Chemical Senses vol. 28 no. 9 © Oxford University Press 2003; all rights reserved.

Influences of Age and Sex on a Microencapsulated Odor Memory Test

Eric S. Choudhury, Paul Moberg and Richard L. Doty

Smell and Taste Center, Department of Otorhinolaryngology: Head and Neck Surgery, Department of Psychiatry, and School of Dental Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Correspondence to be sent to: Richard L. Doty, Ph.D., Director, Smell and Taste Center, University of Pennsylvania Medical Center, 5 Ravdin Building, 3400 Spruce Street, Philadelphia, PA 19104, USA. e-mail: doty@mail.med.upenn.edu

Abstract

While the influences of such variables as age and sex are well established for most standardized tests of odor identification and detection, this is not the case for tests of odor memory. In this study, 231 non-smoking men and women, ranging in age from 10 to 68 years, were administered a standardized 12-item match-to-sample microencapsulated odor memory test (OMT). Anosmics were excluded from the study. Each participant was asked to smell a target odorant after its release from a micro-encapsulated odorant pad and then, after a delay interval of 10, 30 or 60 s, to pick the target from a similarly presented set of four odors, three of which were foils. Backward counting by threes was required during the delay intervals in an effort to minimize semantic rehearsal. Overall OMT scores were higher for women than for men, and decreased, in each sex, as a function of age in a manner similar to the age-related decline observed in tests of odor identification and detection. Performance did not change as a function of delay interval. A significant correlation between the overall OMT test scores and scores on the University of Pennsylvania Smell Identification Test was observed for women, but not for men, in accord with the notion that women may be more likely to employ semantic cues in their strategies to remember odors. The findings are discussed in light of the complexities of the construct of odor memory.

Key words: age, memory, olfaction, sex, smell identification, smoking

Introduction

Advances in psychophysical measurement and the proliferation of easy-to-use standardized tests of olfactory function employing microencapsulation odorant technology have significantly increased our understanding of the sense of smell in humans, including the influences of such factors as age (Doty et al., 1984a), gender (Doty et al., 1985), cigarette smoking (Frye et al., 1990), exposure to toxic agents (Schwartz et al., 1989), and dozens of diseases and medical disorders (Doty, 1997; Doty, 2003a; Murphy et al., 2003). This proliferation has been profound. For example, the 40-odor University of Pennsylvania Smell Identification Test (UPSIT; known commercially as the Smell Identification Test™ or SIT), a test which served as the model for a widely publicized odor survey conducted by the National Geographic Magazine (Gilbert and Wysocki, 1987), has been translated into six languages and administered to thousands of persons throughout the world.

Despite such advances and the continued proliferation of microencapsulated odor tests into the medical and academic communities, it is still not clear to what degree different types of olfactory tests measure the same elements of sensory function, and whether they are differentially sensitive to such variables as subject age and sex. Comparisons of results across tests have been fraught with difficulties, not the least of which is the fact that the reliability of many tests is suspect or unknown (Doty et al., 1995). Moreover, it is not clear to what degree nominally distinct olfactory tests measure independent processes. In one study, for example, nine olfactory tests, including tests of odor detection, identification, discrimination, memory, and suprathreshold intensity and pleasantness perception, were administered to 97 healthy subjects (Doty et al., 1994). A principal components analysis of the correlations among 13 measures derived from these tests revealed four meaningful components. The first component was composed of strong loadings from a range of measures, including ones from tests of odor identification, detection, discrimination, memory, and suprathreshold intensity and pleasantness perception, whereas the second component was largely composed of primary loadings from intensity ratings given to suprathreshold stimuli, whereas the third and fourth components appeared to represent hedonic and response criterion processes, respectively.
Relative to tests of odor identification and detection threshold sensitivity, tests of odor memory have received scant attention, and only one such test is available commercially. The Proustian view of odor memory is that odors are not forgotten to the same extent as other perceptual events (Engen and Ross, 1973). Early support for such a concept came from evidence that as retention intervals are increased, odor memory, but not visual recognition memory, remained virtually unchanged (Engen and Ross, 1973; Engen et al., 1973). While odor memory, as so envisioned, does appear to be reasonably stable over time, its superiority over other forms of sensory memory has been questioned. Thus, there is evidence that some olfactory stimuli are, in fact, more prone to forgetting than some visual and tactile stimuli, and that memory performance largely reflects the richness of the stimulus input, including emotional connotations (Herz, 1998; Larsson and Backman, 1998). Moreover, under normal circumstances, odor memories—like most other memories—are not divorced from associations with semantic and episodic memory systems (Murphy et al., 1991; Larsson and Backman, 1993). Interestingly, reliance upon such systems appears to differ among the sexes. Thus, women tend to outperform men on episodic odor memory tasks, particularly ones tapping verbal abilities, suggesting they may use a richer set of associations than men in both encoding and retrieving odor memories (Larsson et al., 2003).

