Int Arch Occup Environ Health (2006) 79: 299–307
DOI 10.1007/s00420-005-0053-8

REVIEW

Monique A.M. Smeets · Jan H.A. Kroeze
Pamela H. Dalton

Setting occupational exposure limits in humans: contributions from the field of experimental psychology

Received: 22 February 2005 / Accepted: 27 September 2005 / Published online: 20 October 2005

Abstract Psychophysical methods from the field of experimental psychology are evaluated for their utility in the derivation of occupational exposure limits (OELs) for volatile chemicals based on acute sensory irritation in humans. The lateralization threshold method, which involves the localization of trigeminal vapor to the stimulated nostril, is evaluated for its underlying assumptions, reliability and validity. Whole body exposures, on the other hand, which involve the controlled, ambient exposure of human subjects to the irritant at one or a series of concentrations for an extended period are also discussed. It is concluded that the single-organ psychophysical method is largely resistant to response bias is practical and economical. However, its reliability and validity need further assessment. Whole body exposures, while having enhanced ecological validity, are more prone to demand characteristics, response bias, and subject beliefs than the traditional psychophysical procedures. An approach that involves the exposure of only the most sensitive organs such as the eyes and nose, via a mask or facebox, could facilitate the administration and alternation of odorant/irritant stimuli over a wide range of concentrations while enhancing ecological validity.

M. A.M. Smeets (✉)
Faculty of Social and Behavioural Sciences,
Department of Clinical Psychology,
Utrecht University, P.O. Box 80140, 3508 TC,
Utrecht, The Netherlands
E-mail: M.a.m.smeets@fss.uu.nl
Fax: 31-(0)30-2534718

J. H.A. Kroeze
Helmholtz Institute, Department of Psychonomics,
Utrecht University, P.O. Box 80140, 3508 TC,
Utrecht, The Netherlands

P. H. Dalton
Monell Chemical Senses Center, 3500 Market Street,
Philadelphia, PA, 19104 USA

Keywords Occupational exposure · Odor · Psychophysics · Sensory irritation · Trigeminal irritation

Introduction

The expertise and experience of individuals who hail from a variety of disciplines is an important component of the process of evaluating and setting occupational exposure limits (OELs; Feron 1998). Starting from this perspective, we hope to clarify how the discipline of psychology, more specifically experimental psychology, can make a contribution to the process of setting OELs. Given that increasing numbers of exposure studies with human volunteers are being conducted and presented for consideration in the limit-setting process, greater emphasis is placed on the importance of valid and reliable methods aimed at assessing the human perception of, and reactivity to, chemicals used in the workplace. The goal of this paper is to review some useful techniques from the field of experimental psychology and in particular, to point out common artifacts that can be expected during human testing, which may bias the results or interpretation of the data. As the emphasis will be on sensory methods, physiological and cytological methods for the assessment of irritancy will not be addressed.

Contributions from psychophysics: relation between critical effect and absolute threshold

One area within the discipline of experimental psychology, referred to as *psychophysics*, has provided a number of methods that have proven useful for studying chemosensory perception and response in the context of setting OELs. Psychophysics has been defined as the study of "...quantitative relations between changes in physical stimulation and concomitant changes in the reported aspects of sensory experience." (Dember and

Warm 1979, p. 25). In the term psychophysics, the word physics refers to the physical energy of the stimulus, examples of which are decibels for loudness, frequency for pitch or airborne chemical concentration for intensity. The reception of this physical stimulus energy leads to a response, which may entail a perception, a visceral sensation and in some cases, an observable behavioral change. Psychophysics assumes that observable behavior and physical stimulus energy correlate, and that the relationship between the two can be described quantitatively as $R = f(I)$, where R is response and I is intensity of the stimulus. It was of course acknowledged that a number of processes can intervene between reception or perception of the stimulus and the production of the observable response at the level of the brain. However, in the first half of the twentieth century these were difficult to measure and simply not taken into account. More recently the focus in experimental psychology has shifted precisely to what happens at the level of the brain, yet the methods that stem from the earlier psychophysical period are still very practical and currently in wide use.

How can psychophysical methodology be applied in the context of limit setting? Figure 1 displays the dose-effect relationship between a dose/exposure level and the seriousness of a yes/no effect, or magnitude of a specific effect such as the rating of eye irritation (Johanson 2001). The *critical effect* denoted in this figure is the first adverse symptom that appears as dose is increased. It is usually the least serious effect. For many chemical compounds upper airway irritation is the first adverse effect that is experienced at increasing concentrations (Dick and Ahlers 1998; Paustenbach 2001; Triebig 2002). As a consequence, nasal or eye sensory irritation can be considered as critical effects and the level at which they occur is called the *critical limit*. This makes sense from a teleological point of view, as sensations of irritation may protect us from exposure to higher concentrations that might be systemically toxic. Thus, a measure was needed that allows for the determination of the lowest concentration at which humans start to detect sensory irritation. In this paper, a definition of sensory irritation is posed as the subjective *experience* of irritation as opposed to objective or physiological irritation, evidenced by physical reddening or inflammation (Doty et al. 2004). Sensory irritation and physiological irritation may or may not coincide, that is, one may experience irritation in the absence of physical signs of irritation, in which case the sensory irritation threshold would be lower than the physiological irritation threshold. The latter situation may occur as a result of temporal or spatial summation effects (to be discussed more fully later), in which case the individual experiences irritation from, for example, repeated stimulations, each of which cannot be measured separately using objective methodology.

