“U-Sniff” - the international odor identification test for children: an extension of its normative database and study of global reliability*

Abstract

Background: To extend the previous study by Schriever and colleagues from 2018 providing normative data and re-investigating the reliability for U-Sniff test in children in additional countries.

Methodology: A total of 388 children (196 boys, 192 girls) from eight countries (China, Germany, Iran, Netherlands, Norway, Oman, Paraguay, and Russia) participated in this study. The children were recruited from public local schools in those particular countries. The odor identification ability was evaluated using the U-Sniff test, a 12-item odor identification test developed for children. In addition, reliability was examined using test-retest design in the children for each country.

Results: The mean U-Sniff test score across all children was 10.3 ± 1.7 points. Normative data were established. A high test-retest reliability of the U-Sniff test was demonstrated across the eight participating countries.

Conclusions: The U-Sniff test for children exhibits a high test-retest reliability on a global scale.

Key words: U-Sniff test, odor identification, international, olfactory dysfunction, children

Introduction

Olfactory dysfunction is frequent. Several large population-based studies reported that the prevalence of olfactory dysfunction is 19-24%[1-4]. However, the frequency of olfactory dysfunction in children is still unknown. The “Sniffin’ Sticks” test[5] and the University of Pennsylvania Smell Identification Test (UPSIT)[6] are the most frequently used standardized psychophysical tests for evaluating olfactory function in adults. In part because of their limited attention span and unfamiliarity with the odors, it is a challenge to evaluate olfactory function in children using either one of these tests. Previous studies showed that olfactory identification performance on the two tests in children is
lower than that of adults\(^{7-9}\). However, the performance on the “Sniffin’ Sticks” test and UPSIT may not accurately represent the children’s olfactory ability. Therefore, there are several odor identification tests which have been developed for children, such as, an odor identification test based on microencapsulated “Scratch and Sniff” cards\(^ {10}\), the Sydney Children’s Hospital Odor Identification Test (SCHOT)\(^ {11}\), the Lyon clinical Olfactory Test (LCOT)\(^ {12}\), the NIH Toolbox Odor Identification Test\(^ {13}\), the Smell Wheel\(^ {14}\), and the “Sniffin’ Kids” Test\(^ {15}\). To date, however, those tests are not used widely, partly because most of them are not commercially available and they have not been examined in a cross-cultural fashion.

To overcome these issues, Schriever and colleagues\(^ {16}\) developed an international 12-item odor identification test, the U-Sniff test, for children aged 6-8 years. The U-Sniff test shows a high test-retest reliability (r=0.83), and was validated in children with isolated congenital anosmia (ICA). The results showed that the test allows to distinguish children with normosmia from children with ICA with a sensitivity of 100% and specificity of 86%. In addition, normative data were established for 19 countries (Africa: Egypt; America: Canada, Chile, Mexico, United States; Asia: India, Japan; Europe: Czech Republic, Finland, Germany, Greece, Israel, Italy, Poland, Spain, Sweden, Switzerland, Turkey, United Kingdom). In a further study, Gellrich and colleagues\(^ {17}\) provided normative data for the U-Sniff test for a large sample of children and adolescents aged 6-17 years.

However, most of the 19 countries included in the study by Schriever and colleagues\(^ {16}\) are from Europe and America, so it appears necessary to investigate the universality of the U-Sniff test for additional countries. In addition, the previous study has only tested the test-retest reliability of the U-Sniff test in a subgroup of children from Germany. Therefore, it is necessary to also investigate the reliability of the U-Sniff test in other countries. The aim of the current study is therefore, to extend the previous study\(^ {16}\) providing normative data and reliability for U-Sniff test in children in additional countries.

Materials and Methods

Participants

The following eight countries participated in this study (alphabetical order): China, Germany, Iran, Netherlands, Norway, Oman, Paraguay, and Russia. Figure 1 shows a global map of all
countries involved in this study, and in addition, in a different color, the countries participating in the initial study by Schriever and colleagues\(^1\) on the U-Sniff test. Children were recruited from public local schools in those particular countries. With the exception of Oman (n=38) in each country 50 children were tested. Exclusion criteria were as follows: known smell dysfunction, diseases known to have a significant influence on the sense of smell (e.g., renal failure, epilepsy), and acute or chronic rhinosinusitis. A total of 388 children (196 boys, 192 girls), mean age 7.0±0.8 years (range: 6-8 years), were included. Sex distribution between countries did not differ significantly (χ\(^2\)=0.46, p=1.0). Age did not differ significantly between girls (mean: 7.1±0.8 years) and boys (means: 7.0±0.8 years) (t=1.35, p=0.185).

