Chem. Senses 29: 311–317, 2004 DOI: 10.1093/chemse/bjh034

# A Simple and Reliable Method for Clinical Assessment of Odor Thresholds

**Jörn Lötsch[1], Cornelia Lange[2] and Thomas Hummel[2]**

[1]Pharmazentrum Frankfurt, Institute of Clinical Pharmacology, Johann Wolfgang Goethe-University, Theodor Stern Kai 7, D-60590 Frankfurt am Main, Germany and [2]Smell and Taste Clinic, Department of Otorhinolaryngology, University of Dresden Medical School, Fetscherstr. 74, D-01307 Dresden, Germany

*Correspondence to be sent to: Dr Thomas Hummel, Smell and Taste Clinic, Department of Otorhinolaryngology, University of Dresden Medical School, Fetscherstr. 74, 01307 Dresden, Germany. e-mail: thummel@rcs.urz.tu-dresden.de*

## Abstract

We investigated whether presenting of dilutions of phenyl ethyl alcohol at random succession according to the method of constant stimuli can replace the standard procedure of presenting a various number of dilutions in a staircase paradigm. Forty-six men and 44 women, aged 19–76 years, participated in this study. Phenyl ethyl alcohol was diluted in a ratio of 1:2, starting from 4%. Presentation of the odorant followed a three-alternative, temporal forced-choice paradigm with two blanks in addition to the odorant. Twenty dilutions were administered in a randomized order. Odor threshold was obtained by logistic regression of the correct and incorrect identifications of the probe containing the odorant. Thresholds were also calculated on the basis of the first 16 dilution steps only. Results from these procedures were compared with 'gold-standard' threshold assessment employing a three-alternative, temporal forced-choice staircase paradigm with seven reversals using 16 dilutions of phenyl ethyl alcohol. The method of constant stimuli took a shorter and less variable testing time than the staircase technique. The use of 20 dilution steps provided no better results than the use of 16 steps. The method of constant stimuli exhibited a good test–retest reliability ($r = 0.7$; $P < 0.001$) comparable to that of the staircase method and provided unbiased results highly correlated ($r = 0.8$; $P < 0.001$) with those of the staircase technique with similar inter-test variability. Applying 16 dilutions (1:2 steps) of phenyl ethyl alcohol at random succession in a three-alternative, temporal forced-choice paradigm is thus a simple and reliable procedure for the reproducible assessment of odor thresholds that may be contemplated as an alternative to the 'gold-standard' staircase method of clinical odor threshold assessment.

**Key words**: logistic regression, olfaction, smell, staircase technique, threshold

## Introduction

Numerous tests are available for assessment of olfactory function in a clinical environment (for a review, see Doty and Laing, 2003). Partly initiated by the needs of the 'Working Group Olfaction and Gustation' of the German Society for Oto-Rhino-Laryngology, Head and Neck Surgery, the 'Sniffin' Sticks' test (Kobal *et al.*, 1996, 2000; Hummel *et al.*, 1997) has been developed and validated over the last decade. This test relies on pen-like odor dispensing devices and is comprised of three separate tests for odor threshold, odor discrimination and odor identification. For odor identification and odor discrimination, a fixed number of 16 trials are administered. Odor threshold testing follows a triple-forced staircase paradigm (see Doty *et al.*, 1994; Ehrenstein and Ehrenstein, 1999). That is, triplets of pens are presented to the subject, with one pen containing the odorant at a certain dilution, while the two other pens contain only the solvent. The subject's task is to identify the pen that contains the odorant. Two successful identifications in a row or one unsuccessful identification trigger a reversal of the staircase to the next higher or next lower dilution, respectively. Testing is completed after seven reversals have been obtained. This frequently makes odor threshold testing a lengthy procedure of unpredictable duration, especially in patients with hyposmia. Since the subjects' ability or willingness to concentrate strongly influences the quality of the test results, standardized test duration is desirable. We therefore investigated the possibility of replacing the staircase technique with a procedure based on the method of constant stimuli (Fechner 1860) that applied a fixed number of trials (preferably 16) as with the odor discrimination and identification tasks of 'Sniffin' Sticks' test. We investigated whether this provides similar test results at a similar accuracy and repeatability as the well-established staircase paradigm.

