

A comparison of sniff bottle staircase and olfactometer-based threshold tests

J. E. Hayes · A. L. Jinks · R. J. Stevenson

Published online: 27 July 2012
© Psychonomic Society, Inc. 2012

Abstract Olfactometers have been gaining popularity as research tools, but they have yet to replace established testing procedures in a variety of laboratory and clinical settings, including absolute threshold tests. In this research, we designed and operated a simple olfactometer with which to assess threshold. To do this, we used a method-of-adjustment test that was compared to the three-alternative forced choice ascending sniff bottle staircase method, which is currently a standard threshold test procedure. We found that the olfactometer threshold test correlated highly with the staircase method, and that it possessed suitable test–retest reliability. The advantages of the olfactometer threshold test include faster test time and reduced cleaning and reassembly demands. Future use of the olfactometer in olfactory identification and/or detection thresholds amongst odors is also outlined.

Keywords Threshold · Olfactometer · Method of adjustment · Staircase · Sniff bottle

Olfactometers, used to measure varying qualities of olfaction, have been steadily increasing in popularity and function (Johnson & Sobel, 2007). However, olfactometers have yet to replace traditional methods of assessing olfactory

thresholds. Olfactometers have, in general, been limited to laboratory work, but few attempts have been made to incorporate olfactometers in threshold measurement at all (Johnson & Sobel, 2007). The most popular olfactometer threshold test typically implements the method of ascending limits (Smeets et al., 2007), but as a method of psychophysics, it is longer and more complicated to implement than the popular and well-established staircase method (Doty, 1991b). As a result, olfactometer threshold measures have rarely been used in large population studies or widespread clinical applications, despite staircase procedures being somewhat laborious (Doty, 1991a; Johnson & Sobel, 2007).

There are several potential advantages to using olfactometers rather than sniff bottles. The use of the olfactometer not only expands the number of psychophysical methods, it also reduces the potential for researcher error. The olfactometer design also requires less material replacement and subsequent cost after the initial setup; the initial setup, however, is considerably more expensive than traditional threshold tests. The olfactometer only requires cleaning and refilling a single septum port sampling tube and replacing or cleaning inexpensive Teflon tubing, as opposed to current methods that require more extensive reassembly. For sniff bottles, the series of bottles need to be cleaned and serial dilutions remade; Sniffin' Sticks require the entire threshold set to be thrown away after a couple of years; and the booklets for the University of Pennsylvania Smell Inventory Tests are single use only (Doty, 1997).

The application of olfactometers also makes other psychophysical methods possible, including the method of adjustment. This method of adjustment involves participants adjusting a range of stimulus intensities to establish detection of the stimulus. The method of adjustment has, to the authors' knowledge, not yet been attempted as a measure of olfactory threshold, despite being a commonly used psychophysical

J. E. Hayes · R. J. Stevenson
Department of Psychology, Faculty of Human Sciences,
Macquarie University,
Sydney, Australia

A. L. Jinks (✉)
School of Social Science and Psychology,
University of Western Sydney,
Kingswood Campus,
Sydney, Australia
e-mail: a.jinks@uws.edu.au

procedure for assessing detectability (Berglund et al., 2006), as well as, in some instances, being a preferred psychophysical method for other sensory modalities (Rolland et al., 2002). It has also been documented as being significantly faster to run, making it a more efficient tool for large-population analyses (Hirsh & Watson, 1996). This method may also allow for greater precision, given that there is a sliding range of scores as opposed to a geometric series. In addition, the method of adjustment is also a more interactive (and, consequently, more interesting) measurement system for participants, gaining their attention.

In the present investigation, we will attempt to create a threshold test implemented via a simple and inexpensive computationally run olfactometer. To do this, we propose using the method of adjustment as the psychophysical method to test participant's olfactory thresholds and to compare these results to those from a standard staircase procedure. With this investigation we also endeavored to contribute to comparisons between different measurement procedures, for which there has been little previous research (Wise et al., 2008).

Method

Participants

A group of 14 healthy participants (average age=25 years, range=18–38 years, seven males and seven females) were tested. Retests were administered 2–4 weeks after the first test. This study was approved by the University of Western Sydney Human Research Ethics Committee.

