demonstrate a relationship between occupants´ acceptability scores and magnitude estimations of perceived air quality in decipols by trained panels (Bluyssen et al., 1996), may be because of odour adaptation or large variability in occupants' response criteria (Parine, 1996).

In the described method of magnitude estimation of perceived air quality (in decipol), the test subjects are faced with the dual task to judge the acceptability of a presented stimulus

and at the same time are being asked to relate it to an imaginary context situation. Due to person-linked factors judgements of acceptability are likely to increase inter-individual variability compared to judgements of perceived intensity. As a consequence, the resolution power of the scales may be low. The "decipol method with a trained panel" (see section 4.5.3) has been demonstrated to be sensitive enough to reveal effects of interventions on IAQ in a building although the occupants´ scores of acceptability failed to do so (Parine, 1996).

Three categories of air quality are suggested in the ECA-IAQ (1992) report but, besides some statistical considerations, the criteria fulfilment is not specified, e.g., as to accepted variance. According to later studies, it seems that there is too small a difference between the proposed three categories when taking into consideration the resolution power of the methods described. Higher percentages of dissatisfied visitors have been demonstrated in real life situations (Fanger et al., 1988).

Since the exposure-response relation for bioeffluents known from comprehensive binary scaling of acceptability is assumed to be valid for other pollution mixtures rather simple calculations are being made of effects on sensory perception of changed ventilation, reduced emission rates, etc. From the theoretical point of view, it seems difficult to justify that there would be only one exposure-response relation for all the types of pollution being found indoors. Exposure-response relationships have been shown to differ between building materials and the corresponding relationship for human bioeffluents (Knudsen et al., 1997).

10.3 Questionnaire Field Studies of the Sick Building Syndrome

A thorough review and analysis of field studies of the so-called sick building syndrome ("SBS") has been made by the collaborative action between a Task Group within CIB (International Council for Building Research Studies and Documentation) and a Working Group within the ECA-IAQ (Berglund et al., 1996). The majority of the studies on the determinants of the "SBS" have been field investigations. In a few studies, some features of the ideal experiment have been attempted (randomisation, use of controls, blinding or placebo).

In most "SBS" field studies, structured self-administered questionnaires have been used to collect information on the experience of annoyance and symptoms. It is difficult to compare the results between different studies since there are considerable differences in the questions about the type, quantity, intensity and time-frame of the symptoms. However, in most cases the questions have been consistent within each study. Comparison between buildings on different occasions and with different, small, statistically non-representative groups of respondents necessitates procedures for calibration or standardisation of such scales.

A major threat to the validity of most building intervention studies of "SBS" is information bias, which occurs if the intervention is not blinded to the occupants. Other methodological problems in such studies are natural changes with time and the tendency of the participants entered in the study to discontinue their participation; the probability of loss may be related to the effect of treatment.

11 CONCLUSIONS AND RECOMMENDATIONS

This section starts with a brief recapitulation of the purposes for and methods of sensory evaluations relevant to perceived indoor air quality. It continues with a description of for what purposes a given method can be applied. Then follows practical recommendations regarding (a) materials and components testing, (b) IAQ building evaluations, and (c) SBS assessments in building populations.

11.1 Basic Aspects

Good perceived air quality is essential for the well-being of the occupants. However, it should be noted that some harmful air pollutants are not perceived at all (e.g. carbon monoxide and radon) and the sensory effects of some other pollutants are not quantitatively linked with their toxicity. Therefore, perceived air quality cannot be used as the only measure of health effects.

The air pollution loads on indoor spaces typically comprise odorous and irritating compounds. Several approaches are described in the literature and used in practice to study various attributes of perceived air quality indoors. Some are based on measurements of various attributes of odour whereas others are intended to measure a broader range of perceptions. None of these approaches have been intercalibrated .

What method of sensory evaluation is selected will depend on the purpose. Examples of purposes are (a) development of new products or "airscaping" of buildings indoors, (b) the determination and control of permissible sensory loads or absolute sensory levels, (c) comparison and selection of best building materials, installation or building techniques, and (d) characterization of existing materials, spaces, buildings and occupant populations.

The selection of permissible loads or levels may be in the form of a detection measure, for example, number of dilutions needed to reach a criterion for odorlessness or the absolute criterion that emission from building materials should not induce sensory irritation in the average person of the general population or in a special group of sensitive individuals (e.g., persons with hypersensitivity to dust). Permissible level may also be determined from a known exposure-response relationship relating, for example, perceived odour intensity of air to ventilating air rate and ventilation efficiency. A permissible level may also be determined as a maximum scale value on a calibrated scale of perceived odour intensity, or, alternatively, as an opinion poll of percentage dissatisfied persons.

Comparative sensory evaluations would involve building materials, e.g., flooring materials, and the task would be to select the best of them according to one or several criteria (perceived odour intensity, sensory irritation detection, etc.). Comparisons could also be made within a group of materials over time if calibrated measurements are made in the sense that the groups of evaluaters are comparable. In a long term perspective, the development of intelligent systems for such comparisons is desirable and appears theoretically possible.

Many methods are available for determining sensory detectability. Some of these require that a dilution series be made of the original air sample and therefore its odour must be

fairly intense to start with. These methods include, i.a., the method of limit and the method of constant stimulus. They all exist in various versions and are aiming at avoiding various biases. The detectability measures, e.g., absolute thresholds, commonly are determined at an individual level of the panel members and thereafter being assessed as median values for the group. Signal detection is useful for detecting weak odours since the method does not require dilutions of the sample.

Regardless of the detection method employed a forced-choice response mode is needed with presentations which can be separated in time or space. Blanks (for example, 1/3 or more of the exposures being merely dilution air) will make sure the procedure works in a well-controlled way. Blanks are a prerequisite for documenting false alarm rates.

The false alarm rate for detection should be given for the background air (dilution air) at the limit concentration for detection determined in the test, or else at the just barely perceivable dilution concentration determined in the test. The distribution of odour detection frequency responses produced by the panel should be given for all concentrations tested, not only the single median threshold value. The median odour threshold concentration obtained for the group should be checked by evaluating its index of signal detectability (d').

Measures of perceived intensity of odours and sensory irritation are obtained from subjects by various psychological matching and scaling methods, i.a., equal-intensity matching and magnitude estimation. Direct scaling methods are the most common. Preferably, the perceived intensity measures shall be obtained from a calibrated scale (e.g. a master scale) so that measures can be compared between different materials, occasions and laboratories.

The composition of the panel is pending on the purpose of the test. For some purposes the test panel may comprise selected, sensitive and/or trained persons, at other times unselected and naive subjects are preferred. This purpose should be displayed. The panel members should not be staff members of the laboratory. The size of the panel should be statistically relevant to the purpose of the test. The fewer the panel members are, the more individual repetitions of test exposures are needed.

11.2 Materials and Components Testing

The ultimate goal of the sensory evaluation of emissions from building materials is to predict from laboratory evaluations, the consequences of the use of a material for the perceived air quality in actual buildings. Comparing such predictions with established standards or guidelines would provide one criterion for labelling or classification of the material. However, at present the ambition needs to be more restricted.

11.2.1 Models for prediction of real life exposures

Models are required for predicting human reactions to complex, real life exposures from the results of laboratory tests of individual materials. Specifically, these models are needed, e.g., to:

- transform sensory source characterisations in small scale settings into sensory characterisations of indoor air quality in full scale, actual environments;

- predict the relation between sensory responses and air pollutant concentrations and taking into account that these relationships may be non-linear;
- predict the sensory perception of emissions from a combination of sources using measurements made individually for each contributing source; and
- predict occupant responses in buildings using test panel responses in the laboratory.

For the time being, no accepted and validated models for all of the above purposes exist (ECA-IAQ, 1997a). Therefore, only in exceptional cases will laboratory evaluations be useful for an estimate of the consequences of a material emission for the perceived air quality in actual environments. An example of such an exceptional case may be the emission of a flooring material in the very first few days after installation in an otherwise unmodified environment when these emissions are high and predominant.

11.2.2 Criteria for materials and components testing

Any evaluation procedure of material emissions addressing adverse health effects should include sensory tests of odour and sensory irritation. The results of these tests should enable producers to exclude clearly unacceptable materials from being introduced into the market, and consumers to compare the emissions from different materials, or to rank the materials. Primarily, the evaluations should provide results given in terms of detectability and perceived intensity (or attributes closely linked to perceived intensity).

The air emissions of building materials should not give rise to *sensory irritation*. However, due to the large variations of sensitivity in the population (also including hypersensitive persons), the requirement of absence of irritation cannot be satisfied in absolute terms. Therefore, it has sometimes been specified in terms of a defined maximum percentage of those being exposed that report sensory irritation. At present there are no sensory emission standards and no related standard measurement methods. A task force of the World Health Organization (1989) has recommended that sensory irritants should not be present in indoor air in excess of their ED10 detection threshold (the concentration at which 10% of a population detect sensory irritation). Accordingly, emissions from single indoor air pollution sources should cause *considerably* less than 10 % of exposed persons to perceive sensory irritation.