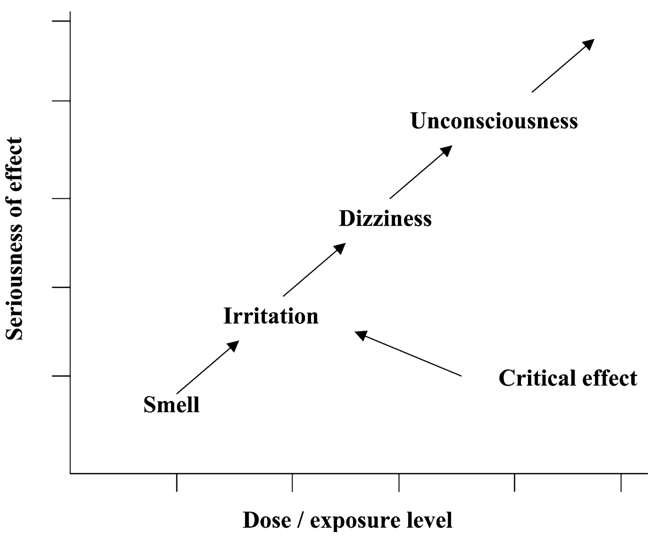


Fig. 1 Dose–effect relationship relating seriousness of an effect to dose/exposure level. Adapted from Johanson 2001

Switching back to psychophysical terminology, and following $R = f(I)$, one would like to be able to identify the lowest concentration/intensity of the stimulus at which a subject first experiences trigeminal irritation. Such a concentration is equivalent to what psychophysicists call the absolute threshold or the stimulus level at which the subject is just capable of perceiving the stimulus, even though the quality of the sensation that is experienced at this threshold may not be considered unpleasant or adverse.

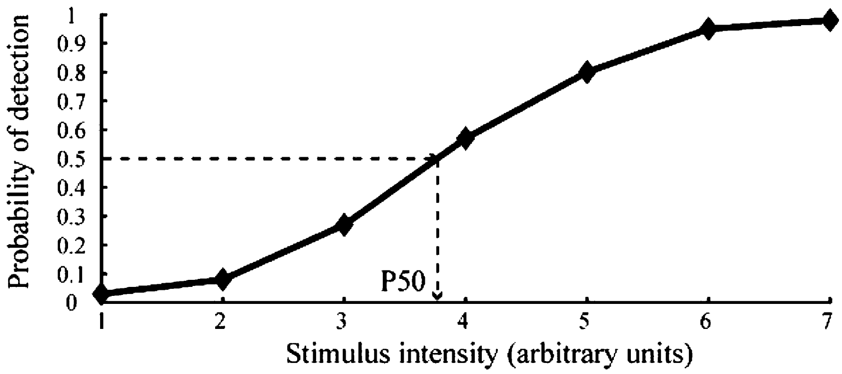
Figure 2 shows the probability of detecting a stimulus, for example, an irritant, at increasing concentrations. The P50 or 50th percentile corresponds to a subject's absolute threshold. It is the concentration that can be perceived by the subject on 50% of presentation trials. Thus, absolute threshold assessment procedures from psychophysics can be employed to determine the critical limit for sensory irritation.

Absolute threshold for detection of sensory irritation and odor

Simply put, the determination of the absolute threshold for sensory irritation would require the presentation of a number of stimuli of varying concentrations interleaved with blank stimuli with no irritancy, accompanied by the instruction to identify the stimulus. In general, thresholds are determined using static olfactometry, employing bottles equipped with nosepieces from which the subject can take a sniff. A series of bottles spanning a range of concentrations is prepared by diluting a mother-concentration in binary or tertiary steps.

The complicating factor when assessing irritation thresholds, however, is the fact that many volatile organic compounds (VOCs) also have an odor, which starts to become perceptible at levels well below the level at which irritation can be detected (Shusterman 2001). Quality and intensity of the odor may affect judgment, causing error and variability in results (Doty et al. 2004). Fortunately, the introduction of the lateralization

Fig. 2 Psychometric function relating one individual's proportion of correct responses to concentration level of stimulus presented. P50 corresponds to the concentration that was correctly detected on 50% of trials



threshold (LT) method for the detection of irritation in the context of odor effectively solved this problem by taking advantage of an important difference in the functional organization of the information processing pathways of the olfactory system, which is responsible for odor sensation, and the trigeminal system, which is responsible for sensory irritation in the eyes, nose and mouth. The technique exploits the fact that odor sensations cannot be lateralized to the left or right nostril, whereas trigeminal stimulation can be localized to a specific area on the skin or mucosal surface. Thus, in practice, subjects who experience a pure olfactory sensation in only one nostril are unable to reliably indicate which nostril was stimulated, regardless of intensity. However, if the concentration of a compound is sufficient to stimulate the free endings of the trigeminal nerve that innervate much of the nasal mucosal surface, subjects are in fact capable of indicating in which nostril they felt the sensation, even if that sensation is embedded in a strong odor (Kobal et al. 1989).