## Materials and methods

### Volunteers and study design

Forty-six men and 44 women, aged 19–76 years (mean ± standard deviation = 49.9 ± 17.3 years), participated in this study. Eighteen of them (mean age 60 ± 10 years) reported the presence of olfactory dysfunction. The investigations were performed in compliance with the Declaration of Helsinki (Summerset West amendment). The study involved odor threshold testing at two different days. At each occasion, both the staircase method and the method of constant stimuli were applied in a sequence which was randomized across subjects; however, the methods were applied in the same succession at both days for an individual subject. The tests were performed birhinally. The interval between threshold tests was 30–60 min and the interval between test days was not longer than 6 weeks.

### Testing of odor thresholds

Odorants were presented in commercially available felt-tip pens ('Sniffin' Sticks'; Burghart GmbH, Wedel, Germany; see Hummel *et al.*, 1997; Kobal *et al.*, 2000). Instead of liquid dye the tampon of the pen was filled with phenyl ethyl alcohol (a rose-like odor) diluted in propylene glycol. Odors were presented in triplets of pens, one containing diluted phenyl ethyl alcohol and two containing only propylene glycol serving as blanks. For odor presentation the cap was removed by the experimenter for ~3 s and the pen's tip was placed 1–2 cm in front of the nostril. The interval between presentations of individual pens of a triplet was ~3 s. Employing a three-alternative, temporal forced-choice paradigm, the subjects had to identify the pen that contained the odorant. Subjects were blindfolded to prevent visual identification of the odor containing pens (for details, see Hummel *et al.*, 1997; Kobal *et al.*, 1996, 2000). Phenyl ethyl alcohol was presented in dilution series made in 1:2 dilution ratios, starting from 4%. Since preliminary experiments indicated that the 16 dilution steps included in the 'Sniffin' Sticks' test (Kobal *et al.*, 1996, 2000; Hummel *et al.*, 1997) may not be sufficient when applying the method of constant stimuli, four more dilution steps (steps 17–20) were added for this study; however, not for the staircase procedure. Thus, 16 (staircase) or 20 (method of constant stimuli) triplets of pens were presented to the subjects in a randomized order at an interval of 20–30 s. In the staircase paradigm two successive correct identifications of the pen containing the odor or one incorrect identification triggered a reversal of the staircase to the next higher or the next lower dilution step, respectively. Seven reversals had to be obtained (Doty *et al.*, 1994; Hummel *et al.*, 1997). When the method of constant stimuli was used, each triplet was presented only once.

### Testing of odor identification and odor discrimination

During the first session assessment of thresholds was followed by tests for odor discrimination and odor identification (see Hummel *et al.*, 1997) to completely assess the subject's objective function. Odor identification was assessed by means of 16 common odors. Using a multiple choice task identification of individual odorants was performed from a list of four descriptors each. The test result was a sum score of the correctly identified odors. For odor discrimination 16 triplets of pens were presented, with two containing the same odorant and the third a different odorant. Subjects had to determine which one of the three odor-containing pens smelled different. The test result was a sum score of successful performances.

### Data analysis

The data analysis used dilution steps rather than true concentrations of the odorant because normative data and previous study results obtained with the 'Sniffin' Sticks' test are available in this form. For comparison with other test procedures, olfactory thresholds are also given as concentrations of phenyl ethyl alcohol when indicated. With the staircase procedure, odor thresholds were obtained as the mean of the last four staircase reversal points of a total of seven reversals. Pathologic odor threshold was defined as a score of ≤5 (Kobal *et al.*, 2000), i.e. phenyl ethyl alcohol cannot be perceived at dilutions greater than dilution step 5, which is equal to a concentration of 0.25% phenyl ethyl alcohol in the pen.

Data obtained by the method of constant stimuli were analyzed twice, once for the set of 20 dilution steps, and again using only the first 16 dilution steps, which are part of the regular 'Sniffin' Sticks' test kit (Kobal *et al.*, 1996, 2000; Hummel *et al.*, 1997). Calculation of thresholds was based upon the assumption that the probability of correct identification of the odorant containing pen is high when the odor is perceived, whereas it is at chance level when the odor is not perceived. Thresholds were calculated by fitting a logistic function (Linschoten *et al.*, 2001) of

$$P(x) = \gamma + (1 - \gamma)\left(\frac{1}{1 + (\alpha/x)^{-\beta}}\right) \qquad (1)$$

to the data using a log-likelihood fitting technique, where $x$ denotes the dilution steps ($1 \geq x \geq 16$ or 20), $\gamma$ the probability of correct identification by chance (0.33 for the three-alternative, temporal forced-choice paradigm), $\alpha$ the odorant dilution step at the halfway point of the probability and $\beta$ the steepness of the function. The probability of correct identification follows a sigmoid curve, which decreases with higher dilutions of the odorant (Figure 1). When the probability reaches 2/3, i.e. half-way between chance and full probability, the threshold was reached, which is given by the value of $\alpha$ in equation (1).