Olfactometer design

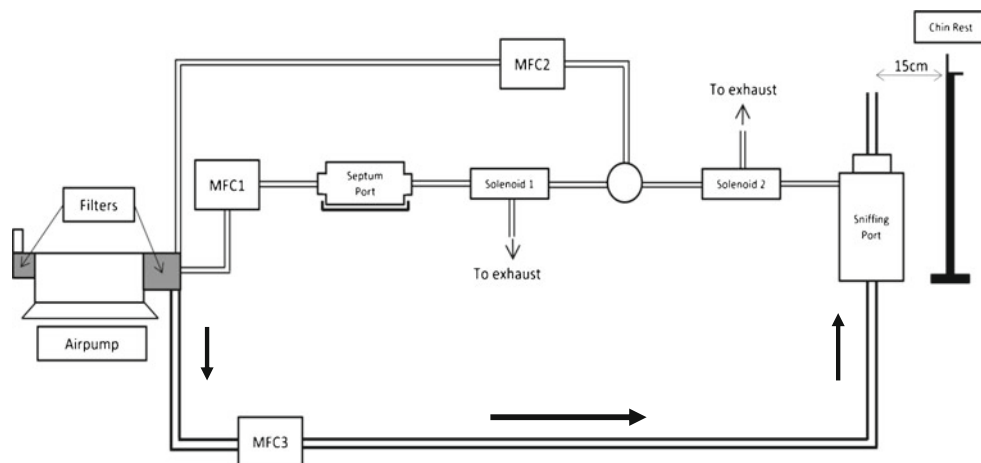
Our olfactometer design is simple and similar to several preexisting olfactometer setups (Davies & Davies, 1999; Johnson & Sobel, 2007; Lorig et al., 1999; Popp et al., 2004). The olfactometer is a “Dravnieks-style” olfactometer

(Fig. 1; Dravnieks, 1975). The maximum odorant concentration using this setup is 4 %, or .04 saturated vapor pressure. Filtered air from an air compressor (Thomas Industries, Inc., Sheboygan, WI) flows into three mass flow controllers (MFCs; Brooks Instrument, PA) via Teflon tubing (Dotmar Engineering, North Rocks, NSW, Australia). The olfactometer procedure is run via Opto 22 software and a SNAPPAC processor. The odorant saturator channel (MFC 1) air has a varied flow of 0–500 ml/min (accurate to 1 ml/min) through a 200×48×48 mm cylindrical glass septum port sampling tube containing *n*-butanol, from which it proceeds to a solenoid (Solenoid 1; General Valve Corp., Fairfield, NJ) that passes the flow either to exhaust or to a splitter that joins with the first air channel (MFC 2). The septum port is assessed to ensure that, at a maximum flow of 500 ml/min, 100 % odor saturation is always obtained, so as to allow us to calculate odorant proportions after air dilution (Dravnieks, 1975; Laing et al., 1994). The septum port is placed in a temperature-controlled water bath to ensure a steady-state temperature of 22 °C as well as humidity control, and this temperature is also maintained in the laboratory as a whole.

Before each method-of-adjustment test, Solenoid 1 directs airflow to exhaust in order to allow fresh air to purge any residual odor from the subsequent Teflon tubing for 15 s. In addition, this time allowed the MFCs to readjust to the new airflow volume, which has been established to take only a couple of seconds. Originally, the odorant saturator channel consisted of two separate MFCs at varied flow rates for a wider range of flow (Prah et al., 1995). However, extensive pilot testing indicated that the flow MFC flow rates varied from 0 to 100 ml/min, and at the lower flow rates odorants could not be detected by participants. Consequently, the previous design possessed no advantage over the current design.

