Development and Validation of the Olfactory Memory Test Battery (OMTB) Based on Odors with High- and Low-Verbalizability

Shu-bin Li, M.S.[a]; Chao Yan, Ph.D[b]; Thomas Hummel, M.D[c]; Lai-quan Zou, Ph.D [a] *

[a] Chemical Senses and Mental Health Lab, Department of Psychology, School of Public Health, Southern Medical University, Guangzhou, Guangdong, China

[b] Key Laboratory of Brain Functional Genomics (MOE&STCSM), Shanghai Changning-ECNU Mental Health Center, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China

[c] Smell & Taste Clinic, Department of Otorhinolaryngology, TU Dresden, Dresden, Germany

* Correspondence to:

Lai-quan Zou. Chemical Senses and Mental Health Lab, Department of Psychology, School of Public Health, Southern Medical University, No. 1023 Shatainan Road, Guangzhou 510515, Guangdong, China. Email: zoulq@smu.edu.cn

Abstract

Odor memory is an important field of clinical research for its distinctive characteristics, which differ from those of other sensory systems. To date, several tests have been implemented for the assessment of odor memory. Despite a range of studies demonstrating the importance of verbal mediation in odor memory, few have distinguished odor memory performance in different odor verbalization levels. We aimed to develop a standardized odor memory test toolbox with one group of odors that are easily verbally identified and the other group of odors that are difficult to identify. The study was conducted with 97 healthy participants. Twenty people were involved in a pre-test for odor preparation. Seventy-seven people took part in the formal test and 50 of them volunteered for a retest after 3 to 7 days. The test contained two odor categories (high- and low-verbalizability odors), each consisting of three subtests (short- and long-term memory and working memory). Satisfactory test-retest reliability of the odor recognition and working memory test were shown in both odor categories. We use both the Sniffin' Test of memory (16 items) and a digit span test as criteria: all subtests were shown to have high correlations with them, indicating that the Olfactory Memory Test Battery has a solid validity. Moreover, people scored significantly better with high- than low-verbalizability odors. A negative age effect on odor memory performance was also found. Results from the present study indicated that odors with high verbalizability produced higher scores in the memory test compared to odors with low verbalizability. In addition, the Olfactory Memory Test Battery is a comprehensive assessment of odor memory with good reliability and validity. It contains high- and low-verbalizability odors and each category has three subtests. They can be used independently or in combination according to clinical and research needs.

Keywords: Olfactory memory; Working memory; Verbalizability; Reliability

Introduction

Odor memory represents an important biological function, and includes storage and retrieval of olfactory information. Both daily life experiences and many research studies have shown the difficulty of remembering and naming a smell (de Wijk et al., 2014; Engen, 1991; Goldman & Seamon, 1992). The peculiar and distinctive system features of odor memory differentiate it from visual and auditory memories (Keller, 2009; Lehn et al., 2013; Zucco, 2003). Moreover, impaired human odor memory ability has been found in some neurodegenerative diseases, such as Parkinson's disease (Eek et al., 2021) and Alzheimer's disease (El Haj et al., 2017).

For its clinical and scientifically recognized value, various odor memory tests have been developed over the decades: the Odor Memory Test (OMT) (Choudhury et al., 2003), the Sniffin' TOM (Sniffin' Test for Odor Memory) (Croy et al., 2015) and the TOM-32 (Sorokowska et al., 2020) are the most common test tools. Twelve-item match-to-sample odors are included in the OMT, a so-called "scratch and sniff test". Each item has four alternative and forced-choice options, comprising the four microencapsulated odors amyl acetate, phenyl ethyl alcohol, peppermint, and peanut. Participants are asked to smell a certain target odor and pick the target from four odors after a delay of 10s, 30s, and 60s. In Sniffin' TOM, the participants have to pick 8 target odors from 16 common odors which are included in the common Sniffin' Sticks test (Hummel et al., 1997) (based on odor presentation in felt-tip-like pens), and in TOM-32 16 target odors are presented together with 16 distractor odors.

Verbal mediation has been controversially discussed in studies on olfactory memory. Most studies confirmed the importance of verbal mediation for episodic olfactory memory (Herz & Engen, 1996; Jehl et al., 1997; Jönsson et al., 2011; Larsson, 1997; Lehrner et al., 1999; Öberg et al., 2002; Seubert et al., 2020). Successful odor recognition and identification are sensitive to correct and consistent odor labeling (Herz & Engen, 1996; Larsson, 1997). Associations between gray matter volume in temporal and prefrontal olfactory areas was found in the most recent neuroimaging study, which compared the potential overlap and differences in brain areas between semantic and episodic odor memory (Seubert et al., 2020). Moreover, it has been reported that odor naming is difficult even when olfactory processing and visual naming are not impaired (Olofsson & Gottfried, 2015); people still performed differently with high- and low-verbalizability odors (Jönsson et al., 2011; Larsson, 1997). Despite the many advantages of previous odor memory tests, the odors used in these tests are all known well from everyday life. Therefore, using these tools it is difficult to disentangle odor memory ability from the memory of verbal labels.

Working memory describes the short-term storage system which is different from the regular short-term memory because of its ability of information maintenance and modification (Atkinson & Shiffrin, 1968; Nee & D'Esposito, 2018). Numerous papers covering different aspects of visual and auditory working memory have been published to date (Atkinson & Shiffrin, 1968; Baddeley, 1992; Kumar et al., 2016; Luck & Vogel, 1997; Nee & D'Esposito, 2018; Vogel et al., 2001). In contrast, very few studies have attempted to discuss olfactory working memory. Dade et al. (2001) found that frontal cortical areas were not modality-specific activated. Jönsson et al. (2011) found that better performance was shown with higher-verbalizability odors. Existing tests only focus on odor recognition memory. As an essential part of olfactory function, olfactory working memory has been less discussed and has not been studied with specific test tools.

The present study aimed to develop a new odor memory test toolbox, including tests for high- and low-verbalizability odors in working memory, short- and long-term memory and to determine its reliability and validity. Giving the potential influence of age and gender on olfactory function that were found in previous studies (Attems et al., 2015; Chen et al., 2021; Hummel et al., 2002; Sorokowski et al., 2019), we also evaluated age- and gender-effect of the OMTB.

Methods

Participants

Ninety-seven participants were recruited from Southern Medical University and surrounding communities. Twenty participants performed intensity-match and discrimination tasks in the odor preparation. Seventy-seven participants (21 males and 56 females, aged 28.1 ± 15.5, average education years 11.9 ± 4.4) took part in the following tasks and performed verbalization investigation tasks. Fifty of these 77 participants (18 males and 32 females, aged 33.3 ± 17.1, average education years 11.9 ± 5.6) volunteered for a retest after 3–7 days. All participants were free of upper respiratory tract infections, nasal diseases, traumatic head injuries, mental illness, other serious illness and drug abuse or alcohol addiction. This program has been approved by the Ethics Committee of Southern Medical University.

Materials

Low-verbalizability odors

We prepared 37 different odors, which can be classified into three categories: 14 floral smells, 12 plant-related smells and 11 fruity smells. In an intensity-match test, odors were diluted with propylene glycol and evaluated by 20 participants on a Likert-type scale from 1 to 9 (1=very weak, 9=very strong). We increased the concentration of odors whose intensity rating was under 5 and decreased the concentration of those above 5. Odors that were insoluble in propylene glycol or where the intensity rating cannot reach 5 even at a pure concentration were excluded. This process was repeated twice until most of the odors had a similar intensity. The dilution ratio of all odors is shown in the Supplement (Table S1).