Data analysis focused on the question whether the results obtained with the method of constant stimuli are similar to those obtained with the staircase procedure that is established as the 'gold standard' in odor threshold testing, i.e.

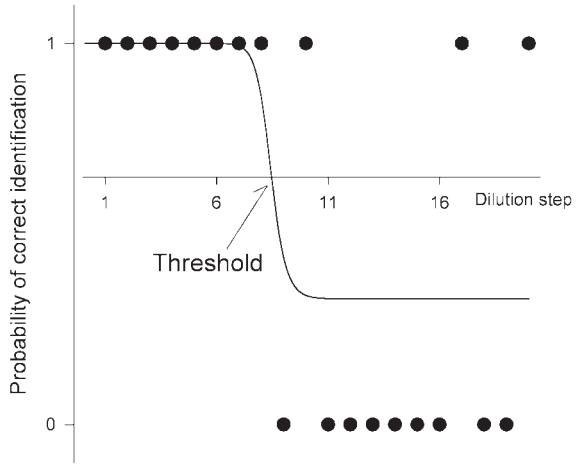


**Figure 1** Typical S-shaped function obtained for stimulus detection in a 3-alternative forced-choice paradigm. Dots present correct (1) or incorrect (0) identifications of a certain concentration of phenyl ethyl alcohol (given in dilution steps starting from 4%). The solid line is the logistic function best-fitting the data using a maximum-likelihood curve-fitting technique (Linschoten *et al.*, 2001). The probability of correct identification decreases when lower concentrations of the odorant (indicated as higher dilution steps) are administered

whether the method of constant stimuli can replace the staircase method in this context. Thus, as previously employed (Bland and Altman, 1986), we assessed whether a new method provides similar results as an established method. Differences between thresholds obtained with the method of constant stimuli and those obtained with the staircase technique were calculated to quantify the bias of the results obtained with the method of constant stimuli. The absolute value of these differences was used to quantify the accuracy of the method of constant stimuli to provide the same test results as the 'gold-standard' staircase procedure. Differences between the second and the first test session were calculated to assess the repeatability between days. Differences between test results were analyzed by means of multivariate analysis of variance for repeated measures (rm-ANOVA), with within-subject factors 'method' (staircase, 16 constant stimuli, 20 constant stimuli) and 'testing day' (1st and 2nd) and between subject factors 'type of subjects' (normals and self-described 'dysfunctional') and 'gender'. *t*-Tests with α-adjustment for multiple comparisons (Bonferroni) were used for *post hoc* comparisons. In addition, correlation analyses were performed were adequate. Statistics were done with SPSS 12.0.1 (SPSS Inc, Chicago, IL). The α-level was set at 0.05. Data are presented as means ± standard deviations). Finally, from the results of the odor threshold, odor identification and odor discrimination tasks, a composite 'TDI score' ('threshold discrimination identification') was derived from the sum of the results obtained in each test (see Cain *et al.*, 1988; Hummel *et al.*, 1997; Kobal *et al.*, 2000; Wolfensberger *et al.*, 2000). Based on the odor discrimination and identification tests, the best possible TDI score is 32 (16 from each test if the subjects identified or discriminated all odors). To this adds the threshold. Pathologic olfactory function was defined as a TDI score of ≤30 (Kobal *et al.*, 2000). A competition of both procedures for best recognition of the 'true' threshold was not possible because this would have required that the 'true' threshold were known by a reference method. The use of olfactory evoked potentials as reference was not possible because their exact quantitative relation to results of subjective olfactory testing has not yet been investigated and it is currently unknown how that limit of 0.25% phenyl ethyl alcohol for separating normal from 'pathologic' thresholds is reflected in the evoked potentials. Another psychophysical test of olfactory function, such as the University of Pennsylvania Smell Identification Test (UPSIT; Doty *et al.*, 1984) was also not suitable as a reference test since it would have provided just a third variant of a test of olfactory function, this time obtained by a test that is possibly less well established in the present study sample of a German population.

