Oral Diseases (2005) 11 (Suppl. 1), 40–44
 1601-5665/05

http://www.blackwellmunksgaard.com

ORIGINAL ARTICLE

Effects of training on odor judges scoring intensity

S Nachnani[1], G Majerus[2], P Lenton[2], J Hodges[2], E Magallanes[1]

[1]*University Health Resources Group Inc., Canterbury Drive, Culver City, CA;* [2]*University of Minnesota, Clinical Dental Research Center, Delaware SE, Minneapolis, MN, USA*

OBJECTIVE: This pilot study was intended to test whether a training protocol improved validity of odor judges (OJs), with or without experience, and whether odorant types differed in error proneness.
METHODS: The OJs (four experienced, two inexperienced) completed a 4-phase training protocol based on the American Society of Testing and Materials standards (ASTM): (i) introduction to sensory scales, n-butanol reference, sniffing techniques; (ii) pretraining measurements; 20 samples of varying intensities of four unpleasant and three pleasant odorants; (iii) exercises assessing quality, intensity, ranking, and matching; and (iv) post-training measurements.
MAIN OUTCOME MEASURES: Subjects' intensity scores were analyzed as the absolute difference from the 'true' intensity (ASTM n-butanol standard) using repeated measures ANOVA.
RESULTS: Training significantly ($P = 0.02$) reduced OJ errors. Experienced and novice judges did not differ in average errors ($P = 0.99$), or in improvement in error from pre- to post-training ($P = 0.94$). Improvement was consistent from pre- to post-training for all odorants except dimethylsulfide for which errors worsened ($P = 0.01$). Unpleasant and pleasant odorants differed ($P = 0.006$) in error. After removing water the effects of water control scores from the pleasant odorants, the difference was not significant ($P = 0.26$).
CONCLUSIONS: The OJs improved in their ability to assess odor intensity irrespective of previous experience. Training is recommended for all OJs prior to research trials.
Oral Diseases (2005) **11** (Suppl. 1), 40–44

Keywords: oral malodor; sensory panels; organoleptic judges; odorants; training

Introduction

Research reports the major offensive odorants responsible for oral malodor (MO) are the volatile sulfur compounds (VSCs) hydrogen sulfide, dimethylsulfide (DMS), and methylmercaptan (Tonzetich, 1971; Schmidt *et al*, 1978; Kostelc *et al*, 1984). The current instruments used to assess oral malodor are the gas chromatograph (GC), sulfide monitors (Halimeter, OralChroma, Brethtron, etc.), and sensory odor judges (OJs). The GC and sulfide monitors quantify and in some cases identify sulfur compounds in concentrations as low as a few p.p.b. (Tonzetich, 1971; Schmidt *et al*, 1978; Rosenberg *et al*, 1991a,b; Yaegaki, 2002; Lenton *et al*, 2004). However, these sensors are only useful for identifying VSCs.

Sensory OJs can not only detect and recognize compounds but also can discriminate complex mixtures, so OJs are considered the gold standard for oral malodor assessment. OJs assess odors for intensity or offensiveness (Allison and Katz, 1919; Rosenberg *et al*, 2001a,b) and for hedonic tone (pleasantness) (Chambers and Wolf, 1996). Currently, the American Dental Association requires that for clinical trials, intensity or hedonic assessments be performed by at least two trained and calibrated OJs blinded to each other (ADA Council on Scientific Affairs, 2003).

Although training and calibration of sensory OJs has been well-documented in other disciplines (foods/beverages, air/water quality, personal care products), few data have been published on methods for training oral malodor judges. To date, most academic groups involved in oral malodor research have trained directly with malodor subjects. However, industry uses training protocols based on American Society for Testing and Materials (ASTM) standards. However, these studies are rarely published due to the proprietary nature of the business. The objective of this pilot study was to test if training OJs, experience in MO assessment and types of odorants improved OJ's intensity scores.

Methods and materials

Subject population

The subject population consisted of four sensory OJs with experience in assessing MO and two without experience. OJs participated in a 4-phase training protocol: (i) introduction to sensory techniques; (ii) pretraining; (iii) sensory training exercises; and (iv) post-training. The training protocol was adapted from the Wheeler *et al* (1981) Guidelines for the Selection, and Training of Sensory Panel Members, Chambers and

Correspondence: Sushma Nachnani, University Health Resources Group, Inc. 5714, Canterbury Drive, Culver City, CA 90230, USA. Tel: 310 645 8576, E-mail: sushman@worldnet.att.net

Wolf (1996) Sensory Testing Methods, Subcommittee E 18.04 on Fundamentals of Sensory (1999) Standard Practices for Referencing Supra threshold Odor Intensity, and Sensory Evaluation Techniques (Meilgaard *et al*, 1999a).