Thus, following this distinction, the LT procedure involves the simultaneous presentation of two stimuli, one stimulus containing clean air and the other stimulus containing the VOC at a pre-determined concentration of the stimulus in dilution (typically in only one of the two bottles of the pair), the other pair containing only the diluent (blank). Subjects are either asked to take a sniff (in active delivery) or allow the stimuli to be flown into their nose (in passive delivery) and then to indicate which nostril experienced the stronger sensation, or to localize the side of stimulus presentation (Wysocki et al. 1997). During odor detection threshold (ODT) measurement, on the other hand, the subject receives two *presentations* of stimuli. Only one presentation (first or second) contains the stimulus, the other contains just clean air. The subject is asked which of the two presentations contains the odorant. Typically, ODTs and LTs are obtained during a single session, in which the ODT is obtained at the outset to prevent adaptation effects by presenting lower concentrations first. A forced-choice procedure, in which the subject is asked to identify the stimulus pair (first or second?) for the ODT threshold, or the location of the stimulus (left or right?) for the LT, is preferred over a yes–no procedure. In the latter procedure the subject simply indicates whether or not the stimulus is experienced. The yes–no method may in some circumstances yield lower thresholds than the forced-choice procedure, suggesting that the subject perceives the stimulus at lower concentrations. This would occur when the subject has a tendency to say "yes" based on the assumption that a stimulus was presented, even though he or she did not really perceive it (a spurious perception). Since yes–no procedures typically present stimuli varying only in concentration, the experimenter will erroneously conclude that the subject correctly perceived the stimulus. To control for guessing, blanks are inserted in the forced choice method, thus enabling the calculation of the false alarm ratio. Data from forced-choice methods can later on be converted to determine a subject's response bias (see Frijters 1981; Kaplan et al. 1978; Klein 2001).

It should be noted that irritation thresholds can also be established for eye irritation using squeeze bottles, such as in a study by Cometto-Muniz and Cain (1991), which involved squeezing the irritant into one eye using a yes–no procedure, or actual LTs using olfactometers, such as that used by Opiekun et al. (2003), and this obviates the concern about odor. Although the free nerve endings of the trigeminal system innervate both the mucosa of the nasal passages and the cornea, the eyes may in many cases be more sensitive to the irritancy of VOCs due to variations in the tear film volume or viscosity or due simply to greater accessibility of the chemical to the corneal receptors than the nasal receptors.

Methods of stimulus presentation

The methods for stimulus presentation most often used are Wetherill and Levitt's up and down staircase (Wetherill and Levitt 1965) and the ascending method of limits (Engen 1971). In the up and down staircase method, stimulus presentations are tailored to a subject's presumed threshold. Higher concentrations are presented only after incorrect responses, and lower concentrations of a stimulus are presented only after a predetermined number (typically three or four) of correct responses have been obtained at one concentration. Switching from (a series of) increasing concentrations to (a series of) decreasing concentrations, and vice versa, is called a reversal. After, for example, five reversals, the absolute threshold can be calculated by averaging the reversal concentrations. In the ascending method of limits procedure, the subject is presented with a sequence of stimuli at increasing concentrations starting with a relatively low (subthreshold) concentration. A number of sequences differing in total number of stimuli presented and starting at different concentrations are presented in this manner, and the absolute threshold is based on the 50th percentile. Some variations on these procedures are made for LT detection, which aim to keep the number and concentrations of stimuli at a minimum for subject safety. For example, frequently, a maximum concentration is determined, above which no stimuli are presented. If the subject cannot lateralize this stimulus, the actual LT is presumed to be higher than this maximum. Obviously, for limit-setting, it is preferable to establish a lowest observed adverse effect level (LOAEL) by establishing a concentration at which sensory irritation does occur.

Although these methods rely on subjects' answers based on some form of introspection, and thus may be perceived as containing a certain degree of subjectivity, the objectivity of these methods is greatly enhanced by the use of a forced-choice procedure, the large number and concentrations of stimuli and repeated presentations. A few examples of ODTs and LTs obtained in the lab comparing previously exposed workers to subjects previously unexposed to a certain compound: (geometric) mean odor and irritation thresholds of 50 ppm and 15,682 ppm, respectively, for acetone were obtained by Wysocki et al. (1997); 10 ppm and 8,874 ppm, respectively, for methyl iso-butyl ketone (MIBK) by Dalton et al. (2000) and 11 ppm and 3,361 ppm, respectively, for isopropyl alcohol by Smeets and Dalton (2002).

Can the lateralization threshold help determine an OEL?

One feasible approach for using LTs to determine an OEL is to calculate the fifth or sixth percentile of a distribution of LTs in a representative population. This number would give an indication of the concentration at or below which intra-nasal irritation was perceived by only 5% of the test population. This level would be low enough to protect the majority of people as they would not yet perceive irritation at that particular level. For example, in Smeets and Dalton (2002) the fifth percentile of LTs for IPA based on a sample of 52 subjects was determined at 400 ppm, which is equivalent to the threshold limit value in the USA, which is 400 ppm (ACGIH 1998). For MIBK, Dalton et al. (2000) reported a fifth percentile of 1,802 ppm, and for acetone, 2,694 ppm, both of which are higher than their current allowable exposures and confirm that the current OEL based on irritation would be sufficiently protective. Alternatively, if LTs from both nasal irritancy as well as ocular irritancy are available, whichever is lower should be considered as the critical limit. Another argument that may play a role is the fact that in the workplace, ocular irritation can be more distracting than nasal irritation.

