CLINICS

CLINICAL SCIENCE

A new cultural adaptation of the University of Pennsylvania Smell Identification Test

Marco Aurélio Fornazieri, Richard L Doty, Clayson Alan dos Santos, Fábio de Rezende Pinna, Thiago Freire Pinto Bezerra, Richard Louis Voegels

Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Otolaryngology, São Paulo/SP, Brazil.

OBJECTIVES: The University of Pennsylvania Smell Identification Test, a test of olfactory function that is widely used by otolaryngologists, geriatricians, and neurologists, has been translated into more than a dozen languages. In some instances, cultural and socioeconomic factors have necessitated changes in the odorant items or the response alternatives to make the test scores congruent with North American norms. The objective of this study was to compare the performance of Brazilian subjects on a new Portuguese language version of the University of Pennsylvania Smell Identification Test with their performance on an earlier Portuguese language version of the test, as well as to assess the influences of gender, age, ethnicity, and economic status on the test scores.

METHODS: Based on pilot data, several response alternatives of the earlier Portuguese language version of the test were altered in an effort to improve test performance. Forty-nine healthy Brazilian volunteers, who represented several economic classes, were tested. The test scores of the study cohort who received the newer version of the test were compared with those of a group of 25 subjects who received the earlier version of the test.

RESULTS: The mean score for the new version [35 (2.1)] was significantly ($p = 0.002$) higher than that for the earlier version [32.5 (3.5)]. Although no apparent influence of socioeconomic status was observed, the female participants outperformed the male participants in the current subject cohort.

CONCLUSION: The changes made in the new cultural adaptation of the Portuguese version of the University of Pennsylvania Smell Identification Test were effective in increasing the average test scores of the participants. Overall, the female subjects outperformed the male subjects on the test.

KEYWORDS: Diagnostic Tests; Human; Olfactory Nerve (I); Olfaction; Olfaction Disorders; Smell.

Fornazieri MA, Doty RL, Santos CA, Bezerra TF, Pinna FR, Voegels RL. A new cultural adaptation of the University of Pennsylvania Smell Identification Test. Clinics. 2013;68(1):65-68.

Received for publication on June 25, 2012; **First review completed on** July 2, 2012; **Accepted for publication on** September 25, 2012

E-mail: marcofornazieri@gmail.com

Tel.: 55 43 3029-1436

INTRODUCTION

Originally published only in English, the University of Pennsylvania Smell Identification Test (UPSIT; commercially known as the Smell Identification Test™, Sensonics, Inc., Haddon Hts., NJ) (1,2) has now been translated into over a dozen languages, including Portuguese. This widely used test, which is considered by many to be the gold standard to which other tests of olfactory function have been compared, is sensitive to the influences of a wide range of variables, including age (3), gender (4), environmental pollution (5), and numerous diseases (6). Interestingly, the American Academy of Neurology now recommends olfactory testing as an aid in diagnosing Parkinson's disease (7).

It is well established that cultural factors, which are sometimes quite subtle, can influence test scores on odor identification tests (8,9). In the case of the UPSIT, cultural and socioeconomic factors have necessitated changes in the odorant items or the response alternatives in a number of foreign-language versions to make the test scores more congruent with North American norms (2). For example, researchers in a pilot study in Taiwan made changes to several UPSIT odors and response alternatives, which markedly and effectively increased the test scores, although some items still caused the total test score to be lower than that observed in the U.S. population (10). In Australia, Mackay-Sim et al. (11), without changing any odors, substituted some alternatives and evaluated the performance of Australians on the test. Even with this adaptation, patients with normal smelling ability had average scores lower than those of the U.S. population. Therefore, those investigators suggested adding a correction factor of two points to the test scores if U.S. norms were to be used.

The cultural adaptation of the UPSIT to another language or culture is not a simple task. In addition to performing the translation and replacing the odors and the response alternatives that are unfamiliar to the local population, the validation and collection of normative data are required.