Introduction to sensory techniques
The OJs received demonstrations on handling samples and sniffing techniques and were briefed on using the odor intensity scale (Table 1A), hedonic tone scale (Table 1B) and standard odor descriptors (McGinley and McGinley, 2002). OJs received a sample reference odorant (n-butanol) that had been determined by study investigators to have an intensity level of 3 for oral malodor. OJs were allowed to sniff and become familiar with the n-butanol intensity level 3.

Pretraining OJ measurements
The OJs independently assessed 20 odorant samples for odor hedonic tone, quality, and intensity. Each sample contained approximately 30 ml of liquid odorant in a 60 ml glass amber bottle with a screw cap. Samples consisted of seven different odorants, one neutral (water), four unpleasant (skatole, putrescine, DMS, and butyric acid) and two pleasant (jasmine, orange) of varying intensities. Study investigators assigned each odorant sample a 'true' intensity score by reference to the 5-point n-butanol study scale (Table 2), adapted from the ASTM 10-point n-butanol scale. OJs assessed odorants at 3-min intervals and were allowed to compare samples to the reference n-butanol, level 3, as needed.

Sensory training exercises
Training consisted of a series of didactic and laboratory practice sessions. Didactic sessions included discussions on the purpose of sensory training, influence of emotional factors, chain of sensory perception (stimulus → sensation and detection → perception and recognition → decision response) and odorant scales, qualities and descriptors. OJ preparation, reducing bias, adaptation and habituation were also discussed. Laboratory sessions included exercises on assessing descriptive qualities, detection and discrimination of similar compounds, memory matching of identical and similar odorants, odorant intensity ranking and relating odorants to the n-butanol reference scale. All laboratory exercises included training, practice, discussion and reassessment of skills. A variety of pleasant (jasmine, orange, pineapple, peppermint, rose, cinnamon, eucalyptus, lemon), unpleasant odorants (garlic, alcohol, hydrogen sulfide, skatole, putrescine, butyric acid, DMS) and neutral odorants (water, pine, spearmint, vanilla, apple, pear) with varying intensities were used during the training exercises.

Table 1 Descriptors for the organoleptic odor intensity and hedonic scales

(A) Odor intensity scale	*(B) Hedonic scale*
1 = no perceived odor	+2 = like very much
2 = faint odor noticed	+1 = like
3 = moderate odor	0 = do not like or dislike
4 = strong odor	−1 = dislike
5 = extremely strong odor	−2 = dislike very much

Table 2 ASTM and study scale relationships to butanol concentrations (p.p.m.)

ASTM 10-point scale (n-butanol)	*Study scale (n-butanol)*
1 = 12	1 = 25
2 = 24	2 = 225
3 = 48	3 = 675
4 = 96	4 = 2025
5 = 194	5 = 6075
6 = 388	
7 = 775	
8 = 1550	
9 = 3100	
10 = 6200	

Post-training OJ measurements
The same 20 samples assessed during pretraining were assessed for the post-training measurement, but in a different random order. OJs assessed samples at 3-min intervals and received a freshly mixed bottle of n-butanol of intensity level 3.

Statistical analysis
Statistical analysis used repeated measures ANOVA, where a 'subject' was an OJ. The dependent variable was the absolute difference between a judge's score and the truth as listed in Table 3 (the absolute difference is the difference with negative signs removed). The fixed effects were experienced *vs* non-experienced, pre- *vs* post-training, gas, and their interactions. A preselected comparison (a 'contrast') was made between pleasant and unpleasant gases using the gas as main effect. Standard errors are derived using error terms from the ANOVA.