From a measurement point of view the WHO criterion would be checkable when applied to a field situation where the number of exposed persons is large. However, in the laboratory setting it is not possible to use large enough samples of test subjects to measure such a low response rate in a reliable way. Other means might be possible, such as extrapolation from a dose-response curve or use of protection factors relative to an exposure with a higher response rate but these means are not discussed further in this report.

In contrast to sensory irritation, *odour* may cause discomfort in some persons but be perceived as indifferent or even pleasant by others. On the other hand, pervasive, strong and long lasting odours may be unacceptable to most people. Furthermore, the presence and/or strength of odorous emissions may be an important argument against the choice of a specific material by specific consumers. The task force of the World Health Organization

(1989) mentioned above has recommended that unwanted odorous compounds should not be present in indoor air in concentrations exceeding the ED50 detection threshold. From a measurement point of view, this WHO criterion for odour detectability may be verified by laboratory testing of single compounds. However, the question of whether a compound has an unwanted odour or not is less easily answered.

The Danish indoor climate labelling procedure determines the period required for all emitted chemicals of a product to gas off to a level below known health and sensory detection limits, based on chemical, toxicological and sensory literature data (Larsen, Nielsen and Wolkoff, 1996). A sensory assessment is performed to check whether the product meets specified votes criteria on category scales of odour intensity and acceptability of air. The required off-gassing period to fully meet the sensory criteria is reported (as days required) in the product standard. If more than 160 days (for some products less) are needed the product will be rejected.

11.2.3 Recommended methods

For materials and components testing the sensory evaluation is recommended to include detectability and/or perceived intensity (or an attribute closely linked to perceived intensity) of odour and/or sensory irritation. The choice of methods and sample of test subjects are pending on the purpose of the test. For quality control purpose, even in the case only perceived odour intensity is being measured the odour threshold concentration of the reference chemical (as measured or as derived from extrapolation from a dose-response relationship) should be tested by the signal detection approach and the hit and false alarm rates be displayed.

Measurements of detectability and/or perceived intensity should not only be made of the test sample but also of the reference chemical being used. The evaluation method and results should be displayed, including false alarms and specification of background air (dilution air) conditions, and adequate basic test requirements should be met. In addition, a qualitative characterization of the test sample may be made at conditions simulated to be typical for the intended use of the material.

The perceived odour intensity scale should be calibrated by use of a reference scale of at least five, or preferably more, concentrations of a reference odorant. The reference standard odorant shall be either acetone (2-propanone), dimethyl monosulfide, hydrogen sulphide, n-butanol or pyridine. The reference standard odorant used shall be traceable and have been compared with at least one other of the alternative reference odorants listed above: as to their absolute odour thresholds and their psychophysical exposure-response relationships for perceived odour intensity. In case another attribute than perceived odour intensity is being used, its psychophysical exposure-response data should be compared to perceived intensity data, e.g., in a plot of attribute-to-attribute comparisons at the same physical exposure levels. Preferably, odour detectability and perceived odour intensity may be measured jointly. All perceived intensity measures should be reported also as concentration equivalents of the reference chemical being used.

The following methods are recommended by the group members. The features of the methods are displayed in Table 1.

A. When high inter-laboratory comparability, resolution power and precision are required (as for the purpose of *labelling materials and products*):

- For the detection of odours
  1. Method of limit
  2. Method of constant stimulus

     (A combination of the two methods may be preferred in some cases. Detection measures also can be obtained by jointly measuring detectability and perceived intensity.)
  3. Signal detection index, d´
     (For checking the odour threshold concentration, as measured or arrived at by extrapolation.)
- For the determination of perceived intensity of odour and/or sensory irritation
  1. Equal-attribute matching
  2. Magnitude estimation with several references
  3. Master scaling
- A qualitative characterization of the test sample

B. When high intra-laboratory comparability and resolution power are required (as for the purpose of *guiding company product development*):

1. All of the methods above
2. Category scaling
3. Methods for multidimensional analysis

11.3 IAQ Evaluations

11.3.1 Criteria for IAQ evaluations

In determining the ventilation rate required for a space from a comfort point of view, a starting point would be to select the desired level of air quality in the ventilated space. Such a decision depends on comfort demands, and on economic and energy conservation considerations as well as on the use of the space. The shifting views on these matters have resulted in widely differing minimum ventilation requirements/guidelines issued by authorities and organisations (ECA-IAQ, 1996).

The sensory pollution load on the air is caused by those pollution sources having an impact on the perceived air quality. The total sensory pollution load in a space is the joint impact by all the sensory loads caused by all the different pollution sources. However, at present information on sensory source strength is available for only a few materials and reliable models are still lacking for prediction from laboratory to field conditions. In the European Database project a database has been developed, which comprises several dozens of sources and information on, i.a., perceived air quality in decipol or percentage of

dissatisfied (Clausen and de Oliveira Fernandes, 1997). A database is also being developed for sensory irritants for use in establishing occupational limits (Schaper, 1992).

The same criteria for adverse sensory effects as given in the WHO document mentioned above (WHO, 1989; see section 11.2.2) may have a bearing also on IAQ evaluations. From a measurement point of view the WHO detectability criteria would be checkable when applied to an IAQ field evalauation when the number of test subjects is large. However, as said before, in the laboratory setting it is not possible to use large enough samples of test subjects to measure low response rates in a reliable way. In the laboratory the 50 percentile value can be determined most reliably.

The established ASHRAE practice, now under revision (ASHRAE, 1996), is that the air of a space can be considered acceptably free from annoying contaminants if 80% of a panel of at least 20 untrained subjects do not express dissatisfaction under representative conditions of use and occupancy (visitor situation). This is a yes-no classification criterion on a nominal scale.

With the exception of a few studies most exposure-response relationships have not been demonstrated for perceived odour intensity but for other attributes (Yaglou et al. 1936; Fanger 1988). In the European Audit project study (Bluyssen et al., 1996), the average perceived air quality level indoors (acceptability) was 5.7 decipol (in physical equivalents 22 ppm acetone (2-propanone)) with a standard deviation of 2.2 decipol, and outdoors 1.9 decipol (4.8 ppm) with a standard deviation of 1.2 decipol. No levels below 2 decipol (5.2 ppm) were encountered.

Perceived odour intensity has been determined for the air of class rooms in school buildings and kitchens by the use of magnitude estimation, equal-intensity matching and master scaling (Berglund, Berglund and Lindvall, 1987). Perceived odour intensity at various ventilation conditions ranged in master scale units from 10 to 30 for class room air (in physical equivalents 20-125 ppb pyridine) and from 40 to 200 for kitchen air (in physical equivalents from 170 ppb pyridine and up) as shown in Figure 8. A criterion for perceived odour intensity has been suggested as the point at which occupant related odour clearly can be separated from the perceived background intensity of a space (Berglund and Lindvall, 1979).

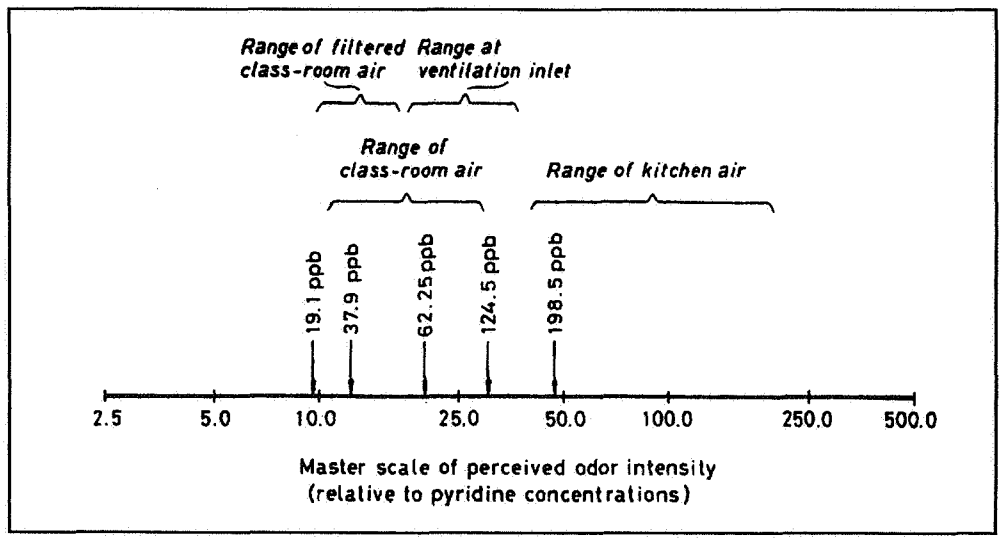

Figure 6. *Perceived odour intensity for different air samples marked on a master scale of odour intensity. The master scale is defined by five concentrations of pyridine (arrows) that were scaled jointly with the air samples (from Berglund, Berglund and Lindvall, 1987).*

No joint comparison has been made of magnitude estimates of attributes expressed in physical equivalents of acetone (2-propanone) as well as pyridine. Therefore, at present the decipol and master scale units cannot be compared.