We put nine components together to form an odor mixture: three with a floral smell, three with a plant-related smell and three with a fruity smell. Twenty odor mixtures were created and stored in brown glass bottles. Supplement (Table S2) shows the "recipe" for these mixtures. In a discrimination test of 20 mixtures, each mixture was compared with others one by one on a Likert-type scale from 1 to 9 (1=very similar, 9=totally different). The six most different odor mixtures were selected for a follow-up experiment (Number 4, 6, 7, 9, 17, 20) and were renamed Odor 7 to 12.

High-verbalizability odors

We chose six daily common odor materials (tea, banana, mint, orange, rose and peach) as high-verbalizability odors, which renamed as Odor 1 to 6 in order.

Procedure

The whole test consisted of two similar parts: high- and low-verbalizability odors. Each part included three subtests, focusing on odor long- and short- term memory and odor working memory. Additionally, all participants had to evaluate the intensity (1-9; 1=not at all, 9=very strong), pleasantness (1-9; 1= very unpleasant, 9=very pleasant) and verbalization (0 = no response or ambiguous response about valence of odor, like "good", "bad" or "disgusting"; 1 = a specific odor category, like "floral", "some kind of candy or air freshener"; 2 = a specific thing) of the odors.

In the long-term memory test, participants were asked to remember a target odor that was presented 1–2cm below both nostrils for 3s. They were expected to pick the target odor after an interval of 60s from three different pseudorandomly presented odors. The interstimulus interval was 8s and participants were required to count backwards aloud by 3 from 270 during the 60s interval for distracting and reducing verbal rehearsal. The short-term memory test had the same rules except the interval was 10s. The order of odor presentation is shown in the Supplement (Table S3-S5).

In the odor working memory test, six odors (high- or low-verbalizability odors) were presented pseudo-randomly in a two-back paradigm. There were 18 trials in total with each odor presented three times. Each stimulus was presented for 3s, with an 8s interstimulus interval. Participants were asked to identify whether the odor they were experiencing was similar to the one before last. They were expected to respond in no more than the interstimulus interval (8s) in cases where the follow-up proceeded normally. Before the formal test, the interviewer would give each participant full and clear instructions and let them have four practice trials to ensure they understood the rules exactly.

In the assessment of validity, Sniffin' TOM (Croy et al., 2015) and digit span tests (Ostrosky - Solís & Lozano, 2006; Schroeder et al., 2012) were conducted. For reliability, 50 participants in the first test session volunteered to have a retest after 3–7 days. All participants were tested individually and were asked to have a break between two tests for at least 3 minutes with the help of a coffee bean smell for dishabituation (Pellegrino et al., 2017).

Statistical Analysis

SPSS 22.0 was used for statistical analysis. Several scores were computed for OMTB: correct items in the short- and long-term memory test, and correct items in the recognition memory and working memory. Pearson correlations and the Bland-Altman analysis [mean difference (MD) and confidence interval (CI)] were used to evaluate validity and test–retest reliability of OMTB. *T*-tests for independent samples were conducted to examine the differences in scores between age, gender and odor groups.

Results

Descriptive statistics

Table 1 shows the valence of the 12 odor materials used in the test, while Table 2 shows the descriptive statistics of test scores. There were statistically significant differences in all OMTB scores between high- and low-verbalizability odors. Except for the T of false alarm rate ($T = 4.26$,

$p < 0.01$), high-verbalizability odors showed higher scores than low-verbalizability odors (all $p < 0.05$)

Test–retest reliability

Satisfactory test–retest reliability was found by both Pearson correlations and the Bland–Altman analysis. According to scatter plots, the results from the two sessions with high- and low-verbalizability odors showed satisfactory test–retest reliability (see Figure 1 and 2, all $p > 0.05$). The Bland–Altman plots (Table 3, Figure 3 and 4) demonstrated an acceptable agreement within thresholds defined by mean $\pm$ 1.96SD from the mean difference. Mean difference between test and retest scores of all subtests in high-verbalizability odors were -0.12, 0.22, 0.10 and -0.20, respectively (95% CI of STM 10s, -2.07 to 1.82; 95% CI of LTM 60s, -2.12 to 2.57; 95% CI of RM, -2.91 to 3.13; 95% CI of 2-back, -4.18 to 3.78). As for low-verbalizability odors, mean difference of all subtests were 0.18, -0.35, -0.04 and -0.71 (95% CI of STM 10s, -2.75 to 3.12; 95% CI of LTM 60s, -3.05 to 2.35; 95% CI of RM, -4.02 to 3.94; 95% CI of 2-back, -6.24 to 4.82).

Validity test

We used Sniffin' TOM and digit span tests to test the validity of the OMTB (Table 4). The total score of Sniffin' TOM had significant correlations with all scores of the OMTB test for the high-verbalizability odors (all $p < 0.01$). Except for the long-term odor memory score, the total score of TOM-16 had significant correlations with other scores of the OMTB test for low-verbalizability odors (all $p < 0.05$). Scores of digit span test also had strongly positive correlations with OMTB scores (all $p < 0.01$).

Age and gender effect

Older people showed worse performance in short- and long-term memory, recognition and working odor memory for both high- and low-verbalizability odors (all $p < 0.01$) (Figure 5 and 6). Most odor memory tasks showed no significant gender effect ($p > 0.05$), except that women had better performance than men on short-term memory tasks in low-verbalizability odors and 2-back working memory tasks in high-verbalizability odors (STM10s: $t = -2.19$, $p = 0.03$; 2-back: $t = -2.03$, $p = 0.046$).

Discussion

Our study introduced a relatively comprehensive olfactory memory test tool to balance some of the weaknesses of previous test tools. We used two kinds of odor materials (high- and low-verbalizability odors), which had similar valence. Each subtest reflected a different aspect of odor memory, which means that they can be used separately depending on the circumstances and needs.

The OMTB was found to have credible reliability and validity. All test scores showed significant correlations between the two test sessions. The olfactory recognition and working memory scores were shown to be stable with coefficients ranging from 0.5 to 0.8. The most standardized odor tests also showed similar reliability coefficients (Croy et al., 2015; Doty et al., 1996; Hsieh et al., 2017; Kobal et al., 1996; Sorokowska et al., 2020). However, the reliability of odor short- and long-term memory was not satisfactory, which can be explained by the insufficient number of items (only six). Therefore, we suggest that the recognition memory score (the sum of short- and long-term memory scores) is a better evaluation indicator in clinical and research assessment. The standardized odor memory test TOM-16 (Croy et al., 2015) and the digit span test, which is known for verbal working memory evaluation (Conklin et al., 2000; Reisberg et al., 1984; Woods et al., 2011), were tested in the validity test. Our results showed that all OMTB scores significantly correlated with TOM-16 and digit span scores.

Consistent with other studies, participants had better olfactory memory performance with high-verbalizability odors than with low-verbalizability odors, which is further testimony to verbal mediation in olfactory memory (Herz & Engen, 1996; Jönsson et al., 2011; Moss et al., 2019).

Furthermore, the age effect of our test tool was significant. Both *t*-test and correlation analysis showed that older participants had lower odor memory scores in all subtests than younger participants (Backman, 1997; Doty, 1989; Hummel et al., 2002; Sorokowska et al., 2015; Zhang & Wang, 2017; Hummel & Oleszkiewicz, 2020). Age-related deterioration of olfactory (Backman, 1997; Doty, 1989; Hummel et al., 2002; Sorokowska et al., 2015; Zhang & Wang, 2017; Hummel & Oleszkiewicz, 2020) and cognitive function (Grady & Craik, 2000; Hidalgo et al., 2019; Prull et al., 2006; Rhodes et al., 2019) resulting partly from the gradual decline of prefrontal cortex engagement could be a reason for this finding (Choudhury et al., 2003; Croy et al., 2015; Lehrner et al., 1999; Yaple et al., 2019). They are also the distinctive features of Alzheimer's disease (Bastin et al., 2019; Jahn, 2013; Jung et al., 2019; Lee et al., 2020; Murphy, 2019; Murphy et al., 1990; Roberts et al., 2016; Serby et al., 1991), suggesting that our toolbox could possibly be used in the clinical identification of mild cognitive impairment.