Results

Overall, training significantly reduced (improved) OJ errors (Table 3, item 1; $P = 0.02$). Previous experience

Table 3 Average absolute error for selected groups of measurements, pre- and post-training

	Average absolute error		
	Pretraining	*Post-training*	*s.e. of average absolute error*[a]
Overall	1.13	0.77	0.09
Experienced	1.13	0.76	0.15
No experience	1.13	0.78	0.10
Butyric acid	1.33	0.83	0.23
Citrus (orange)	1.29	0.53	0.19
Dimethylsulfide	0.97	1.58	0.23
Water	0.42	0.08	0.23
Floral (jasmine)	1.24	0.61	0.19
Putrescine	1.33	0.81	0.19
Skatole	1.11	0.72	0.19

[a]Applies to both pre- and post-testing.

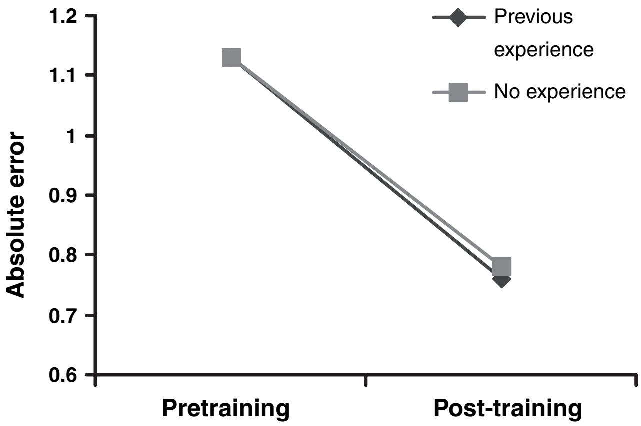

Figure 1 Relationship between odor judges previous experience and improvement (reduction in errors) from pre- to post-training, for all odorants combined

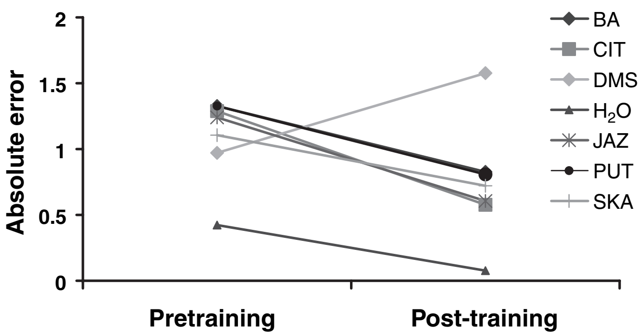

Figure 2 Average pre- and post-training odor judge scores for each odorant

was not associated with the size of the errors (experienced 0.94, no experience 0.95; $P = 0.99$). Experience was also not associated with the improvement (reduction in errors) from pre- to post-training (Figure 1; Table 3, item 2; $P = 0.84$).

Judges improved from pre- to post-training for all odorants except DMS (gas-by-training interaction, $P = 0.01$), for which they worsened (Figure 2; Table 3, item 3). Judges tended to have greater errors scoring unpleasant odorants (unpleasant 1.09, pleasant 0.70; $P = 0.006$). However, after removing the water score from the pleasant odorants, the difference was no longer significant (pleasant without water 0.93; $P = 0.26$). Figure 2 shows why; water had much lower errors both pre- and post-training than any other compound.

Discussion

Sensory odor assessment attempts to measure, analyze, and interpret the olfactory system's reaction to a perceived odor. It is difficult to assess olfactory sensations as there is no standard unit to quantify or measure odors as there is with other senses (e.g. hearing). There are approximately 17 000 odorous compounds (Harper, 1972) for which there is no international standardized terminology.

Individuals vary considerably in recognition of odors (Pangborn, 1981) and perception of pleasantness (Cain and Johnson, 1978). Rabin (1988) demonstrated that attaching verbal labels to olfactory stimuli significantly improved OJs ability to recognize odors. To facilitate odor assessment in this study, OJs were given a list of seven distinct odorant descriptors, which was made available during all phases of training and measurement. Although this study did not evaluate odor quality recognition, all of the study judges indicated that listing the descriptors helped them to quickly recognize/describe an odorant, which then enabled them to focus on odor intensity.

The OJs were allowed to rate how much they liked or disliked a sample. Study investigators believed that giving OJs the opportunity to rate odorant pleasantness could help to control the subjective emotional attachment a judge could have to a particular odorant, as described by Cain (1988). Stevenson *et al* (1998) have also suggested sweetness of an odorant can enhance odor intensity perception. Therefore, OJs were instructed to break down their assessment into four distinct steps, odor stimulation → rating odor pleasantness → odor description → odor intensity, similar to Shiffman's (1996) chain of sensory perception. Investigators believed these steps could enhance the learning of the intensity scale by training judges to address odor character and pleasantness before focusing on intensity. All OJs, even experienced OJs, indicated these steps helped them to focus on each aspect independently and quickly.