Odour exposure may also be expressed by measures based on dilution factors related to defined odour levels. One such dilution measure is the odour unit which relates to the odour threshold. Although obtained by the use of the human sense of smell, odour units are exposure measures only, not measures of perceived odour intensity.

11.3.2 Recommended methods

Sensory IAQ evaluations are recommended to include detectability and/or perceived intensity (or an attribute closely linked to perceived intensity) of odour and/or sensory irritation. Measurements of perceived intensity should not only be made of the air sample being tested but also of the reference chemical being used. Air samples should be measured on-line and not as grab-sampling since the losses are expected to be significant from, i.a., adsorption and chemical reactions.

The perceived odour intensity scale should be calibrated by use of a reference scale of at least five, preferably more, concentrations of a reference odorant. The reference standard odorant shall be either acetone (2-propanone), dimethyl monosulfide, n-butanol or pyridine (in IAQ evaluations hydrogen sulfide cannot easily be handled in a safe way to suit as a reference chemical). The reference standard odorant used shall be traceable and have been compared with at least one other of the alternative reference odorants listed above: as to their absolute odour thresholds and their psychophysical exposure-response relationships for perceived odour intensity. In case another attribute than perceived odour intensity is being used, its psychophysical exposure-response data should be compared to perceived intensity data, e.g., in a plot of attribute-to-attribute comparisons at the same physical exposure levels. All perceived intensity measures should be reported also as concentration equivalents of the reference chemical being used.

The following methods are recommended by the group (see also Table 1).

A. When high inter-investigation comparability, resolution power and precision are required (as for the purpose of *auditing buildings*):

- For the detection of odours

  Signal detection index, d´
- For the determination of perceived intensity of odour and/or sensory irritation
  1. Magnitude estimation with several references
  2. Master scaling

B. When moderate intra-investigation comparability and resolution power are required (as for the purpose of comparisons of sources and interventions within a building):

1. All of the methods above
2. Category scaling

3. Magnitude estimation with memory reference
4. Magnitude estimation with one reference
5. Descriptor profiling
6. Classification (yes-no)

11.4 Population Response Studies in Buildings

11.4.1 Criteria for population response studies in buildings

As criteria for acceptability and annoyance, WHO (WHO, 1987) uses the nuisance threshold level, being defined as the concentration at which not more than a small proportion of the population (less than 5%) experiences annoyance for a small part of the time (less than 2%). However, for value judgements, like dissatisfaction and annoyance, experience from surveys in the outdoor environments indicates that prevalence rates below 10% cannot be distinguished from the background "noise" in the response data (Lindvall and Radford, 1973). This will hamper the application of the WHO-criterion for nuisance.

There are no agreed upon criteria for when occupants of a building should be judged as suffering from "SBS", or when a building should be classified as a "sick building" (Jaakkola, 1998).

11.4.2 Recommended methods

In population response studies in buildings, the choice of procedure is likely to depend on the size of the building being studied and the amount of time available for the study. The scope is greater the larger the buildings, the larger the populations and the longer the studies.

The population response evaluation can include measures of environment perceptions and body perceptions (symptoms) as well as perceived quality and value judgements. The evaluation method and results should be displayed, and adequate basic test requirements should be met. When perceived intensities and value judgements are being measured calibration should be strived for. Thereby, physical as well as memory references may be used, within the same modality or with a different one. If calibration cannot be made, the scales at least should be standardised by transforming the empirical response distributions to a standard distribution of z-scores.

The following methods are recommended by the group (see also Table 1).

A. When high inter-investigation comparability, resolution power and precision are required (as for the purpose of *comparing buildings or occupant groups*):

For the determination of perceived intensity of odour and/or sensory irritation

1. Magnitude estimation with several references
2. Category scaling
   (Calibration needed, e.g. after Thurstonian scaling, see sections 4.2 and 4.4)

B. When moderate intra-investigation comparability and resolution power are required (as for the purpose of *comparing systems and interventions within a building*):

1. All of the methods above
2. Category scaling
   (At least standardisation to z-scores is needed)
3. Magnitude estimation with one reference
4. Descriptor profiling
   (A thoroughly developed measuring instrument is needed)
5. Classification (yes-no)

In case another attribute than perceived intensity of odour (or perceived air quality) is being used, e.g. perceived degree of acceptability, its psychophysical exposure-response data should be compared to perceived intensity data achieved in an additional experiment. The purpose is to allow for a comparison between the attributes. All perceived odour intensity measures should be reported also as concentration equivalents of the reference odorant chemical being used.

The control group of subjects should be selected in the same building if possible, double blinding be made whenever possible and a short recall period be used in order to maximise recall accuracy.

Magnitude estimation scales of any attribute should be calibrated by use of a reference scale of at least three "concentrations" of a reference. The reference does not have to belong to the same modality (see sections 4.3 and 4.5.4); intra- as well as cross-modality comparisons are possible. When possible, all perceived magnitude estimates should be reported also as physical equivalents of the reference being used (mass concentrations, sound pressure levels, etc.). For odours, the reference should be either acetone (2-propanone), dimethyl monosulfide, n-butanol or pyridine (not hydrogen sulfide, see above). The reference standard odorant used shall be traceable and have been compared with at least one other of the alternative reference odorants listed above as to their psychophysical exposure-response relationships for perceived odour intensity.

In case another attribute than perceived odour intensity is being used, its psychophysical exposure-response data should be compared to perceived intensity data, e.g., in a plot of attribute-to-attribute comparisons at the same physical odorant exposure levels. All perceived odour intensity measures should be reported also as concentration equivalents of the reference chemical being used.

To establish a satisfactory comparability of controls it is recommended that:

- the control group(s) is selected in the same building if possible, otherwise in a nearby building with similar design and workforce;
- the unit is defined that is to be used as basis of matching or statistical control (e.g. people vs. rooms vs. buildings);
- where possible, experimental and control groups are matched for basic characteristics of

the starting conditions (gender, environment and health measures, etc.). If the groups are sufficiently large it may be possible to control for these variables in the statistical analysis, without matching;

- measurement artefacts are controlled for, e.g., some groups are used which are not blind to the interventions, in order to assess the effect of being aware of the changes.
- within-subject comparisons are made wherever possible, with repeated measures in the "after-change" condition in order to identify any carry-over of response between conditions;
- cross-over design is used where possible, with or without control groups depending on whether the change can be limited to part of the study population;
- monitoring of all relevant variables is made in the same time period, e.g., occupant response, indoor environment and extraneous effects such as a change in staff management. This will often mean monitoring the environment first, then seeking retrospective occupant evaluations of the period monitored;
- all factors and all samples are being assessed at the same time of day and week; and
- a short recall period is used in order to maximise recall accuracy.

11.5 Proposed Research and Development

In the following are given a few recommendations for future research and development:

1. Comprehensive measurement system for testing building materials that include all important aspects of the perception of indoor air. The system should include relevant attributes such as detectability, perceived air quality, perceived strength, and specific sensations (e.g., cooling, warmth, freshness/staleness, appeal, preference, intrusiveness, irritation).
2. Study of what subjects really are evaluating when requested to judge the acceptability of a stimulus either if it appears in an imaginary prescribed situation, or in a real situation.
3. Models for predicting sensory effects in the field situation from the results of laboratory investigations of source emission rates, including sensory as well as chemical loads, as basis for calculating required ventilation rates. Special attention should be devoted to studies of the interaction principles within and between similar (homogeneous) and dissimilar (heterogeneous) percepts of compounds and material emissions.
4. Signal detection approaches to be used for materials testing after an off-gassing period has passed.
5. Psychophysical models for IAQ-related hedonics.
6. Practical systems for comparing sensory evaluation methods and attributes.
7. Sensory profiling of human test subjects and means for the characterization and selection of test panels and members.
8. Means for making test panels consistent in performance over time in spite of drop-outs among the panel members.

9. Establishment of the relation between the results of non sensory methods for the detection of odours and human responses.
10. Establishment of a prestandardization working group aiming at CEN standards for sensory evaluation of IAQ, addressing prenormative issues such as comparison of reference chemicals, detailed prescriptions on equipments and conditions for exposure and stimulus presentation, test protocols, interlaboratory calibration, panel selection, and quality assurance.

REFERENCES

Alarie, Y. (1984). Establishing threshold limit values for airborne sensory irritants from an animal model and the mechanisms of olfaction of sensory irritants. *Adv Environ Toxicol*, 8, 153-164.