There was no effect of the factor gender in most of our odor memory tasks. Although it has been reported that women outperformed men in olfactory tasks (Choudhury et al., 2003; Croy et al., 2015; Doty & Agrawal, 1989), many studies had similar results with us that gender-effect on olfactory function was not significant (Kern et al., 2014; Öberg et al., 2002; Sorokowska et al., 2015). A recent meta-analysis reported that the effect sizes of gender ranged from g = 0.08 to g = 0.3 which was quite weak (Sorokowski et al., 2019) and this could explain the inconsistency between previous work and our results. However, the present results should be viewed with caution because of the relatively small sample of men.

Limitations of the present study should be addressed. First, the test–retest reliability of the OMTB for testing short- and long-term odor memory with low-verbalizability odors was not as good as other OMTB scores due to the limited number of items and high difficulty level of the tasks. Second, the sample size of the present study may not have been large enough. In addition, only healthy participants were included, which means that the validity of the OMTB is unknown for clinical assessment. Therefore, our future research will focus on norm-referenced scores and the validation of the test in a clinical context.

In conclusion, the present study introduced the OMTB as a new odor memory test; high-verbalizability odors showed higher scores than low-verbalizability odors. The OMTB is reliable and valid. It contains three subtests, each of which can be used separately, making it convenient for clinical and research usage. However, the combination of long- and short-term memory tests are recommended as recognition memory tests. Although awaiting further explorations it can be expected to be of use in the clinical identification of mild cognitive impairment and olfactory cognitive assessment.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. In K. W. Spence & J. T. Spence (Eds.), *Psychology of Learning and Motivation* (Vol. 2, pp. 89-195). Academic Press. https://doi.org/https://doi.org/10.1016/S0079-7421(08)60422-3

Attems, J., Walker, L., & Jellinger, K. A. (2015). Olfaction and Aging: A Mini-Review. *Gerontology*, *61*(6), 485-490. https://doi.org/10.1159/000381619

Backman, M. L. L. (1997). Age-related differences in episodic odour recognition: the role of access to specific odour names. *Memory*, *5*(3), 361-378.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559. https://doi.org/10.1126/science.1736359

Bastin, C., Delhaye, E., Moulin, C., & Barbeau, E. J. (2019). Novelty processing and memory impairment in Alzheimer's disease: A review. *Neurosci Biobehav Rev*, *100*, 237-249. https://doi.org/10.1016/j.neubiorev.2019.02.021

Chen, Z., Xie, H., Yao, L., & Wei, Y. (2021). Olfactory impairment and the risk of cognitive decline and dementia in older adults: a meta-analysis. *Braz J Otorhinolaryngol*, *87*(1), 94-102. https://doi.org/10.1016/j.bjorl.2020.07.009

Choudhury, E. S., Moberg, P., & Doty, R. L. (2003). Influences of age and sex on a microencapsulated odor memory test. *Chemical senses*, *28*(9), 799-805.

Conklin, H. M., Curtis, C. E., Katsanis, J., & Iacono, W. G. (2000). Verbal working memory impairment in schizophrenia patients and their first-degree relatives: evidence from the digit span task. *Am J Psychiatry*, *157*(2), 275-277. https://doi.org/10.1176/appi.ajp.157.2.275

Croy, I., Zehner, C., Larsson, M., Zucco, G. M., & Hummel, T. (2015). Test–Retest Reliability and Validity of the Sniffin' TOM Odor Memory Test. *Chemical senses*(3), 173-179.

Dade, L. A., Zatorre, R. J., Evans, A. C., & Jones-Gotman, M. (2001). Working memory in another dimension: functional imaging of human olfactory working memory. *Neuroimage*, *14*(3), 650-660.

de Wijk, R. A., Schab, F. R., & Cain, W. S. (2014). Odor identification. In *Memory for odors* (pp. 29-46). Psychology Press.

Doty, R. L. (1989). Influence of Age and Age-Related Diseases on Olfactory Functiona. *Annals of the New York Academy of Sciences*, *561*(1), 76-86. https://doi.org/https://doi.org/10.1111/j.1749-6632.1989.tb20971.x

Doty, R. L., & Agrawal, U. (1989). The shelf life of the University of Pennsylvania smell identification test (UPSIT). *The Laryngoscope*, *99*(4), 402-404.

Doty, R. L., Marcus, A., & Lee, W. W. (1996). Development of the 12-item Cross-Cultural Smell Identification Test (CC-SIT). *Laryngoscope*, *106*(3 Pt 1), 353-356. https://doi.org/10.1097/00005537-199603000-00021

Eek, T., Larsson, M., & Dizdar, N. (2021). Odor Recognition Memory in Parkinson's Disease: A Systematic Review. *Front Aging Neurosci*, *13*, 625171. https://doi.org/10.3389/fnagi.2021.625171

El Haj, M., Gandolphe, M. C., Gallouj, K., Kapogiannis, D., & Antoine, P. (2017). From Nose to Memory: The Involuntary Nature of Odor-evoked Autobiographical Memories in Alzheimer's Disease. *Chem Senses*, *43*(1), 27-34. https://doi.org/10.1093/chemse/bjx064

Engen, T. (1991). *Odor sensation and memory*. Greenwood Publishing Group.

Goldman, W. P., & Seamon, J. G. (1992). Very long-term memory for odors: Retention of odor-name associations. *The American journal of psychology*, 549-563.

Grady, C. L., & Craik, F. I. M. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology*, *10*(2), 224-231. https://doi.org/https://doi.org/10.1016/S0959-4388(00)00073-8

Herz, R. S., & Engen, T. (1996). Odor memory: Review and analysis. *Psychonomic bulletin & review*, *3*(3), 300-313.

Hidalgo, V., Pulopulos, M. M., & Salvador, A. (2019). Acute psychosocial stress effects on memory performance: Relevance of age and sex. *Neurobiol Learn Mem*, *157*, 48-60. https://doi.org/10.1016/j.nlm.2018.11.013

Hsieh, J. W., Keller, A., Wong, M., Jiang, R.-S., & Vosshall, L. B. (2017). SMELL-S and SMELL-R: Olfactory tests not influenced by odor-specific insensitivity or prior olfactory experience. *Proceedings of the National Academy of Sciences*, *114*(43), 11275-11284.

Hummel, T., Heilmann, S., & Murphy, C. (2002). Age-related changes of chemosensory functions. *Olfaction, taste, and cognition*, 441-456.

Hummel, T. & Oleszkiewicz, A. (2020) Age-related changes of chemosensory function. In Fritzsch, B. (ed) The senses: : A Comprehensive Reference. Academic Press, pp. 717-726

Hummel, T., Sekinger, B., Wolf, S. R., Pauli, E., & Kobal, G. (1997). 'Sniffin'sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chemical senses*, *22*(1), 39-52.

Jahn, H. (2013). Memory loss in Alzheimer's disease. *Dialogues Clin Neurosci*, *15*(4), 445-454. https://doi.org/10.31887/DCNS.2013.15.4/hjahn

Jehl, C., Royet, J. P., & Holley, A. (1997). Role of verbal encoding in shortand long-term odor recognition. *Perception & Psychophysics*, *59*(1), 100-110. https://doi.org/10.3758/BF03206852

Jönsson, F. U., Møller, P., & Olsson, M. J. (2011). Olfactory working memory: effects of verbalization on the 2-back task. *Memory & cognition*, *39*(6), 1023-1032.

