Development of an international odor identification test for children – the "U-Sniff" test

Authors: Valentin A. Schriever,MD[1,2], Eduardo Agosin, PhD[3], Aytug Altundag, MD[4], Hadas Avni, MD[5], Helene Cao Van, MD[6], Carlos Cornejo, PhD[7], Gonzalo de los Santos, MD[8], Gad Fishman, MD[9], Claudio Fragola, MD[8], Marco Guarneros, PhD[10], Neelima Gupta, MD[11], Robyn Hudson, PhD[10], Reda Kamel, MD[12], Antti Knaapila, PhD[13,14], Iordanis Konstantinidis, MD[15], Basile N. Landis, MD[16], Maria Larsson, PhD[17], Johan N. Lundström, PhD[18,19], Alberto Macchi, MD[20], Franklin Mariño-Sánchez, MD[8,21], Lenka Martinec Nováková, PhD[22,23], Eri Mori, MD[24], Joaquim Mullol, MD[21], Marie Nord, MSc[17], Valentina Parma, PhD[18,25], Carl Philpott, MD[26,27], Evan J. Propst, MD[28], Ahmed Rawan, MD[12], Mari Sandell, PhD[29], Agnieszka Sorokowska, PhD[30,31], Piotr Sorokowski, PhD[30], Lisa-Marie Sparing-Paschke[1], Carolin Stetzler, MD[1], Claudia Valder[32], Jan Vodicka, MD[33], Thomas Hummel, MD[1]

Affiliation:
[1]Smell & Taste Clinic, Department of Otorhinolaryngology, Medizinische Fakultät Carl Gustav Carus, Technische Universität, Dresden, Germany; [2]Abteilung Neuropädiatrie, Medizinische Fakultät Carl Gustav Carus, Technische Universität, Dresden, Germany; [3]Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile; [4]Istanbul Smell and Taste Center, Istanbul, Turkey; [5]Pediatric Feeding Disorders clinic, Dana Children's Hospital, Tel Aviv Medical Center, Tel Aviv, Israel; [6]Pediatric Otolaryngology Unit, Department of Otorhinolaryngology, Head and Neck Surgery, University Hospitals of Geneva, Geneva, Switzerland; [7]Escuela de Psicología, Pontificia Universidad Católica de Chile, Santiago, Chile; [8]Unidad de Rinología y Cirugía de Base de Cráneo. Servicio de Otorrinolaringología. Hospital Universitario Ramón y Cajal. Madrid, Spain; [9]Pediatric Otolaryngology, Dana Children's Hospital, Tel-Aviv Medical Center, Sackler School of Medicine, Tel-Aviv, Israel; [10]Departamento de Biología Celular y Fisiología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico; [11]Department of Otorhinolaryngology, University College of Medical Sciences and GTB Hospital, Delhi, India; [12]Department of Otorhinolaryngology, Head and Neck Surgery,Cairo University, Cairo, Egypt; [13]Food Chemistry and Food Development, Department of Biochemistry, University of Turku, Turku, Finland; [14]Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland; [15]Otorhinolaryngology Department of Aristotle University, Thessaloniki, Greece; [16]Rhinology-Olfactology Unit, Department of Otorhinolaryngology Head and Neck Surgery, University Hospital of Geneva, Switzerland; [17]Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Stockholm, Sweden; [18]Monell Chemical Senses Center, Philadelphia, PA, United States; [19]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; [20]ORL Clinica, Università of Insubriae Varese, ASST Settelaghi, Italy; [21]INGENIO, Immunoal.lèrgia Respiratòria Clínica i Experimental (IRCE), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hospital Clínic, CIBERES. Barcelona, Catalonia, Spain; [22]Department of Anthropology, Faculty of Humanities, Charles University, Czech Republic; [23]National Institute of Mental Health, Klecany, Czech Republic; [24]Department of Otorhinolaryngology, the Jikei University School of Medicine, Tokyo, Japan; [25]International School for Advanced Studies (SISSA), Neuroscience Area, Trieste, Italy; [26]Smell & Taste Clinic, James Paget University Hospital, Gorleston, UK; [27]Norwich Medical School, University of East Anglia, Norwich, UK; [28]Department of Otolaryngology – Head & Neck Surgery, Hospital for Sick Children, University of Toronto, Canada; [29]Functional Foods Forum, University of Turku, Turku, Finland; [30]Smell and Taste Research Lab, Institute of Psychology, University of Wroclaw, Poland; [31]Department of Psychotherapy and Psychosomatic Medicine, TU Dresden, Dresden, Germany; [32]Frey and Lau, Henstedt-Ulzburg, Germany; [33]Department of Otorhinolaryngology and Head and Neck Surgery, Pardubice Hospital, Faculty of Health Studies, University of Pardubice, Czech Republic

Corresponding Author: Valentin A. Schriever, Abteilung Neuropädiatrie, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Fetscherstrasse 74, 01307 Dresden, Germany; phone +49-351-458-2230, valentin.schriever@uniklinikum-dresden.de

Keywords: olfaction, children, smell testing, "Sniffin' Sticks", olfactory dysfunction

Funding: Funding was provided to VAS from the Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden (MedDrive) and the Else Kröner-Fresenius-Stiftung; to MG by the Postdoctoral Fellowship Program of Universidad Nacional Autónoma de México (DGAPA-UNAM); to AS by the Polish Ministry of Science and Higher Education (scientific grant Iuventus Plus # IP2014 043773); to AK and MS from the Academy of Finland; to LMN by Ministry of Education, Youth, and Sports (MEYS) of the Czech Republic under the NPU I program, by the PROGRES program Q22 at the Faculty of Humanities, Charles University within the Institutional Support for Long-Term Development of Research Organizations from MEYS, and by the Czech Science Foundation (GA17-14534S); to TH by the Deustche Forschungsgemeinschaft (DFG HU 411/181); to JV by by a grant of the Technology Agency of the Czech Republic TA04011114; to ML and MN by the Swedish Foundation for Humanities and Social Sciences (M14-0375:1)

Conflict of interest: None of the authors have a financial relationship relevant to this article to disclose. None of the authors have a conflict of interest to disclose.