ASHRAE (1996). Standard 62-1989R, Proposed: Ventilation for acceptable indoor air quality, Appendix C: Air quality guidelines - Informative, Atlanta, USA: ASHRAE, August 1996.

ASTM (1968). Manual on sensory testing methods. ASTM E-18, American Society for Testing and Materials, USA, ASTM special technical publication 434.

ASTM (1981). Standard practices for referencing suprathreshold odour intensity. American Society for Testing and Materials, USA, Annual book for ASTM standards E544-76 (reapproved 1981), pp.32- 44.

ASTM (1984). Standard test method for estimating sensory irritancy of airborne chemicals. ASTM standard E981-84, American Society for Testing and Materials, USA.

ASTM (1997). Standard practices for referencing suprathreshold odor intensity. E 544-75, reapproved 1993, Vol 18.07, pp. 23-32.

Axel, R. (1995) The molecular logic of smell. *Scientific American*, October 1995, 130- 137.

Baird, J.C. and Noma, E. (1978). Fundamentals of scaling and psychophysics. ISBN 0-471-04169-6. New York: John Wiley & Sons.

Baird, J.C., Berglund, B., Berglund, U. and Lindvall, T (1990). Symptom patterns as an early warning signal of community health problems. *Environment International*, 16, 3-9.

Beck, L. H., Kruger, L., and Calabresi, P. (1954). Observations of olfactory intensity. I. Training procedure, methods, and data for two aliphatic homologous series. *Ann NY Acad Sci*, 58, 225-238.

Berglund, B. (1974) Quantitative and qualitative analysis of industrial odours with human observers. *Ann NY Acad Sci*, 237, 35-51.

Berglund, B., Berglund, U., Jonsson, E. and Lindvall, T. (1977). On the scaling of annoyance due to environmental factors. *Environ Psychol Nonverbal Behaviour*, 2, 83-92.

Berglund, B., Berglund, U. and Lindvall, T. (1971). On the principle of odour interaction. *Acta Psychologica*, 35, 255-268.

Berglund, B., Berglund, U. and Lindvall, T. (1974). Measurement of rapid changes of odour concentration by a signal detection approach. *J Air Pollut Contr Assoc*, 24,162-164.

Berglund, B., Berglund, U. and Lindvall, T. (1976). Psychological processing of odour mixtures. *Psychol Review*, 83, 432-441.

Berglund, B., Berglund, U. and Lindvall, T. (1978). Olfactory self- and cross-adaptation: Effects of time of adaptation on perceived odour intensity. *Sensory Processes*, 2, 191-197.

Berglund, B., Berglund, U. and Lindvall, T. (1984). Characterisation of indoor air quality and "sick buildings". *ASHRAE Transactions*, 90, 1045-1055.

Berglund, B., Berglund, U. and Lindvall, T. (1986). Theory and methods for odour evaluation. *Experentia*, 42, 280-287.

Berglund, B., Berglund, U. and Lindvall, T. (1987). Measurement and control of annoyance. In H.S. Koelega (ed.) Environmental Annoyance: Characterisation, Measurement and Control. Elsevier.

Berglund, B., Berglund, U., Lindvall, T. and Nicander-Bredberg, H. (1982). Olfactory and chemical characterisation of indoor air. Towards a psychophysical model for air quality. *Environment International*, 8, 327-332.

Berglund, B., Berglund, U., Lindvall, T. and Svensson,L.T. (1973). A quantitative principle of perceived intensity summation in odour mixtures. *J Exp Psychol*, 100, 29-38.

Berglund, B., Berglund, U. and Lindvall, T. (1975). Scaling of annoyance in epidemiological studies. Proc. International Symposium "Recent Advances in Assemssment of the Health Effects of Environmental Pollution", vol. I, pp. 119-137. Luxembourg: Commission of the European Communities.

Berglund, B., Jaakola, J., Raw, G. and Valbjørn, O. (1996). Sick building syndrome: The design of intervention studies. CIB Report 199. Rotterdam: International Council for Building Research Studies and Documentation.

Berglund, B and Lindvall, T. (1979). Olfactory evaluation of indoor air quality. In: Fanger and Valbjørn (Eds.), Proceedings of *INDOOR CLIMATE*, Danish Building Research Institute, Copenhagen, pp. 141-157.

Berglund, B. and Lindvall, T. (1982). Olfaction. In I. Andersen and D. Proctor (eds.), The Nose: Upper Airway Physiology and the Atmospheric Environment. Amsterdam: Elsevier, pp. 279-305.

Berglund, B. and Lindvall, T. (1986). Sensory reactions to "sick buildings". *Environment International*, 12, 147-159.

Berglund, B. and Lindvall, T. (1990). Sensory criteria for healthy buildings. In: D.S. Walkinshaw (ed.), Proceedings of *INDOOR AIR '90*, 5th International Conference on Indoor Air Quality and Climate, Toronto, Canada, vol. 5, pp 65-78.

Berglund B. and Lindvall T. (1992). Questionnaires in exposure and effect assessment in the field. In H. Knoeppel and P. Wolkoff (eds.) Chemical, Microbiological, Health and Comfort Aspects of Indoor Air Quality - State of the Art in SBS. Dordrecht / Boston / London: Kluwer Academic Publishers, pp. 263-273.

Berg-Munch, B., Clausen, G. and Fanger, P.O. (1986). Ventilation requirements for the control of body odour in spaces occupied by women, *Environmental International*, 12, 195-199.

Bernstein, I.H., Garbin, C.P. and Teng, G.K (1988). Applied Multivariate Analysis. New York: Springer-Verlag.

Bluyssen, P., Kondo H., Pejtersen,J., Gunnarsen, L., Clausen, G. and Fanger, P.O. (1989). A trained panel to evaluate perceived air quality. Proceedings of *CLIMA 2000*, Sarajevo, vol. 3, pp 25-30.

Bluyssen, P.M. (1990). Air quality evaluated by a trained panel. Ph.D. thesis, Laboratory of Heating and Air Conditioning, Technical University of Denmark. October 1990.

Bluyssen, P.M. and Fanger, P.O. (1991). Addition of olfs from different pollution sources. *Indoor Air*, 4, 414-421.

Bluyssen, P.M., de Oliveira Fernandes, E., Groes, L., G., Fanger, P.O., Valbjørn, O., Bernhard, C.A. and Roulet, C.A. (1996). European Audit project to optimise indoor air quality and energy consumption in office buildings. *Indoor Air*, 6,221-238.

Bluyssen, P.M. and Cornelissen H.J.M. (1997). Addition of olfs: simple or not? In: Prodeedings of *HEALTHY BUILDINGS/IAQ '97*, September 27-October 2, Washington DC, USA, vol. 2, pp. 213-218.

Bluyssen, P.M. and Cornelissen, H.J.M. (1998). Addition of sensory pollution loads: simple or not that is the question. Submitted to *ASHRAE Transactions*.

Bluyssen, P.M. and Walpot, J.I. (1993). Sensory evaluation of perceived air quality - a comparison of the threshold and the decipol method. In: J.J.K. Jaakola, R. Ilmarinen and O. Seppänen (eds.), Proceedings of *INDOOR AIR '93*, 6th International Conference on Indoor Air Quality and Climate, Helsinki, vol. 1, pp. 65-70.

Brown, K., MacLean, C.M. and Robinette, R.R. (1968). The distribution of the sensitivity to chemical odours in man. *Human Biol*, 40, 456-472.

Bruckman, H.W.L. et al. (1994), Kunststof CO-sensor, TNO-Industrie, TNO-rapport 0795/U94.

Butcher, S.S. and R.J. Charlson (1972). An Introduction to Air Chemistry. New York and London: Academic Press.

Cain, W.S. (1969). Odour intensity: Differences in the exponent of the psychophysical function. *Perception & Psychophysics*, 6, 349-354.

Cain W.S. (1977). Odour magnitude: coarse vs. fine grain. *Perception & Psychophysics*, 22, 545-549.

Cain, W.S. (1992). Irritation versus odour. NIVA-Course on Sick Building Syndrome - Measurement and Evaluation. Noordwijk aan Zee, The Netherlands, 11-16 October 1992.

Cain, W.S. and Cometto-Muniz, J.E. (1993). Irritation and odour: symptoms of indoor air pollution. In: J.J.K. Jaakola, R. Ilmarinen and O. Seppänen (eds.), Proceedings of *INDOOR AIR '93*, 6th International Conference on Indoor Air Quality and Climate, Helsinki, vol. 1, pp. 21-31.