Jung, H. J., Shin, I. S., & Lee, J. E. (2019). Olfactory function in mild cognitive impairment and Alzheimer's disease: A meta-analysis. *Laryngoscope*, *129*(2), 362-369. https://doi.org/10.1002/lary.27399

Keller, A. (2009). Odor memories: the first sniff counts. *Curr Biol*, *19*(21), R988-989. https://doi.org/10.1016/j.cub.2009.09.046

Kern, D. W., Wroblewski, K. E., Schumm, L. P., Pinto, J. M., Chen, R. C., & McClintock, M. K. (2014). Olfactory function in Wave 2 of the National Social Life, Health, and Aging Project. *J Gerontol B Psychol Sci Soc Sci*, *69 Suppl 2*(Suppl 2), S134-143. https://doi.org/10.1093/geronb/gbu093

Kobal, G., Hummel, T., Sekinger, B., Barz, S., Roscher, S., & Wolf, S. (1996). "Sniffin' sticks": screening of olfactory performance. *Rhinology*, *34*(4), 222-226.

Kumar, S., Joseph, S., Gander, P. E., Barascud, N., Halpern, A. R., & Griffiths, T. D. (2016). A Brain System for Auditory Working Memory. *J Neurosci*, *36*(16), 4492-4505. https://doi.org/10.1523/jneurosci.4341-14.2016

Larsson, M. (1997). Semantic factors in episodic recognition of common odors in early and late adulthood: a review. *Chemical senses*, *22*(6), 623-633.

Lee, Y. H., Bak, Y., Park, C. H., Chung, S. J., Yoo, H. S., Baik, K., . . . Lee, P. H. (2020). Patterns of olfactory functional networks in Parkinson's disease dementia and Alzheimer's dementia. *Neurobiol Aging*, *89*, 63-70. https://doi.org/10.1016/j.neurobiolaging.2019.12.021

Lehn, H., Kjønigsen, L. J., Kjelvik, G., & Håberg, A. K. (2013). Hippocampal involvement in retrieval of odor vs. object memories. *Hippocampus*, *23*(2), 122-128.

Lehrner, J. P., Glück, J., & Laska, M. (1999). Odor identification, consistency of label use, olfactory threshold and their relationships to odor memory over the human lifespan. Chemical Senses, 24(3), 337-346.

Lehrner, Walla, P., Laska, M., & Deecke, L. (1999). Different forms of human odor memory: a developmental study. *Neuroscience Letters*, *272*(1), 17-20. https://doi.org/https://doi.org/10.1016/S0304-3940(99)00566-2

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281. https://doi.org/10.1038/36846

Moss, A., Miles, C., Elsley, J., & Johnson, A. J. (2019). Olfactory working memory: exploring the differences in n-back memory for high and low verbalisable odorants. *Memory*, *27*(10), 1319-1344. https://doi.org/10.1080/09658211.2019.1653469

Murphy, C. (2019). Olfactory and other sensory impairments in Alzheimer disease. *Nat Rev Neurol*, *15*(1), 11-24. https://doi.org/10.1038/s41582-018-0097-5

Murphy, C., Gilmore, M. M., Seery, C. S., Salmon, D. P., & Lasker, B. R. (1990). Olfactory thresholds are associated with degree of dementia in Alzheimer's disease. *Neurobiology of aging*, *11*(4), 465-469. https://doi.org/https://doi.org/10.1016/0197-4580(90)90014-Q

Nee, D. E., & D'Esposito, M. (2018). The Representational Basis of Working Memory. *Curr Top Behav Neurosci*, *37*, 213-230. https://doi.org/10.1007/7854_2016_456

Öberg, C., Larsson, M., & Bäckman, L. (2002). Differential sex effects in olfactory functioning: the role of verbal processing. *Journal of the International Neuropsychological Society*, *8*(5), 691-698.

Olofsson, J. K., & Gottfried, J. A. (2015). The muted sense: neurocognitive limitations of olfactory language. *Trends in cognitive sciences*, *19*(6), 314-321.

Ostrosky-Solís, F., & Lozano, A. (2006). Digit Span: Effect of education and culture. *International Journal of Psychology*, *41*(5), 333-341. https://doi.org/10.1080/00207590500345724

Pellegrino, R., Sinding, C., De Wijk, R., & Hummel, T. (2017). Habituation and adaptation to odors in humans. *Physiology & Behavior*, *177*, 13-19.

Prull, M. W., Dawes, L. L. C., Martin Iii, A. M., Rosenberg, H. F., & Light, L. L. (2006). Recollection and familiarity in recognition memory: Adult age differences and neuropsychological test correlates. *Psychology and Aging*, *21*(1), 107-118. https://doi.org/10.1037/0882-7974.21.1.107

Reisberg, D., Rappaport, I., & O'Shaughnessy, M. (1984). Limits of working memory: The digit digit-span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(2), 203-221. https://doi.org/10.1037/0278-7393.10.2.203

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: a meta-analysis. *Psychon Bull Rev*, *26*(5), 1529-1547. https://doi.org/10.3758/s13423-019-01649-y

Roberts, R. O., Christianson, T. J., Kremers, W. K., Mielke, M. M., Machulda, M. M., Vassilaki, M., . . . Petersen, R. C. (2016). Association Between Olfactory Dysfunction and Amnestic Mild Cognitive Impairment and Alzheimer Disease Dementia. *JAMA Neurol*, *73*(1), 93-101. https://doi.org/10.1001/jamaneurol.2015.2952

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable Digit Span: a systematic review and cross-validation study. *Assessment*, *19*(1), 21-30. https://doi.org/10.1177/1073191111428764

Serby, M., Larson, P., & Kalkstein, D. S. (1991). The nature and course of olfactory deficits in Alzheimer's disease. *The American Journal of Psychiatry*, *148*(3), 357-360. https://doi.org/10.1176/ajp.148.3.357

Seubert, J., Kalpouzos, G., Larsson, M., Hummel, T., Bäckman, L., & Laukka, E. J. (2020). Temporolimbic cortical volume is associated with semantic odor memory performance in aging. *Neuroimage*, *211*, 116600.

Sorokowska, A., Sabiniewicz, A., & Larsson, M. (2020). TOM-32–An extended test for the assessment of olfactory memory. *Journal of Neuroscience Methods*, *344*, 108873.

Sorokowska, A., Schriever, V. A., Gudziol, V., Hummel, C., Hähner, A., Iannilli, E., . . . Hummel, T. (2015). Changes of olfactory abilities in relation to age: odor identification in more than 1400 people aged 4 to 80 years. *Eur Arch Otorhinolaryngol*, *272*(8), 1937-1944. https://doi.org/10.1007/s00405-014-3263-4

Sorokowski, P., Karwowski, M., Misiak, M., Marczak, M. K., Dziekan, M., Hummel, T., & Sorokowska, A. (2019). Sex Differences in Human Olfaction: A Meta-Analysis. *Front Psychol*, *10*, 242. https://doi.org/10.3389/fpsyg.2019.00242

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92-114. https://doi.org/10.1037/0096-1523.27.1.92

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., . . . Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of*

Clinical and Experimental Neuropsychology, *33*(1), 101-111. https://doi.org/10.1080/13803395.2010.493149

Yaple, Z. A., Stevens, W. D., & Arsalidou, M. (2019). Meta-analyses of the n-back working memory task: fMRI evidence of age-related changes in prefrontal cortex involvement across the adult lifespan. *Neuroimage*, *196*, 16-31.

Zhang, C., & Wang, X. (2017). Initiation of the age-related decline of odor identification in humans: A meta-analysis. *Ageing Res Rev*, *40*, 45-50. https://doi.org/10.1016/j.arr.2017.08.004

Zucco, G. M. (2003). Anomalies in cognition: olfactory memory. *European Psychologist*, *8*(2), 77.