Cognitive and psychological factors can influence OJs perception, such as unfamiliarity with the odor or lack of training in articulating the sensation in words or numbers (Leibowitz and Post, 1980; Meilgaard *et al*, 1999b). Study investigators adapted a 10-point ASTM odorant intensity scale (n-butanol) to a 5-point intensity scale (Table 2). To familiarized OJs with the study scale, training exercises included intensity rating by comparing sample odorants to the study reference odorants n-butanol. The purpose was to eliminate OJs tendency to distribute ratings evenly over the available scale or to avoid extreme responses, the 0's and 5's (Parducci, 1965). OJs were instructed to use the n-butanol intensity reference odorants and odorant descriptors whenever they felt the need to verify or confirm their assessment. The purpose was to aid in memorizing this static scale and anchor OJs to a physical reference. Judges reported that the use of the n-butanol reference helped when a sample was very unpleasant.

An OJ can have different responses to the same stimulus on different occasions because of over stimulation of the olfactory receptors causing the phenomena of adaptation, cross-adaptation, habituation and/or fatigue (Cain and Johnson, 1978; Brinkman *et al*, 1980; Dalton and Wysocki, 1996; Rawson, 1999). A variety of distinct odorant types (floral, fruity, amines, acids, medicinal, etc.) were used to eliminate adaptation and cross-adaptation. OJs assessed samples at 3-min intervals to control fatigue. All OJs reported that the 3-min intervals seemed long that they became bored while waiting. Investigators are not sure if this loss of focus during testing had any effect on the study results.

As repeated sniffing past the first sniff rarely provides additional information (Laing, 1983), recovery time probably could be shortened.

All study OJs improved in their assessment of odor intensity from pre- to post-training for all the odorants, except for DMS for which errors became worse. Investigators mixed fresh solutions of DMS at both measurement occasions. As DMS volatizes easily, the repeated opening of the sample containers may have diluted the sample. Also sample containers were glass and sulfur-containing compounds can be reactive to glass (Sulyok *et al*, 2001).

Intensity scores differed between pleasant and unpleasant odorants when water scores were included with the pleasant odorants. However, there was no significant difference when the water score was removed from the pleasant odorant group. This suggests that offensive compounds may not be needed to train malodor judges, or that a pleasant reference odorant could be used during training.

This study was limited in that it only included six OJs. Also, only two pleasant odorants were used and there was no control group that did not receive training. True oral malodor is a mixture of compounds and this study used only pure compounds. Further studies are needed to justify the assumption that training on these odorants is transferable to other odorants. Follow-up studies should include training and assessment with mixtures of compounds and with subjects having oral malodor and identification of a reference odorant for use in studies of people with oral malodor.

Conclusions

The OJ training protocol in this pilot study improved subjects' ability to assess odor intensity irrespective of previous experience and training and is recommended for all sensory OJs prior to conducting research trials. Because study OJs did not score pleasant and unpleasant odorants differently perhaps a pleasant reference odorant could be identified and used in OJ training rather than using human subjects with oral malodor. This could reduce cost and time involved in training OJs. Few academic groups have developed training programs and as there is no standardized training protocol established and published in the oral malodor field this pilot study could possibly be used as a basis to develop a standardized training program in this field.

References

ADA Council on Scientific Affairs (2003). Products used in the management of oral malodor. *JADA* **134:** 209–214.

Allison VC, Katz SH (1919). An investigation of stenches and odours for industrial purposes. *J Ind Eng Chem* **11:** 336–338.

Brinkman D, Goetzinger J, Cotton F *et al* (1980). Human evaluations of warning odorants in propane. In: Kuznicki J, Rutkiewic E, Johnson R, eds. *Selected sensory methods: problems and approaches to measuring hedonics*. Standard Technical Publications, ASTM: West Conshohocken, PA, **773**, pp. 49–59.

Chambers E, Wolf MB (eds) (1996). *Sensory testing methods*, 2nd edn. ASTM Manual Series, West Conshohocken, PA, **26**.

Leibowitz H, Post R (1980). Capabilities and limitations of the human being as a sensor. In: Kuznicki J, Rutkiewic E, Johnson R, eds. *Selected sensory methods: problems and approaches to measuring hedonics*. Standard Technical Publication, ASTM: West Conshohocken, PA, **773**, pp. 3–10.