Limitations of the lateralization threshold: assumptions

The utility of the LT in the context of setting OELs depends on the reliability and validity of its method, and stability of results between research groups. Before dealing with these issues, the validity of the assumption on which the concept rests, which holds that volatile stimuli cannot be localized by odor alone should be first addressed. This assumption is largely based on a publication by Von Skramlik in 1925, and a short communication by Kobal et al. (1989) entitled "Is there directional smelling?" Using a dynamic olfactometer, Kobal et al. (1989) demonstrated that with vanillin and hydrogen sulfide, which they assume to be pure olfactants, no localization was possible. Using the pure irritant carbon dioxide and menthol, a compound with mixed olfactory and trigeminal characteristics, localization was nearly perfect. They concluded that "directional smelling exclusively mediated by the olfactory nerve does not exist" (Kobal et al. 1989, p. 131).

However, one should be cautious when drawing general conclusions concerning a sensory system based only on a few examples. Another problem has to do with ecological validity of these findings. When trying to detect the direction of an odor in the environment, the organism will sniff while moving its head. This will cause changes in turbulence and concentration between the two nostrils, information that might be used by the system to localize the source of the odor. Fechner's law, $\Delta I/I = k$, in which I is the intensity of the stimulus and ΔI is the increase or decrease required to perceive a difference, and k a constant, predicts that the higher the stimulus intensity, the greater the increase or decrease required for the difference between the consecutive stimuli to be noticed. An evaluation of direction smelling would require an investigation into just noticeable differences in odor intensity between nostrils, as the authors themselves acknowledge. Preferably, such an investigation should involve low overall concentrations within the range of naturally occurring odors.

Recently, a study was published presenting findings that are relevant to the topic of localization based on olfaction (Porter et al. 2005). Using a nasal mask consisting of two compartments permitting the delivery of a separate air stream to each nostril, Porter et al. (2005) presented both pure and trigeminal odorants (e.g., phenyl ethyl alcohol (PEA) and propionic acid) into one nostril and clean air into the other. Their finding that subjects correctly identified the location—left or right—of stimulation 70% of the time suggests that humans are capable of localizing pure olfactants. With respect to the LT measurement procedure, which also involves the simultaneous but separate presentation of clean air and an odorant/irritant to either nostril, this would imply that (1) humans may, under some circumstances and delivery methods, be capable of identifying the location of a pure odorant such as PEA or vanillin, and that (2) the ability to lateralize an irritant in such a setting may not be based on trigeminal stimulation only. If these implications hold true, it would make the LT procedure in the eyes more suitable than the LT derived from nasal presentation for the assessment of a trigeminally based critical limit. Further study of the assumptions underlying its use seem therefore desirable.

Reliability and validity of the LT procedure

Very few findings have been published with regard to the test–retest reliability of the LT assessment procedure. In a study of $n=32$, with 6–25 days between test and retest, Frasnelli and Hummel (2005) reported a test–rest correlation of $r=0.41$, $p=0.022$ for linalool, and $r=0.48$, $p=0.006$ for menthol. In a study by Shusterman et al. (2005) the test–retest reliability for *n*-propanol was found to be $r=0.75$, using bottled stimuli, and $r=0.50$ for CO_2, using a dynamic olfactometer. Testing sessions were separated by at least 1 day. Wysocki et al. (2003) in a study evaluating trigeminal sensitivity across the adult lifespan, found test–retest reliability for *n*-butanol thresholds of $r=76$ when sessions were separated by 1 h. More readily available data on the reliability on similar parameters such as odor thresholds also show variability over time: for example, using PEA, Doty et al. (1995) found a test–retest reliability of $r=0.88$ for threshold measurements separated by on average, 2 weeks, but $r=0.49$ for butanol over the same period. Partly, such differences may be based on instrumentation—changes in the measuring instrument over time—for example, in the case of temperature differences between the first and second test, which may affect the vapor concentrations of the chemical in the bottle. On the other hand, one would expect individual thresholds to vary over time as a result of natural fluctuations. In contrast to what the term ‘‘threshold’’—which stems from classical psychophysics—suggests, one no longer assumes an absolute threshold that only yields an odor percept when it is exceeded. There are day-to-day, and even circadian, physiological variations in olfactory thresholds probably in relation to receptor sensitivity, such that, at best, a range of concentrations can be identified within which individual thresholds can be located, and one does not expect this situation to be different for LTs. Perfect correlations between individual thresholds over time are therefore unlikely. However, since such variations will vary randomly across individuals, the mean group threshold should remain the same.

