

Olfactory assessment using the NIH Toolbox

Pamela Dalton, PhD,
MPH
Richard L. Doty, PhD
Claire Murphy, PhD
Robert Frank, PhD
Howard J. Hoffman, MA
Christopher Maute, EdM
Michael A. Kallen, PhD,
MPH
Jerry Slotkin, PhD

Correspondence to
Dr. Dalton:
pdalton@pobox.upenn.edu

ABSTRACT

The human olfactory system provides us with information about our environment that is critical to our physical and psychological well-being. Individuals can vary widely in their ability to detect, recognize, and identify odors, but still be within the range of normal function. Although several standardized tests of odor identification are available, few specifically address the issues in testing very young children, most of whom are likely to be unfamiliar with many of the odor stimuli used in adult tests and have limited ability to read and identify labels to select among choices. Based on the format of the San Diego Odor Identification Test and the delivery system of the University of Pennsylvania Smell Identification Test, we developed 2 versions of an odor identification test using standardized odor stimuli in a scratch-and-sniff format in which participants match 5 (children) or 9 (adults) odors to pictures representing the odor source. Results from normative testing and validation showed that for most participants, the test could be completed in 5 minutes or less and that the poorer performance among the youngest children and the elderly was consistent with data from tests with larger numbers of items. Expanding on the pediatric version of the test with adult-specific and public health-relevant odors increased the ecological validity of the test and facilitated comparisons of intraindividual performance across developmental stages. *Neurology*® 2013;80 (Suppl 3):S32-S36

GLOSSARY

B-SIT = Brief Smell Identification Test; **SDOIT** = San Diego Odor Identification Test; **UPSIT** = University of Pennsylvania Smell Identification Test.

The human olfactory system allows us to detect odors, to recognize and discriminate odor qualities, and to identify the sources of odors in our world. Humans are capable of detecting and discriminating thousands of different odorant molecules, many at extremely low concentrations (i.e., parts per billion or trillion). Our sense of smell provides us with information about our air, water, and food that is critical to our health and safety, nutrition, and psychological well-being.

The process of olfaction is initiated when volatile chemicals stimulate olfactory receptor neurons located on a relatively small patch of specialized epithelial tissue high in the nasal cavity.¹ These sensory neurons have axons that travel as the olfactory nerve (cranial nerve I) to terminate in the olfactory bulb. In turn, the olfactory bulb projects more centrally and contributes inputs for higher cortical processing, which results in olfactory perception.² Odorants reach the olfactory receptors in 2 ways: they can enter the nostrils during normal inhalation (orthonasal route) or travel from the back of the oral cavity to the olfactory receptors via the nasal pharynx (retronasal route). The perception of food flavor involves a combination of olfactory activation caused by odorous compounds released into the nasopharynx retronasally through chewing, drinking, and deglutition, and the blending of taste (salty, sour, bitter, sweet, umami) and oral somatosensory sensations (texture, heat, cold).³ Nasal blockage and swelling can prevent odors from entering the retronasal stream, resulting in a shift of flavor toward blandness. Reported taste loss is more typically a loss in the perception of food flavor due to loss of retronasal olfactory function than to a decrement in taste perception, per se.

From the Monell Chemical Senses Center (P.D., C.M.), Philadelphia; Smell and Taste Center and Department of Otorhinolaryngology: Head and Neck Surgery (R.L.D.), Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; San Diego State University (C.M.), San Diego, CA; Ohio University (R.F.), Athens, OH; Epidemiology and Statistics Program (H.J.H.), National Institute on Deafness and Other, Communication Disorders, National Institutes of Health, Bethesda, MD; and Department of Medical Social Sciences (M.A.K., J.S.), Northwestern University, Chicago, IL.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Often patients do not recognize the presence of a less-than-total loss of smell until they are formally tested. As with the other major senses, quantitative testing of olfactory function is essential to determine the validity of a patient's chemosensory complaint, characterize the specific nature of the problem, accurately monitor changes in function over time (including influences of pharmacologic, surgical, or immunologic interventions), detect malingering, and establish compensation for disability. Multiple factors can compromise smell function, including allergies, bacterial and viral infections, head injuries, sinonasal disease, and environmental exposures to toxins and pollutants.⁴ Importantly, smell dysfunction is one of the hallmark "preclinical" signs of neurodegenerative diseases such as Alzheimer disease and Parkinson disease.⁴⁻⁷