Abstract:

Objective: The measurement of olfactory function in children is challenging and at present, there is no test that is commonly used. Our objective was to assess olfactory function in children and to create and validate an odor identification test to diagnose olfactory dysfunction in children, which we called the "Universal-Sniff (U-Sniff)" test.

Study design: This is a multicenter study involving 19 countries. The "U-Sniff" was developed in three phases including 1760 children age 5-7 years. Phase 1: Identification of potentially recognizable odors; Phase 2: Selection of odorants for the odor identification test; and Phase 3: Evaluation of the test and acquisition of normative data. Test—retest reliability was evaluated in a subgroup of children (n=27) and the test was validated using children with congenital anosmia (n=14).

Results: Twelve odors were familiar to children and therefore included in the "U-Sniff". Children scored a mean±SD of 9.88±1.80 points out of 12. Normative data was obtained and reported for each country. The "U-Sniff" demonstrated a high test—retest reliability (r_{27}=0.83, $p<0.001$) and enabled discrimination between normosmia and children with congenital anosmia with a sensitivity of 100% and specificity of 86%.

Conclusion: The "U-Sniff" is a valid and reliable method of testing olfaction in children and can be used internationally.

Approximately 20% of people have a reduced sense of smell and 5% have functional anosmia [1-3]. The incidence of olfactory dysfunction is assumed to be lower in children and adolescents than in adults [4], but reliable data to support this hypothesis are lacking. This may be due in part to difficulties performing olfactory testing in children. Anosmia in children may be congenital (among others: isolated disorder or Kallmann syndrome [5]) or acquired secondarily, such as from head trauma, adenoid hypertrophy or cystic fibrosis [6-9].

Many tests for evaluating olfactory function have been developed over the past few decades [10-13] because of an increasing appreciation of the importance of olfaction in everyday life. People with olfactory dysfunction experience an increased frequency of hazardous events, such as food poisoning or failure to detect smoke [14], and have an overall decreased quality of life [15]. Olfactory function is most commonly evaluated orthonasally both clinically and for research purposes using the University of Pennsylvania Smell Identification Test (UPSIT) [11] and the "Sniffin' Sticks" battery - especially the odor identification subtest of the "Sniffin' Sticks" [10]. In addition to orthonasal olfactory assessment, measurements for retronasal olfactory testing such as using the "candy smell test" and the "taste powders" are available [16, 17]. The range of stimuli for retronasal olfactory testing is limited due to simultaneous gustatory stimulation in a sweet (sorbitol) candy, and odors such as fish or cut grass cannot be used [16]. Even though both, the "Sniffin' Sticks" and the UPSIT test have been used in children as young as 5 years of age, they are suboptimal for evaluating olfaction in young children. In both odor identification tests an increase in test performance are observed from childhood over adolescence into adulthood [18, 19]. However, the increment of performance is not due to actual increase in olfactory function. Children and adults perform equally well on olfactory threshold testing but children's performance is lower than adults on odor identification tasks [20, 21], which may be attributed to "odor learning" [20, 22-24]. Odorants used in identification tests might not be familiar to children, additionally the complexity of an olfactory test is considerable. For example, odor identification tests are commonly administered using a four alternative forced-choice paradigm - (4AFC), i.e. the presented odor has to be identified with the help of four descriptors [11, 25]. These

descriptors usually are presented in writing, which may not be optimal for children. To overcome these shortcomings, odor identification tests were developed for children [21, 26-31]. However, only two tests have gained use, namely the "Smell Wheel" and the Sydney Children's Hospital Odor Identification Test "SCHOT" [27, 28]. The "Smell Wheel" has been used to evaluate olfactory function in children with a tracheostomy and the "SCHOT" has been used to study children with cystic fibrosis, otitis media, renal disease, and following bone marrow transplantation [32-36]. These tests have not been used commonly likely because they were developed for children from a single country and are not translatable across cultures, [27, 28] and most tests are not commercially available.

Cultural background also is of importance in odor identification. To counter this, the CC-SIT (Cross-Cultural Smell Identification Test) was developed for adults, which is based on the UPSIT [37]. Today several country-specific, modified versions of the UPSIT and the "Sniffin' Sticks" odor identification test are used, e.g. in Brazil, China, South Korea, Turkey and Egypt [38-42]. Due to the child's development in odor learning, it is plausible that especially for children, the cultural background has substantial impact on odor identification tasks.

The aim of this multicenter study was to develop and validate an international odor identification test for children, which we called the Universal Sniff test ("U-Sniff") – to enable the discrimination between normosmia and a reduced sense of smell with high sensitivity and specificity. We hypothesize that the study design enables the development of an odor identification test for children, can be used internationally, but which odor identification scores might differ might countries.

Materials & methods:

This study was performed in accordance with the Declaration of Helsinki on Biomedical Studies Involving Human Subjects. This study was approved by the local Ethics Committee of the Medical Faculty at the TU Dresden (EK 150042014, EK 383092015) and additionally by individual ethics

committees of participating centers. Study details were explained to children and their parents/legal guardians and oral and/or written consent was obtained where required. In addition, children provided assent. The study was divided into three phases: Phase 1 – Identification of potentially recognizable odors items; Phase 2 – Selection of odorants for the odor identification test; Phase 3 – Evaluation of the test and acquisition of normative data.

Laboratories and clinics from the following countries participated: Europe: Czech Republic, Finland, Germany, Greece, Italy, Poland, Spain, Sweden, Switzerland, Turkey, and United Kingdom (UK) (only Phase3); America: Canada, Chile, Mexico, and the United States of America (USA). In addition Egypt (Phases 2 and 3), India, Israel and Japan contributed to this study.