With respect to validity of the LT procedure, there are various approaches. Irritation thresholds collected in anosmics, who have subnormal or no sense of smell, should be comparable to LTs collected in ‘‘normosmics’’ who have a normal sense of smell, provided the trigeminal nerve is intact in the anosmics and the anosmia is complete. Contradictory findings have been reported. In a series of studies by Cometto-Muñiz et al. (1998, 1998a, b) no significant differences were found between anosmics and normosmics in concentrations that yielded irritancy employing forced-choice procedures using homologous alcohols as the stimuli (Cometto-Muñiz and Cain 1998) or terpenes (Cometto-Muñiz et al. 1998a; b). However, in studies by Hummel et al. (1996) and Kendal-Reed et al. (1998), normosmics reported chemesthetic sensations at lower concentrations than did anosmics. These findings may point towards integration between the olfactory and trigeminal systems that facilitate localization of the chemical stimulus.

Studies that involve the collection of irritation thresholds in the eyes, not the nose, also speak to the validity of the nasal LT method as the cornea of the eyes are innervated by the trigeminal nerve as well as the nose. To this end, Cometto-Muñiz and Cain (1991) employed squeeze bottles that contained a 25 ml conical measuring chamber the rim of which could be placed around the eye. A squeeze of the bottle delivered a puff of vapor directly to the eye. Bottles containing a blank or a chemical were consecutively presented to either the left or right eye, and the subject had to decide which bottle produced the stronger ocular sensation. Results they obtained from studies comparing eye thresholds between normosmics and anosmics show that nasal localization thresholds agree well with eye irritation thresholds (Cometto-Muñiz and Cain 1998; Cometto-Muñiz et al. 1998b*)*. Opiekun et al. (2003) applied the lateralization concept to the eyes. Using extremely low air flow to prevent irritation from mechanical stimulation they, as opposed to the previous finding, found ocular thresholds to be lower than the nasal ones. Other approaches towards studying irritation thresholds comprise quantification of the number of eyeblinks in response to airborne irritants using (filtered) EMG recordings from a single eye (Kieswetter et al. 2005) or counting based on videorecorded sessions (Klenø and Wolkoff 2004). So far, no studies have compared nasal LTs to these measures.

Temporal and spatial summation

An obvious question that presents itself is how thresholds for irritancy based on single-organ testing are related to the sensitivity of the organism, or, in other words, how can such results be used in the derivation of an OEL? Although there is little research of immediate relevance to this issue, two theoretical predictions can be made. One prediction concerns the issue of *spatial summation*, or the dependency of sensitivity and sensory magnitude on the areal extent of the stimulus (Marks 1974; Sherrick and Cholewiak 1986).

During LT testing, the subjects always receives the stimulus in only one nostril. In real life exposures, on the other hand, people will always experience stimulation of both nostrils simultaneously, as well as ocular stimulation, as the trigeminal free nerve endings innervate the mucosa in all of these regions. During bilateral stimulation the stimulated surface is larger than during monolateral stimulation, which will affect sensation. In visual perception research it has been found that complete spatial summation occurs when the product of stimulus area and threshold intensity is constant, upto a certain critical area. In other words, the *larger* the space that is stimulated, the *lower* the stimulus intensity needed to enable detection. This constancy has been referred to as Ricco's law (Vassilev et al. 2003). Such principles seem to apply to other sensations, such as heat pain, as well and explain why many parents use their elbow to gauge the temperature of the baby's bath water rather than just their finger: in the latter case more heat is required to enable heat detection of the small finger area, resulting in underestimation of the temperature and a screaming (if not burnt!) baby. Findings from Green's (1990) study on spatial summation of irritation from capsaicin applied to the forearm, indeed suggest summation for detection of chemical stimuli around threshold.

The temporal analog of Ricco's law is known as Bloch's law, which holds that the longer the stimulation, the lower the intensity needed to reach the threshold, upto a certain maximum, at which a plateau is reached. Using ammonia, Cometto-Muñiz and Cain (1984) demonstrated that the longer the pungent stimulus was inhaled (upto 4.5 s) the greater is its perceived magnitude. Using the odourless CO_2, Wise et al. (2004) and ammonia (Wise et al. 2005) showed that the nasal trigeminal system can detect progressively weaker stimuli by integration over time, an example of temporal summation. Based on Bloch's law, one would predict *higher* irritation thresholds during intra-nasal testing than would be the case for multiple organ or ambient testing. It is desirable that studies are conducted to determine how well Ricco's and Bloch's laws apply to the sense of irritation, and to establish the relevant constants. These, then, can be applied to correct thresholds, which have been collected using less labor-intensive methods (i.e., monolateral stimulation) to estimate irritation thresholds with increased ecological validity.

In summary, the LT procedure seems to be a promising and practical method in the context of OEL-setting based on sensory irritation, for example, by the trigeminal nerve. LTs are easy to administer and relatively inexpensive to obtain in humans (when compared with chamber studies). On the other hand, it is still too early to grant LTs a central role in limit setting, as issues with the underlying assumptions of localizability of the chemical stimulus, reliability, validity, temporal and spatial summation still need further attention. It is important that studies are undertaken to address these issues to prevent this potentially useful method to fall by the wayside.