A major goal of the present study was to evaluate the influences of age and gender on a standardized test of short-term odor memory that employs microencapsulation technology for odorant delivery. Another goal was to determine whether such influences, if present, are related to odor identification ability, as measured by the UPSIT (Doty et al., 1984b). If women, in fact, employ a more multifaceted strategy for remembering odors than men (e.g. by relying more heavily on cues related to odor identification), one would expect stronger relationships between odor memory and odor identification test scores in women than in men.

### Materials and methods

#### Subjects

Data were collected from 231 subjects who had never smoked cigarettes. Seventy-five were male (mean age ± SD = 33.21 ± 12.74) and 156 were female (mean age ± SD = 44.81 ± 17.05). The distribution of sexes and ages within the six age categories used in this study are presented in Table 1. Most of the subjects were recruited from advertisements placed in a local newspaper or on campus bulletin boards, and were paid ~$10.00/h for their participation. None were anosmic, as evidenced by UPSIT scores > 18 (Doty, 1995). Persons who reported a history of head trauma were excluded from participation. Written informed consent was obtained from the participants in accordance with the University’s Committee for the Study of Human Beings.

#### Experimental Design and Statistical Analyses

Each subject was administered a standardized microencapsulated odor memory test and the UPSIT, bilaterally, during the same test session. All testing occurred in a 2.44 × 2.60 × 2.44 m climate-controlled room at the Smell and Taste Center that was designed specifically for chemosensory testing. Only the examiner and examinee were present during test administration. The odor memory test was a 12-item, single-target, four-alternative, forced-choice test with 10, 30 and 60 s delay intervals, commercially known as the Odor Memory Test™ (OMT; Sensonics, Haddon Heights, NJ) (Doty, 2003c). This test, based upon the Peterson and Peterson (1959) match-to-sample paradigm, is a 12-item version of the nine-item odor memory test developed by Bromley and Doty (1995). A unique feature of the OMT, which is pictured elsewhere (Doty, 2003b), is the ease of presentation of the stimuli, since the test administrator must simply scratch open a set of microencapsulated odors with a pencil for the subject to sample at the appropriate time points. The microencapsulated odorants used in this test—amyl acetate (EA; Purified Grade, Fisher Scientific,
Fairlawn, NJ), phenyl ethyl alcohol (PEA; P-6134, Sigma Chemical Company, St Louis, MO), peppermint (MINT; peppermint oil, 3M Corporation, Minneapolis, MN) and peanut (PNUT; a mixture of mainly pyrazines from Arcade Marketing, New York, NY)—have been previously employed in the UPSIT and were chosen to be of similar intensity (see figure 1 in Doty et al., 1984b), making the qualitative stimulus dimension the factor to be discriminated.

