UNIVERSITY OF CINCINNATI

Date: July 24, 2009

I, Jason M. Bailie,

hereby submit this original work as part of the requirements for the degree of:

Doctor of Philosophy

in Psychology

It is entitled:

The Influence of Odorant Intensity on Odor Identification in Older Adults

Student Signature: Jason M. Bailie

This work and its defense approved by:

Committee Chair: Robert A. Frank, Ph.D.

Paula K. Shear, Ph.D.

Steven R. Howe, Ph.D.

Approval of the electronic document:

I have reviewed the Thesis/Dissertation in its final electronic format and certify that it is an accurate copy of the document reviewed and approved by the committee.

Committee Chair signature: Robert A. Frank

The Influence of Odorant Intensity on Odor Identification in Older Adults

A dissertation submitted to the

Graduate School
of the University of Cincinnati

in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

in the Department of Psychology
of the College of Arts and Sciences

By

Jason M. Bailie

August 2009

B.A., San Diego State University
M.A., University of Cincinnati

Committee Chair: Robert A. Frank, Ph.D.
Committee Member: Paula K. Shear, Ph.D.
Committee Member: Steven R. Howe, Ph.D.

Abstract

Although it is well established that olfactory functioning diminishes with normal aging, the basis for this effect is not well understood. Research investigating age-related hyposmia has typically found that measures of odor identification and odor threshold (the two most common psychophysical methods of olfactory assessment) are only moderately correlated. Though this finding provides evidence that the tests are assessing a common source of variance, it also suggests confounding factors are influencing performance on these tests and thereby obscuring the basis for age-related loss of olfactory abilities. The present study investigated the impact of three factors on the relationship between measures of odor identification and threshold in older adults: the poor reliability of threshold tests, the influence of cognitive deficits on olfactory performance, and the lack of standardization of the intensity of odor identification tests. The study attempted to control for the influence of poor odor threshold reliability through incorporation of a four-test composite threshold estimate. The influence of odor intensity was investigated by developing an odor identification test that employed multiple intensity levels. Finally, performance was assessed on a broad battery of neuropsychological measures to control for the effects of cognitive impairments. It was predicted that the relationship between odor threshold and identification would be substantially improved as a result of a more accurate threshold measurement. Specifically, a composite threshold would be a better predictor of odor identification than an estimate based on a single test. Further, it was hypothesized that in addition to odor sensitivity (threshold), both odor intensity and cognitive functioning would be significant predictors of identification ability in an older adult sample. This would not be the case for a sample of young adults who were included as a control group. The study included 45 older adults and 40 young adults. The results of a generalized linear model revealed that odor threshold level, odor intensity, and cognitive functioning were all significant predictors of odor identification ability. Further, a predicted interaction was observed between threshold and odor intensity in which the ability of threshold to predict odor identification was strongest for low intensity odors. However, despite efforts to control confounding sources of variance, threshold at best accounted for only 27% of the variance in odor identification performance. The results suggest that diminished olfactory sensitivity is just one factor that explains older adults' performance on odor identification tests. This study suggests their performance is also impacted by the intensity of the odorants used in the identification test, an effect that needs to be taken into account when odor identification tests are employed to characterize age-related hyposmia or when performance on such a test is compared to other measures of olfaction. This study also brings to focus some methodological issues that need to be addressed in future research on age-related hyposmia, especially with regard to the use of computer technology in olfactory assessment.

Acknowledgments

First, I want to thank my mentor, Robert Frank, for all of his help over the years. He has been a truly gracious and motivational mentor providing not only excellent guidance in the development of my career but also serving as an exceptional role model. I would also like to express gratitude to my co-mentor, Paula Shear. She always provided the time to help me make the difficult decisions which in the end shaped my time at UC as well as my future. My thanks to Steve Howe who served as a terrific example to me over the years. I greatly appreciate him giving his time to serve on my dissertation. The technical and manual support needed for this project was tremendous and truly a group effort. Thus, I would like to extend my warmest appreciation to my colleagues at the Smell and Taste lab for all of their support and assistance.

Finally, on a personal note, this project marks the end of my graduate training. This was a very arduous journey but perhaps one of the most influential of my life. I could not have succeeded without the loving support of my family and friends; most importantly among those being my wife. She has been a guiding light for many years and had to endure so many sacrifices for me to have this accomplishment. I am forever indebted to her for the love and support she gave as well as being one of my most influential sources of inspiration.

Table of Contents

List of Tables and Figures

Chapter 1

Background and Significance

As with other sensory domains, olfactory functioning diminishes with normal aging. Older adults have been shown to have increased absolute thresholds for a variety of stimuli, decreased perceived strength of suprathreshold odors, a decline in their ability to name odors (odor identification), as well as deficits in odor memory (Gilbert, Pirogovsky, Ferdon, & Murphy, 2006; Murphy, 1995; Schiffman, 1997; Stevens & Cain, 1987). Odor-induced sniff suppression (the tendency to reduce sniffing behavior in the presence of an odor) has also been observed to be abnormal in older adults (Frank et al., 2006). The decline in olfactory functioning is believed to begin as early as the fourth decade of life becoming more pronounced and prevalent after the sixth decade with further progression thereafter (Cain et al., 1995; Doty, Shaman, & Dann, 1984; Frank et al., 2006; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). Recent estimates place the prevalence of olfactory impairment in older adults at 19.1 to 24.5% (Bramerson, Johansson, Ek, Nordin, & Bende, 2004; Murphy et al., 2002).