DOI: 10.6061/clinics/2013(01)OA10

Moreover, numerous factors must be considered, including the gender, age, health, education, and socioeconomic level of the subjects. Importantly, even within a culture, alterations in the viability of a given odorant item can change over time. For example, synthetic motor oil does not have the same odor as traditional motor oil, suggesting that lack of knowledge of this UPSIT odor may influence some norms based upon younger cohorts.

The present study examined the performance of healthy Brazilian subjects on a new Portuguese language version of the UPSIT. Specifically, scores from this new test were compared with those obtained in a pilot study using an earlier Portuguese version of the test (12). We examined whether individuals from different economic strata perform differently on the test. The ultimate goal of our ongoing program, of which this study is a part, is to develop a Brazilian Portuguese UPSIT with scores that can be directly classified on the basis of North American norms (2) and without a correction factor (11).

MATERIALS AND METHODS

Subjects

The study population was composed of two groups of Brazilians. The first group, which was administered an earlier Portuguese version of the UPSIT, consisted of 25 volunteers [13 men; 12 women; mean (SD) age, 32.4 (11.5) years] (12). The second group was composed of 49 Brazilian volunteers who represented a broader spectrum of ages, economic classes, and professions than the first group [21 men; 28 women; mean (SD) age, 30.4 (8.9) years]. Individuals were excluded from participation if they had a current upper respiratory infection or any history of head trauma or neurological or psychiatric disease. All the subjects reported that they believed they had normal olfactory function. The volunteers in both groups were divided into three socioeconomic classes: individuals earning a monthly income greater than $3,000 (Class A), individuals earning a monthly income between $1,201 and $3,000 (Class B), and individuals earning $1,200 or less per month (Class C). All the subjects provided informed written consent. The study was reviewed and approved by the institution's ethics committee under Protocol 0359/09. The two study groups did not differ significantly with respect to the demographic variables of age, gender, smoking behavior, ethnicity, and income (Table 1).

Table 1 - Study group demographics. All the *p*-values reflect uncorrected chi-square p values except for the *p*-value for age, which represents the *p*-value from an independent t-test.

Variable	Category	Group 1	Group 2	*p*-value
		N (%)	N (%)	
Gender	Male	13 (52%)	21 (42.9%)	
	Female	12 (48%)	28 (57.1%)	0.455
Age (years)		32.4±11.5	30.4±8.9	0.573
Smoking		1 (4%)	5 (10.2%)	0.355
Race	White	13 (52%)	32 (65.3%)	
	Nonwhite	12 (48%)	17 (34.7%)	0.267
Income	Class A	6 (24%)	12 (24.5%)	
	Class B	12 (48%)	16 (32.7%)	
	Class C	7 (28%)	21 (42.9%)	0.368

University of Pennsylvania Smell Identification Test

The UPSIT consists of four booklets, each with 10 pages. Microencapsulated ''scratch and sniff'' odorant strips are positioned on brown strips that are located at the bottom of each page, resulting in a total of 40 odorants (4). The subject releases the odor by scratching the strip with a pencil tip in a standardized manner. He or she then indicates the smell that is perceived by choosing a name from a set of four odor descriptors located just above the odorized strip (1). The number of correctly identified odors serves as the test score. A response is required for each odor even if no smell is perceived (i.e., the test is a forced-choice test). This procedure enables the detection of malingering based on improbable responses and increases the likelihood that a subject will pay close attention to the released odorant. The UPSIT is strongly correlated with odor threshold tests, and the magnitude of these correlations is limited by the reliability of the threshold test that is being evaluated (2,5).

Experimental procedures

In the earlier Portuguese test, we had replaced odors unfamiliar to certain people with odors more familiar to Brazilians. For example, the smell of root beer was replaced by the scent of a rubber tire. In the new version, we changed the response alternatives to words better known by the local population or to words that were less likely to be misconstrued as the target stimulus. For example, the correct answer for Item 22, namely, ''popcorn,'' was replaced by "rubber" because, to most of the subjects, the popcorn odor smelled more like rubber than popcorn. The other substitutions or changes that were made are shown in Table 2.