Prior to Phase 1, a pilot study was conducted whereby investigators from each contributing country submitted names of odor items that they felt would be well known to children in their country. A list of 42 odor items was generated. Items (n=36) that were most common to all countries are listed in Figure 1 and were subsequently used in Phase1.

Phase 1 - Identification of potentially recognizable odors items

Participants: A total of 324 children with age ranging from 5 to 7 years from 17 participated. Each country interviewed 20 participants, except Finland n=17 and Canada n=7. The mean age was 5.9±0.3 (SD) years. Slightly more girls (52.4%) than boys (47.6%) were included, but the difference was not statistically significant ($X^2_{(df=1)}$=0.57, p=0.45). There was no difference in sex distribution across countries ($X^2_{(df=13)}$=13.47, p=0.41). However, the sex of children from three countries (India, Israel and Japan) was not recorded.

Material: Photographs of each of the 36 odor items generated in the pilot phase of this study were presented to the children (Figure 1). For each item a photograph representing the item was chosen. The majority of photographs were produced in the Smell and Taste Clinic in Dresden, Germany, and a few, copyright-free photographs were acquired from the internet.

Procedure: Children were tested individually in a quiet room. The task was explained verbally to each child and one photograph at a time was shown to each child. Children were asked the following questions: 1) Do you know what this is? (recorded as yes/no) 2) How does it smell? (responses written by the investigator).

Phase 2 - Selection of odorants for the odor identification test

Participants: A total of 495 children aged 6 to 8 years from 18 countries were included. Thirty children were tested from each country, except Egypt n=28, Turkey n=26, Finland n=25, USA n=25, Greece n=21, Czech Republic n=9. The mean age was 6.3±0.5 years. There was an equal number of girls (n=241) and boys (n=254; $X^2_{(df=1)}$=0.58 p=0.45) and there was no difference in sex distribution across countries ($X^2_{(df=17)}$=14.98, p=0.60).

Material: Based on results from Phase 1, 17 odor items were used to create an odor identification test (Figure 1). Appropriate odorants were selected by a panel of experienced investigators to represent the visual items. Pen-like "Sniffin' Sticks" were used for odorant presentation. Pens were filled with 4 mL of each odorant and numbered 1 to 17. For details about the odorants see Table 1. Odor identification was cued using a 4AFC procedure. Four descriptors (one target and three distractors) were used for each odor. One related and two unrelated items were chosen as distractors: e.g. target: strawberry, distractors: flower (related), butter, cheese (unrelated) (Figure 2). Photographs of odor items (from Phase1) with additional words were used as descriptors.

Procedure: The study was described in detail to each child and his/her parents/legal guardians. Each child was tested individually in a quiet, well-ventilated room. Odorants were presented one at a time by removing the cap from the "Sniffin' Stick" and holding it 2-3 cm in front of the nose for 3 seconds. Children were asked to identify the odor from the four given descriptors, which were shown to the children before odor presentation. If unsure, children were allowed to smell the odor again.

Phase 3 - Evaluation of the test and acquisition of normative data

Participants: A total of 927 children aged 6 to 8 years from 19 countries participated. Fifty children were tested in each country, except UK n=41 and Canada n=36. The mean (SD) age was 6.9±0.8 years. There was a significant difference in age across countries ($F_{(df=18)}$=24.22, p<0.001). There was no significant difference in sex distribution (girls, n=467, 51.7%; boys, n=436, 48.3%; $X^2_{(df=1)=}$1.06, p=0.30) or sex distribution across countries ($X^2_{(df=18)}$=14.2, p=0.71). The sex of 24 children was not provided.

Material: The newly created 12-item odor identification test, the "U-Sniff" test, was used based on results of Phase 2 (Figure 1). As before, "Sniffin' Sticks" were presented in a 4AFC procedure. To increase contrast between descriptors[43] based on the results from Phase 2, the following descriptors were changed; a related distractor was changed to a non-related distractor: target: apple (orange changed to biscuit); target: onion (chocolate changed to strawberry and fish changed to banana); target: orange (lemon changed to flower); target: peach (strawberry changed to coffee) (Table 2).

Procedure: The task was explained to each participant and his/her parents/legal guardians and the "U-Sniff" test was administered in a similar manner as in Phase 2. The sum of correct answers was computed as the odor identification score.

Test—retest reliability

A group of 27 children in Germany (17 girls, 10 boys; mean (SD) age 6.8±0.7 years) was tested twice using the same testing procedure in two separate sessions. The minimum interval between sessions was two days.

Test validation in olfactory disorders

Fourteen children (8 girls, 6 boys, mean (SD) age 14.2±3.1 years, range 6-17 years) with isolated congenital anosmia (ICA) from Germany were included for test validation. These children were previously tested in our Smell & Taste Clinic using the original "Sniffin' Sticks" test [18] (olfactory threshold, odor discrimination and odor identification) and were diagnosed as having ICA. Children

living in Dresden (n=3) were retested using the new "U-Sniff" test in our Smell & Taste Clinic and the test was mailed to children not living in Dresden (n=11) with detailed instructions to be administered by their parents.

Data analysis

Analyses were performed using IBM Statistical Package for the Social Sciences version 23.0 (SPSS Inc., Chicago, IL, USA) with significance set at $p<0.05$. Non-parametric tests were used to analyze data from Phase1 due to the nature of the underlying data. An ANOVA with country, odor and sex was performed. Bonferroni-corrected post-hoc tests were used for multiple pairwise comparisons. Chi-square tests were used to evaluate the sex distribution of the populations. The test was designed as a clinical screening test, meaning it had to distinguish between normal olfactory function and olfactory dysfunction. Therefore, the 10th percentile was used as a cut-off based on existing tests [11, 44]. Receiver operator characteristic curve (ROC) was used in conjunction with the Youden index (Y=sensitivity+specificity-1) [45] to define the highest sensitivity and specificity of the new "U-Sniff" test. In addition, Pearson correlation was used to analyze test—retest reliability.