Cavalini, PM. (1992) It's an ill wind that brings no good: Odour annoyance and the dispersion of odorant concentrations from industries. Ph.D. thesis. Groningen, 1992

Cavalini, P.M., Koeter-Kemmerling, L.G. and Pulles, M.P.J. (1991). Coping with odour annoyance and odour concentration: Three field studies. *J Environ Psychol*, 11, 123-142

CEN (1995). Odour concentration measurement by dynamic olfactometry. TC 264, WG2 Final Draft pr EN, Document 063/e. Brussels: European Committee for Standardization (CEN).

Clark, C.R. (1984). The effects of noise on health. In: D.M. Jones and A.J. Chapman (eds.) Noise and Society. Chichester: Wiley, pp. 11-124

Clausen, G., Carrick, L., Fanger, P.O., Kim, S.W., Poulsen, T. and Rindel, J.H. (1993). A comparative study of discomfort caused by indoor air pollution, thermal load and noise. *Indoor Air*, 3 (4), 255-262.

Clausen, G., Pejtersen, J., and Bluyssen, P.M. (1993). Final Research Manual of European Audit project to optimise Indoor Air Quality and Energy consumption in Office Buildings. Edited by Technical University of Denmark and TNO-Building and Construction Research, Copenhagen and Delft.

Clausen, G. and de Oliveira Fernandes, E. (1997). European data base on indoor air pollution sources in buildings. Final report contract JOU2-CT93-0343, volume 1. European Commission: Joule II Programme, Energy Conservation and Utilisation,.

Cometto-Muniz, J.E. and Cain, W.S. (1995) Relative sensitivity of the ocular trigeminal, nasal trigeminal and olfactory systems to airborne chemicals. *Chemical Senses*, 20, 191-198.

Cometto-Muniz, J.E., Cain,W.S. and Hudnell,H.K. (1997). Agonistic sensory effects of airborne chemicals in mixtures: Odour, nasal pungency, and eye irritation. *Perception & Psychophysics*, 59, 665-674.

Corbit, T.E. and Engen, T. (1971). Facilitation of olfactory detection. *Perception and Psychophysics*, 10, 433-436.

Cornsweet, T. (1962). The stair-case method in psychophysics. *Amer J Psychol*, 75, 485

Costa, P.T. and McCrae, R.R. (1992). NEO PI-R (Revised NEO Personality Inventory). Professional Manual. © Psychological Assessment Resources Inc., P.O. Box 998, Odessa, Florida 33556, U.S.A.

Craik KH. (1987) Environmental perception and environmental annoyance: issues of measurement and interpretation. In: H.S. Koelega (ed.), Environmental Annoyance: Characterisation, Measurement and Control. Amsterdam: Elsevier, pp. 45-49

Devos, M., Patte, F., Ronault, J., Laffort, P., and van Gemert, L.J. (1990). Standardised human olfactory thresholds. New York: IRL Press at Oxford University Press.

Doty, R., ed. (1995). Handbook of olfaction and gustation. New York: Dekker.

ECA-IAQ (1992). Guidelines for ventilation requirements in buildings. European Concerted Action "Indoor Air Quality and Its Impact on Man" (ECA-IAQ), Report no 11. Luxembourg: Office for publications of the European Communities.

ECA-IAQ (1996). Indoor air quality and the use of energy in buildings. European Collaborative Action "Indoor Air Quality and Its Impact on Man" (ECA-IAQ), Report No. 17, EUR 16367 EN, ISBN 92-827-6347-1. Luxembourg: Office for Official Publications of the European Communities.

ECA-IAQ (1997a). Evaluation of VOC emissions from building products. Solid flooring materials. European Collaborative Action "Indoor Air Quality and Its Impact on Man" (ECA-IAQ), Report No. 18, EUR 17334 EN, ISBN 92-828-0348-8. Luxembourg: Office for Official Publications of the European Communities.

ECA-IAQ (1997b). Total volatile organic compounds (TVOC) in indoor air quality investigations. European Collaborative Action "Indoor Air Quality and Its Impact on Man" (ECA-IAQ), Report No. 19, EUR 17675 EN, ISBN 92-828-1078-X. Luxembourg: Office for Official Publications of the European Communities.

Ekman,G., Berglund B., Berglund, U. and Lindvall T. (1967). Perceived intensity of odour as a function of time of adaptation. *Scand J Psychol*, 8, 177-186.

Ema, K. et al. (1989). Odour-sensing system using a quartz-resonator sensor array and neural-network pattern recognition. *Sensors and Actuators*, 18, 291-296.

Engen, T. (1972). Psychophysics. I. Discrimination and detection. In J.W. Kling and L.A. Riggs Woodworth and Schlosberg's experimental psychology, vol. l: Sensation and perception. New York: Holt, Rinehart and Winston, pp.11-46

Fanger, P.O. (1988). Introduction of the olf and the decipol units to quantify air pollution perceived by humans indoors and outdoors. *Energy and Buildings*, 12, 1-6.

Fanger, P.O., Lauridsen, J., Bluyssen, P.M., and Clausen, G. (1988). Air pollution sources in offices and assembly halls, quantified by the olf unit. *Energy and Buildings*, 12, 7-19.

Fanger, P.O. and Berg-Munch, B. (1983). Ventilation and body odour. Proc. Management of Atmospheres in Tightly Enclosed Spaces. Atlanta: ASHRAE, pp. 45-50.

Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin*, 87, 564-567.

Gardner, J.W. and P.N. Bartlett (eds.) (1992). Sensors and Sensory Systems for an Electronic Nose. NATO ASI Series, vol. 212.

Gardner, J.W., et al. (1990). The design of an artificial olfactory system, In: D. Schild (ed.), Chemosensory Information Processing, vol.H39, Springer-Verlag, pp.l31-173.

Garriga-Trillo, A. (1985). Funcion psicofisica y medida de la sensibilidad olfativa. Unpublished thesis, Universidad Autonoma de Madrid.

Garriga-Trillo, A. (1987). Olfactory psychophysics: Sensitivity measures. In E.E. Roskam and R. Suck (eds.), Progress in Mathematical Psychology - l. Amsterdam: North Holland.

Garriga-Trillo, A., Borg, G., Baird, J. and Niñòn, P. (1994). Personality and psychophysical judgements. Symposium conducted at the 23rd International Congress of Psychology (Abstracts, SY ASS (9) 0, p.86), Madrid, Spain.

Garriga-Trillo, A. (1996). How could you say that magnitude estimates are correct? An application to calibration studies. In: S. Masin (ed.), Fechner Day 96, Proc. 12th Annual Meeting of the International Society for Psychophysics, Padua, pp. 257-262.

Garriga-Trillo, A. (1997). Probability distributions for new sensitivity measures. In A. Preiss (ed.) Fechner Day 97. Poznan: International Society for Psychophysics, pp. 257-262.

Gescheider, G. (1985). Psychophysics: Method, Theory and Application. Hillsdale, N.J.: LEA.

Green, D..M. and Swets, J.A. (1966). Signal Detection Theory and Psychophysics. New York: Wiley.

Guirao, M. (1987). Estimaciones numericas de brillo y luminosidad con y sin rangos fijos. *Revista Latinoamericana de Psicologia*, 19, 297-317.

Gunnarsen, L. and Fanger, P.O. (1992). Adaptation to indoor air pollution. *Environment International*, 18, 43-54.

Gunnarsen, L., Nielsen, P.A. and Wolkoff, P. (1994). Design and characterisation of the CLIMPAQ, Chamber for laboratory investigation of materials, pollution and air quality. *Indoor Air*, 4, 56-62.

Guski R. (1987) Lärm. Wirkungen unerwünschter Geräusche. Bern: Huber.

Hangartner, M., Hartung, J. and Voorburg, J.H. (1985). Recommendations on olfactometric measurements. *Environ Technology Letters*, 6, 415-420.

Hempel-Jørgensen, A., Kjærgaard, S.K., Mølhave, L., (1997). Integration in human eye irritation. *Internat Arch Occup Environ Health*, 69, 289-294.

Heustache H., Allois J., Bonnefond C., Jacquot A., Lorant A. and Mevel V. (1988), *Analysis*, 16, 37.

ISO 6658 (1985). Sensory analysis - Methodology - General Guidance. ISO TC 34 (Agricultural food products), SC 12 (Sensory analysis).

Iwashita, G. and Kimura, K. (1995). Addition of olfs from common indoor air pollution sources measured with Japanese subjects. CIB Working Group WG7 "Indoor Climate", Milan, Conseil International de Bâtiments.

Jaakkola, J.J. (1998). The office environment model: A conceptual analysis of the Sick Building Syndrome. *Indoor Air*, Suppl. 4, 7-16.

Jensen B., Wolkoff P. and Wilkins C.K.(1993), Characterization of Linoleum. Part 2: Preliminary odour evaluation. In: K. Saarela, P. Kalliokoski and O. Seppänen (eds.), Proceedings of *INDOOR AIR '93*, 6th International Conference on Indoor Air Quality and Climate, Helsinki, vol. 2, pp. 449-454.