RESULTS

The changes in the response alternatives resulted in a higher overall mean UPSIT test score on the revised version of the test compared with the earlier version (respective means (SD), 35 (2.1) and 32.5 (3.5); t-test $p = 0.002$). The primary reason for this increased score was improved performance on the items whose response alternatives were changed (Table 2). Chi-square tests revealed that the percentage of correct answers was significantly higher for the test items that were modified, except for Items 32 and 34 (Table 3). To most of the subjects, Item 22 smelled more like rubber than popcorn; thus, providing a response alternative of ''rubber'' increased the rate of correct identification from 24 to 76%.

No significant correlations between the UPSIT scores on the revised version and either the subject's age or the time spent taking the test were present [respective Pearson r values (*p*-values) = 0.06 (0.681) and 0.17 (0.25)], nor where there any differences between the UPSIT scores with respect to social class [respective mean (SD) values for A, B, and C = 34.3 (2.5), 35.5 (1.8), and 34.9 (1.8); *p*-values>0.35). However, the women outperformed the men (respective male and female means (SDs) = 35.7 (2.1) and 33.9 (1.7); $p = 0.003$).

DISCUSSION

Compared with the previous version of the test, the revised Portuguese UPSIT described in this paper resulted in test scores that were closer to the North American norms. For the current version, the average test score was 35, which falls within the normal range, unlike the prior version's

Table 2 - Changes in the test alternatives in the earlier and newer UPSIT versions.

Item No.	Odor	Alternative	Earlier Version (English)	Earlier Version (Portuguese)	New Version (English)	New Version (Portuguese)
4	Cherry	A	Beer	Cerveja	Fish	Peixe
		B	Honey	Mel	Lemon	Limão
		C	Vanilla	Baunilha	Garlic	Alho
		D	Cherry	Cereja	Cherry	Cereja
19	Chocolate	A	Garlic	Alho	Garlic	Alho
	(Honey Bread)	B	Chocolate	Chocolate	Honey Bread	Pão de Mel
		C	Tire	Pneu	Tire	Pneu
		D	Pepper	Pimenta	Pepper	Pimenta
21	Flower	A	Flower	Flor	Perfume	Perfume
	(Perfume)	B	Spaghetti	Espaguete	Clove	Cravo
		C	Coconut	Coco	Gasoline	Gasolina
		D	Beer	Cerveja	Smoke	Fumaça
22	Rubber	A	Popcorn	Pipoca	Rubber	Borracha
	(Popcorn)	B	Soap	Sabão	Pineapple	Abacaxi
		C	Dog	Cachorro	Pizza	Pizza
		D	Spaghetti	Espaguete	Mint	Hortelã
24	Tire	A	Tire	Pneu	Tire	Pneu
		B	Watermelon	Melancia	Watermelon	Melancia
		C	Banana	Banana	Banana	Banana
		D	Smoke	Fumaça	Honey Bread	Pão de Mel
25	Pickles	A	Pineapple	Abacaxi	Pineapple	Abacaxi
		B	Cucumber	Pepino	Pickles	Picles
		C	Tire	Pneu	Watermelon	Melancia
		D	Pepper	Pimenta	Flower	Flor
32	Grass	A	Peppermint	Menta	Honey Bread	Pão de Mel
		B	Apple	Maçã	Apple	Maçã
		C	Grass	Grama	Grass	Grama
		D	Strawberry	Morango	Strawberry	Morango
34	Pine	A	Wood	Madeira	Wood	Madeira
		B	Smoke	Fumaça	Baby Powder	Talco de Bebê
		C	Flower	Flor	Bubble Gum	Chiclete
		D	Orange	Laranja	Grape	Uva
37	Soap	A	Soap	Sabão	Soap	Sabão
		B	Pepper	Pimenta	Pepper	Pimenta
		C	Baby Powder	Talco de Bebê	Orange	Laranja
		D	Peanut	Amendoim	Peanut	Amendoim