Results:

Phase 1**:** All children were able to perform the task. To select the most highly recognizable odor items, items were ranked according to the children's answers for each country separately. The most highly recognizable odor item for children from a specific country was assigned a ranking score of 36, the second a ranking score of 35, and so forth to the least recognizable odor item which was assigned a ranking score of 1. Averaging the ranking scores of all 17 countries, "chocolate" was identified as the most recognizable odor and "rubber" as the least known to children (Table 3). The top 20 most recognizable odor items were selected for further analysis. From most to least recognizable, these items were: chocolate, apple, strawberry, banana, flower, biscuit, orange, lemon, milk, honey, coffee, fish, cola, tomato, onion, cheese, cigarette, butter, cut grass, and peach. Because we were

unable to create suitable odorants for milk, cola and cigarette, they were excluded from the study. There was a significant difference in ranking between the 17 selected odor items ($F_{(df=16)}$=6.70, p<0.001). Although ranking of single odor items varied greatly between countries, there was no significant difference when average odor item rankings were compared ($F_{(df=16)}$=0.70, p=0.79). A detailed description of the ranking analysis can be seen in Table 3. These 17 odors were used for Phase 2 as seen in Figure 1.

Phase 2**:** All children understood the task and were able to complete testing. No children were excluded from the study. The mean (SD) percentage of correct odor identification across all 17 odors was 73.4±14.9%. There was a significant difference between countries on mean odor identification ($X^2_{(df=17)}$=125.1, p<0.001) and across odors ($X^2_{(df=16)}$=673.4, p<0.001). Figure 3 displays the mean identification (and 95% confidence interval) score for each odor. Lemon was most commonly identified correctly across countries and honey was least commonly identified across countries.

Odor selection for Phase 3: We selected only those odors that were identified greater than 2/3 (>66%) of the time in Phase 2. The following odors were selected: lemon, banana, coffee, flower, strawberry, fish, cut grass, orange, onion, apple and peach. Although there was a significant difference in correct identification of single odors between countries (except for lemon and flower) the mean correct identification of the selected 12 odors was >66% in all countries (mean 79.3±6.6% range: 67.5 – 92.2%).

Phase 3: All 927 children who participated in Phase 3 were able to complete the task. No children were excluded from the study. The mean (SD) odor identification score across all children was 9.88±1.80 points (range 2 - 12 points). The range of mean scores across countries varied from 8.2 to 11.2 points with a main effect of country ($F_{(df=18)}$=4.94, p<0.001), meaning that odor identification scores differed significantly across countries. In addition, a main effect of sex was observed with girls (mean 10.1±1.6 points) scoring higher on the "U-Sniff" test than boys (9.7±1.9 points) ($F_{(df=1)}$=7.85, p=0.005), but no main effect of age was found ($F_{(df=2)}$=0.66, p=0.52). Further analysis showed no

interaction between country and sex, country and age, or age and sex on the odor identification score (all p>0.1). Chi-square analysis for single odors revealed a significant difference in correct odor identification of single odors between countries (all p<0.001). In accordance with the results from Phase 2, all odors except "butter" (64%) were identified on average >66% of the time (range 64.0 – 90.8%).

Countries were grouped into continents to obtain normative data. On average, higher odor identification scores were reached in European compared with American countries ($t_{(df=725)}$=2.21, p=0.028).

Europe: The mean (SD) odor identification score of European countries was 10.2±1.7 points (range 9.3 – 11.2 points). A significant difference of odor identification scores across European countries was observed ($F_{(df=10)}$=5.50, p<0.001); however, Bonferroni adjusted post hoc-tests revealed that only Italy (higher scores) and the Czech Republic (lower scores) were significantly different from each other. No other comparisons of odor identification scores between European countries reached significance. To define the cut-off between normal olfactory function and a reduced sense of smell, the 10th percentile of data distribution was used as a criterion. Across European countries, the 10th percentile on odor identification score was 8 points (Table 4). When analyzing the 10th percentile cut-off for each country individually, only cut-offs for the Czech Republic (7 points) and Poland (6 points) were lower (Table 4).

America: The mean odor identification score of the American countries was 9.8±1.8 points (range: 9.4 – 10.0 points). A main effect of country on the odor identification test was found ($F_{(df=3)}$=2.78, p=0.042) but Bonferroni adjusted post hoc tests showed no significant difference for odor identification scores across the four American countries. When the 10th percentile criterion was applied, a score below 7 points on the "U-Sniff" test would indicate a reduced sense of smell (Table 4).

Other countries: Children in Egypt scored a mean (SD) 9.0±1.7 points on the "U-Sniff" test. The cut-off

between normal olfactory function and a reduced sense of smell would be 8 points. Children in India scored a mean (SD) 8.2±1.9 points on the "U-Sniff" test and a score below 5 points would indicate a reduced olfactory function. In Israel, children scored on average 8.7±2.3 points on the "U-Sniff" test. The 10th percentile cut-off would be 6 points. The average odor identification score for children from Japan was 10.8±1.0 points resulting in a 10th percentile cut-off of <9 points (Table 4).

Test—retest reliability: In the subgroup of 27 German children undergoing test—retest, the mean (SD) interval between tests was 57.6±68.0 days (range: 2 – 229 days). Scores from the first (10.15±2.33 points) and second tests (10.26±2.12 points) did not differ significantly ($t_{(df=26)}$=0.44, p=0.66). A strong positive correlation between odor identification scores from the first and second testing was observed (r_{27}=0.83, p<0.001) (Figure 4).