This study was performed according to the Declaration of Helsinki on Biomedical Studies Involving Human Subjects, and was approved by the Ethics Committee of the Medical Faculty of the University of Dresden Medical School and additionally by the ethics committees of each participating center. All the children gave their assent and their parents or legal guardians gave oral or written consent, according to the local ethics regulation, to participate following detailed explanations on aims and potential risks of the study.

### U-Sniff test

The study consisted of two phases. In phase 1, the U-Sniff test was used to evaluate the odor identification ability in children. The odor identification test includes 12 odors (12 items; apple, banana, butter, coffee, cut grass, fish, flower, lemon, onion, orange, peach, and strawberry), presented by using the “Sniffin’ Stick”. The descriptors (targets and distractors) were showed in the Table 1. Each item was presented separately by uncapping the “Sniffin’ Stick” and keeping the “Sniffin’ Stick” approximately 2cm for 2-3 seconds beneath the child’s nose. Children were asked to identify each odor with the help of four descriptors. Before odor presentation, the descriptors (pictures and words) were shown and read to the children, and then, after odor presentation, the children were asked to identify the smell.

This was based on previous work showing that presenting the descriptors first leads to better results than presenting the odor first\(^1\). If uncertain, the child was allowed to smell the odor up to three times. No immediate feedback as to the accuracy of the responses was given to the children.

In phase 2, the children from phase 1 were tested a second time using the “U-Sniff” test. The interval between the first and the second testing ranged from 2 to 93 days (10.31±13.85 days). The procedure was the same as in Phase 1 of the study.

### Statistical analyses

The IBM SPSS 23.0 (SPSS Inc, Chicago, IL, USA) was used for all analyses. Significance levels were set at p < 0.05. The sex distribution of the participants was evaluated using χ\(^2\) test. Analysis covariance (ANCOVA) followed by Bonferroni post-hoc tests were used to examine the differences in olfactory function in different countries. In line with the study by Schriever and colleagues\(^1\), with the 10th percentile was used as the criterion...
to discriminate between normosmia and reduced olfactory function in this study. In addition, test-retest reliability was calculated using Pearson correlation.

**Results**

**Phase 1**

The mean U-Sniff test score across all children at Phase 1 was 10.3±1.7 points (range 2-12 points), and the mean scores across countries ranged from 9.4 to 11.2 points. The rank order of the U-Sniff test score was as following: Iran, Germany, Russia, Netherlands, Norway, China, Paraguay and Oman, sorted from lowest to highest score (Table 2). There was a significant difference on the U-Sniff test scores across countries (F=6.43, p<0.001). Bonferroni adjusted post-hoc tests showed that the performance of children from Germany was poorer than that from Oman and Paraguay (ps<0.05); the performance of Children from Iran was poorer than that from China, Norway, Oman and Paraguay (ps<0.05); and the performance of Children from Russia was poorer than that from Oman (p<0.05). All other comparisons were not significant.

There was a main effect of sex on the U-Sniff test scores, showing that the scores in girls (mean 10.5±1.6 points) were higher than that of boys (mean 10.2±1.7 points) (F=4.02, p=0.046). In addition, a significant effect of age on the U-Sniff test scores was found (F=13.49, p<0.001) with older children scoring higher than younger children. Post-hoc analysis showed a significant difference in odor identification score between children age 6 years (mean 9.7±2.0 points) and older children (7 years: mean 10.4±1.5 points; 8 years: mean 10.8±1.4 points) (ps<0.01) but not between children age 7 and 8 years (p=0.33). There was no interaction effect between country and sex, country and age, or age and sex on the U-Sniff test score. Single odors were correctly identified between 73 – 97% across all children.

---

Table 2. Normative data of the U-Sniff test for children by country.