average score of 32. Nonetheless, the mean test score was below the average that was expected based on the North American normative data; the mean female score corresponded to the 16th percentile of scores of North American women in the same age group, and the mean male score corresponded to the 27th percentile of scores of men in the same age group (12). Assuming that the cohort of individuals tested in this study truly had normal olfactory function, additional modifications to the test appear to be necessary to increase the test scores to the levels observed in the North American normative data set. Alternatively, given the cultural changes that have occurred since the determination of the North American norms during the early 1980s, revised norms may be required for the North American test to equate the test scores. Indeed, a revision of the North American UPSIT norms is presently underway. Although the present study demonstrated significant improvements in most of the modified test items (Table 2), certain test items in the modified group could be refined further (e.g., grass and wood). Changes in these and other items that yield a lower percentage of correct responses would most likely adjust the test scores to values closer to the current North American norms.

Table 3 - Comparison between the percentages of correct answers according to group.

Item No.	Group 1 (n = 25)	Group 2 (n = 49)	
	%	%	Chi-Square *p*-value
4	60	92	0.002*
19	52	82	0.015*
21	36	90	<0.001*
22	24	76	<0.001*
24	68	100	<0.001*
25	36	76	0.002*
32	76	65	0.484
34	72	71	0.856
37	52	88	0.002*

One function of the UPSIT is to provide an olfactory diagnosis by comparing a patient's scores with the normal score from normative tables, adjusted for gender and age (15). Such age- and gender-adjusted comparisons are necessary because women tend to outperform men and olfactory function declines with age, with a lesser decline in women (3,15). With a culturally adapted UPSIT in which the average scores of the Brazilian population is similar and comparable to the American population, the physician can determine, without calculation, the patient's percentile rank relative to normal individuals of equivalent age and gender.

As in our previous study, no statistically significant differences in UPSIT scores between the various economic classes appeared. In an earlier Portuguese UPSIT study, Silveira-Moriyama et al. (13) observed a poorer performance of subjects who had a lower socioeconomic status and

suggested that this finding may have been due to the subjects' increased exposure to occupational dusts, their unfamiliarity with the test items, or their levels of education. The difference between this result and the results of the present study could be attributed to several factors. First, we used an improved version of the UPSIT, in which some of the items were presumably less difficult and did not require participants to make subtle distinctions. Second, our study group was significantly younger than the study group used in the study by Silveira-Moryama et al. (mean ages = 32.5 and 54.7 years, respectively). Third, our sample size was smaller than that of Silveira-Moriyama et al.'s study (n = 49 and 88, respectively); thus, it is possible that our statistical power to detect a socioeconomic effect was limited. Supporting this possibility, the regression analysis that was performed in the 1984 UPSIT development study, in which over 1,400 subjects were tested, revealed a significant, albeit weak, effect of education (1). Finally, the two studies' samples may have differed in other ways, although the subjects of both studies came from the same catchment area of the same hospital. Regardless of the reasons for these differences, it is clear that the present version of the Portuguese UPSIT is greatly improved over the earlier versions and provides test scores that closely approximate the North American norms.

The new cultural adaptation of the UPSIT in Portuguese was effective in increasing the average scores of volunteers compared with the average scores derived from an earlier Portuguese version of the test. Women outperformed men on this new version, as was the case in the earlier version. Revisions to the test are ongoing to ensure that Brazilian test scores fall within the UPSIT normative values that are obtained in North America.

Conflicts of Interests: Professor Richard L. Doty is President and a major shareholder of Sensonics, Inc., the manufacturer of the UPSIT.