There are a variety of ways to assess olfactory capabilities and olfactory loss in humans, which were reviewed by Doty,⁸ but most are not consistent with the requirements of the NIH Toolbox initiative to be an "off the shelf," brief and inexpensive test, suitable for use in ages 3 to 85 years.⁹ Odor detection is evaluated by measuring the lowest concentration of an odorant at which an individual a) can just detect the odor's presence, or b) can discriminate it from a sample of odorless air. The concentration at which this occurs is considered the detection threshold for that compound. This value can be compared with normative data obtained on other individuals to determine whether the measured sensitivity falls within a normal range. Although threshold testing is a sensitive way to measure differences in olfactory function, threshold tests take a great deal of time to administer. Odor identification, however, is evaluated by presenting individuals with a variety of recognizable odorants at supra-threshold concentrations (i.e., above the detection level for most subjects) and requiring them to choose the correct identity of the odor from a set of possible names or pictures. Performance can be compared with norms applicable to healthy young adults, or alternatively, to age- and sex-adjusted norms. Depending on how the odor is presented, identification tests can be brief to administer and may not require specialized equipment or training of the examiners.

Our literature review found a number of different tests that have been used to evaluate olfaction in both adults and children. However, there were very few brief measures that were validated for the evaluation of olfactory ability across the lifespan and that could be used by nonspecialists in a variety of testing venues. After an extensive literature review, 2 tests were identified as suitable for further investigation or development: the 12-item cross-cultural version of the University of Pennsylvania Smell Identification Test (termed the Brief Smell Identification Test [B-SIT]),¹⁰ a test using microencapsulated "scratch-and-sniff" odors, and the San Diego Odor Identification Test (SDOIT),¹¹ a test comprising 8 everyday food or household items that are presented for sniffing in opaque plastic jars. Both of these tests, which are well correlated with one another,¹² have been used to test children and adults, although the SDOIT was specifically developed for children 5 to 12 years of age. Only 2 items (chocolate and cinnamon) from the 12 items on the B-SIT overlapped with those in the SDOIT, thus providing valuable data on the best odors to use when developing a brief olfactory assessment to be used for individuals aged 3 to 85 years.

Two studies were initially conducted in the development of the test and are described in detail in an earlier publication.¹³ In the first study, 369 children (aged 3–17 years) and 277 adults (parents) were tested. In the second study, 50 children (aged 3–4 years) and 43 adults were given a revised version of the test with 8 odors that had been judged to be more representative of the odor source and more familiar to children. As expected, identification accuracy increased with age. Children as young as 3 years were able to understand the instructions and perform the test, but took longer than did older children and adults. From the results of these studies, we developed 2 versions of the test, using the items depicted in table 1.

Using these 2 versions of the test, a large-scale study was then conducted at multiple sites in the United States to provide normative data for the test and to validate it against previously established tests of olfactory function: the University of Pennsylvania Smell Identification Test (UPSIT)¹⁴ and the 12-item B-SIT.¹⁰ Items 1 to 6 were administered to children between

Table 1 Odorants used in the NIH Toolbox Odor Identification Test for children (nos. 1-6 only) and adults

Cinnamon ^a	1
Play-Doh	2
Chocolate	3
Lemon	4
Popcorn	5
Bubble gum	6
Flower	7
Coffee	8
Natural gas	9
Smoke	10

^a Because of concerns about odorant stability, cinnamon odor was subsequently removed from both versions of the test.

the ages of 3 and 9 years and items 1 to 10 were administered to individuals between the ages of 10 and 85 years. The first 6 items were selected based on their familiarity to young children. The final 4 items included odorants that were not only more familiar to adults (i.e., coffee), but also included 2 odorants that are signals of potential hazard (wood smoke and natural gas) and for which there is evidence of significant age-related declines.¹⁵