Johansson, I. (1990). Volatile organic compounds in indoor air. Stockholm: IMM, Karolinska Institute. IMM-report 6/90. (English summary).

Jones, F.N. (1956). A forced-choice method of limit. Amer J Psychol 69, 672.

Karpe P. (1995). Analyse physico-chimiques des phases volatiles émises par les revêtements de sol et de mur. Caractérisation des composés odorants et des odeurs des matériaux. Thèse, University of Paris 6, December 1995.

Karpe P., Kirchner S. et Rouxel P. (1995a). Sensory evaluation of wall and floor coverings using odour intensity method. Proceedings of *HEALTHY BUILDINGS '95*, International Conference on Healthy Building in Mild Climate, Milan, 11-14 September 1995, vol. 2, pp. 935-941.

Karpe P., Kirchner S. et Rouxel P. (1995b). Thermal desorption-gas chromatography-mass spectrometry-flame ionization detection-sniffer multi-coupling : a device for the determination of odorous volatile organic compounds in air. *J Chromatography* A, 708, 105-114.

Kastka J, Noack R. and Mau U. (1986) Längs- und Querschnittsuntersuchungen zur Geruchsbelästigung bei Katastererhebungen des Landes Nordrhein-Westfalen. In: Verein Deutscher Ingenieure (Hrsg.) Düsseldorf: VDI-Verlag, pp. 255-270.

Kastka J. (1976) Untersuchungen zur Belästigungswirkung der Umweltbedingungen Verkehrslärm und Industriegerüche. In: Kaminski G (Hrsg.) Umweltpsychologie: Perspektiven - Probleme - Praxis. Stuttgart: Klett, 187-223.

Khiari D., Brenner L., Burlingame G.A. and Suffet I.H. (1997). *Wat Sci Tech*, 25, 97.

Kirk R.L. and Stenhouse N.S. (1953). Ability to smell solutions of potassium cyanide. *Nature,* 171, 698-699.

Knudsen, H.N., Clausen, G. and Fanger, P.O. (1997). Sensory characterisation of emissions from materials. *Indoor Air*, 7(2), 107-115.

Kobal, G. and Hummel, T. (1991). Olfactory evoked potentials in humans. In T.V. Getchell et al. (eds.) Smell and taste in health and disease. New York: Raven Press, pp. 255-276.

Koelega, H.S. (ed.) (1987). Environmental Annoyance: Characterisation, Measurement and Control. Amsterdam: Elsevier.

Koelega H.S. and Köster E.P. (1974). Some experiments on sex differences in odour perception. *Ann NY Acad Sci*, 237, 234-246.

Krantz, D.H., Luce, R.D., Suppes, P. and Tversky, A. (1971). Foundations of Measurement, vol. I. New York: Academic Press.

Kroenke K. and Price R. (1993) Symptoms in the community: Prevalence, classification and psychiatric comorbidity. *Arch Intern Med*; 153: 2474-2480

Kruger, L., Feldzamen, A. N. and Miles, W. R. (1955) A scale for measuring supra-threshold olfactory intensity. *Amer J Psychol*, 68, 386-395.

Larsen, A., Nielsen, P. and Wolkoff, P. (1996). Labelling of emission from building products: First year status. In: Proceedings of *INDOOR AIR '96*, 7th International Conference on Indoor Air Quality and Climate, Nagoya, Japan, vol. 1, pp 353-358.

Lawless, H.T. and Klein, B.P. (Eds.) (1991). Sensory Science Theory and Applications in Foods. New York: Marcel Dekker, Inc.

Last, J.M. (Ed.) (1988). A Dictionary of Epidemiology. Oxford University Press, New York.

Lauridsen, J., Pejtersen, J., Mitric, M., Clausen, G., and Fanger, P.O. (1988). Addition of olfs for common indoor pollution sources. Proceedings of *HEALTHY BUILDINGS '88*. Stockholm, Sweden: Swedish Council for Building Research, vol. 3, pp. 189-195.

Lazarus R. and Launier R. (1978). Stress-related transactions between person and environment. In: L.A. Pervin and M. Lewis (eds.), Internal and External Determinants of Behaviour. New York, Plenum Press, pp. 287-327

Lindvall, T. (1970) On sensory evaluation of odorous air pollutant intensities. *Nordisk Hygienisk Tidskrift*, Suppl. 2., 1-181.

Lindvall, T. (1973) Sensory measurement of ambient traffic odour. *J Air Poll Contr Ass*, 23, 697-700

Lindvall, T. (1974) Monitoring odorous air pollution in the field with human observers. *Ann NY Acad Sci*, 237, 247-260

Lindvall, T. and Radford, T.P. (1973) Measurements of annoyance due to exposure to environmental factors. *Environmental Research*, 6, 1-36

Lindvall ,T. and Svensson L.T. (1974) Equal unpleasantness matching of malodorous substances in the community. *J Appl Psychol*, 59, 264-269.

Lodge, J.P. (ed.) (1989). Methods of Air Sampling and Analysis. 3rd edition. Intersociety Committee, Lewis Publishers, Inc.

Lord, F.M. and Novick, M.R. (1968). Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley.

Luce, R.D., Bush, R.R. and Galanter, E. (eds.) (1963) Handbook of mathematical psychology. New York: John Wiley and Sons, Inc.

Meilgaard, M., Civille, G.V. and Carr, B.T. (1991). Sensory Evaluation Techniques. Boca Raton: CRC Press.

Mølhave, L., Bach, B., and Pedersen, O.F. (1986). Human reactions to low concentrations of volatile organic compounds. *Environment International*, 12, 167-175.

Mølhave, L., Hempel-Jorgensen, A. and Kj(rgaard, S.K. (1993). Integration and adaptation in eye irritation. In: J.J.K. Jaakola, R. Ilmarinen and O. Seppänen (eds.), Proceedings of *INDOOR AIR '93*, 6th International Conference on Indoor Air Quality and Climate, Helsinki, vol. 1, pp. 35-40.

Mølhave, L., Jensen, J.G. and Larsen, S. (1991). Subjective reactions to volatile organic compounds as air pollutants. *Atmosph Environm*, 25, 1283-1293.

Mølhave, L. (1997). The ethics of using human panels in the indoor air sciences. *Indoor Air*, Suppl. 4, 87-95.

Nakamoto, T., Fukunishi, K. and Moriizumi, T. (1990). Identification capability of odour sensor using quartz-resonator array and neural-network pattern recognition. Sensors and Actuators, B 1, pp.473-476.

Narems, L. (1985). Meaningfulness and the Erlanger program of Felix Klein. Paper presented at the European Mathematical Group Meeting in Osnabrück.

National Academy of Sciences. (1979). Odours from Stationary and Mobile Sources. Washington, D.C.: National Academy Press.

Neotronics Scientific (1994), Neotronics Olfactory Equipment, Issue 2. Essex, UK.

Neutra, R., Lipscomb, J., Satin, K. and Shusterman, D. (1991). Hypotheses to explain the higher symptom rates observed around hazardous waste sites. *Environmental Health Perspectives*, 94, 31-38.

Nielsen, G.D., Hansen, L.F. and Wolkoff, P. (1997) Chemical and biological evaluation of building material emissions. II. Approaches for setting indoor air standards or guidelines for chemicals. *Indoor Air*, 7, 17-32.

Noma, E., Berglund, B., Berglund, U., Johansson, I. and Baird, J.C. (1988). Joint representation of physical locations and volatile organic compounds in indoor air from a healthy and a sick building. *Atmospheric Environment*, 22, 451-460.

NRC (National Research Council) (1992). Multiple Chemical Sensitivities - a Workshop, Working Group 1: Research protocol for clinical evaluation. Washington, D.C.: National Academy Press, p. 6.

Oseland, N.A. and Raw, G. (1993). Perceived air quality: a discussion of the new units. *Build Serv Eng Res Tech*, 14, 137-141.

Pangborn, R.M., Berg, H.W., Roessler, E.B. and Webb,A.D. (1964). Influence of methodology on olfactory response. *Percept.Mot Skills*, 18, 91-103.

Parine, N. (1996). Sensory evaluation of indoor air quality by building occupants versus trained and untrained panels. *Indoor Built Environ*, 5, 34-43.

Patte F., Etcheto M. and Laffort P. (1975). Selected and standardised values of suprathreshold odour intensities for 110 substances. *Chem Sens Flav*, 1, 282-305.

Ramalho O. (1998). CSTB Report EAE/SB N° 98007, 42p.

Ramalho O., Regoui C. & Kirchner S. (1997). Discrimination of paint emission using an electronic nose, Proceedings of the 4th International Conference on Characterization and Control of Emissions of Odors & VOCs, Octobre 1997, Montreal (Canada), pp. 409-418.

Roberts, F.S. (1976). Discrete Mathematical Models, with Applications to Social, Biological and Environmental Problems. Englewood Cliffs, N.J.: Prentice Hall, Inc.