In this test, a target odorant was released by scratching the odorized label that was then presented to the subject for sampling. After a given delay interval, four microencapsulated odorants (three of which were foils) were similarly released and presented at ~5 s intervals. The subject’s task was to report which odor in the odor response set was the same as the target stimulus. During the delay interval, each subject was required to count aloud backwards by three from 280 to minimize verbal rehearsal (cf. Engen et al., 1973; Murphy et al., 1991). The presentation order of the stimuli was counterbalanced such that (i) all target odorants occurred an equal number of times at each delay interval; (ii) each target odorant was presented at a given delay interval once in each of the four possible response positions (a, b, c and d); and (iii) all four odorants were presented in the first, second, and third segments of the test. The presentation of a given stimulus or delay interval never followed that of another stimulus or delay interval more than twice in succession. The order of delay intervals was as follows for the 12 items of the test: 10, 30, 30, 60, 10, 60, 30, 60, 30, 10 and 10 s. The target odors (TO) and the response set choices (RC) for each of these 12 items were as follows (correct response in italics): (1) TO: PEA; RC: EA, PNUT, PEA, MINT; (2) TO: MINT; RC: PEA, MINT, EA, PNUT; (3) TO: PNUT; RC: EA, PEA, MINT, PNUT; (4) TO: EA; RC: EA, PNUT, PEA, MINT; (5) TO: MINT; RC: MINT, EA, PEA, PNUT; (6) TO: PEA; RC: EA, PEA, PNUT, MINT; (7) TO: PNUT; RC: MINT, EA, PNUT, PEA; (8) TO: EA; RC: PNUT, MINT, EA, PEA; (9) TO: MINT; RC: PEA, EA, PNUT, MINT; (10) TO: PEA; RC: PEA, MINT, EA, PNUT; (11) TO: EA; RC: PNUT, PEA, MINT, EA; and (12) TO: PNUT; RC: EA, PNUT, MINT, PEA. The number of items correctly answered at each delay interval served as the dependent measure. In all cases, the UPSIT was administered after the odor memory test to minimize potential transference of verbal categorization to the latter test.

Results

In general, an age-related decline in the total test score was noted for both men and women from the fourth to the seventh decades of life; men exhibited lower scores than women at all ages (Figure 1). A sex (M, F) by age group (10–19, 20–29, 30–39, 40–49, 50–59 and 60–69 years) by delay interval (10, 30 and 60 s) analysis of variance (ANOVA) with repeated measures on the last factor found significant main effects of sex \(F(1,219) = 11.59, P < 0.001\) and age group \(F(5,219) = 8.32, P < 0.000\), but no significant effects of delay interval or any interactions among factors (all \(P > 0.29\)). For men and women combined, paired comparisons revealed statistically significant differences between the mean OMT scores of the subjects within the following age groups: 30–39 and 60–69 \((P = 0.0001)\), 40–49 and 50–59 \((P = 0.019)\), and 40–49 and 60–69 \((P = 0.01)\). Although analogous differences were noted within each sex separately, the small sample sizes in some age groups made valid statistical comparisons problematic. For the women, the \(P\) values of the aforementioned respective age group comparisons were 0.004, 0.099 and 0.009. For the men, the only \(P\) value that approached significance at the 0.05 \(\alpha\) level was that for the comparison between the 30–39 and 60–69 age groups \((P = 0.067)\).

To determine whether an association was present between the OMT and UPSIT scores, we computed Pearson product moment correlations between these measures for men and women separately. The correlation was statistically significant for the women \((r = 0.48, P < 0.000)\), but not for the men \((r = -0.025, P = 0.83)\), in accord with the notion that women may be more likely to employ identification or semantic cues in remembering odors than men. However, it appears unlikely that this association explains the aforementioned sex or age effects, since the addition of the UPSIT as a covariate had no appreciable influence on the findings, despite being statistically significant \([\text{UPSIT } F(1,218) = 15.87, P < 0.000; \text{sex } F(1,218) = 12.77, P < 0.001; \text{age group } F(5,218) = 8.56, P < 0.000; P_s > 0.26 \text{ for all other factors and interactions}]\).

Discussion

The present study demonstrates that the overall OMT score is systematically influenced by both sex and age. However, no sex- or age-related influences were observed for the delay interval component of the test, and no differences in
performance across delay intervals were present. The latter finding is in keeping with those from other odor memory studies showing similarly stable binasal performance across both short- and long-term delay intervals of young and elderly men and women (Engen and Ross, 1973; Engen et al., 1973; Murphy et al., 1991; Lehrner, 1993). This may not be the case, however, with some types of patients (e.g. those with hyposmia secondary to severe head trauma or temporal lobe epilepsy) and with long-term memory tasks that associate unfamiliar verbal labels with unfamiliar odors (Dempsey and Stevenson, 2002). Moreover, tests of uninasal function, as indexed by left and right sides of the nose separately, may also tell a different story, conceivably reflecting the need for additive input from both sides of the nose to more clearly encode the initial stimulus and maintain optimal performance across delay intervals (Bromley and Doty, 1995).