## Results

### Test duration

Administration of the staircase technique required 7.9 ± 1.7 min for odor threshold testing. The method of constant stimuli with 20 triplets of pens required 7.5 ± 0.7 min. The estimate of the test duration of the method of constant stimuli with 16 dilution steps was 6 ± 0.6 min (calculation by division of the duration of testing required for the method of constant stimuli by 20 and multiplication of the result with 16; Figure 2). In addition, compared to the staircase paradigm, test duration with the method of constant stimuli exhibited considerably less interindividual variation (coefficients of variation of 22 and 9%, respectively). The difference in test duration was statistically significant (rm-ANOVA main effect 'method': df = 2,152, $F = 122.4$, $P < 0.001$; significant differences to the staircase method for both 16 and 20 dilution steps: *post hoc* *t*-tests $P < 0.001$). In addition, a significant interaction 'method' by 'type of subjects' ($F = 5.1$, $P < 0.01$) was found. Specifically, while subjects with self-reported olfactory dysfunction needed ~0.5 min longer for the staircase testing, they were tested ~0.2 min faster with the method of constant stimuli. Other significant main effects or interactions were not found with the rm-ANOVA.

### Thresholds

Interday repeatability of the test procedures is shown in Figure 3; the differences between days and the variance (coefficients of variation) of these differences are presented in Table 1. Range (Figure 3*B*) and magnitude (Figure 3*C*) of the differences between days were similar among methods (Table 1). Differences of the results obtained with the method of constant stimuli to the results obtained with the staircase technique are presented in Figure 4 and Table 1. The correlation between thresholds obtained at days 1 and 2

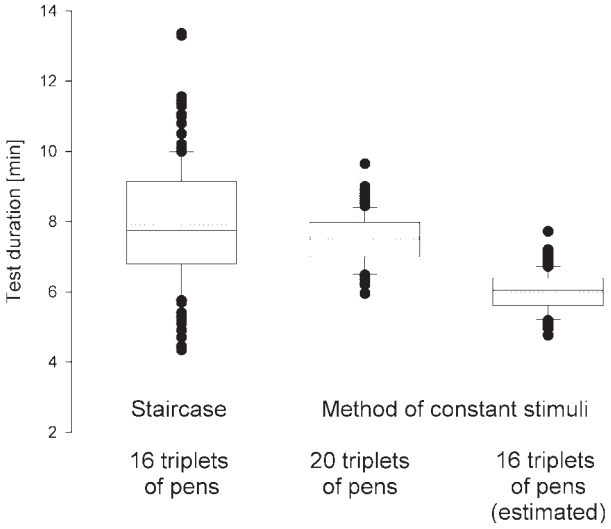


**Figure 2** Duration of the tests with the staircase technique (left) where variable numbers of triplets were presented to the subjects and the method of constant stimuli where 20 triplets of pens were administered. The estimated test duration is also shown if the method of constant stimuli would have been performed with 16 instead of 20 triplets. The lower boundary of the box indicates the 25th percentile, a line within the box the median, a dotted line within the box the mean and the upper boundary of the box indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles, respectively. In addition, outliers are shown as dots.

was high with $r = 0.82$, $P < 0.001$ for the staircase paradigm, and $r = 0.77$ and 0.79 ($P < 0.001$) for the method of constant stimuli analyzed for all 20 and for the first 16 dilutions steps, respectively. The rm-ANOVA revealed no significant main effects for the factors 'method' ($P = 0.54$), 'testing day' ($P = 0.95$) and 'gender' ($P = 0.28$) and no significant interactions. This indicates that similar thresholds were obtained with all methods.

Subjects with self-described olfactory dysfunction had significant higher thresholds than subjects who reported themselves as being normosmic (rm-ANOVA factor 'type of subjects': df = 1,86, $F = 60.7$, $P < 0.001$). Specifically, subjects who self-reported olfactory dysfunction exhibited mean thresholds of $5.2 \pm 2.1$ dilution steps corresponding to 0.22% phenyl ethyl alcohol, whereas subjects who estimated their olfactory function as being normal exhibited mean thresholds of $11.3 \pm 3.4$ corresponding to 0.0033% phenyl ethyl alcohol. The mean threshold of 5.2 dilution steps is almost at the limit of pathologic threshold. This owes to the fact that the self-estimate of poor olfactory function was only supported by the tests in 10 cases while in the other eight subjects olfactory threshold was found to be within the normal range. This emphasizes that ratings of olfactory function are unreliable in healthy, untrained subjects as previously demonstrated (Landis *et al.*, 2003).