Roberts, F.S. (1979). Measurement and Utility. Encyclopaedia of Mathematics and Its Applications. Reading, Mass.: Addison-Wesley.

Russel, J.A. and Pratt, G. (1980) A description of the affective quality attributed to environments. *J Personality Social Psychol*, 38, 311-322

Santos, A.M.B. and Gunnarsen, L. (1997). Trade-off between temperature and noise, air velocity and window area during chamber tests. Proceedings of *HEALTHY BUILDINGS/IAQ '97*, September 27-October 2, 1997, Washington DC, USA, vol. 2, pp. 41-46

Sävenhed R., Borén H. and Grimvall A. (1985). Stripping analysis and chromatographic sniffing for the source identification of odorous compounds in drinking water. *J. Chromatography*, 328, 219-231.

Schaper, M. (1992). Development of a database for sensory irritants and its use in establishing occupational limits. *Am Ind Hyg Assoc J*, 54, 488-544.

Steinheider, B. and Winneke, G. (1993) Industrial odours as environmental stressors: Exposure-annoyance-associations and their modification by coping, age and perceived health. *J Environ Psychol*, 13, 353-363

Steinheider, B., Winneke, G. and Schlipköter, H.W. (1993) Somatische und psychische Wirkungen intensiver Geruchsimmissionen. Eine Fallstudie aus der Substratherstellung für die Champignonzucht. *Staub - Reinhaltung der Luft*, 53, 425-431

Steinheider B. (1997) Die Wirkungen von Industriegerüchen als Umweltstressoren. Wiesbaden: Deutscher Universitäts-Verlag.

Stevens, J. (1992). Applied Multivariate Statistics for the Social Sciences. Hillsdale, N.J.: LEA.

Stevens, S.S. (1951). Mathematics, Measurement and Psychophysics. In S.S. Stevens (Ed.), Handbook of Experimental Psychology. New York: Wiley.

Stevens, S.S. (1957). On the psychophysical law. *Psychol Review*, 64, 153-181.

Stevens, J. C., Mack, J. D., and Stevens, S. S. (1960). Growth of sensation on seven continua as measured by force of handgrip. *J Exp Psychol*, 59, 60-67.

Stevens, J. C. and Stevens, S. S. (1960) Warmth and cold dynamics of sensory intensity. *J Exp Psychol*, 60, 183-192.

Stone, H. and Bosley, J.J. (1965). Olfactory discrimination and Weber's Law. *Percept Mot Skills*, 20, 657-665.

Sundell, J. (1994). On the association between building ventilation characteristics, some indoor environmental exposures, some allergic manifestations and subjective symptom reports. Indoor Air, Suppl. 2, 1-148.

Sundell, J. and Lindvall, T. (1993) Indoor air humidity and the sensation of dryness as risk indicators of SBS. *Indoor Air*, 3, 382-390.

Suppes, P. and Zinnes, P. (1963). Basic measurement theory. In R.D. Luce, R.R. Bush and E. Galanter (eds.), Handbook of Mathematical Psychology, vol. 1. New York: John Wiley and Sons, Inc., pp. 1-76.

Svensson, L.T. and Lindval, T. (1974). On the consistency of intramodal intensity matching in olfaction. *Perception & Psychophysics*, 16, 264-270.

Svensson L.T. and Szczygiel K. (1974). A matching method of successive approximations and its instrumentation. *Behaviour Research Methods & Instrumentation*, 6, 13-18.

Tanner,W. and Swets,J. (1954). A decision making theory of visual detection. *Psychol Rev*, 61, 401.

Taylor A.J. and Mottram D.S. (1990). Composition and odour of volatiles from autoxidised methyl arachidonate. J. Sci. Food Agric., 50, 407-417

Thurstone, L.L. (1927, 1994). A law of comparative judgment. *Psychol Rev*, 34, 273-286, 1927 (reedited Psychol Rev, 101, 266-270, 1994).

Torgerson, W.S. (1958). Theory and Methods of Scaling. New York: Wiley.

Van de Pol, F.J.R. (1989). Issues of Design and Analysis of Panels. Amsterdam, Sociometric Research Foundation.

VDI (1986). Olfactometry, odour threshold determination, fundamentals. VDI Richtlinie 3881, Blatt 1/part 1. Düsseldorf: Verein Deutscher Ingenieure.

Wargocki, P., Clausen, G. and Fanger, P.O. (1996). Field study on addition of indoor air sensory pollution sources. Proceedings of *INDOOR AIR '96*, 7th International Conference on Indoor Air Quality and Climate, Nagoya, Japan, July 21-26; vol. 4, pp. 307-312.

Wenzel, B. (1949). Differential sensitivity in olfaction. *J Exp Psychol*, 39, 129-143.

Wessely, S. (1997) Psychological, social and media influences on the experience of somatic symptoms. Paper presented at the ESF-Workshop "Cognitive Factors as Mediators of Environmental Effects on Health", September 15-17, Ottrott (France).

Whissel-Buechy, D. and Amoore J.E. (1973). Odour-blindness to musk: simple recessive inheritance. *Nature*, 242, 271-273.

WHO (1987). Air quality guidelines for Europe. WHO Regional Publications, European Series No. 23. Copenhagen: World Health Organisation, Regional Office for Europe.

WHO (1989). Indoor air quality: Organic pollutants. EURO Reports and Studies No. 111. Copenhagen: World Health Organisation, Regional Office for Europe.

Winneke, G. and Kastka J. (1987) Comparison of odour annoyance data from different industrial sources: problems and implications. In: H.S. Koelega (ed.): Environmental Annoyance: Characterisation, Measurement and Control. Amsterdam: Elsevier, pp. 129-137

Winneke, G., Neuf, M. and Steinheider, B. (1996) Separating the impact of exposure and personality in annoyance response to environmental stressors, particularly odours. *Environment International*, 22, 73-81

Winneke, G. and Neuf, M. (1992) Psychological response to environmental stressors: trait or state? *Applied Psychology*, 41, 257-267

Yaglou, C.P., Riley, E.C. and Coggings, D.I. (1936). Ventilation requirements. *ASHVE Transactions*, 42, 133-162.

Yaglou, C.P. and Witheridge, W.N. (1937). Ventilation requirements (Part 2). *ASHVE Transactions*, 43, 423-436.

VOCABULARY

For the purpose of this document the following vocabulary is being used.

Adaptation: Adaptation is the property of sensory receptors to adjust their sensitivity in response to the level of exposure (i.e. the higher the exposure the lower becomes the sensitivity and vice versa). Adaptation typically takes place during the first minutes of exposure. Habituation is when people after repeated exposures get used to the exposure by high level processing of the sensory information. The sensory information is still present but it is disregarded.

Additivity of effects: A condition that exists if the effect of changing the level of an independent variable, A, is not influenced by the level of another independent variable, B, and vice versa. This definition can be extended to more than two variables. It is important to observe that the property of additivity depends on the scales in which the measured value and the independent variables are expressed.

Annoyance: In indoor air sciences (IAS), a term describing a mood state which may be associated with environmental stress. It is in IAS referring to problematic, provocational, or displeasing environmental exposures.

Climate chamber: In IAS a room with one or more controlled atmospheric conditions or elements. A climate chamber is often used in IAS for human or animal exposure experiments or for source emission characterization. (Bluyssen et al.1997; WG10 report).

Comfort: (1) A state of optimal health, i.e., optimal physical and psychological well-being of humans considering all relevant covariables. (2) A state of being free from annoyance.

Common chemical sense: A set of senses including n. →Trigeminus, and non-myelinated nerves in the skin areas and in the eyes, face, and in part of the nose and mouth cavity. Mostly found in the facial skin, the mucosal membranes of eyes, nose, and mouth. These nerves respond in an unspecific way to chemical or physical stimulation. The stimulation is generally perceived as itching, feeling or dryness or pain.

Decipol: Unit of perceived air quality (see chapter 4.5.3).

Epithelium: The cellular covering of internal and external surfaces of the body, including the lining of vessels and other small cavities. It consists of cells joined by small amounts of cementing substances. Epithelium is classified into types on the basis of the number of cell layers and the shape of the superficial cells.

Hedonic tone: The quality of a feeling of pleasure following a sensory perception.

Kallman syndrome (Olfacto-genital syndrome): Reduced function of genital glands often including anosmia (i.e. the lack of the sense of smell).

Mucosa: The membrane covered with epithelium that lines the tubular organs of the body.

Multiple Chemical Sensitivity (MCS): Sensitivity to chemicals consisting in symptoms or signs related to chemical exposures at levels tolerated by the population at large that is

distinct from such well recognized hypersensitivity phenomena as IgE-mediated immediate hypersensitivity reactions, contact dermatitis, and hypersensitivity pneumonitis. Sensitivity may be expressed as symptoms and signs in one or more organ systems. Symptoms and signs wax and wane with exposures (NRC, 1992).