So far, the focus has been on determining irritation in a single sensory (nasal or ocular) mucosa. Another approach to studying the organism's response to airborne irritants is that of whole body exposures in which the participant receives an ambient exposure to a volatile chemical in a controlled environment. Numerous endpoints can be monitored over the duration of the exposure, typically involving ratings of odor and irritation intensity and health symptoms at suprathreshold concentrations, and physiological and cytological endpoints related to irritancy of in most cases the upper respiratory pathways and eyes (Kieswetter et al. 2005; Smeets et al. 2002; Van Thriel et al. 2005). This approach has several advantages over the LT method just described, as it allows for the assessment of effects due to adaptation, habituation or perhaps even sensitization of the trigeminal and olfactory systems over the course of time, which may alter the subject's reactivity. A study by Cain (1974), for example, showed increased irritation to formaldehyde over time upto some asymptote, followed by a decrease as seen with adaptation (see also Hempel-Jørgensen et al. 1999).

There can also be disadvantages to this approach. First, there is less control of artifacts, as many objective psychophysical approaches cannot be used. During a session a subject is typically exposed to a single concentration, as opposed to a series or alternating concentrations, which would be the case in the threshold procedures previously discussed. Similarly, forced-choice procedures cannot be used, because they require multiple presentations to a variety of concentrations as well as blanks. Thus, when using sensory measures, whole body exposure sessions generally rely on the use of rating scales, such as seven-point scales or visual analog scales, to evaluate effects from one or more suprathreshold exposures, which methods are most prone to artefacts known as *demand characteristics*. This refers to the fact that subjects are never neutral to the experimental situation. For example, participants may presume that appropriate precautions for their safety must have been taken (Orne 1969), which may cause them to underreport. So, for example, if an odor is present, subjects may hypothesize about its experimental purpose, and respond accordingly by over- or understating certain responses so as to "help" or sabotage the experimenter. This was demonstrated by Knasko and Gilbert (1990), who found a higher proportion of participants reporting symptoms when they believed they were exposed to a unpleasant odorant, than when they believed they were exposed to a pleasant odorant. It is important to note that in all cases no actual odor was presented and that participants were only led to believe that an odor had been dispersed in the room. This research underscores the necessity of a placebo or no-odor (clean air) control in whole body exposure studies for determining demand characteristic effects. In a similar vein, Dalton et al. (1997) exposed participants to an actual, but unfamiliar, odor. Subjects received different characterizations of the odor in three independent conditions, as either positive (i.e., harmless, natural), negative (i.e., potentially harmful, chemical) and neutral (i.e., just an odor stimulus). Instructions affected not only the number and intensity of symptom effects reported, but also perceived intensity of odor and irritation experienced from the odor over time. These findings emphasize the importance of including an exposure to an odor (not an irritant), preferably with hedonics or presumed effects that resemble that of the experimental odor/irritant of interest, to assess the impact of such factors independent of irritancy effects.

Another artifact of interest in studies to determine OELS based on irritation is response bias. In addition to biases related to psychophysical procedures, other types of biases can alter the experimental results. For example, subjects may have a tendency to use one end of the rating scale, or only the center categories. Poulton (1989) has given an overview of such scaling biases and methods to avoid or minimize them. Likewise, subjects may not quite understand which attribute of the stimulus they are supposed to rate and thus confuse sensations. This is particularly true, when a stimulus contains multiple dimensions, but the subject is only allowed to rate one (i.e., irritancy or odor intensity), leading to a phenomenon known as 'halo dumping' (Clark and Lawless 1994). For example, Dalton et al. (2000) were surprised to find much higher ratings for the intensity of odor and irritation of MIBK than acetone at comparable concentrations than might be expected given their relative potency. When interviewing the subjects after testing, it appeared they had experienced the smell of MIBK as unfamiliar and annoying. It seemed as if the intensity ratings had been influenced by the hedonics of the odorants, or that there had been some confusion about what they were supposed to rate. These effects largely disappeared after the researchers included a separate rating of annoyance and adjusted instructions. Apparently, the inclusion of a rating of affect or liking can be used to unload the affective component of attribute ratings such as of intensity. These types of procedures are common to sensory science and consumer testing, which typically focus on the sensory testing of novel food products (Meilgaard et al. 1987). Here, an important distinction is made between affective testing (e.g., how much do you like this product?) and the more analytical testing of product attributes, such as intensity, creaminess and saltiness. Typically, affective ratings are never obtained from the same subjects or during the same session as product attribute ratings. Alternatively careful training, which facilitates the recognition and discrimination of the relevant sensations that are under investigation (for an interesting example in the context of odor and trigeminal sensation, see Doty et al. 1978), sensations and familiarity with the scaling instrument can all contribute to the collection of laboratory data, which will correlate well with an individual's subjective experience in the workplace.

Conclusion

In this paper we have discussed how methods derived from experimental psychology may contribute to the derivation of OEL limits based on sensory irritation in humans. We have described how threshold procedures from the domain of psychophysics can be applied to determining critical limits in single organs based on both odor and irritation. Thanks to the practical and economic benefits of these procedures, such thresholds can be important tools for determining critical effects based on irritancy, but the underlying assumption that localization of the chemical stimulus relies exclusively on sensations of irritancy, and not smell, needs additional empirical substantiation. Furthermore, additional studies need to be performed to evaluate the reliability and validity of this method as well as to evaluate issues concerning temporal and spatial summation.