### Diagnosis of pathologic olfactory function

Using the staircase technique, a pathologic threshold (dilution step ≤ 5 corresponding to a threshold of not less than

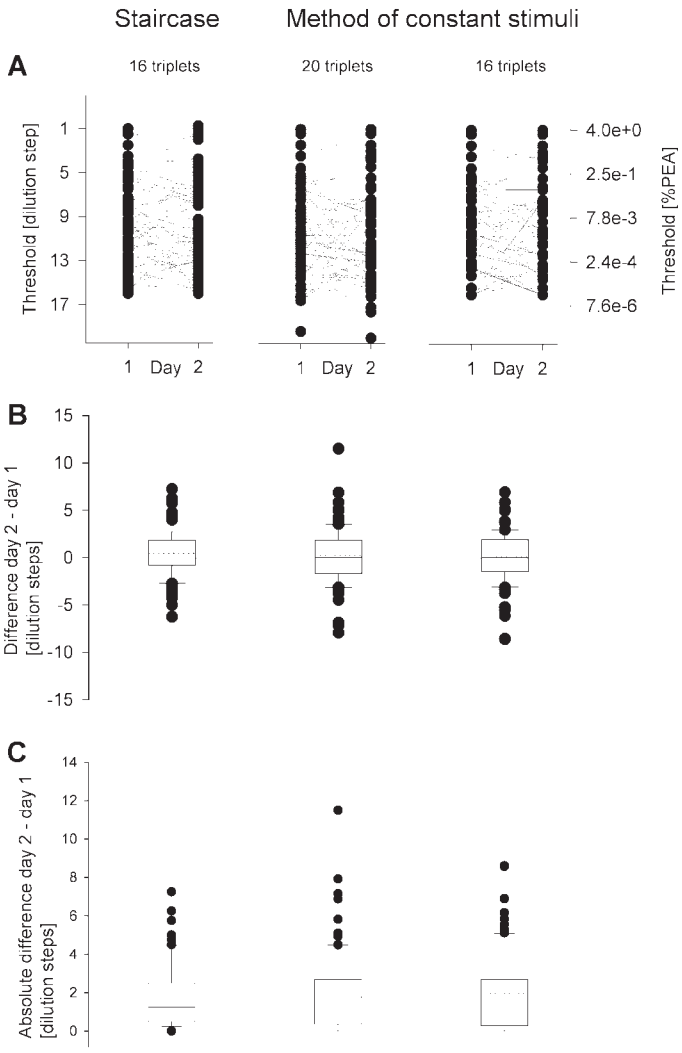


**Figure 3** Interday repeatability of different procedures for assessment of odor thresholds: Solid horizontal lines indicate the median, and dotted horizontal lines indicate the arithmetic mean. The staircase paradigm and presentation of phenyl ethyl alcohol at different concentrations in randomized succession. **(A)** Thresholds of individuals obtained at days 1 and 2 are connected with lines to demonstrate their inter-day differences. The left ordinate shows the threshold in dilution steps, the right ordinate the thresholds in percent phenyl ethyl alcohol concentration (%PEA) in the 'Sniffin' Sticks' pen. **(B)** Differences between day 2 and day 1, demonstrating the inter-day bias of the thresholds. **(C)** Absolute values of the differences between day 2 and day 1, demonstrating the inter-day accuracy of the thresholds.

0.25% phenyl ethyl alcohol) was detected 24 times during the 180 tests (90 participants, two tests per subject). The diagnosis disagreed between testing days in 2 subjects (thresholds at 4.5 and 6.5 dilution steps, or at 4.25 and 6.75 dilution steps at day 1 and day 2, respectively). The performance of the method of constant stimuli to detect pathological thresholds is presented in Table 2. The disagreeing results were seen mostly with thresholds around the cut-off at dilution step 5 (Figure 5), i.e. thresholds that were around the limit between 'normal' and 'pathologic', except for one threshold

**Table 1** Differences and absolute differences in thresholds between days obtained with the staircase technique and the method of constant stimuli and differences and absolute differences of thresholds obtained with the method of constant stimuli compared to thresholds obtained with the staircase technique

| | | Staircase | Method of constant stimuli | |
|---|---|---|---|---|
| | | 16 dilution steps | 20 dilution steps | 16 dilution steps |
| Correlation between days (Pearson's *r*)[a] | | 0.82 | 0.77 | 0.79 |
| Difference between days | Mean (dilution step) | 0.5 | 0.2 | 0.1 |
| | CV (%) | 25.7 | 30.5 | 29.0 |
| Absolute difference between days | Mean (dilution step) | 1.8 | 2.0 | 2.0 |
| | CV (%) | 17.6 | 21.7 | 20.3 |
| Correlation with staircase (Pearson's *r*) | | 1 | 0.85 | 0.84 |
| Difference to staircase | Mean (dilution step) | – | 0.2 | –0.1 |
| | CV (%) | – | 23.8 | 28.5 |
| Absolute difference to staircase | Mean (dilution step) | – | 1.7 | 1.7 |
| | CV (%) | – | 15.8 | 16.7 |

CV, coefficient of variation.
[a]Correlations were always statistically significant at $P < 0.001$.