As in the previous study, the odorants were microencapsulated and placed onto small individual cards. After scratching the card and sniffing from each sample, respondents were asked to choose the best match for each odor from 4 pictures, representing the correct and 3 distractor odor sources. Before the actual test, to ensure familiarity with the pictured odor sources, children younger than 10 years were asked to point to the correct picture (targets and distractors) when the experimenter named them. If they made an error, they were shown the correct answer and retested on only those items. This is an important step because odor identification

tasks depend not only on the ability to smell, but familiarity with the odor stimulus, which will vary with age and experience.

PARTICIPANTS For the normative study, 1,446 children (aged 3–9 years; 726 females) were administered the 6-item pediatric version of the NIH Toolbox Odor Identification Test, and 2,884 individuals (aged 10–85 years; 1,629 females) were administered the 10-item NIH Toolbox Odor Identification Test. The pediatric sample included Hispanic or Latino (45.8%) and non-Hispanic or Latino (52.1%) individuals (2.1% had missing data on this point). The adult sample included Hispanic or Latino (24.4%) and non-Hispanic or Latino (73.8%) individuals (1.8% missing). For evaluating test-retest reliability, 480 individuals (aged 10–85 years) and 106 children (aged 3–9 years) were retested 1 week later. For the validation study, 445 of those individuals (ages 10–69 only) were also administered the 40-item UPSIT. Because the 12-item B-SIT was composed of items in the larger UPSIT, the B-SIT data were extracted for analysis as well.

RESULTS As shown in table 2, we found clear evidence of differences in identification accuracy, both as a function of the odor item as well as a function of age. The pattern of lower performance among the youngest children and the elderly is consistent with data from tests containing a larger number of items.¹⁴ Test-retest reliability was significant at $p < 0.001$ for all groups, although relatively low because of the small number of test items (Pearson 0.575 and 0.454 for ages 10–85 and 3–9 years, respectively). One item in particular, cinnamon, seemed to be poorly identified by all age groups and, because of concerns about odor quality, was subsequently excluded from both the pediatric and adult versions of the test.

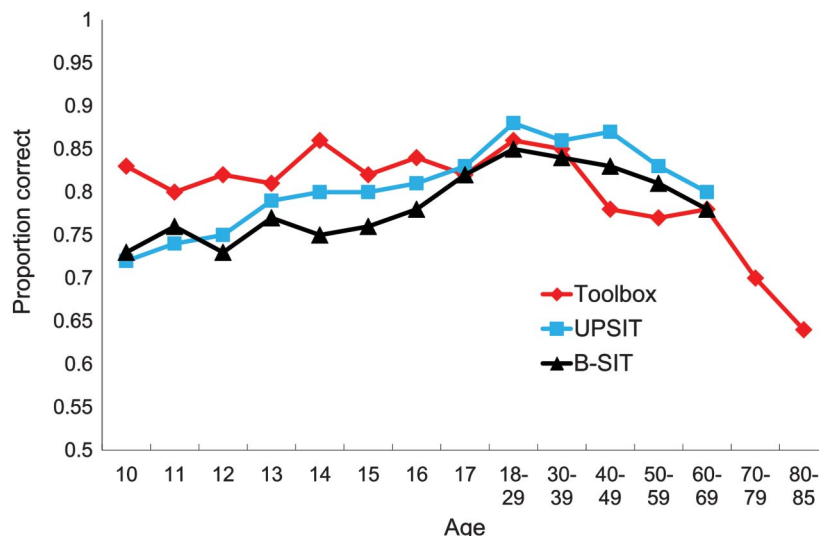
The figure depicts the mean overall test scores in the validation sample, comparing the NIH Toolbox Odor Identification Test with the 40-item UPSIT and the 12-item B-SIT. Across the age bands, performance on the Toolbox test tracks the performance of the other 2 tests fairly closely. Importantly, the Toolbox test does seem to identify the age-related losses in olfactory acuity