MFC 2 is fixed at a flow rate of 300 ml/min and is used to mix and push the odorant stream toward the sniffing port. The combined flow of MFCs 1 and 2 then passes through another solenoid (Solenoid 2) which directs the flow to exhaust or onward to be combined further with the airflow

Fig. 1 [COMP: Align lines and arrows within parens. Cf. pdf p. 15.] Illustration of the olfactometer setup. Thicker lines (—) represent 1.5 cm-diameter Teflon tubing, whereas thinner lines (—) depict 0.5 cm-diameter Teflon tubing. Arrows (→) represent the direction of airflow, with the thickness of the arrow depicting the degree of relative airflow volume



channel from MFC 3. Immediately following the fresh air purge of Solenoid 1 to exhaust is followed by a purge of Solenoid 2, which removes the buildup of pressurized and odorized air in the septum port, as well as air flowing directly towards Solenoid 2. At the beginning of the next test, the airflow through Solenoid 2 is redirected and combined with airflow from Solenoid 3 at 11,700 ml/min. This further dilutes the odorized air as it passes through a tube directed into the headspace. The combined maximum total flow for the olfactometer is 12,500 ml/min.

Procedure

The olfactometer thresholds for *n*-butanol were compared with thresholds implemented via sniff bottles in a single-staircase, three-alternative forced choice method. The odor was presented with 18 dilutions of a geometric series, with a maximum 4 % odor concentration that was halved for each subsequent dilution. Each sniff bottle contained 20 ml of solution, with the solvent for dilution being water. The solvents and their concentrations corresponded to those in typical olfactory tests (Doty, 1991a, b; Doty & Kobal, 1995).

Three bottles were presented every 10 s within a 30 s time frame. The bottle was placed 15 cm under both nostrils and squeezed three times. Beginning at the lowest concentration, participants were required to identify which bottle of the three contained the odor, the remaining two bottles containing odorless blanks. An incorrect selection meant that the next bottle in the series contained a stronger odor concentration than that used in the previous presentation. Alternately, two successive correct selections lowered the odor concentration in the target bottle. The participant was blindfolded during the test, but not for a practice run for each series, in which only the highest concentration was used. Operationally, this staircase method is similar to the threshold measure for Sniffin' Sticks (Gudziol et al., 2006).

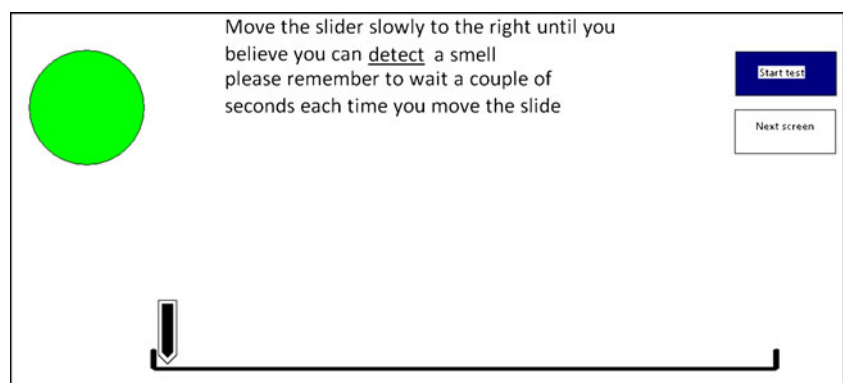
The olfactometer threshold test commenced 10 min after the staircase procedure. The participants were presented with instructions both on computer screen and by the researcher about how to conduct the experiment, including a no-odor trial run. Each method of adjustment test was displayed on screen (Fig. 2). In order to be unaffected by the odorized air purge, participants were instructed to sit away from the sniffing port when the “red light” icon was displayed, and an exhaust fan was situated above headspace at all times. When the “green light” icon was displayed, participants rested their heads on the chinrests, 15 cm from the sniffing port and began the test. The level of odor concentration was determined by the participants, by using their mouse to move a marker along a slide scale on the computer screen. The participants were encouraged to only move the marker in small, incremental steps, as indicated by test administrator, and once the odor was detected, to finely move the slider to the point of absolute threshold. Participants were instructed to wait a few seconds at each increment for the MFC to adjust to the new airflow setting. This process was repeated five times, and a score was derived as the mean average of the five results.

Both the sniff bottle threshold testing and the olfactometer threshold testing were timed, for later comparison. Both of the times included waiting times between trials.