Test validation: A group of 14 children who were previously diagnosed in our Smell & Taste Clinic as having ICA using the standard "Sniffin' Sticks" test including olfactory threshold, odor discrimination and odor identification testing, were investigated. These anosmic children previously scored a mean (SD) of 10.80±3.30 points of a possible maximum of 48 points on the standard "Sniffin' Sticks" test (sum of threshold, discrimination and identification test) and now scored a mean (SD) of 3.57±1.83 points on the "U-Sniff" test. Odor identification sores differed significantly between patients with ICA and the German study population ($t_{(df=62)}$=13.7, p<0.001) (Figure 5a). In addition only one patient scored 7 points, while all other children with ICA scored ≤6 points. A ROC analysis to distinguish between ICA and healthy controls by means of the "U-Sniff" test showed an area under the curve (AUC) of 0.99 (p<0.001) (Figure 5b). By using the highest Youden-index, a sensitivity of 100% and a specificity of 86% to confirm a normal sense of smell was reached when a cut-off of ≥6 points was used. When a cut-off of ≥8 points was used the sensitivity was 92% and the specificity was 100%, respectively.

Discussion:

We developed an international odor Identification test for children – the "U-Sniff" test. Normative data were generated and the test's validity and test—retest reliability were evaluated.

We included children aged 6 to 8 years in this study. Previous studies have demonstrated that the ability to identify odors increases with age in children [26, 30, 46, 47]. This is due to an ongoing process of odor learning [20, 23, 24] rather than an actual increase in olfactory function[46]. Our aim during the development of the "U-Sniff" test for children was to only include odors that are well known and able to be correctly identified by a majority of children from different cultures around the world. To avoid the bias of age-dependent odor learning, we included young children who were old enough to understand and perform an odor identification task. Recently, Cavazzana and colleagues reported that odor identification tasks in children younger than 5 years of age to be unreliable [48], considering previous studies [39, 46]. Although some studies have claimed that children as young as 3-4 years are able to perform an odor identification task [28-30], results must be evaluated with caution. In the study by Dzaman and colleagues, parents were allowed to explain the odor descriptors to the children, which may promote response bias [29]. Odors such as cinnamon and Play-Doh in the NIH-Toolbox were poorly identified by children with odor identification, reaching almost chance-level [30]. Taken together, these studies suggest that measurements of odor identification in children younger than 5 or 6 years of age are unreliable. The upper age limit of 8 years in our study was chosen to minimize the possible influence of age on odor identification scores during the development of the "U-Sniff" odor identification test. Therefore no age difference in odor identification was observed in the current study. Because we only studied this small age range future studies should aim to test a broader age range in a more systematic manner, including both younger children and adolescents, to determine age dependent results of the "U-Sniff" test.

Of 927 children in Phase 3, we found a difference in odor identification score between girls and boys.

The literature regarding sex differences with respect to odor identification in children is contradictory. Richman et al. [26] and others [21, 47] found that girls outperformed boys, but Sorokowska et al. [49], and others [27, 29, 31], found no difference between girls and boys on an odor identification test. Differences in these studies might result from the different age ranges of study populations as well as from use of different tests.

Odor identification scores differed significantly across countries despite the fact that all odorants were selected with data from all countries. The final odorant selection was based on average scores including all countries. Different cultural backgrounds might account for the difference in odor identification scores across countries. This is in line with previous studies in adults that have shown differences in odor identification scores across countries using the "Sniffin' Sticks" 16-item odor identification test [38, 50]. Even though the odor identification scores differed significantly across countries in our study, children were able to perform the test in all countries and test scores of 69 – 93% correct identification are comparable with previous studies in children, i.e. NIH-toolbox 72%, Dzaman et al. 76%, Van Spronsen et al. 62% and "SCHOT" 85-89% [27, 29, 30, 47].

Odor items for our "U-Sniff" test for children were selected in three phases, after piloting the possible odor items. First, familiar items were chosen by using photographs of odor items. In Phase 2, the most well-known odorants were selected for an odor identification test and only the items identified correctly in more than 66% of cases were chosen for inclusion. Previous studies used a slightly higher cut-off for including odors in an odor identification test. Hummel and colleagues used 75% in the development of the "Sniffin' Sticks" odor identification test [10]. The same criterion was used by Dzaman et al. in the development of a pediatric smell test [29]. We chose a lower criterion of 66% for odor identification because this test was developed internationally in a young population of children and therefore lower average odor identification scores are expected. The ranking of familiar odor items in Phase 1 did not completely match the final odor identification scores, e.g. chocolate was ranked as the most recognized odor item based on its photograph but was only the 13th most commonly identified odor in Phase 2 and did not meet the criterion for inclusion in the final "U-Sniff"

test version. A similar phenomenon was observed in children using the original "Sniffin' Sticks" 16-item odor identification test. In a study population of 537 children age 6 - 17 years, the item "apple" was only identified 34% of the time [31]. This is surprising because apple is a fruit commonly consumed by children. The difference between knowing the odor item and correct odor identification of the same might result from suboptimal odorant selection or poor - fitting picture-odor concept [43]. This is speculative, however, because no congruency rating of the odor – picture combination was measured in this study.

Test—retest reliability of the "U-Sniff" test was evaluated in a subgroup of 27 German children and found to be highly reliable (r=0.83). This was more highly reliable than the majority of other pediatric smell tests ("Sniffin' Kids" (r=0.44), NIH-Toolbox (r=0.45), "Smell Wheel" (r=0.70), "SCHOT" (r=0.98))[27, 28, 30, 31]. Its reliability is in the same range of the standard odor identification test in adults such as the UPSIT (r=0.92) or the "Sniffin' Sticks" odor identification test (r=0.88)[11, 51].