As shown in Figure 1, women generally outperformed men on the OMP across a wide age span, a finding similar to that observed in studies of odor identification, as well as in some studies of odor detection threshold (LeMagnen, 1952; Köster and Koelega, 1976; Doty et al., 1986). The reasons for this sex difference are probably multiple. First, detection thresholds for at least two of the odorants employed in the OMT (i.e. amyl acetate and phenyl ethyl alcohol) have been reported to be lower in women than in men (e.g. Koelega, 1970, 1994; Doty, 1986; Doty et al., 1986), implying that women may perceive weak odors as more salient. The reliability of this influence, however, is somewhat questionable, as odor threshold sex differences are more labile than analogous auditory and tactile threshold sex differences (for review, see Velle, 1987) and are not observed in all studies (e.g. Deems and Doty, 1987; Betchen and Doty, 1998). Secondly, women appear to have more acute suprathreshold odor perception than men. Thus, they rate a wide range of suprathreshold odorants as more intense or stronger than do men, including various body odors (for review, see Doty, 1986). In conjunction with greater basal sensitivity, this suprathreshold enhancement would be expected to aid in providing not only a more salient stimulus for women to encode, but in possibly providing greater resistance to adaptation or habituation. The OMT requires that a relatively large number of stimuli be smelled (i.e. $5 \times 12 = 60$ total) in a short period of time, with only $5$ s intervening between the presentation of the response alternatives, suggesting the possibility of susceptibility to such processes. Thirdly, there appears to be a fundamental difference between men and women in their reliance upon semantic strategies for both encoding and recalling odor memories (Hertlitz et al., 1999; Lewin et al., 2001; Larson et al., 2003), and there is considerable evidence that such strategies enhance performance on odor memory tests (Rabin and Cain, 1984; Walk and Johns, 1984; Lyman and McDaniel, 1990; Lehrner, 1993). For example, Oberg et al. (2002) reported that when familiar odors (i.e. ones that could be identified) were employed, females outperformed males. When unfamiliar odors were used, this sex difference disappeared. The present observation of a significant correlation between UPSIT and OMT scores in women, but not in men, is congruent with the general notion that women are more likely than men to employ semantic strategies in encoding and remembering odors. However, this association does not completely account for the sex difference on the OMT, since the sex difference effect was not meaningfully altered statistically by including UPSIT scores as a covariate.

The second major point illustrated in Figure 1 is that total OMT scores decreased with age in both males and females. The age-related decline in performance on the OMT appears to be first statistically meaningful between the fifth (i.e. 40–49 years) and sixth (i.e. 50–59 years) decades of life, and roughly parallels age-related decreases noted in tests of odor identification and detection (Doty et al., 1984a,b; Deems and Doty, 1987), as well as in some measures of auditory, visual, gustatory, and somatosensory function (Pitts, 1982; Jerger, 1973; Hinchcliffe, 1962). These age-related associations may be responsible, in part, for why the OMT loads strongly on the same principal component in a principal components analysis as odor identification and detection tests (Doty et al., 1994). Like the sex difference observed in the present study, the age-related changes in OMT scores presumably reflect multiple factors. Among the physical correlates of aging in the olfactory system proper are decreased numbers of olfactory receptor cells, decreased numbers of olfactory glomeruli, altered vascularity within the olfactory epithelium, loss of neurotrophic factors, decreased mitotic activity within the neuroepithelium, and increased viscosity of mucus (for review, see Doty, 2001). Interestingly, there is now strong support for the observation of Krmpotic-Nemanić (1969) of an age-related decline in the number of patent cribriform plate foramina, reflecting appositional bone growth that, in effect, pinches off olfactory nerves as they course from the nasal epithelium through the ethmoid bone to the olfactory bulb (Kalmeay et al., 1998). As with sex differences, the type of strategy employed in remembering odors may contribute to the age-related effect. Thus, many elderly have fewer semantic resources to rely on during odor encoding and recall, contributing to difficulties in remembering odors (Schemper et al., 1981). Interestingly, however, some declines in such semantic resources may be due to poorer integration of information secondary to decrements in hearing and vision (Baltes and Lindenberger, 1997).