Statistics

The results were run via SPSS PASW Statistics 18 (IBM). Bivariate regressions compared each participant's sniff bottle threshold results with the mean average threshold results from the olfactometer in order to assess similarities between the scores. Test–retest reliability was assessed via correlation analysis. Differences between the different procedures' times were analyzed via one-way ANOVA. The interrater reliability for both tests was measured via an intraclass correlation of the individual method-of-adjustment ratings.

Fig. 2 A screenshot of the Opto 22 software during a method-of-adjustment threshold test. The volunteer is required to slide the scale to the right (increasing the concentration of the odor) until an odor can be detected, at which stage an odorant concentration value is recorded for that test



Results

In this experiment, we found a strong correlation between the sniff bottle threshold scores and the averaged method-of-adjustment scores for participants, $r(12) = .778$, $p = .001$ (Fig. 3). This indicates that the scores derived from the method-of-adjustment test were comparable to the results derived from the standard threshold test procedure.

Test–retest reliability for the method-of-adjustment threshold test was assessed and found to be acceptable, $r(12) = .808$, $p < .001$. The staircase tests were also strongly correlated with each other, $r(12) = .694$, $p = .006$. Intraclass correlation was assessed throughout both testing phases and was moderately significant, $r(12) = .497$, $p < .001$. These three analyses suggest that the method-of-adjustment threshold test is both valid, as shown by its strong correlation to the staircase test, and reliable, due to repeated tests by participants conferring very similar results, as well as good interrater reliability.

The difference between the times taken for the two tests was significant, $F(1, 26) = 25.3$, $p < .001$ (see Fig. 4). On average, the method-of-adjustment test took 5 min, 55 s ($SD = 2$ min, 22 s), while the staircase method took 12 min, 53 s ($SD = 4$ min, 36 s).

Discussion

In this experiment, we hypothesized that a simple olfactometer design, incorporating a method-of-adjustment procedure, would produce olfactory threshold results comparable to those from more established methods. We found that this procedure was strongly correlated with a standard sniff bottle staircase method and possessed several distinct advantages.

To begin with, after the initial setup, very little cleaning or preparation is required when replacing the odorant, and the septum port can be easily detached. Teflon tubing is easy

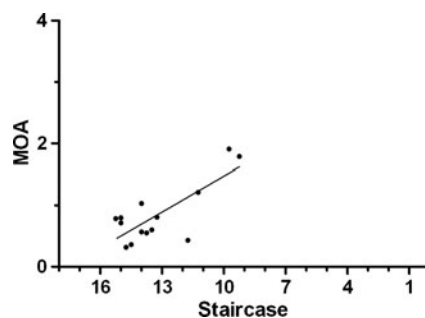


Fig. 3 Correlation between a method-of-adjustment (MOA) threshold test and a staircase threshold test. Sniff bottle concentrations (numbered across the x -axis) were presented to each individual in a staircase method during a time trial test. A strong correlation could be found between the sniff bottle threshold scores and the averaged MOA scores for the participants [$r(12) = .778$, $p = .001$]. The slope of the linear line of best fit is $slope = 0.12$, with a y -intercept of 2.13 ($r^2 = .61$)

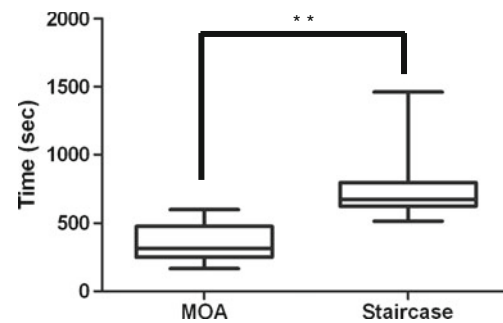


Fig. 4 A comparison of the times taken to complete an olfactometer method-of-adjustment (MOA) threshold test and a staircase threshold test. This boxplot compares the mean levels (the line in each box) of the time taken to complete both the MOA and staircase threshold tests across all volunteers ($n = 14$), with the whiskers representing the minimum and maximum times recorded for all volunteers during either test. The difference between the times taken for the tests was significant, $F(1, 26) = 25.3$, $p < .001$; on average, the MOA test took 355 ± 122 s, whereas the staircase method took 773 ± 276 s

to find and inexpensive, and it can also be cleaned thoroughly. The availability of all parts of the olfactometer means that it can be established easily and, with the exception of the MFCs and the electronic equipment, inexpensively (Johnson & Sobel, 2007).