**Table 2** Detection of pathologic thresholds and pathologic global olfactory function (TDI score, a sum of the scores obtained with the tests of odor threshold, odor discrimination and odor identification) by the different analyses of the method of constant stimuli

| | Staircase | | Method of constant stimuli | | | |
|---|---|---|---|---|---|---|
| | | | 20 dilution steps | | 16 dilution steps | |
| | Threshold | TDI | Threshold | TDI | Threshold | TDI |
| Correctly identified pathologic scores | 24 | 11 | 17 | 10 | 17 | 10 |
| Percentage correctly identified pathologic scores | 100 | 100 | 71 | 91 | 71 | 91 |
| Mean score if pathologic score was not detected | – | – | 7 | 24 | 7.7 | 24 |
| Greatest threshold if pathologic score was not detected | – | – | 8.4 | 32 | 13 | 32 |
| Percentage of wrongly detected pathologic scores | 0 | 0 | 3 | 0 | 3 | 0 |
| Mean staircase threshold when wrongly detected as pathologic | – | – | 6.7 | – | 6.7 | – |
| Maximum staircase threshold when wrongly detected as pathologic | – | – | 8 | – | 8 | – |
| Percentage of tests pathologic/non-pathologic disagreeing between days | 2 | – | 4 | – | 4 | – |
| Mean difference if disagreeing between days | 2.3 | – | 2.6 | – | 2.6 | – |
| Mean staircase threshold if disagreeing between days | 5.5 | – | 5.3 | – | 5.3 | – |

The staircase technique served as reference for correct or incorrect detection. Note that odor identification and odor discrimination were tested only at the first study day and TDI scores are available only for that day.

that was 4 (0.5% phenyl ethyl alcohol) with the staircase method and 13 (0.0098% phenyl ethyl alcohol) with the method of constant stimuli.

### Relations to age

Older subjects had higher odor thresholds, i.e. thresholds at lower dilution steps, which was seen in significant correlations between age and thresholds for all tests and data analyses (Pearson's ρ of –0.23 to –0.36, $P > 0.05$ to $P < 0.001$).

## Discussion

Compared to the staircase technique the randomized presentation of different odor concentrations in order to test odor thresholds can be performed within a shorter period of time