Whole body exposures, while having better external validity, may also elicit so-called demand characteristics and biases that can complicate the interpretation of the data. These artifacts are less likely to occur in psychophysical threshold procedures, which are characterized by enhanced internal validity, thus strengthening the validity of causal inferences. Ideally, a human model of irritation employed for the derivation of OELs would integrate the advantages of both approaches. It seems that much hinges on the ability to provide a large number of stimulus exposures varying over a wide range on concentrations, alternated with blanks combined with a forced-choice response procedure. This is expensive, time-consuming and difficult to achieve with ambient, controlled exposure in exposure chambers. Since OELs serve to protect the most sensitive organs, a solution may involve restriction of exposure to, for example, the eyes, possibly including the nose, by administering the airborne stimulus through a mask or facial box. As the exposure volume inside the mask is smaller than in an exposure chamber, concentrations can be more easily switched to determine dose–response effect relationships, and available instruments such as olfactometers can be used. Likewise, the exposure may also involve longer-lasting exposures to single concentrations of the chemical and can be combined with physical activity or cognitive performance as often seen in the work simulation study.

Acknowledgements Preparation of this paper was supported by NWO grant 452–03-334 to MAMS, and by RO1 DC-03704 to PHD.

References

American Conference of Governmental and Industrial Hygienists (1998) Threshold limit values for chemical substances and physical agents

Cain WS (1974) Perception of odor intensity and the time-course of olfactory adaptation. ASHRAE Transactions 80:53–75

Sherrick CE, Cholewak RW (1986) Cutaneous sensitivity. In: Boff KF, Kaufman L, Thomas JP (eds) Handbook of Perception and Human Performance. Vol. I: Sensory Processes and Perception. Wiley Interscience, New York, pp. 12-11–12-13

Clark CC, Lawless HT (1994) Limiting response alternatives in time-intensity scaling: an examination of the halo-dumping effect. Chem Senses 19:583–594

Cometto-Muňiz JE, Cain WS (1984). Temporal integration of pungency. Chem Senses 8:315–327

Cometto-Muňiz JE, Cain WS (1990). Threshold for odor and nasal pungency. Physiol Behav 48:719–725

Cometto-Muňiz JE, Cain WS (1991). Nasal pungency, odor, and eye irritation thresholds for homologous acetates. Pharmacol Biochem Behav 39:983–989

Cometto-Muňiz JE, Cain WS (1998) Trigeminal and olfactory sensitivity: comparison of modalities and methods of measurement. Int Arch Occup Environ Health 71:105–110

Cometto-Muňiz JE, Cain WS, Abraham MH, Kumarsingh R (1998a). Sensory properties of selected terpenes. Thresholds for odor, nasal pungency, nasal localization, and eye irritation. Ann NY Acad Sci 855:648–651

Cometto-Muňiz JE, Cain WS, Abraham MH, Kumarsingh R (1998b).Trigeminal and olfactory chemosensory impact of selected terpenes. Pharmacol Biochem Behav 60:765–770

Dalton P, Wysocki CJ, Brody MJ, Lawley HJ (1997) The influence of cognitive bias on the perceived odor, irritation and health symptoms from chemical exposure. Int Arch Occup Environ Health 69:407–417

Dalton PH, Dilks DD, Banton MI (2000) Evaluation of odor and sensory irritation thresholds for methyl isobutyl ketone in humans. Am Ind Hyg Assoc J 61:340–350

Dember WN, Warm JS (1979) Psychology of Perception, 2 ed. Holt, Rinehart and Winston, New York

Dick RB, Ahlers H (1998) Chemicals in the workplace: incorporating human neurobehavioral testing into the regulatory process. Am J Ind Med 33:439–453

Doy RL, Brugger WE, Jurs PC, Orndorff MA, Snyder PJ, Dale Lowry L (1978) Intranasal trigeminal stimulation from odorous volatiles: Psychometric responses from anosmic and normal humans. Physiol & Behav 20:175–185

Doty RL, Cometto-Muňiz JE, Jalowayski AA, Dalton P, Kendal-Reed M, Hodgson M (2004) Assessment of upper respiratory tract and ocular irritative effects of volatile chemicals in humans. Crit Rev Toxicol 34:85–142

Doty RL, McKeown DA, Lee WW, Shaman P (1995). A study of the test-retest reliability of ten olfactory tests. Chem Senses 20:645–656

Engen T (1971) Psychophysics I Discrimination and detection. In: Kling J, Riggs L (eds) Woodworth's & Schlosberg's Experimental Psychology, Holt, Rinehart & Winston, New York

Feron VJ (1998) Recommending health-based exposure limits in the national and international arena: a personal view. In: Bal R, Halffman W (eds) The politics of chemical risk: Scenarios for a regulatory future. Kluwer Academic Publishers, Dordrecht, pp 121–129

Frasnelli J, Hummel T (2005) Intranasal trigeminal threshold in healthy subjects. Environ Toxicol Pharmacol:575–580

Frijters JER (1981) Expanded tables for conversion of a proportion of correct responses (Pc) to the measure of sensory difference (d') for the triangular method and the 3-alternative forced choice procedure. J Food Sci 47:139–143

Green BG (1990) Spatial summation of chemical irritation and itch produced by topical application of capsaicin. Percept Psychophys 48:12–18

Hempel-Jørgensen A, Kjaergaard SK, Mølhave L, Hudnell HK (1999) Time course of sensory eye irritation in humans exposed to N-butanol and 1-Octene. Arch Environ Health54:86–94