Table 1. Valence of high-and low-verbalizality odors

	High-verbalizality odors (n = 77)					
	Odor1	Odor2	Odor3	Odor4	Odor5	Odor6
Intensity	4.09 ± 1.89	6.48 ± 1.49	6.64 ± 1.64	5.58 ± 1.70	6.13 ± 1.96	6.35 ± 1.50
Pleasantness	5.74 ± 1.65	5.91 ± 1.84	5.92 ± 1.94	6.75 ± 1.52	5.29 ± 2.02	6.70 ± 1.62
Verbailizibility	1.35 ± 0.84	1.77 ± 0.63	1.78 ± 0.60	1.74 ± 0.57	1.34 ± 0.84	1.52 ± 0.70
	Low-verbalizability odors (n = 77)					
	Odor7	Odor8	Odor9	Odor10	Odor11	Odor12
Intensity	4.30 ± 1.91	5.50 ± 1.91	4.64 ± 1.94	6.35 ± 1.68	6.01 ± 1.95	5.87 ± 1.71
Pleasantness	5.08 ± 1.79	4.53 ± 1.73	4.92 ± 1.88	5.16 ± 2.09	3.52 ± 1.86	5.70 ± 1.87
Verbailizibility	1.04 ± 0.87	1.08 ± 0.79	0.83 ± 0.85	1.05 ± 0.84	1.07 ± 0.91	1.05 ± 0.81

Table 2. Descriptive statistics of test scores

	High-verbalizability odors (n=77)	Low-verbalizability odors (n=77)	*T*	*P*
STM10s	5.09 ± 1.11	4.35 ± 1.24	-3.89	<0.01
LTM60s	4.92 ±1.28	3.61 ±1.34	-6.20	<0.01
RM	10.01 ± 2.17	7.96 ± 2.11	-5.95	<0.01
Correct items of WM	12.92 ± 2.50	10.49 ± 2.45	-6.09	<0.01
P(H)	0.77 ± 0.20	0.63 ± 0.25	-3.63	<0.01
P(FA)	0.31 ± 0.20	0.44 ±0.20	4.26	<0.01

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s), WM = working memory, H = hit, FA = false alarms, P(H) = hit rate, P(FA) = false alarm rate.

Table 3. Test-retest reliability by Bland-Altman analysis (MD and CI) and Pearson Correlations

		MD	CI (M ± 1.96SD)	*r*	*p*
High-verbalizability odors	STM10s	-0.12	-2.07~1.82	0.65	<0.01
	LTM60s	0.22	-2.12~2.57	0.65	<0.01
	RM	0.10	-2.93~3.13	0.79	<0.01
	2-back	-0.20	-4.18~3.78	0.76	<0.01
Low-verbalizability odors	STM10s	0.18	-2.75~3.12	0.36	<0.01
	LTM60s	-0.35	-3.05~2.35	0.47	<0.01
	RM	-0.04	-4.02~3.93	0.56	<0.01
	2-back	-0.71	-6.24~4.82	0.41	<0.01

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s), MD = mean difference of Bland-Altman analysis, CI = confidence interval of Bland-Altman analysis.

Table 4. Validity test with TOM-16 and digit span test

	High-verbalizability odors			
	STM10s	LTM60s	RM	Correct items of WM
Correct items of TOM-16	0.35**	0.42**	0.44**	0.26*
Digit span_T	0.48**	0.33**	0.43**	0.50**
Digit span_P	0.41**	0.32**	0.38**	0.43**
Digit span_N	0.51**	0.34**	0.45**	0.52**
	Low-verbalizability odors			
	STM10s	LTM60s	RM	Correct items of WM
Correct items of TOM-16	0.42**	0.21	0.38	0.27*
Digit span_T	0.36**	0.33**	0.40**	0.45**
Digit span_P	0.25*	0.29*	0.30**	0.38**
Digit span_N	0.43**	0.32**	0.44**	0.45**

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s), WM = working memory, Digit span_T = total score of digit span test (the sum of positive and negative order), Digit span_P = the score of digit span test in positive order, Digit span_N = the score of digit span test in negative order. *$p < 0.05$, **$p < 0.01$.

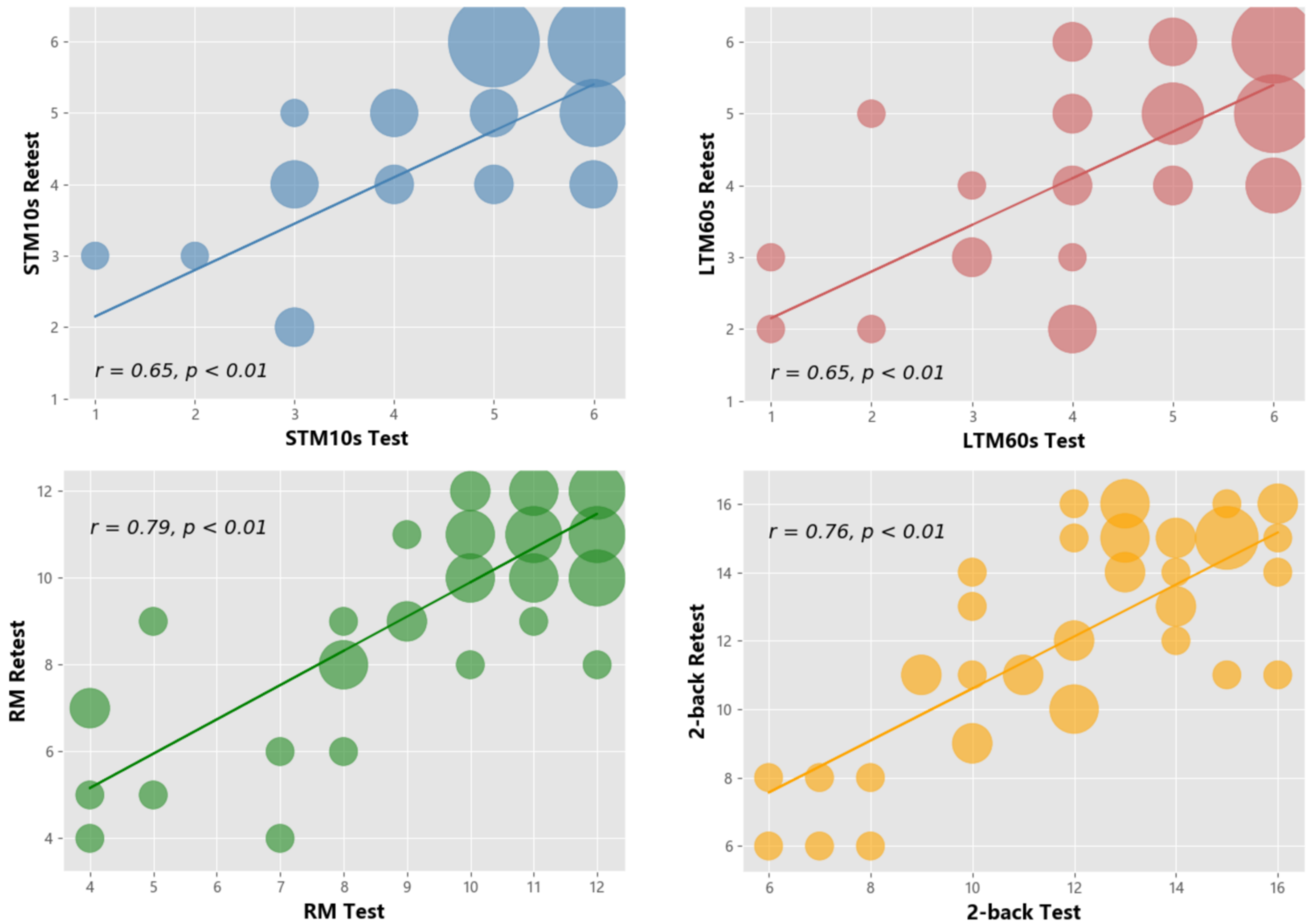

Figure 1. Test-retest reliability of all scores in high-verbalizability odors

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s).