It is doubtful that odor memory can be conceptualized as a monolithic construct or that a test such as the OMT is, in fact, solely measuring such a construct. For the most part, odors are used to signify and identify environmental animate or inanimate objects, reflected by the fact that nearly all odors are primarily classified according to an object referent (peach, lemon, apple, pizza, strawberry, motor oil, leather, peppermint, fecal matter, medicine-like,
et al.) or, in cases where such referents are not available, according to hedonic or broad semantic categories (disgusting, pleasant, unpleasant, fresh, green, lively, bland). [The object itself may determine the odor employed in describing it. Morrot et al. (2001) showed that odors assigned to wines by wine tasters are largely represented by objects that have a similar color to the wine. When white wines, which received such descriptors as honey, lemon, grapefruit, and peanut, were artificially colored red, wine tasters switched their descriptions to reflect objects associated with red (e.g. prune, bilberry, cherry, cedar, violet, cinnamon, etc.).] Hence, when a subject smells an odor, there is a strong tendency to rapidly associate the odor with an object or hedonic property (i.e. to identify the source of the odor, if known) and to put into memory this association. Once this has occurred, then what is recalled later need not be the actual sensation of an odor, but simply the recollection of having smelled an odor with a name reflecting the name or source of the odor (e.g. lemon), a recollection cued by the presentation of the stimulus. Hence, often the subject is remembering, in effect, ‘I smelled a lemon or an odor that smells like lemon’ and later, when lemon odor is presented, ‘I recall having smelled this stimulus, i.e. lemon’. During the retention interval, knowledge of what a lemon smells like was always present in long-term memory. This is quite a different paradigm than that initially employed by Ebbinghaus (1913) for memories of nonsense syllables, in which relatively unique entities are put into memory and later recalled without much confounding from other associations, and may explain the close correspondence between multidimensional space obtained by using odors alone with that obtained by using just the concepts of odors (Carrasco and Ridout, 1993). Wilson and Stevenson (2003) have argued that, in fact, all odors are initially encoded as ‘objects’ in the anterior piriform cortex, reflecting a synthesis of feature-detector information from the olfactory bulb mitral cells. They hypothesize that ‘odor perception is wholly dependent on the integrity of this <piriform> memory system.’

Recently, Larsson (2002) has pointed out the complexity of odor memory paradigms and suggested that a fruitful approach to examining memory for odors may be that outlined by Schacter and Tulving (1994). From this perspective, memory systems can be divided into non-declarative (i.e. perceptual representation and procedural) and declarative (i.e. semantic, work, and episodic) classes, with the latter being more influenced by such factors as age and sex. Procedural memory would be involved, for example, in the development of conditioned odor or taste aversions, whereas episodic memory would be involved in attempting to remember a specific odor over time, as explicitly assessed in the present work.

While the present data demonstrate that a microencapsulated odor test of nominal odor memory is sensitive to sex and age, when no delay interval influences are present the test could be construed as mainly an odor discrimination test. This begs the question as to when no delay interval effects are present or when odor memory effects are assumed without explicit variation in delay intervals (e.g. Hudry et al., 2003), whether such a test provides any additional information about an individual’s ability to smell than that provided by tests of odor identification, discrimination, or detection threshold. This question is particularly salient in light of correlations among these types of tests, and the fact that they exhibit rather robust loadings, in a principal components analysis, on the same principal component (Doty et al., 1994). Since tests of odor memory, discrimination, identification and detection are not mutually exclusive and are interdependent upon common elements (e.g. sensitivity, various forms of memory including, in some cases, the ability to identify overtly or covertly the stimuli), it is conceivable that their relative efficacies reflect idiosyncratic factors (e.g. number and types of odors, time between sampling epochs, degree to which attention is engaged). Such factors would be expected to influence, in turn, their fidelity of measurement (as indexed, for example, by reliability). Future research is needed to disentangle the perceptual and cognitive elements involved in various types of olfactory tests and to determine under what circumstances distinct sets of olfactory traits, such as odor detection, identification and memory, can be clearly differentiated.

Acknowledgements

We thank Kara-Lynne Kerr, Jessica Neff, Cheryl Brown, Isabelle Tourbier, Al Hahn, Catherine Balderston, David Roalf and Kiana Owzar for the testing of a number of the subjects of the study. This research was supported, in part, by the following grants from the National Institutes of Health, Bethesda, MD: PO1 DC 00161, RO1 DC 04278, RO1 DC 02974, RO1 AG17496 and RO1 MH 63381.

References


Accepted October 20, 2003