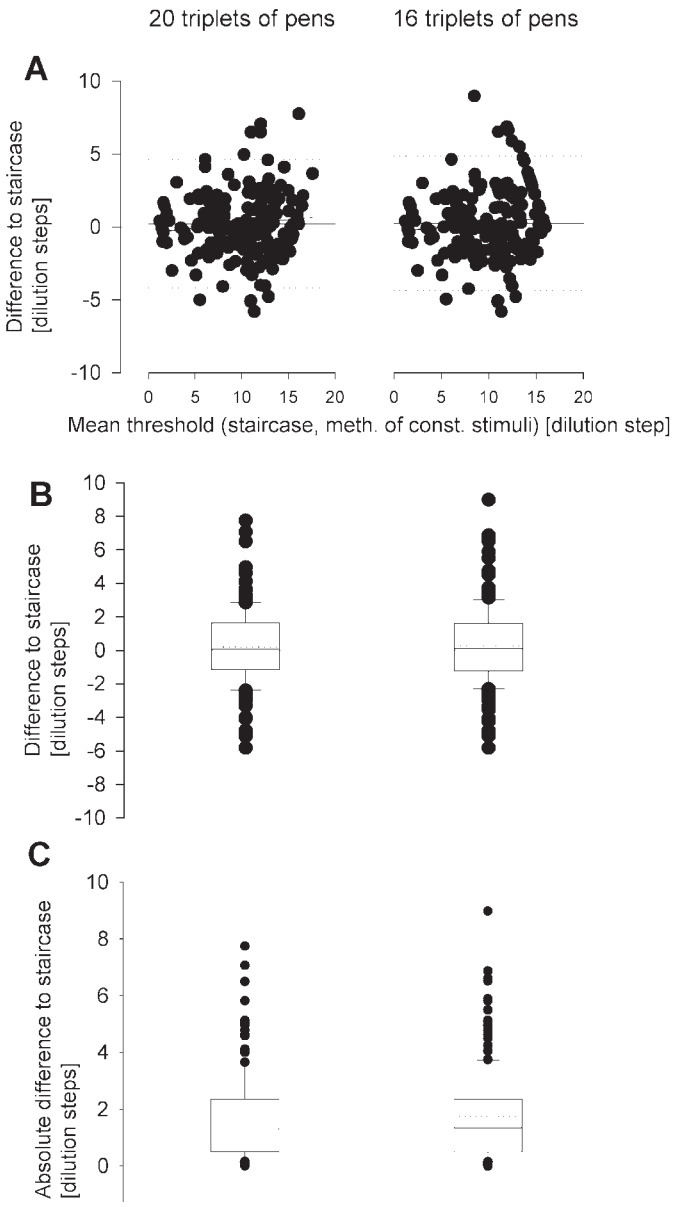


**Figure 4** Comparison of results from the random presentation of PEA at different concentrations with the staircase technique for odor threshold testing. **(A)** Differences of thresholds obtained with the method of constant stimuli to those obtained with the staircase technique, plotted against the mean of both values (Bland and Altman, 1986). This indicates that the method of constant stimuli performed equally well at all threshold levels. **(B)** Differences of the method of constant stimuli to the staircase technique, showing the absence of an overall bias with the method of constant stimuli. **(C)** Absolute differences of the thresholds obtained with the method of constant stimuli to those obtained with the staircase technique, showing the accuracy of the method of constant stimuli.

and the time needed to test is more standardized. This is an advantage especially in situations where subjects have difficulties to concentrate, or when the effects of drugs on olfactory function are investigated when subjects are sedated (Lötsch *et al.*, 2001), or in any situation where routine testing is performed, e.g. in clinical environments (Cain *et al.*, 1983; Leopold and Bartels, 2002).

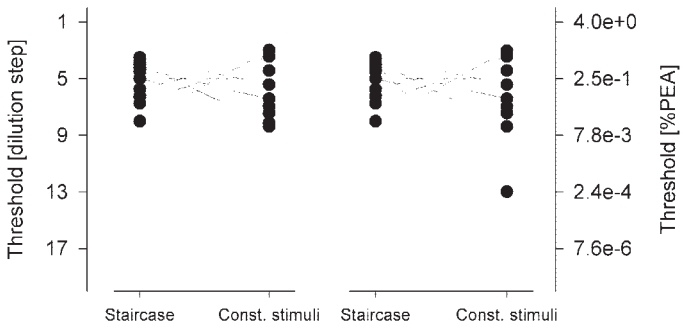


**Figure 5** Thresholds that disagreed between the staircase and the constant stimuli methods in the diagnosis of a pathologic threshold, i.e. a threshold equal to or higher than 0.25% phenyl ethyl alcohol (dilution step 5). The left ordinate shows the threshold in dilution steps, the right ordinate the thresholds in percentage phenyl ethyl alcohol concentration (%PEA) in the 'Sniffin' Sticks' pen.

The method of constant stimuli, however, can only be considered as an alternative to the staircase technique when it results in similar thresholds as the staircase technique. The method of constant stimuli exhibited a test–retest reliability similar to that of the staircase technique. Thresholds obtained by the method of constant stimuli were highly correlated with those obtained through the staircase technique. Moreover, differences to the staircase technique did not exceed the between-days differences that had been obtained through the staircase technique. This suggests that the observed differences reflect the variance in olfactory function (Stevens and Dadarwala, 1993; Frasnelli *et al.*, 2002) and both methods performed equally well in its assessment. The method of constant stimuli detected three-quarters of the pathologic thresholds according to the staircase technique as reference. Disagreement of the diagnosis of pathologic threshold occurred when the threshold was close to 0.25% phenyl ethyl alcohol that marked the limit between normal and non-normal ('pathologic') odor threshold (Kobal *et al.*, 2000). Again, this appears to reflect the variance in olfactory function rather than a false test result. When computing the TDI score, which is a composite score from the results of tests for odor threshold, odor identification and odor discrimination (Kobal *et al.*, 2000; Wolfensberger *et al.*, 2000), almost all cases of pathologic olfactory function were diagnosed when odor thresholds had been obtained with the method of constant stimuli. Similarly, the false alarm rate of the method of constant stimuli was only 7% when taking the results of the staircase method as the reference and this again occurred when the threshold in the staircase method was not far from 0.25% phenyl ethyl alcohol. Importantly, these differences between the two methods did not result in a false diagnosis of pathologic olfactory function by the TDI score.