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample size</th>
<th>Means</th>
<th>Range</th>
<th>10th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across all countries</td>
<td>388 (girls: 192; boys:196)</td>
<td>10.32±1.70</td>
<td>2-12</td>
<td>8</td>
</tr>
<tr>
<td>China</td>
<td>50 (girls: 25; boys:25)</td>
<td>10.60±1.13</td>
<td>8-12</td>
<td>9</td>
</tr>
<tr>
<td>Germany</td>
<td>50 (girls: 25; boys:25)</td>
<td>9.94±1.95</td>
<td>3-12</td>
<td>7</td>
</tr>
<tr>
<td>Iran</td>
<td>50 (girls: 25; boys:25)</td>
<td>9.36±2.16</td>
<td>3-12</td>
<td>7</td>
</tr>
<tr>
<td>Netherlands</td>
<td>50 (girls: 27; boys:23)</td>
<td>10.12±1.49</td>
<td>6-12</td>
<td>9</td>
</tr>
<tr>
<td>Norway</td>
<td>50 (girls: 26; boys:24)</td>
<td>10.40±1.64</td>
<td>4-12</td>
<td>8</td>
</tr>
<tr>
<td>Oman</td>
<td>38 (girls: 18; boys:20)</td>
<td>11.21±0.99</td>
<td>8-12</td>
<td>10</td>
</tr>
<tr>
<td>Paraguay</td>
<td>50 (girls: 25; boys:25)</td>
<td>11.04±1.12</td>
<td>7-12</td>
<td>10</td>
</tr>
<tr>
<td>Russia</td>
<td>50 (girls: 25; boys:25)</td>
<td>10.10±1.10</td>
<td>2-12</td>
<td>8</td>
</tr>
</tbody>
</table>

Notes. All the children were born in that particular country. The performance in the children from Germany was poorer than that from Oman and Paraguay (ps<0.05); the performance in the children from Iran was poorer than that from China, Norway, Oman and Paraguay (ps<0.05); and the performance in the children from Russia was poorer than that from Oman (p<0.05).

Table 3. Test-retest correlation.

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample size</th>
<th>Correlation coefficient</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across all countries</td>
<td>388 (girls: 192; boys:196)</td>
<td>r=0.71</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>China</td>
<td>50 (girls: 25; boys:25)</td>
<td>r=0.69</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Germany</td>
<td>50 (girls: 25; boys:25)</td>
<td>r=0.72</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Iran</td>
<td>50 (girls: 25; boys:25)</td>
<td>r=0.61</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Netherlands</td>
<td>50 (girls: 27; boys:23)</td>
<td>r=0.63</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Norway</td>
<td>50 (girls: 26; boys:24)</td>
<td>r=0.80</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Oman</td>
<td>38 (girls: 18; boys:20)</td>
<td>r=0.74</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Paraguay</td>
<td>50 (girls: 25; boys:25)</td>
<td>r=0.27</td>
<td>p=0.055</td>
</tr>
<tr>
<td>Russia</td>
<td>50 (girls: 25; boys:25)</td>
<td>r=0.73</td>
<td>p&lt;0.001</td>
</tr>
</tbody>
</table>

Note. All the children were born in that particular country.
Phase 2
The mean U-Sniff test score across all children at phase 2 was 10.5±1.7 points (range 3-12 points). Mean odor identification scores differed significantly between phase 1 and phase 2 (t=2.48, p=0.013) with higher odor identification scores at the second visit. Although the difference in odor identification score between phase 1 and 2 reached statistical significance, it has to be pointed out, that the mean difference between odor identification score of phase 1 and 2 was only 0.2 points.

Test-retest reliability
The U-Sniff test scores from phase 1 and phase 2 showed a positive correlation (r=0.71, p<0.001). Except for children examined in Paraguay (r=0.27, p=0.055), odor identification scores from phase 1 and phase 2 significantly correlated for each individual country with correlations ranging from 0.61 to 0.80 (Figure 2, Table 3).

10th percentile
The 10th percentile of the U-Sniff test score distribution was used as the criterion to define the cut-off between normosmia and olfactory dysfunction. Across all countries, the 10th percentile of the U-Sniff test score was 8 points. For individual countries, the 10th percentile cut-off scores ranged between 7 and 10 points (Table 2).

Discussion
The current study extends the previous work by Schriever and colleagues providing normative data and test-retest reliability for the U-Sniff test for children between age 6 to 8 years in eight additional countries.