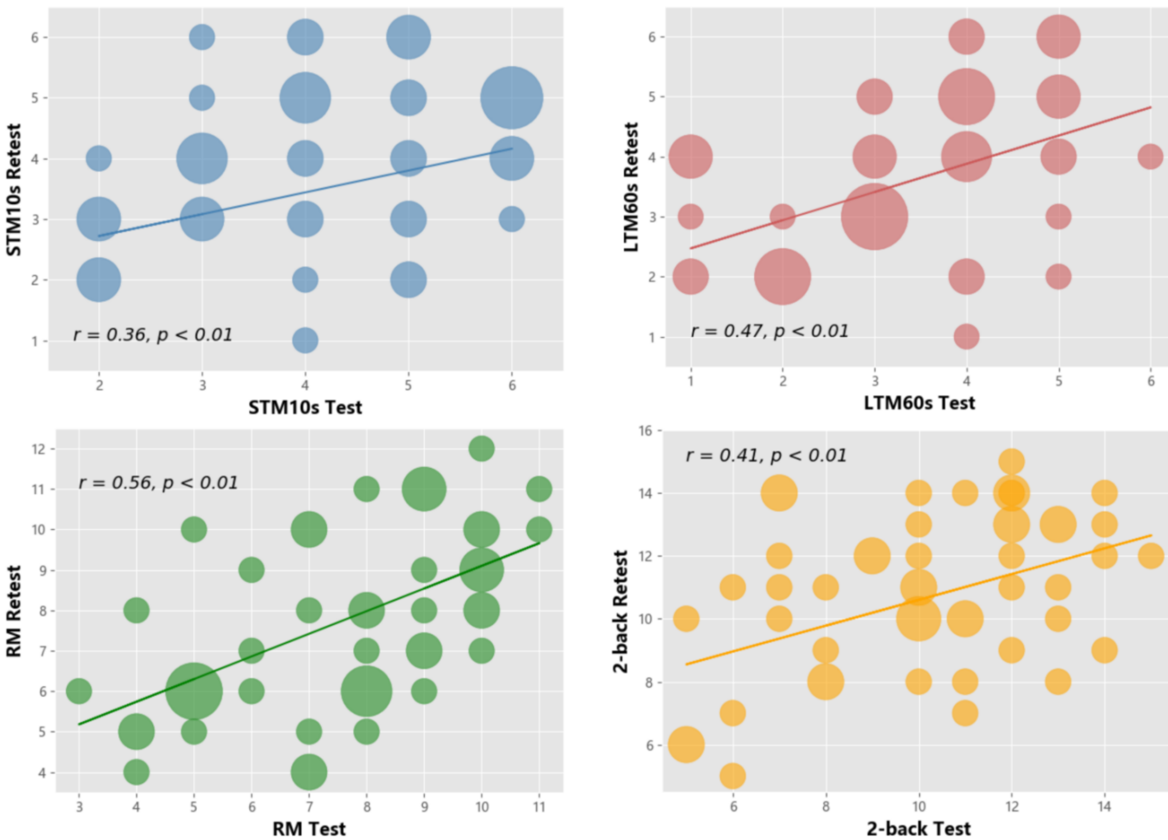

Figure 2. Test-retest reliability of all scores in low-verbalizability odors

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s).

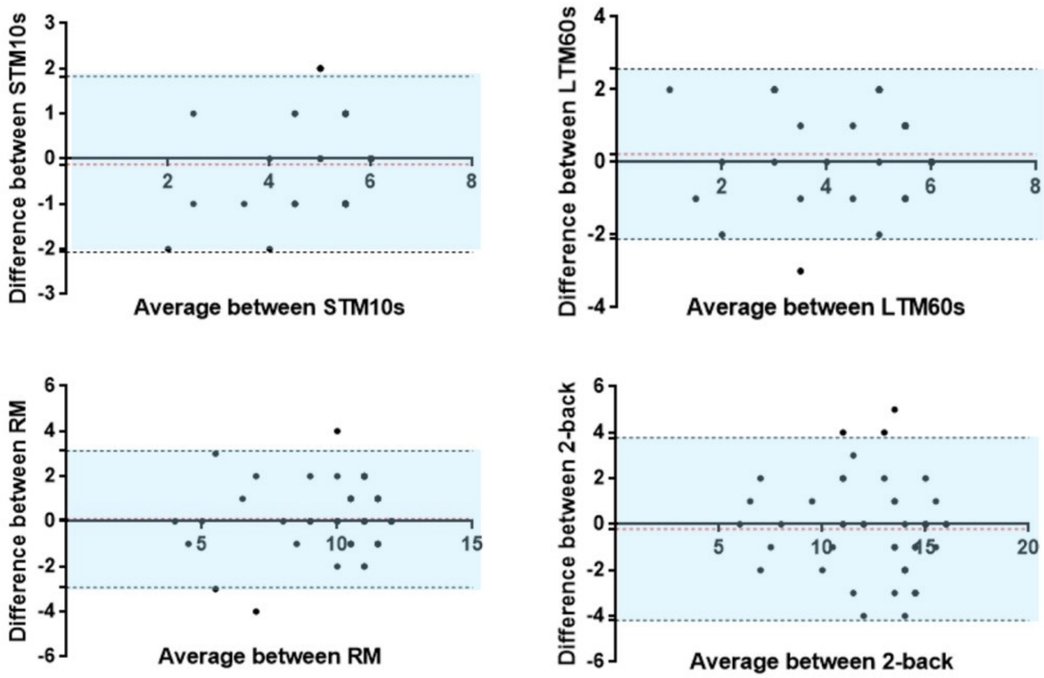

Figure 3. The Bland-Altman plot between test and retest in high-verbalizability odors

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s).

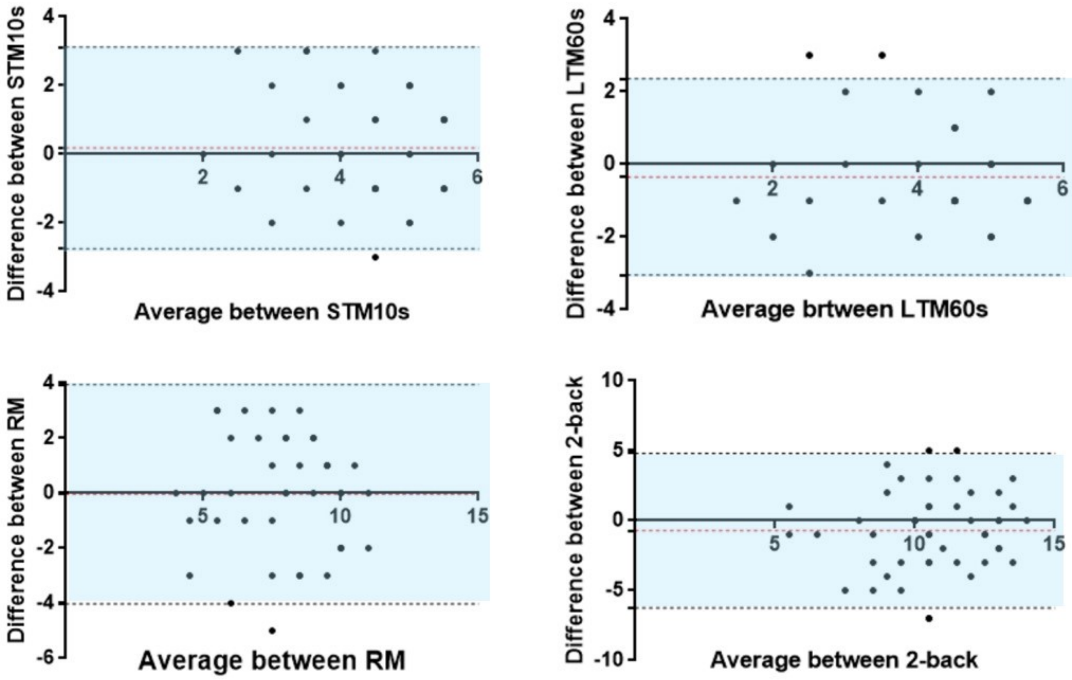

Figure 4. The Bland-Altman plot between test and retest in low-verbalizability odors

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s).