Before entirely relying on the alternative procedure of odor threshold assessment one has to keep in mind that the staircase technique has been tested intensively, its clinical usefulness has been demonstrated and that it has been

shown to provide results in accordance to other established test of olfactory function (Kobal *et al.*, 1996, 2000; Hummel *et al.*, 1996) — all of which still needs to be established for the method of constant stimuli.

Thus, based on the present results we propose to use 16 dilution steps, applied in triplets of pens with two blanks at a randomized succession of dilutions as an alternative, time effective procedure of odor threshold testing.

## Acknowledgements

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG HU441/2-1) to T.H.

## References

**Bland, J.M.** and **Altman, D.G.** (1986) *Statistical methods for assessing agreement between two methods of clinical measurement*. Lancet, 1, 307–310.

**Cain, W.S., Gent, J., Catalanotto, F.A.** and **Goodspeed, R.B.** (1983) *Clinical evaluation of olfaction*. Am. J. Otolaryngol., 4, 252–256.

**Cain, W.S., Gent, J.F., Goodspeed, R.B.** and **Leonard, G.** (1988) *Evaluation of olfactory dysfunction in the Connecticut Chemosensory Clinical Research Center*. Laryngoscope, 98, 83–88.

**Doty, R.L.** and **Laing, D.G.** (2003) *Psychophysical measurement of human olfactory function, including odorant mixture assessment*. In Doty, R.L. (ed.), Handbook of Olfaction and Gustation, Marcel Dekker, New York, pp. 203–228.

**Doty, R.L., Shaman, P., Kimmelman, C.P.** and **Dann, M.S.** (1984) *University of Pennsylvania Smell Identification Test: a rapid quantitative olfactory function test for the clinic*. Laryngoscope, 94, 176–178.

**Doty, R.L., Smith, R., McKeown, D.A.** and **Raj, J.** (1994) *Tests of human olfactory function: principal components analysis suggests that most measure a common source of variance*. Percept. Psychophys., 56, 701–707.

**Ehrenstein, W.H.** and **Ehrenstein, A.** (1999) *Psychophysical methods*. In Windhorst, U. and Johansson, H. (eds), Modern Techniques in Neuroscience Research. Springer, Berlin, pp. 1211–1241.

**Fechner, G.T.** (1860) Elemente der Psychophysik. Breitkopf & Härtel, Leipzig.

**Frasnelli, J., Livermore, A., Soiffer, A.** and **Hummel, T.** (2002) *Comparison of lateralized and binasal olfactory thresholds*. Rhinology, 40, 129–34.

**Hummel, T., Sekinger, B., Wolf, S., Pauli, E.** and **Kobal, G.** (1997) *'Sniffin' Sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold*. Chem. Senses, 22, 39–52.

**Kobal, G., Hummel, T., Sekinger, B., Barz, S., Roscher, S.** and **Wolf, S.** (1996) *'Sniffin' sticks': screening of olfactory performance*. Rhinology, 34, 222–226.

**Kobal, G., Klimek, L., Wolfensberger, M., Guziol, H., Temmel, A., Owen, C.M., Seeber, H., Pauli, E.** and **Hummel, T.** (2000) *Multicenter investigation of 1036 subjects using a standardized method for the assessment of olfactory function combining tests of odor identification, odor discrimination, and olfactory thresholds*. Eur. Arch. Otorhinolaryngol., 257, 205–211.

**Landis, B.N., Hummel, T., Hugentobler, M., Giger, R.** and **Lacroix, J.S.** (2003) *Ratings of overall olfactory function*. Chem. Senses, 28, 691–694.

**Leopold, D.A.** and **Bartels, S.** (2002) *Evaluation of olfaction*. J. Otolaryngol., 31, S18–S23.

**Linschoten, M.R., Harvey, L.O., Jr, Eller, P.M.** and **Jafek, B.W.** (2001) *Fast and accurate measurement of taste and smell thresholds using a maximum-likelihood adaptive staircase procedure*. Percept. Psychophys., 63, 1330–1347.

**Lötsch, J., Darimont, J., Skarke, C., Zimmermann, M., Hummel, T.** and **Geisslinger, G.** (2001) *Effects of the opioid remifentanil on olfactory function in healthy volunteers*. Life Sci., 69, 2279–2285.

**Stevens, J.C.** and **Dadarwala, A.D.** (1993) *Variability of olfactory threshold and its role in assessment of aging*. Percept. Psychophys., 54, 296–302.

**Wolfensberger, M., Schnieper, I.** and **Welge-Lussen, A.** (2000) *Sniffin' Sticks: a new olfactory test battery*. Acta Otolaryngol., 120, 303–306.

*Accepted February 28, 2004*