The olfactometer threshold procedure is approximately twice as fast as the staircase method, which would allow the researcher to conduct further analyses in less time. For instance, thresholds for various odors could be tested and then comparatively assessed. The computational operation of the system also reduces the potential for researcher error. For instance, the standardized positioning of headspace may allow for more accurate results, as compared to the repeatedly approximated distances required by sniff bottles and other forms of olfactory testing.

The design of this olfactometer shares similarities with those in previous research. The designs implemented by Popp et al. (2004); Lorig et al. (1999) share many similarities, including the method by which odorant dilutions are calculated, as well as the use of multiple successive air streams to ensure odor dilution. In addition, Lorig et al. noted that the odor can be delivered to the participant in a variety of ways with little alteration required; in most previous research a face mask has been used, but modification to a sniff port into headspace would be trivial. The main differentiator between the basic designs of these olfactometers is that MFCs were used in our research, whereas both Lorig et al. and Popp et al. used flow meters to determine odor intensity; the constant modification of the odor dilution for threshold tests would probably make flow meters untenable. Johnson & Sobel's (2007) considerably more sophisticated olfactometer also implemented MFCs by means of calculating concentrations. However, they also commented that threshold testing using their olfactometer could be compromised by cross-odor contamination (Johnson & Sobel,

2007). The simplicity of the present setup means that some features of Johnson and Sobel's olfactometer were not available, but in the instance of threshold testing, multiple odorants could be readily integrated without cross-odor contamination by enlarging the design to contain only odor-specific airstreams. The cost of the MFCs and the associated electronic equipment is considerable; the complete setup, not including ventilation and temperature control, was approximately \$15,000 AU. This expense is for the initial startup, however; after that, maintenance and upkeep is comparable to (or less than) that for other olfactory kits.

The method of adjustment is sometimes considered a less accurate measure than other methods of psychophysics. However, some points of this particular implementation should be considered. To begin with, the results clearly demonstrate that the two procedures created very similar results. Second, threshold tests typically classify an individual's scores on the basis of a range of potential results. For example, many clinical trials classify participants' scores as groups within a range, such as anosmic or normosmic (Koskinen et al., 2004); the olfactometer design could readily implement these delineations, owing to its large range. Provided with proper instructions, participants are capable of making minute concentration changes to establish absolute thresholds. The participants' behavior during the test can be analyzed by looking at the time taken to undergo the test, the distances between odor concentration increases, and general observation by the researcher; thus, participants can be screened appropriately.

Future research should investigate different odorants. Variance in the chemicals tested will affect saturated vapor pressure as well as the time needed to purge the Teflon tubing between tests. However, provided that proper calibration is performed, theoretically any odorant can be easily tested via the olfactometer. Varying the slide scale ranges, or changing the design of the scale within a trial may offer more protection against antagonistic participants. The same slide scale could also be readily applied to odor identification ability: A participant's familiarity with an odor would be tested, and he or she would then be given a set of choices as to what odor was present. In this way, the relationship between threshold detecting and identification could be analyzed dynamically. Various odors could be assessed, allowing for detailed analyses of phenomena such as selective odor anosmias.

In summary, this research has shown that olfactometer-based threshold testing is tenable, and that there are several distinct advantages to using such a setup. The benefits of this setup include the speed and accuracy of the computer-based design, which eliminates many difficulties that occur when implementing other established methods. Olfactometers have enormous potential for widespread clinical application, and future investigation using these methods could produce more detailed and accurate analyses into a participant's olfactory capacity.

Author note The authors thank David Laing for his advice and assistance in the construction of the olfactometer and his direction of the study. We also thank Jeremy Barton for his assistance in designing the olfactometer program.