Odour units: 1 odour unit is the amoung of an odour carrier which - diluted in 1 m^3 of neutral air - just provokes an odour sensation according to the definition of the odour threshold.

Olf: A unit for the emission rate of indoor air pollutants. One olf is defined as the emission rate causing the same level of dissatisfaction as bioeffluents from one seated person.

Perceived environmental quality: The mental evaluation of the environment (e.g. indoor air) made by humans with respect to sensory experience and/or acceptability/annoyance. Other psychological aspects may be included, such as attractivesness and induced mood.

Perceived air quality (PAQ): The air quality as rated by humans in subjective evaluations (cfr. measured air quality).

Perceived indoor air quality (PIAQ): Indoor air quality as rated by humans in subjective evaluations (cfr. measured air quality). *PIAQ* is often rated with terms such as: need of additional ventilation, stuffiness.

Potency: The power of an agent to produce an effect (often measured as the slope of the dose-response-curve at ED_{50}).

Prevalence: The total numbe of cases of a specific disease in a given population at a certain time.

Receptor: (1) A molecule on the surface or within a cell that recognizes and binds with specific molecules, producing some effect in the cell; e.g., the cell-surface receptors of immunocompetent cells that recognize antigens, complement components, or lymphokines, or those of neurons and target organs that recognize neurotransmitters or hormones.
(2) A sensory nerve ending that responds to various stimuli.

Reference scale: Measurement scale for the evaluation of sensory stimuli established by the evaluations of one or several reference stimuli by panels/individuals.

Sensitivity: In sensory analysis, the ability to perceive, identify, and/or differentiate, qualitatively and/or quantitatively, one or more stimuli by means of the sense organs.

Sensory evaluation: Assessment of perceived environmental quality by the use of human observers and their senses of olfaction and somesthesia (the common chemical sense). The methods applied are generally psychological and taken from the subdiscipline of psychophysics.

Sensory modality: A specific sensory entity such as taste or odour intensity

Sick Building Syndrome (SBS): "Sick building" syndrome is used to describe a building in which a significant number (more than 20 per cent) of building occupants report illness perceived as being building-related. The complaints are characterized by a range of

symptoms including, but not limited to eye, nose, and throat irritation, dryness of mucous membranes and skin, nosebleeds, skin rash, mental fatigue, headache, cough, hoarsness, wheezing, nausea and dizziness.

Symptom: Any morbid phenomenon or departure from the normal function, appearance, or sensation, experienced by the patient and indicative of disease.

Syndrome: The aggregate of signs and symptoms associated with any morbid process, and constituting together the picture of the disease.

Synapse: The junction between two neurons or between a neuron and an effector organ where neural impulses are transmitted by neurotransmitters, i.e. by chemicals such as acetylcholine or norepinephrine.

Trigeminus (n. Trigeminus): The fifth cranial nerve; it arises in the pons (one of the basic subdivisions of the brain stem) and is composed of sensory and motor fibers. It has three divisions: ophthalmic, maxillary, and mandibular. The ophthalmic division supplies sensory fibers to the skin of the upper eyelid, side of the nose, forehead, and anterior half of the scalp. The maxillary division carries-sensory impulses from the mucous membranes of the nose, the skin of the cheek and side of the forehead, and the upper lip and upper teeth. The mandibular division carries sensory impulses from the side of the head, chin, mucous membranes of the mouth, lower teeth, and anterior two-thirds of the tongue.

z-scores: Probability values obtained by the linear transformation of a normal (Gaussian) distribution with a mean value of zero and a standard deviation of one.

Appendix: Members of the ECA-IAQ Steering Committee

AUSTRIA

Peter **TAPPLER**, Donauuniversität Krems, Zentrum für Bauen und Umwelt, Krems

Aleksander **ZILBERSZAC**, Bundesministerium für Gesundheit, Sport und Konsumentenschutz, Wien

BELGIUM

Francis **SARTOR**, Scientific Institute of Public Health - Louis Pasteur, Brussels

Peter **WOUTERS**, Belgian Building Research Institute, Brussels

DENMARK

Geo **CLAUSEN**, Centre for Indoor Environment and Energy, Technical University of Denmark, Lyngby

Lars **MØLHAVE** (vice chairman), Institute of Environmental & Occupational Medicine, Aarhus Universitet, Aarhus

FINLAND

Matti J. **JANTUNEN**, Division of Environmental Health, National Public Health Institute, Kuopio

Esko **KUKKONEN**, Ministry of the Environment, Physical Planning and Building Department, Helsinki

FRANCE

Anne **SERRE**, Ministère du Travail et des Affaires Sociales, Direction Generale de la Santé, Paris

Christian **COCHET**, CSTB, Centre Scientifique et Technique du Bâtiment, Marne la Vallée

Fabien **SQUINAZI**, Ministère de l'Emploi et de la Solidarité, Paris

GERMANY

Antonius **KETTRUP**, GSF - Forschungszentrum Umwelt und Gesundheit m.b.H., Institut für ÖkologischeChemie, Oberschleissheim

Bernd **SEIFERT** (chairman), Institut für Wasser-, Boden- und Lufthygiene des Umweltbundesamtes, Berlin

GREECE

Panayotis **SISKOS**, Laboratory of Analytical Chemistry, University of Athens, Athens

Matheos **SANTAMOURIS**, Department of Applied Physics, University of Athens, Athens

IRELAND

James P. **McLAUGHLIN**, Department of Physics, University College, Belfield, Dublin

Patrick A. **WRIGHT**, EOLAS, Glasnevin, Dublin

ITALY

Francesco **BOCHICCHIO**, Istituto Superiore di Sanità, Roma

Marco **MARONI**, International Centre for Pesticide Safety, Busto Garolfo

Giovanni **ZAPPONI**, Istituto Superiore di Sanità, Roma

THE NETHERLANDS

Bert **BRUNEKREEF**, Department of Epidemiology & Public Health, Agricultural University, Wageningen

Cor J.M. **VAN DEN BOGAARD**, Inspectorate for the Environment., VROM, The Hague

NORWAY

Jan **HONGSLO**, Department of Environmental Medicine, National Institute of Public Health, Oslo

Sten Olaf **HANSSEN**, NTNU, Norges teknisk-naturvitenskapelige universitet, Institutt for klima- og kuldeteknikk, Trondheim

PORTUGAL

Edoardo **DE OLIVEIRA FERNANDES**, Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto, Porto

SWEDEN

Birgitta **BERGLUND**, Department of Psychology, University of Stockholm, Stockholm

Thomas **LINDVALL**, Institute of Environmental Medicine, Karolinska Institute, Stockholm

Björn **LUNDGREN**, Swedish National Testing and Research Institute, Borås

SWITZERLAND

Heinz **ROTHWEILER**, Stadt Kloten, Umwelt und Gesundheit, Kloten

Christian **MONN**, Institut für Hygiene und Arbeitsphysiologie der ETH, Zürich

UNITED KINGDOM

Paul T. **HARRISON**, MRC Institute for Environment & Health, University of Leicester, Leicester

Linda **SMITH**, Department of the Environment, Transport and the Regions, Chemicals and Biotechnology Division, London

WORLD HEALTH ORGANIZATION

Michal **KRZYZANOWSKI**, WHO - European Centre for Environment and Health, Bilthoven

EUROPEAN COMMISSION

Maurizio **DE BORTOLI** (secretary), Joint Research Centre, Environment Institute, Ispra, (Varese)

Helmut **KNÖPPEL**, Joint Research Centre, Environment Institute, Ispra, (Varese)

European Commission

EUR 17675 – European Collaborative Action 'Indoor Air Quality and Its Impact on Man': Sensory Evaluation of Indoor Air Quality

The Steering Committee

Luxembourg: Office for Official Publications of the European Communities

1999 – 100 pp. – 21.0 x 29.7 cm

Environment and quality of life series

ISBN 92-828-5699-2

This report presents background to and advice on methodologies for sensory evaluation of indoor air quality (IAQ). The report gives a short introduction to sensory mechanisms and responses and to the theory of measurement underlying sensory evaluations and discusses in detail available sensory evaluation techniques. After a critical methodological analysis of some recently published documents on IAQ, sensory methods best suited for the evaluation of material emissions and of IAQ and for population response studies are recommended. Also non-sensory techniques for the evaluation of odour and mucosal irritation are briefly discussed. However, it is concluded that, at present, human subjects are indispensable in the measurement of perceived indoor air quality.

The proposed methods will enable designers, manufacturers, chemical and ventilating engineers, consumers, building and health authorities, and other decision makers to compare and select appropriate building materials, furnishings etc. Thereby the design, supply and control for good perceived air quality in indoor spaces will be easified which will lower the costs and minimize waste of energy.