Hummel T, Barz S, Lotsch J, Roscher S, Kettenmann B, Kobal G (1996) Loss of olfactory function leads to a decrease of trigeminal sensitivity. Chem Senses 21(1):75–80

Johanson G (2001) Basic concepts in toxicological risk assessment. Paper presented at the Occupational exposure limits - approaches and criteria, In: Proceedings from a NIVA course, Uppsala, Sweden

Kaplan H, MacMillan N, Creelman C (1978) Tables of d' for variable-standard discrimination paradigms. Behav Res Meth Instr 10:796–813

Kendal-Reed M, Walker J, Morgan WT, LaMacchio M, Lutz RW (1998) Human responses to proprionic acid. I . Quantification of within- and between-participant variation in perception by normosmics and anosmics. Chem Senses 23:71–82

Kieswetter E, Von Thriel C, Schäper M, Blaskwewicz M, Seeber A (2005) Eye blinks as indicator for sensory irritation during constant and peak exposures to 2-ethylhexanol. Environ Tox Pharm 19:531–541

Klein SA (2001) Measuring, estimating, and understanding the psychometric function: A commentary. Percept & Psychophys 63:1421–1455

Klenø J, Wolkoff P (2004) Changes in eye blink frequency as a measure of trigeminal stimulation by exposure to limonene oxidation products, isoprene oxidation products and nitrate radicals. Int Arch Occup Environ Health 77:235–243

Knasko SC, Gilbert AN (1990) Emotional state, physical well-being, and performance in the presence of feigned ambient odor. J Appl Social Psychol 20:1345–1357

Kobal G, Van Toller S, Hummel T (1989) Is there directional smelling? Experientia 45:130–132

Marks LE (1974) Sensory processes: The new psychophysics. Academic Press, New York

McGuire WJ (1969) Suspiciousness of experimenter's intent. In: Rosenthal R, Rosnow RL (eds) Artifact in Behavioral Research, Academic Press, New York, pp 13–57

Meilgaard M, Civille GV, Carr BT (1987) Sensory Evaluation Techniques. CRC Press, Boca Raton

Opiekun RE, McDermott R, Dalton PH (2003) Gender differences and nasal integration studies using an ocular exposure device for detection or irritation thresholds: The T.I.D.E. System. Poster presented at Achems Conference, Sarasota, Fl

Orne MT (1969) Demand Characteristics and the Concept of Quasi-Controls. In: Rosenthal R, Rosnow RL (eds) Artifact in Behavioral Research. Academic Press, New York, pp 143–179

Paustenbach D (2001) Approaches and considerations for setting occupational exposure limits for sensory irritants: report of recent symposia. Am Ind Hyg Assoc J 62:697–704

Poulton EC (1989) Bias in quantifying judgments. Lawrence Erlbaum, Hove, UK

Porter J, Anand T, Johnson B, Khan RM, Sobel N (2005) Brain mechanisms for extracting spatial information from smell. Neuron:581–592

Sherrick CE, Cholewak RW (1986) Cuteanous sensitivity. In: KR Boff, L Kaufman, and Thomans, JP (eds) Handbook of perception and human performance. Vol I: Sensory processes and perception. New York: Wiley Interscience, pp 12–11 – 12–13

Shusterman D (2001) Odor-associated health complaints: competing explanatory models. Chem Senses 26:339–343

Shusterman D, Murphy MA, Balmes J (2003) Differences in nasal irritant sensitivity by age, gender, and allergic rhinitis status. Int Arch Occup Environ Health 76:577–583

Smeets M, Dalton P (2002) Perceived odor and irritation of isopropanol: a comparison between naive controls and occupationally exposed workers. Int Arch Occup Environ Health 75:541–548

Triebig G (2002) Chemosensory irritation and the setting of occupational exposure limits. Int Arch Occup Environ Health 75:281–282

Vassilev A, Mihaylova MS, Racheva K, Zlatkova M, Anderson RS (2003) Spatial summation of S-cone ON and OFF signals: effects of retinal eccentricity. Vision Res 43:2875–2884

Van Thriel C, Kieswetter E, Schäper M, Blaskewicz M, Golka K, Seeber A (2005) An integrative approach considering acute symptoms and intensity ratings of chemosensory sensations during experimental exposures. Environ Tox Pharm 19:589–598

Wetherill GB, Levitt H (1965) Sequential estimation of points on a psychometric function. Br J Math Stat Psychol 18:1–10
Wise PM, Radil T, Wysocki CJ (2004) Temporal integration in nasal lateralization and nasal detection of carbon dioxide. Chem Senses 29:137–142
Wise PM, Canty TM, Wysocki CJ (2005) Temporal integration of nasal irritation from ammonia at threshold and supra-threshold levels. Toxicol Sci 87:223–31
Wysocki CJ, Dalton P, Brody MJ, Lawley HJ (1997) Acetone odor and irritation thresholds obtained from acetone-exposed factory workers and from control (occupationally unexposed) subjects. Am Ind Hyg Assoc J 58:704–712
Wysocki CJ, Cowart BJ & Radil T (2003) Nasal trigeminal chemosensitivity across the adult lifespan. Percept & Psychophys 65:115–122