Several odor identification tests have been developed for children to distinguish between normosmia and olfactory dysfunction. Only two tests, the "Sniffin' kids" odor identification test and the test developed by Richman and colleagues, have been validated by including anosmic children during test development [26, 31]. The "U-Sniff" test was validated by including 14 children with diagnosed ICA into the study. Children with ICA scored significantly lower on the "U-Sniff" test than the control group. In addition, it was possible to distinguish between normal sense of smell and anosmia with high sensitivity and specificity. The test validation does not allow a separation between anosmia and hyposmia. To increase the number of anosmic children in our test validation, the age range of this population was 6-17 years. Previous studies have reported an increase in odor identification score with age [27, 28, 39, 52]. Such an increase is not expected in children with ICA. In fact, no correlation between age and odor identification score was observed in children with ICA in the current study (ρ=-0.37, p=0.194). Therefore the difference in age range between the study populations should not affect the study outcome.

By including a large study population of 927 children in Phase 3, we are able to present normative data for children aged 6-8 years. We chose the 10th percentile as a cut-off, as the 10th percentile is commonly used to separate normosmia from a reduced sense of smell in olfactory testing, [11, 18, 26, 31]. Although, the 10th percentile value varied across 19 countries studied, country-specific values distinguished between normosmia and anosmia with high sensitivity and specificity. ROC analysis also was conducted. By using the highest Youden-index a cut-off of 6 points lead to the highest sensitivity and specificity to distinguish between normosmia and a reduced sense of smell. Comparing the two cut-off criteria, the 10th percentile (≥8 points) lead to a higher specificity but slightly lower sensitivity than the ROC analysis (≥6 points) to of confirm normosmia. Due to the lower frequency of olfactory dysfunction in children [4] this cut-off should be used as an orientation rather than a fixed value. Scores must be considered in regard to the whole clinical appearance of the patient, e.g. medical history, including subjective reporting of the sense of smell, and other relevant investigations.

Limitations of this study are that most of the 19 countries included are from Europe and America, and therefore it is necessary to study the generalizability of this test to the rest of the world. In addition, odor identification has been shown to be associated with verbal fluency of children [21] and the current study did not investigate verbal fluency of participants. Compared with the UPSIT and the "Sniffin' Sticks" extended odor identification test with 40 and 16 odor items, respectively, the final version of the "U-Sniff" test with only 12 items seems rather short [10, 11]. However, previously it has been proven that 12 items are sufficient for an odor identification test, e.g. CC-SIT, "Sniffin' Sticks" 12-item odor identification test [37, 53]. Considering the close range of odor identification scores across countries, the reliability of the "U-Sniff" test was only tested within the German sub-population. Future studies should investigate the reliability of the "U-Sniff" test in additional countries. Odor items were suggested based on the experience of participating researchers and therefore, it might be possible that other odor items also would have been suitable for inclusion. The majority of children with ICA were tested at home with the test being administered by their parents. Although this

method was not validated in the current study, previous research has demonstrated no difference in regard to odor identification scores between self-administered and examiner conducted tests using the "Sniffin' Sticks" in adults population [54] and the "Smell Wheel in children [28].

The 12-item "U-Sniff" international odor identification test for children demonstrates a high test-retest reliability and was validated by including children with ICA. This test offers an efficient method of distinguishing with high sensitivity and specificity children with normosmia from those with a reduced sense of smell.

Abbreviations:

UPSIT	University of Pennsylvania Smell Identification Test
AFC	Alternative forced choice
CC-SIT	Cross Cultural Smell Identification Test
U-Sniff	Universal-Sniff
ICA	Isolated congenital anosmia
ROC	Receiver operator characteristics
SCHOT	Sydney Children's Hospital Odor Identification Test

Acknowledgments:

We thank the following colleagues for assistance with the data collection: Edith Schriever and Marie-Luise Gruhn (Germany), Sarah Selvadurai (Canada), Annachiara Cavazzana, Kristen Gregory, Jovanna Pope and Christiane Wesarg (USA), Alejandra Aguirre, Javier Alonso and Sandra Dominguez (Spain), Ricardo Morales and Marcela Osorio (Chile); the City of Turku for allowing data collection in a day-care center, all children and their parents / legal guardians for participating in this study. None of the individuals listen in the acknowledgments has a financial relationship relevant to this article to disclose. None of the individuals listen in the acknowledgments have a conflict of interest to disclose.

References

[1] Temmel AF, Quint C, Schickinger-Fischer B, Klimek L, Stoller E, Hummel T. Characteristics of olfactory disorders in relation to major causes of olfactory loss. Arch Otolaryngol Head Neck Surg. 2002;128:635-41.

[2] Landis BN, Konnerth CG, Hummel T. A study on the frequency of olfactory dysfunction. Laryngoscope. 2004;114:1764-9.

[3] Bramerson A, Johansson L, Ek L, Nordin S, Bende M. Prevalence of olfactory dysfunction: the skovde population-based study. Laryngoscope. 2004;114:733-7.

[4] Oozeer NB, Forbes K, Clement AW, Kubba H. Management of paediatric olfactory dysfunction: how we do it. Clin Otolaryngol. 2011;36:494-9.

[5] Ottaviano G, Cantone E, D'Errico A, Salvalaggio A, Citton V, Scarpa B, et al. Sniffin' Sticks and olfactory system imaging in patients with Kallmann syndrome. Int Forum Allergy Rhinol. 2015;5:855-61.

[6] Schriever VA, Studt F, Smitka M, Grosser K, Hummel T. Olfactory function after mild head injury in children. Chem Senses. 2014;39:343-7.

[7] Altundag A, Salihoglu M, Cayonu M, Tekeli H. Clinical assessment of olfactory functions in children who underwent adenotonsillectomy during pre- and post-operative period. Int J Pediatr Otorhinolaryngol. 2014;78:1138-42.