Table 2 Percent of individuals in each age group correctly identifying the odorant^a

Age group, y (no. tested)	Cinnamon	Lemon	Play-Doh	Bubble gum	Chocolate	Popcorn	Coffee	Smoke	Natural gas	Flower
3-4 (420)	35.2	55.5	32.8	62.8	53.9	49.0	—	—	—	—
5-9 (1,026)	59.9	85.7	59.7	81.2	74.1	69.3	—	—	—	—
10-17 (1,638)	48.7	94.8	73.4	93.6	68.8	79.7	94.9	69.7	97.0	82.6
18-64 (1,000)	34.2	96.3	64.1	94.6	74.5	84.2	95.9	75.5	97.9	92.5
65-85 (246)	28.5	86.6	25.6	77.6	66.7	82.9	91.9	65.0	88.6	80.1

^a Note that 6 of the odorants were administered to all subjects; the 10-odor test was only administered to subjects aged 10 years and older.

Figure Validation of the Toolbox test vs UPSIT and B-SIT



Mean proportion correct score for the Toolbox test (10 items) compared with the UPSIT (40 items) and the B-SIT (12 items). B-SIT = Brief Smell Identification Test; UPSIT = University of Pennsylvania Smell Identification Test.

that begin in the fifth decade of life, consistent with the UPSIT and B-SIT. Although the 4-alternative matching test of an odor to a picture can be understood and completed by children as young as 3 years, the lower performance among this group clearly reflects the unfamiliarity of some of the stimulus items to some children as well as the potential developmental trajectory of olfactory function. It is worth noting, however, that overall performance for children 10 to 17 years of age is better on this test than on the UPSIT or B-SIT (figure) ($t_{(277)} = 4.2$, $p = 0.00$ and $t_{(277)} = 3.04$, $p = 0.003$ for the UPSIT and B-SIT, respectively). This suggests that the pictorial response format of the Toolbox test and the selection of odor items that are more familiar to children separates the assessment of olfactory ability from the potential confounds of cognitive or developmental influences.

CONCLUSION Both versions of the NIH Toolbox Odor Identification Test seem to meet the requirements and goals of the NIH Toolbox initiative to develop brief, inexpensive, easy-to-administer assessments that can detect change over time and in a wide age range of individuals. It can be easily and quickly administered to even very young children and the elderly and, based on feedback from this study, is uniformly enjoyed by the test subjects.

Expanding on the pediatric version of the test by adding more adult-specific and public health-relevant odors not only increases the ecological validity of the test, but also allows comparisons of intraindividual performance across developmental stages. Importantly, there is evidence that the test in its present form can distinguish impairments in olfactory function due to age-related

losses and may therefore be able to distinguish olfactory dysfunction in other age groups as well.

AUTHOR CONTRIBUTIONS

Dr. Dalton: study concept and design, data interpretation, and primary authorship of the manuscript. Dr. Doty: data interpretation and critical revision of the manuscript for important intellectual content. Dr. Murphy, Dr. Frank, Mr. Hoffman: data interpretation and critical revision of the manuscript for important intellectual content. Mr. Maute: participation in the study concept and design and revision of the manuscript for important intellectual content. Dr. Kallen: data analysis and interpretation. Dr. Slotkin: participation in the study concept and design, study supervision, and revision of the manuscript for important intellectual content.

ACKNOWLEDGMENT

The authors acknowledge the contributions of Julie A. Mennella, PhD, to this effort.

STUDY FUNDING

This study is funded in whole or in part with Federal funds from the Blueprint for Neuroscience Research, National Institutes of Health, under contract no. HHS-N-260-2006-00007-C and with support from the US Army Research Office under grant no. W911NF-11-1-0087.