Subcommittee E 18.04 on Fundamentals of Sensory (1999). Standard practices for referencing suprathreshold odor intensity. *Annual book of ASTM standards*. ASTM E: West Conshohocken, PA, **11.01**, pp. 544–599.

Wheeler J, Hoersch H, Mahy H *et al* (eds) (1981). *Guidelines for the selection and training of sensory panel members*. Standard Technical Publication, ASTM: West Conshohocken, PA, 758.

Cain W (1988). Olfaction, perception and motivation. In: Atkinson R, Stevens SS, eds. *Stevens' handbook of experimental psychology*. New York, USA: Wiley, pp. 409–459.

Cain W, Johnson F (1978). Liability of odor pleasantness; influence of mere exposure. *Perception* **7:** 459–465.

Dalton P, Wysocki C (1996). The nature and duration of adaptation following long-term odor exposure. *Percept Psychophysics* **58:** 781–792.

Harper R (1972). *Human senses in action*. Churchill Livingston: London, UK, 358 pp.

Kostelc J, Preti G, Zelson P *et al* (1984). Oral odors in early experimental gingivitis. *J Periodont Res* **19:** 303–312.

Laing DJ (1983). Natural sniffing gives optimum odor perception for humans. *Perception* **12:** 99–117.

Lenton P, Majerus G, Levitt M *et al* (2004). Comparison of new breath-testing device to establish assessment methods. *J Dent Res* **83** (special issue A)**:** No. 3330 (abstract).

McGinley C, McGinley M (2002). Odor testing biosolids for decision making. *Proceedings at Water Environment Federation Specialty Conference*, Austin, USA. http://www.fivesenses.com/Documents/Library/37%20Odor%20Testing%20Biosolids%20for%20Decision%20Making.pdf (Last accessed 30 January 2005).

Meilgaard M, Civille G, Carr BT (1999a). *Sensory evaluation techniques*, 3rd edn. CRC Press: Boca Raton.

Meilgaard M, Civille G, Carr BT (1999b). *Sensory evaluation techniques*, 3rd edn. CRC Press: Boca Raton, pp. 1–14.

Pangborn R (1981). Individuality in response to sensory stimuli. In: Solms J, Hall R, eds. *Criteria of food acceptance. How man chooses what he eats*. Forster-Verlag: Zurich, Switzerland, 177 pp.

Parducci A (1965). Category judgement: a range-frequency model. *Psycho Rev* **72:** 407–418.

Rabin M (1988). Experience facilitates olfactory quality discriminations. *Percept Psychophysics* **44:** 532–540.

Rawson N (1999). Cell and molecular biology of olfaction. *Quint Inter* **30:** 335–341.

Rosenberg M, Septon I, Eli I *et al* (1991a). Halitosis measurement by an industrial sulphide monitor. *J Periodontol* **62:** 487–489.

Rosenberg M, Kulkarni G, Bosy A *et al* (1991b). Reproducibility and sensitivity of oral malodor measurements with a portable sulphide monitor. *J Dent Res* **70:** 1436–1440.

Schmidt NF, Missam SR, Taber WJ *et al* (1978). Correlation between organoleptic mouth-odor rating and levels of volatile sulfur compounds. *Oral Surg* **45:** 560–567.

Shiffman HR (1996). *Sensation and perception. An integrated approach*, 4th edn. John Wiley & Sons: New York, USA.

Stevenson R, Boaks R, Prescott J (1998). Changes in odor sweetness resulting from implicit learning of a simultaneous odor sweetness association: an example of learned synesthesia. *Leaning Motiv* **29:** 113–132.

Sulyok M, Haberhauer-Troyer C, Rosenberg E *et al* (2001). Investigation of the storage stability of selected volatile sulfur compounds in different sampling containers. *J Chromatogr A* **917:** 367–374.

Tonzetich J (1971). Direct gas chromatographic analysis of sulfur compounds in mouth air of man. *Arch Oral Biol* **16:** 587–597.

Yaegaki K (2002). The present status of halitosis treatment: Future of dental clinics predicted based on present North American dentistry. In: Ito K, Hanada K, eds. *The Quintessence Year Book*. Quintessence Publications: Tokyo, Japan, pp. 35–43.