[8] Mueller CA, Quint C, Gulesserian T, Temmel AF, Hummel T. Olfactory function in children with cystic fibrosis. Acta Paediatr. 2007;96:148-9.

[9] Konstantinidis I, Triaridis S, Triaridis A, Petropoulos I, Karagiannidis K, Kontzoglou G. How do children with adenoid hypertrophy smell and taste? Clinical assessment of olfactory function pre- and post-adenoidectomy. Int J Pediatr Otorhinolaryngol. 2005;69:1343-9.

[10] Hummel T, Sekinger B, Wolf SR, Pauli E, Kobal G. 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. Chem Senses. 1997;22:39-52.

[11] 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.

[12] Cain WS, Gent JF, Goodspeed RB, Leonard G. Evaluation of olfactory dysfunction in the Connecticut Chemosensory Clinical Research Center. Laryngoscope. 1988;98:83-8.

[13] Cardesin A, Alobid I, Benitez P, Sierra E, de Haro J, Bernal-Sprekelsen M, et al. Barcelona Smell Test - 24 (BAST-24): validation and smell characteristics in the healthy Spanish population. Rhinology. 2006;44:83-9.

[14] Santos DV, Reiter ER, DiNardo LJ, Costanzo RM. Hazardous events associated with impaired olfactory function. Arch Otolaryngol Head Neck Surg. 2004;130:317-9.

[15] Croy I, Nordin S, Hummel T. Olfactory disorders and quality of life--an updated review. Chem Senses. 2014;39:185-94.

[16] Renner B, Mueller CA, Dreier J, Faulhaber S, Rascher W, Kobal G. The candy smell test: a new test for retronasal olfactory performance. Laryngoscope. 2009;119:487-95.

[17] Heilmann S, Strehle G, Rosenheim K, Damm M, Hummel T. Clinical assessment of retronasal olfactory function. Arch Otolaryngol Head Neck Surg. 2002;128:414-8.

[18] Hummel T, Kobal G, Gudziol H, Mackay-Sim A. Normative data for the "Sniffin' Sticks" including tests of odor identification, odor discrimination, and olfactory thresholds: an upgrade based on a group of more than 3,000 subjects. Eur Arch Otorhinolaryngol. 2007;264:237-43.

[19] Doty RL, Shaman P, Applebaum SL, Giberson R, Siksorski L, Rosenberg L. Smell identification ability: changes with age. Science. 1984;226:1441-3.

[20] Cain WS, Stevens JC, Nickou CM, Giles A, Johnston I, Garcia-Medina MR. Life-span development of odor identification, learning, and olfactory sensitivity. Perception. 1995;24:1457-72.

[21] Monnery-Patris S, Rouby C, Nicklaus S, Issanchou S. Development of olfactory ability in children: sensitivity and identification. Dev Psychobiol. 2009;51:268-76.

[22] Lehrner JP, Gluck J, Laska M. Odor identification, consistency of label use, olfactory threshold and their relationships to odor memory over the human lifespan. Chem Senses. 1999;24:337-46.

[23] Fjaeldstad A, Sundboll J, Niklassen A, Ovesen T. Odor Familiarity and Identification Abilities in Adolescents. Chem Senses. 2017.

[24] Mullol J, Alobid I, Marino-Sanchez F, Quinto L, de Haro J, Bernal-Sprekelsen M, et al. Furthering the understanding of olfaction, prevalence of loss of smell and risk factors: a population-based survey (OLFACAT study). BMJ Open. 2012;2.

[25] Kobal G, Hummel T, Sekinger B, Barz S, Roscher S, Wolf S. "Sniffin' sticks": screening of olfactory performance. Rhinology. 1996;34:222-6.

[26] Richman RA, Post EM, Sheehe PR, Wright HN. Olfactory performance during childhood. I. Development of an odorant identification test for children. J Pediatr. 1992;121:908-11.

[27] Laing DG, Segovia C, Fark T, Laing ON, Jinks AL, Nikolaus J, et al. Tests for screening olfactory and gustatory function in school-age children. Otolaryngol Head Neck Surg. 2008;139:74-82.

[28] Cameron ELD, R.L. Odor identification testing in children and young adults using the smell wheel. Int J Pediatr Otorhinolaryngol. 2013;77:346-50.

[29] Dzaman K, Zielnik-Jurkiewicz B, Jurkiewicz D, Molinska-Glura M. Test for screening olfactory function in children. Int J Pediatr Otorhinolaryngol. 2013;77:418-23.

[30] Dalton P, Doty RL, Murphy C, Frank R, Hoffman HJ, Maute C, et al. Olfactory assessment using the NIH Toolbox. Neurology. 2013;80:S32-6.

[31] Schriever VA, Mori E, Petters W, Boerner C, Smitka M, Hummel T. The "sniffin' kids" test - a 14-item odor identification test for children. PLoS One. 2014;9:e101086.

[32] Kennedy WP, Lewis CP, Stow J, Sobol SE. A Critical Period in Postnatal Neuroplasticity of Olfaction: A Pediatric Tracheostomy Model. JAMA Otolaryngol Head Neck Surg. 2016;142:127-31.

[33] Laing DG, Armstrong JE, Aitken M, Carroll A, Wilkes FJ, Jinks AL, et al. Chemosensory function and food preferences of children with cystic fibrosis. Pediatr Pulmonol. 2010;45:807-15.

[34] Armstrong JE, Laing DG, Wilkes FJ, Laing ON. Olfactory function in Australian aboriginal children and chronic otitis media. Chem Senses. 2008;33:503-7.

[35] Armstrong JE, Laing DG, Wilkes FJ, Kainer G. Smell and taste function in children with chronic kidney disease. Pediatr Nephrol. 2010;25:1497-504.

[36] Cohen J, Laing DG, Wilkes FJ. Taste and smell function in pediatric blood and marrow transplant patients. Support Care Cancer. 2012;20:3019-23.