References

- Berglund, B., Lindvall, T., & Nordin, S. (2006). Environmentally induced changes in sensory sensitivities. *Annals of the New York Academy of Sciences*, 641, 304–321.
- Davies, C. W., & Davies, S. (1999). Prediction of olfactory response based on age, gender and smoking habits. *Journal of Medical Engineering & Technology*, 23, 73–76.
- Doty, R. L. (1991a). Olfactory system. In T. V. Getchell, R. L. Doty, L. M. Bartoshuk, & J. B. Snow Jr. (Eds.), *Smell and taste in health and disease* (pp. 175–203). New York, NY: Raven Press.
- Doty, R. L. (1991b). Psychophysical measurement of odor perception in humans. In D. G. Laing, R. L. Doty, & W. Breipohl (Eds.), *The human sense of smell* (pp. 95–134). Berlin, Germany: Springer.
- Doty, R. L. (1997). Studies of human olfaction from the University of Pennsylvania Smell and Taste Center. *Chemical Senses*, 22, 565–586.
- Doty, R. L., & Kobal, G. (1995). Current trends in the measurement of olfactory function. In R. L. Doty (Ed.), *Handbook of olfaction and gustation* (pp. 191–225). New York, NY: Dekker.
- Dravnieks, A. (1975). Instrumental aspects of olfactometry. In D. G. Moulton, A. Turk, & J. W. Johnston (Eds.), *Methods in olfactory research* (pp. 1–62). London, U.K.: Academic Press.
- Gudziol, V., Lötsch, J., Hähner, A., Zahnert, T., & Hummel, T. (2006). Clinical significance of results from olfactory testing. *Laryngoscope*, 116, 1858–1863. doi:10.1097/01.mlg.0000234915.51189.cb
- Hirsh, I., & Watson, C. (1996). Auditory psychophysics and perception. *Annual Review of Psychology*, 47, 461–484.
- Johnson, B., & Sobel, N. (2007). Methods for building an olfactometer with known concentration outcomes. *Journal of Neuroscience Methods*, 160, 231–245.
- Koskinen, S., Vento, S., Malmberg, H., & Tuorila, H. (2004). Correspondence between three olfactory tests and suprathreshold odor intensity ratings. *Acta Otolaryngologica*, 124, 1072–1077. doi:10.1080/00016480410015776
- Laing, D. G., Eddy, A., Francis, G. W., & Stephens, L. (1994). Evidence for the temporal processing of odor mixtures in humans. *Brain Research*, 651, 317–328.
- Lorig, T. S., Elmes, D. G., Zald, D. H., & Pardo, J. V. (1999). A computer-controlled olfactometer for fMRI and electrophysiological studies of olfaction. *Behavior Research Methods, Instruments, & Computers*, 31, 370–375. doi:10.3758/BF03207734
- Popp, R., Sommer, M., Müller, J., & Hajak, G. (2004). Olfactometry in fMRI studies: Odor presentation using nasal continuous positive airway pressure. *Acta Neurobiologiae Experimentalis*, 64, 171–176. doi:10.1080/00016480410015776
- Prah, J. D., Sears, S. B., & Walker, J. C. (1995). Modern approaches to air dilution olfactometry. In R. L. Doty (Ed.), *Handbook of olfaction and gustation* (pp. 227–256). New York, NY: Dekker.
- Rolland, J. P., Meyer, C., Arthur, K., & Rinalducci, E. (2002). Method of adjustments versus method of constant stimuli in the quantification of accuracy and precision of rendered depth in head-mounted displays. *Presence: Teleoperators and Virtual Environments*, 11, 610–625. doi:10.1162/105474602321050730
- Smeets, M. A., Bulsing, P. J., van Rooden, S., Steinmann, R., de Ru, J. A., Ogink, N. W., & Dalton, P. H. (2007). Odor and irritation thresholds for ammonia: A comparison between static and dynamic olfactometry. *Chemical Senses*, 32, 11–20. doi:10.1093/chemse/bjl031
- Wise, P., Bien, N., & Wysocki, C. (2008). Two rapid odor threshold methods compared to a modified method of constant stimuli. *Chemosensory Perception*, 1, 16–23.