The mean odor identification score on the U-Sniff test across all children at Phase 1 was 10.3±1.7 points, which is similar to the results in the previous studies (9.88 ± 1.80 points)(14) and (10.10 ± 1.68 points)(15). The odorants of the U-Sniff test were selected according to the average scores across 17 different countries as the same as the study by Schriever and colleagues. In line with the results of the study by Schriever and colleagues, the U-Sniff test scores differed significantly across the eight countries in the present study, which might be due to different cultural backgrounds. For example, the accuracy of tropical or subtropical fruits (such as banana and lemon) in tropical or subtropical countries (such as Oman and Paraguay) were higher than those in other countries (banana: 98%-100% vs. 74%-90%; lemon: 94%-97% vs. 78%-88%). The accuracy of odor identification in the U-Sniff test in all the eight countries ranged from 78% to 93% in consistent with the previous study(14), and other olfactory tests in children, such as SCHOT (range 88%-91%)(15), NIH-Toolbox (range 48%-72%)(16), Dzaman et al. (62%-90%)(19), and van Spronsen et al. (58%-82%)(20).

There was a small but significant difference of the U-Sniff test scores between girls and boys, with girls scoring higher than boys in the present study. The sex differences in olfactory function were also found in the previous studies(12, 16, 20), while others could not observe a significant difference in olfactory performance between girls and boys(11, 14, 19). The different findings might be due to differences in age of the examined population. It can be concluded, that in cases of sex differences in odor identification performance, girls outperformed boys, but never the other way around. We also observed an age difference in the U-Sniff test, with children at age 7 or 8 years scoring higher than children at age 6, but not between children age 7 and 8 years. The results are consistent with the majority of previous studies, which showed that odor identification abilities increase with age in children(15, 13, 20, 21). During the development of children, odor learning and verbal development seem to be related to odor identification ability(17, 22, 23). However, the results were inconsistent with the previous study by Schriever and colleagues(16), who found no age difference in the U-Sniff test score. A possible reason for this discrepancy may be the different samples including children from different countries with a variety of cultural backgrounds.

The present study found a high test-retest reliability of the U-Sniff test in eight countries (r=0.71). With the exception of Paraguay (r=0.27), the test-retest reliability ranged from 0.61 to 0.80, showing similar or slightly higher reliability when compared to other olfactory tests in children (SCHOT [r = 0.98](15), NIH-Toolbox [r = 0.45](16), Smell Wheel [r = 0.70](14), and Sniffin' Kids [r = 0.44](15)). The reason why the test-retest reliability was low in Paraguay maybe that several children who scored 7-9 in phase 1 performed perfectly (scored 11-12) in phase 2, which resulted that the score range was narrow (from 10 to 12) in phase 2. Even though the test-retest reliability in Paraguay was low, children from Paraguay scored on average high on the odor identification test in both phases (phase 1: 11.04±1.12; phase 2: 11.18±0.77). The 10th percentile of the U-Sniff test score distribution as the criterion was used to discriminate between normal and reduced olfactory function in the present study. Across all countries, the 10th percentile on the U-Sniff test score was 8 points. For each country individually, the 10th percentile cut-off scores ranged between 7 and 10 points, which was similar to the previous study by Schriever and colleagues. Previous studies commonly used the 10th percentile as a cut-off to discriminate between normal and reduced olfactory function in adults(6, 7) and children(16, 17). However, the 10th percentile as a cut-off criterion might result in false-positive diagnosis of reduced olfactory function, because the frequency of reduced olfactory function in children is still unknown and might be lower than 10 percent(17). Therefore, the cut-off score at 10th percentile on the U-Sniff test for each country in the present study must be considered together with the whole clinical appearance of the child, such as medical history, including self-reported olfactory function, and other
Conclusions
In conclusion, the 12-item U-Sniff test for children has a high test-retest reliability in the eight countries investigated in the current study, and can be used globally, as now shown for a total of 27 countries.

Acknowledgements
We thank all the children who participated in this study, and thank Mr. Rui Jiang for his help in drawing the figures.

Supported by the National Natural Science Foundation of China (31700963 [to LQZ]); the Natural Science Foundation of Guangdong Province, China (2017A030310069; 2019A1515012135 [to LQZ]); the Medical Science and Technology Foundation of Guangdong Province, China (201906546 [to LQZ]); the Russian Science Foundation (16-1510312 [to VV]); the Deutsche Forschungsgemeinschaft (DFG HU411/18-1 [to TH]); and the Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden (MedDDrive) (to VS).

Authorship contribution
TH and VAS designed the study and formulated the research question. LQZ, AD, RA, SKK, MA, MAB, SB, BS, IB and VV conducted the data collection examination. LQZ, AD and VAS analyzed the data. LQZ wrote the first draft of the manuscript. LQZ, TH and VAS contributed to the interpretation and discussion of the results and commented on the draft. All authors have read and approved the final manuscript.

Conflict of interest
The authors declare no competing interests.

References