[37] Doty RL, Marcus A, Lee WW. Development of the 12-item Cross-Cultural Smell Identification Test (CC-SIT). Laryngoscope. 1996;106:353-6.

[38] Oleszkiewicz A, Taut M, Sorokowska A, Radwan A, Kamel R, Hummel T. Development of the Arabic version of the "Sniffin' Sticks" odor identification test. Eur Arch Otorhinolaryngol. 2016;273:1179-84.

[39] Bastos LO, Guerreiro MM, Lees AJ, Warner TT, Silveira-Moriyama L. Effects of age and cognition on a cross-cultural paediatric adaptation of the Sniffin' Sticks Identification Test. PLoS One. 2015;10:e0131641.

[40] Jiang RS, Su MC, Liang KL, Shiao JY, Wu SH, Hsin CH. A pilot study of a traditional Chinese version of the University of Pennsylvania Smell Identification Test for application in Taiwan. Am J Rhinol Allergy. 2010;24:45-50.

[41] Altundag A, Tekeli H, Salihoglu M, Cayonu M, Yasar H, Kendirli MT, et al. Cross-culturally modified University of Pennsylvania Smell Identification Test for a Turkish population. Am J Rhinol Allergy. 2015;29:e138-41.

[42] Cho JH, Jeong YS, Lee YJ, Hong SC, Yoon JH, Kim JK. The Korean version of the Sniffin' stick (KVSS) test and its validity in comparison with the cross-cultural smell identification test (CC-SIT). Auris Nasus Larynx. 2009;36:280-6.

[43] Gudziol V, Hummel T. The influence of distractors on odor identification. Arch Otolaryngol Head Neck Surg. 2009;135:143-5.

[44] Kobal G, Klimek L, Wolfensberger M, Gudziol H, Temmel A, Owen CM, et al. Multicenter investigation of 1,036 subjects using a standardized method for the assessment of olfactory function combining tests of odor identification, odor discrimination, and olfactory thresholds. Eur Arch Otorhinolaryngol. 2000;257:205-11.

[45] Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3:32-5.

[46] Hummel T, Bensafi M, Nikolaus J, Knecht M, Laing DG, Schaal B. Olfactory function in children assessed with psychophysical and electrophysiological techniques. Behav Brain Res. 2007;180:133-8.

[47] van Spronsen E, Ebbens FA, Fokkens WJ. Olfactory function in healthy children: normative data for odor identification. Am J Rhinol Allergy. 2013;27:197-201.

[48] Cavazzana A, Wesarg C, Schriever VA, Hummel T, Lundstrom JN, Parma V. A Cross-Cultural Adaptation of the Sniffin' Sticks Olfactory Identification Test for US children. Chem Senses. 2017;42:133-40.

[49] Sorokowska A, Schriever VA, Gudziol V, Hummel C, Hahner A, Iannilli E, et al. Changes of olfactory abilities in relation to age: odor identification in more than 1400 people aged 4 to 80 years. Eur Arch Otorhinolaryngol. 2015;272:1937-44.

[50] Tekeli H, Altundag A, Salihoglu M, Cayonu M, Kendirli MT. The applicability of the "Sniffin' Sticks" olfactory test in a Turkish population. Medical science monitor : international medical journal of experimental and clinical research. 2013;19:1221-6.

[51] Haehner A, Mayer AM, Landis BN, Pournaras I, Lill K, Gudziol V, et al. High test-retest reliability of the extended version of the "Sniffin' Sticks" test. Chem Senses. 2009;34:705-11.

[52] Dalton P, Mennella JA, Maute C, Castor SM, Silva-Garcia A, Slotkin J, et al. Development of a test to evaluate olfactory function in a pediatric population. Laryngoscope. 2011;121:1843-50.

[53] Hummel T, Konnerth CG, Rosenheim K, Kobal G. Screening of olfactory function with a four-minute odor identification test: reliability, normative data, and investigations in patients with olfactory loss. Ann Otol Rhinol Laryngol. 2001;110:976-81.

[54] Mueller CA, Grassinger E, Naka A, Temmel AF, Hummel T, Kobal G. A self-administered odor identification test procedure using the "Sniffin' Sticks". Chem Senses. 2006;31:595-8.

Table 1: Details of the odorants used to create the "U-Sniff" odor identification test.

Table 2: Four alternative descriptors given to identify an odor are shown for each target odor (written in bold).

Table 3: Mean ± standard deviation (SD) odor item ranking and range determined from Phase 1 of study.

Table 4: Normative data of the "U-Sniff" test for children by country. Mean ± standard deviation (SD) odor identification score, range and the cut-off to distinguish between normosmia and a reduced sense of smell by using the 10^{th} percentile are shown.

Figure legends

Figure 1: Phases of odor item selection, the odor items and odors used in Phase 1 (36 items), Phase 2 (17 odors) and Phase 3 (12 odors) are displayed.

Figure 2: Descriptors for odor identification, an example for the usual descriptors for odor identification (strawberry being the target) is displayed.

Figure 3: Odor identification Phase 2, percent correct identification of odors used in Phase 2 (mean + 95% CI interval). Twelve odors (dark grey) were identified correctly in 66% of cases and were selected for the "U-Sniff" test for children. The dashed line indicates the 66% odor identification.

Figure 4: Reliability of the "U-Sniff" test, test—retest reliability (r=0.83) of the "U-Sniff" test is shown. The size of dots represents the number of participants.

Figure 5: Comparison of areas for anosmia and control children, a) Mean ± standard deviation (SD) odor identification scores for children with ICA (light gray) and controls (dark gray). Children with isolated congenital anosmia (ICA) scored significantly lower than controls ($t=13.7$, $p<0.001$). b) receiver operating characteristic curve (ROC) analysis distinguishing between children with ICA and. Area under the curve (AUC) was 0.99.