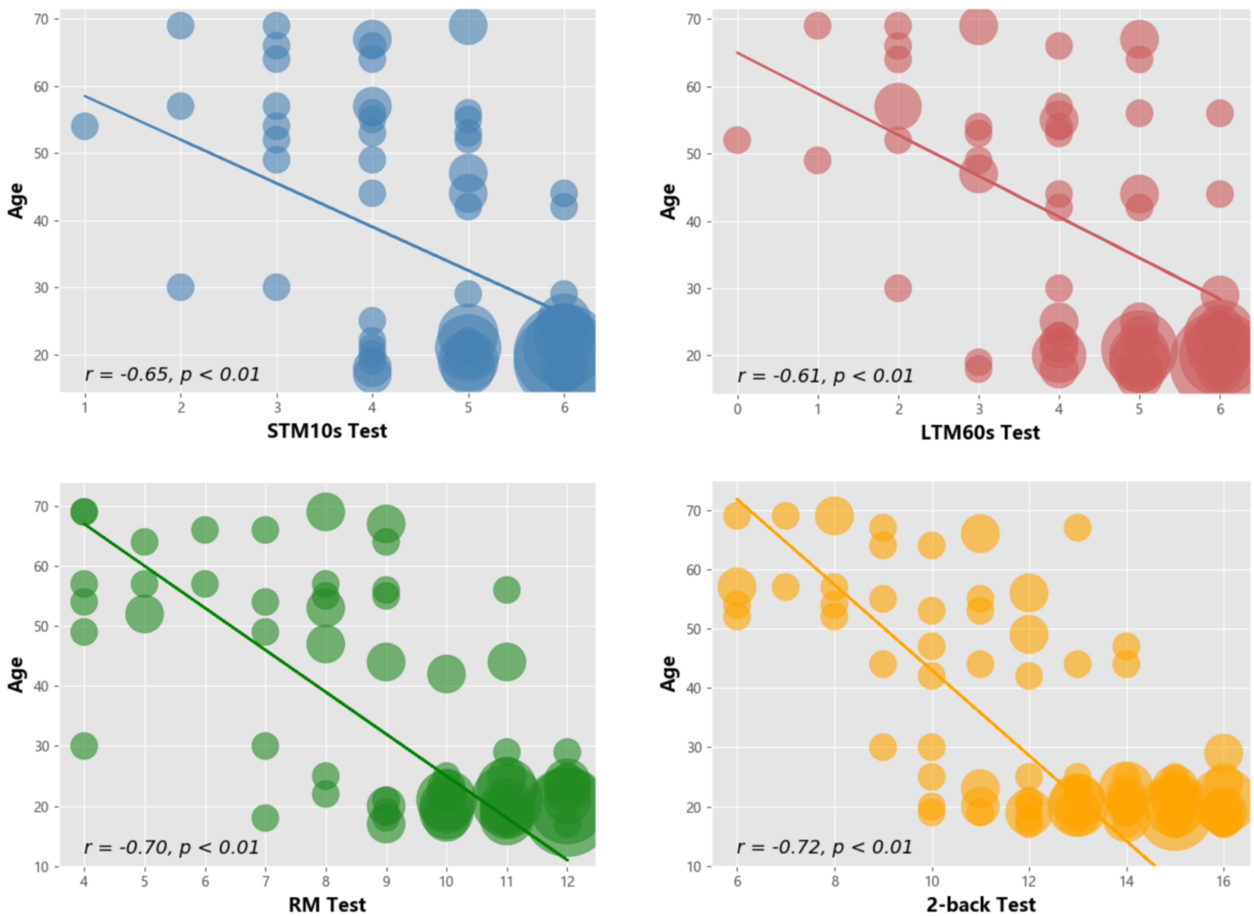

Figure 5. Correlations between age and all OMTB scores in high-verbalizability odors

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s).

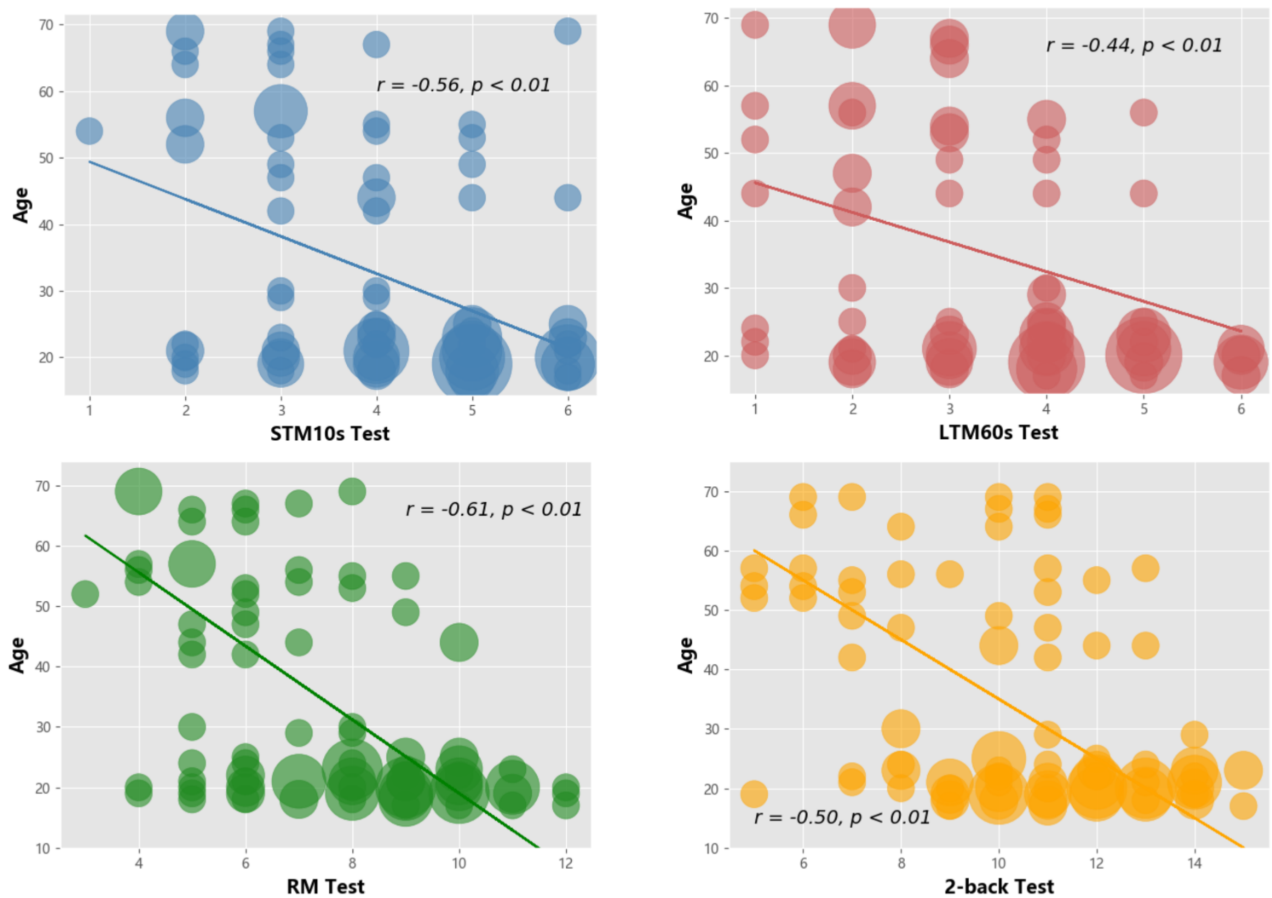

Figure 6. Correlations between age and all OMTB scores in low-verbalizability odors

Notes: STM10s = the sum of short-term memory scores, LTM60s = the sum of long-term memory scores, RM = recognition memory (the sum of STM10s and LTM60s).