ACKNOWLEDGMENTS

This work was supported by Programa de Apoio à Pós-Gradução of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

AUTHOR CONTRIBUTIONS

Fornazieri MA executed the research, Doty RL co-wrote the manuscript, Santos CA collected data, Bezerra TF and Pinna FR performed data analysis, and Voegels RL coordinated the study.

REFERENCES

1. 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(3):489-502, http://dx.doi.org/10.1016/0031-9384(84)90269-5.
2. Doty RL. The Smell Identification TestTM Administration Manual, Sensonics, Inc., Haddon Heights, New Jersey 1995.
3. Doty RL, Shaman P, Applebaum SL, Giberson R, Siksorski L, Rosenberg L. Smell identification ability: changes with age. Science. 1984;226(4681):1441-3, http://dx.doi.org/10.1126/science.6505700.
4. Doty RL, Applebaum SL, Zusho H, Settle RG. Sex differences in odor identification ability: a cross-cultural analysis. Neuropsychologia 1985;23(5):667-72.
5. Calderon-Garciduenas L, Franco-Lira M, Henriquez-Roldan C, Osnaya N, Gonzalez-Maciel A, Reynoso-Robles R, et al. Urban air pollution: influences on olfactory function and pathology in exposed children and young adults. Experimental and toxicologic pathology : official journal of the Gesellschaft fur Toxikologische Pathologie. 2010;62(1):91-102, http://dx.doi.org/10.1016/j.etp.2009.02.117.
6. Hawkes CH, Doty RL. The Neurology of Olfaction. Cambridge: Cambridge University Press. 2009, 244 pp.
7. Suchowersky O, Reich S, Perlmutter J, Zesiewicz T, Gronseth G, Weiner WJ, et al. Practice Parameter: Diagnosis and prognosis of new onset Parkinson disease (an evidence-based review) Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology. 2006;66(7):968-75, http://dx.doi.org/10.1212/01.wnl.0000215437.80053.d0.
8. Shu CH, Yuan BV, Lin SH, Lin SZ. Cross-Cultural application of the ''Sniffin' Sticks'' odor identification test. Am J Rhinol. 2007;21(5):570-3.
9. Doty RL, Marcus A, Lee WW. Development of the 12-item Cross-Cultural Smell Identification Test. Laryngoscope. 1996;106(3 Pt 1):353-6, http://dx.doi.org/10.1097/00005537-199603000-00021.
10. 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. American Journal of Rhinology and Allergy. 2010;24(1):45-50, http://dx.doi.org/10.2500/ajra.2010.24.3388.
11. Mackay-Sim A, Doty RL. The University of Pennsylvania Smell Identification Test: Normative adjustment for Australian subjects. Australian Journal of Oto-Laryngology. 2001;4:174-7.
12. Fornazieri MA, Pinna Fde R, Bezerra TF, Antunes MB, Voegels RL. Applicability of the University of Pennsylvania Smell Identification Test (SIT) in Brazilians: pilot study. Brazilian journal of otorhinolaryngology. 2010;76(6):695-9, http://dx.doi.org/10.1590/S1808-86942010000600004.
13. Silveira-Moriyama L, Azevedo AMS, Ranvaud, Barbosa ER, Doty RL, Lees AJ. Applying a new version of the Brazilian-Portuguese UPSIT smell test in Brazil. Arq Neuropsiquiatr. 2010;68(5):700-5, http://dx.doi.org/10.1590/S0004-282X2010000500005.
14. Doty RL. Studies of human olfaction from the University of Pennsylvania Smell and Taste Center. Chem Senses. 1997;22(5):565-86, http://dx.doi.org/10.1093/chemse/22.5.565.
15. Katotomichelakis M, Balatsouras D, Tripsianis G, Tsaroucha A, Homsioglou E, Danielides V. Normative Values of Olfactory Function Testing Using the 'Sniffin Sticks'. Laryngoscope. 2007;117(1):114-20, http://dx.doi.org/10.1097/01.mlg.0000246518.79894.7e.