DISCLOSURE

P. Dalton received grant funding from the Department of Defense (ARO) and the NIH (R01 DC 03275 and a subproject in P50 DC 006760) and received research support from Altria Inc., International Flavors & Fragrances, Reckitt Benckiser, Cadbury, and SC Johnson Company. She has been a consultant to Johnson & Johnson Co., Reckitt Benckiser, and an expert witness for legal proceedings to Premium Standard Farms (McGuire Woods, attorneys), and the City of Philadelphia. She has received honoraria from Columbia University, Mount Sinai Medical Center, and the International Fragrance Association. R. Doty receives grant funding from the NIH (R01 MH 59852; R01 MH 63381) and the Department of Defense (USAMRAA W81XWH-09-1-0467). He is a consultant to NIH grants R01 AG041795 and U54 HD028138, and a mentor on NIH K01 MH090548-01. He is President and major shareholder of Sensonics, Inc., a manufacturer and distributor of tests of taste and smell. Over the last 2 years, he has been a consultant to Intelligent Beauty, Pfizer Inc., PBS Television, NIH, Western Medical Assessments, and a witness or consultant for legal proceedings. During this time, he received publishing royalties

from Cambridge University Press, Informa, and Johns Hopkins University Press and honoraria from the American Academy of Oral Medicine, Australasian Association for ChemoSensory Science, Columbia University, Harvard University, Hospital ABC (Mexico City), Merck Pharmaceuticals, Monash University, and Mt. Sinai Medical Center. C. Murphy is funded by NIA grant R01 AG04085-24 and NIDCD grant DC02064-14. R. Frank, H. Hoffman, C. Maute, M. Kallen, and J. Slotkin report no disclosures. Go to Neurology.org for full disclosures.

Received August 8, 2012. Accepted in final form November 26, 2012.

REFERENCES

1. DeMaria S, Ngai J. The cell biology of smell. *J Cell Biol* 2010;191:443–452.
2. Murthy V. Olfactory maps in the brain. *Ann Rev Neurosci* 2011;34:233–258.
3. Ruijschop RMAJ, Boelrijk AEM, de Graaf C, Westerterp-Plantenga MS. Retronasal aroma release and satiation: a review. *J Agric Food Chem* 2009;57:9888–9894.
4. Doty RL. The olfactory system and its disorders. *Semin Neurol* 2009;29:74–81.
5. Wilson RS, Arnold SE, Schneider JA, et al. Olfactory impairment in presymptomatic Alzheimer's disease. *Ann NY Acad Sci* 2009;11:730–735.
6. Doty RL. Olfactory dysfunction in Parkinson disease. *Nat Rev Neurol* 2012;8:329–339.
7. Ponsen MM, Stoffers D, Booi J, van Eck-Smit BL, Wolters ECh, Berendse HW. Idiopathic hyposmia as a preclinical sign of Parkinson's disease. *Ann Neurol* 2004;56:173–181.
8. Doty RL. Office procedures for quantitative assessment of olfactory function. *Am J Rhinol* 2007;21:460–473.
9. Gershon RC, Cella D, Fox NA, Havlik RJ, Hendrie HC, Wagster MV. Assessment of neurological and behavioural function: the NIH Toolbox. *Lancet Neurol* 2010;9:138–139.
10. Doty RL, Marcus A, Lee WW. Development of the 12-item cross-cultural smell identification test (CC-SIT). *Laryngoscope* 1996;106:353–356.
11. Murphy C, Anderson JA, Markison S. Psychophysical assessment of chemosensory disorders in clinical populations. In: Kunihara KSN, Ogawa H, editors. *Olfaction and Taste XI*. Tokyo: Springer-Verlag; 1994:609–613.
12. Krantz EM, Schubert CR, Dalton DS, et al. Test-retest reliability of the San Diego Odor Identification Test and comparison with the Brief Smell Identification Test. *Chem Senses* 2009;34:435–440.
13. Dalton P, Mennella JA, Maute C, et al. Development of a test to evaluate olfactory function in a pediatric population. *Laryngoscope* 2011;121:1843–1851.
14. Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. *Physiol Behav* 1984;32:489–502.
15. Wysocki CJ, Gilbert AN. National Geographic Smell Survey: effects of age are heterogeneous. *Ann NY Acad Sci* 1989;561:12–28.