Supplement tables

Table S1. Dilution ratio of all odorants

CAS	Dilution ratio	Rotation speed	Rotation time (min)	Olfactory classification
106-24-1	1:25	300	3	Floral
105-53-3	1:52	300	3	Fruity
111-70-6	1:60	300	3	Grass
108-29-2	1:18	300	3	Floral
7452-79-1	1:200	300	3	Fruity
1128-08-1	1:72	300	3	Floral
10599-70-9	1:360	300	3	Grass
1125-21-9	1:6	300	3	Grass
2305-25-1	1:28	300	3	Fruity
488-10-8	1:40	300	3	Floral
106-30-9	1:200	300	3	Fruity
60-12-8	3:4	300	3	Floral
1193-79-9	1:76	300	3	Grass
106-27-4	1:200	300	3	Fruity
101-84-8	1:80	300	3	Grass
110-38-3	1:60	300	3	Fruity
693-54-9	1:310	300	3	Grass
39255-32-8	1:250	300	3	Fruity
122-70-3	1:50	300	3	Floral
606-45-1	1	300	3	Fruity
134-20-3	1:50	300	3	Floral
821-55-6	1:200	300	3	Grass
627-90-7	1:70	300	3	Fruity
106-21-8	1:90	300	3	Floral
122-03-2	1:270	300	3	Grass

97-53-0	1:140	300	3	Floral
85-91-6	1:40	300	3	Floral
928-96-1	1:100	300	3	Grass
140-67-0	1:140	300	3	Grass
142-19-8	1:140	300	3	Fruity
112-66-3	1:110	300	3	Floral
98-86-2	1:300	300	3	Grass
80-54-6	1:1	300	3	Floral
104-76-7	1:110	300	3	Floral
143-08-8	1:100	300	3	Floral
119-84-6	1:120	300	3	Grass

Table S2. Recipe of mixture odorants

Number	Fruity odorants			Grass odorants			Floral odorants		
1	606-45-1	2305-25-1	627-90-7	119-84-6	10599-70-9	101-84-8	80-54-6	112-66-3	106-24-1
2	110-38-3	142-19-8	105-53-3	122-03-2	98-86-2	1125-21-9	134-20-3	85-91-6	122-70-3
3	106-27-4	106-70-3	39255-32-8	140-67-0	693-54-9	1193-79-9	106-21-8	104-76-7	488-10-8
4	606-45-1	110-38-3	106-27-4	111-70-6	821-55-6	928-96-1	1128-08-1	97-53-0	108-29-2
5	2305-25-1	142-19-8	106-70-3	119-84-6	122-03-2	140-67-0	80-54-6	134-20-3	106-21-8
6	627-90-7	105-53-3	39255-32-8	10599-70-9	98-86-2	693-54-9	112-66-3	85-91-6	104-76-7
7	110-38-3	106-27-4	7452-79-1	101-84-8	1125-21-9	1193-79-9	106-24-1	122-70-3	488-10-8
8	142-19-8	106-70-3	106-30-9	122-03-2	140-67-0	111-70-6	134-20-3	106-21-8	1128-08-1
9	606-45-1	105-53-3	39255-32-8	98-86-2	693-54-9	821-55-6	85-91-6	104-76-7	97-53-0
10	2305-25-1	106-27-4	7452-79-1	1125-21-9	1193-79-9	928-96-1	122-70-3	122-70-3	108-29-2
11	627-90-7	106-70-3	106-30-9	119-84-6	140-67-0	111-70-6	106-21-8	1128-08-1	143-08-8
12	606-45-1	142-19-8	39255-32-8	10599-70-9	693-54-9	821-55-6	104-76-7	97-53-0	60-12-8
13	2305-25-1	105-53-3	7452-79-1	101-84-8	101-84-8	928-96-1	80-54-6	488-10-8	108-29-2
14	627-90-7	106-27-4	106-30-9	119-84-6	122-03-2	111-70-6	112-66-3	1128-08-1	143-08-8
15	606-45-1	105-53-3	106-30-9	10599-70-9	98-86-2	821-55-6	106-24-1	97-53-0	60-12-8
16	606-45-1	105-53-3	106-27-4	101-84-8	1125-21-9	928-96-1	80-54-6	134-20-3	108-29-2

17	2305-25-1	106-27-4	106-70-3	119-84-6	98-86-2	1193-79-9	112-66-3	85-91-6	143-08-8
18	627-90-7	106-70-3	39255-32-8	10599-70-9	1125-21-9	111-70-6	106-24-1	122-70-3	1128-08-1
19	110-38-3	39255-32-8	7452-79-1	101-84-8	140-67-0	821-55-6	80-54-6	85-91-6	488-10-8
20	142-19-8	7452-79-1	106-30-9	122-03-2	693-54-9	928-96-1	112-66-3	122-70-3	1128-08-1

Table S3. Presentation order of odor recognition memory (high-verbalizibility odors)

10s				60s			
Trial	TO	EO	Answer	Trial	TO	EO	Answer
1	Odor 1	Odor 4 Odor 1 Odor 3	2	7	Odor 1	Odor 4 Odor 1 Odor 5	2
2	Odor 2	Odor 5 Odor 6 Odor 2	3	8	Odor 2	Odor 2 Odor 5 Odor 6	1
3	Odor 3	Odor 3 Odor 1 Odor 2	1	9	Odor 3	Odor 4 Odor 2 Odor 3	3
4	Odor 4	Odor 6 Odor 5 Odor 4	3	10	Odor 4	Odor 2 Odor 4 Odor 1	2
5	Odor 5	Odor 3 Odor 5 Odor 4	2	11	Odor 5	Odor 3 Odor 5 Odor 6	2
6	Odor 6	Odor 1 Odor 6 Odor 2	2	12	Odor 6	Odor 1 Odor 3 Odor 6	3

Note: TO = target odor, EO = experiment odor.

Table S4. Presentation order of odor recognition memory (low-verbalizibility odors)

10s				60s			
Trial	TO	EO	Answer	Trial	TO	EO	Answer
1	Odor 7	Odor 10 Odor 7 Odor 9	2	7	Odor 7	Odor 10 Odor 7 Odor 11	2
2	Odor 8	Odor 11 Odor 12 Odor 8	3	8	Odor 8	Odor 8 Odor 11 Odor 12	1
3	Odor 9	Odor 9 Odor 7 Odor 8	1	9	Odor 9	Odor 10 Odor 8 Odor 9	3
4	Odor 10	Odor 12 Odor 11 Odor 10	3	10	Odor 10	Odor 8 Odor 10 Odor 7	2
5	Odor 11	Odor 10 Odor 11 Odor 9	2	11	Odor 11	Odor 9 Odor 11 Odor 12	2
6	Odor 12	Odor 7 Odor 12 Odor 8	2	12	Odor 12	Odor 7 Odor 9 Odor 12	3

Note: TO = target odor, EO = experiment odor.

Table S5. Presentation order of odor working memory

High-verbalizibility odors		Low-verbalizibility odors	
Trial	Odor number	Trial	Odor number
1	1	1	7
2	2	2	8
3	1	3	7
4	3	4	9
5	4	5	10
6	3	6	9
7	5	7	11
8	6	8	12
9	5	9	11
10	4	10	10
11	1	11	7
12	4	12	10
13	2	13	8
14	3	14	9
15	2	15	8
16	6	16	12
17	5	17	11
18	6	18	12