Aging is associated with dramatic pathologic changes in the olfactory bulb, specifically in the epithelium and mucosa (Bhatnagar, Kennedy, Baron, & Greenberg, 1987; Chen, Getchell, Sparks, & Getchell, 1993; Paik, Lehman, Seiden, Duncan, & Smith, 1992). In older adults, there is an overall decrease in the surface area of the olfactory epithelium (Paik, Lehman, Seiden, Duncan, & Smith, 1992). There is a decrease in innervations within the olfactory mucosa (Chen, Getchell, Sparks, & Getchell, 1993) as well as a decreased number of mitral cells within the olfactory bulb (Bhatnagar, Kennedy, Baron, & Greenberg, 1987). Furthermore, Kovacs and colleagues found that 86% of a non-demented older adult sample had neurofibrillary tangles and about one-third had amyloid plaques in the olfactory bulb (Kovacs, Cairns, & Lantos, 1999).

These findings strongly suggest that an important factor contributing to age-related olfactory deficits is dysfunction of the primary olfactory sensory organs.

Imaging studies also suggest that olfactory changes associated with aging are caused by abnormalities in early sensory processing. A PET study by Kareken and colleagues (2003) revealed the presence of age-related metabolic changes in the brain. In this study, the largest difference in brain activation between older adults and younger adults across multiple olfactory tasks was seen in the gyrus rectus/medial orbital gyrus (Kareken, Mosnik, Doty, Dzemidzic, & Hutchins, 2003). However, when Kareken's group statistically adjusted their analysis to account for behavioral measures of olfactory sensitivity (a proxy measure of early olfactory processing), the difference in metabolism between older adults and their younger counterparts was eliminated. This finding supports the idea that deficits in initial olfactory processing can affect how this information is analyzed in later olfactory projection areas.

Taken together, this information provides the foundation for a hypothesis that impairment in olfactory sensitivity is the primary dysfunction causing age-related hyposmia. Conceptualized in this way, age-related hyposmia would negatively impact performance on measures of odor threshold, odor discrimination, odor naming, and odor memory. As outlined above, this formulation of age-related hyposmia has strong empirical support and is consistent with the common-sense assumption that "complex" or "higher-level" cognitive processing of sensory information is reliant on normal sensory input.

If age-related loss of sensory input is the primary mechanism affecting how an older adult perceives the olfactory world, it would follow that there should be strong correlations between various measures of olfactory ability that are all presumably reliant on olfactory sensitivity (Doty, Smith, McKeown, & Raj, 1994). Unfortunately, behavioral data are not entirely consistent with this conclusion and suggest substantial amounts of variance between measures are not accounted for with traditional techniques.

Studies commonly report a moderately strong correlation between measures of odor identification and odor threshold in samples that include older adults. A notable exception compared performance on the University of Pennsylvania Smell Identification Test (UPSIT) to a threshold test and revealed a correlation of 0.89 (Doty, Shaman, & Dann, 1984). This was a small and relatively young sample (n = 64, M = 42.41, SD = 18.93) and the correlation was recalculated to 0.79 once a large number of anosmic outliers were removed. A later study by the same research group utilizing the same measures and a larger sample (n = 97 healthy adults, M = 45.84, SD = 20.17) revealed a less impressive correlation between the UPSIT and three different odor threshold techniques, $|r|$[1] ranging from 0.41 to 0.63 (Doty, Smith, McKeown, & Raj, 1994). Dulay and colleagues (2007) assessed odor identification using the UPSIT and an odor threshold test in 142 elderly participants (M = 77.1, SD = 8.5, range 56-93) and found a similar correlation, $|r|$ = 0.49 (Dulay, Gesteland, Shear, Ritchey, & Frank, 2007).

The Sniffin' Sticks olfactory battery is another widely used measure of olfaction that includes both tasks of odor threshold and odor identification. In this battery's normative sample (n = 104, mean age 49.5 years, SD = 18.5; age range 18-84 years) the identification and threshold tests correlated 0.54 (Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). This same sample was also assessed with the Connecticut Chemosensory Clinical Research Center Test (CCCRC), which is an olfactory battery consisting of an odor identification test and an odor threshold test.

[1] Pearson correlation coefficients are presented as absolute values. The scale of threshold tests are not consistent across studies which results in variations between positive and negative correlations between tests of threshold and identification.

On this battery, the correlation between odor identification and threshold was still weaker ($|r|$ = 0.29). Between battery correlations were also found to be a similar strength, $|r| = 0.24$ and 0.38. This is also consistent with a larger study (*n* = 2,076; age range 6 to 90 years; *M* = 35.2, *SD* = 16.2) that used the Sniffin' Sticks battery and reported a correlation of only 0.28 between threshold and identification tests (Lotsch, Reichmann, & Hummel, 2008).

The strength of correlations found across these studies suggests that diminished odor sensitivity only accounts for a portion of the variance in age-related changes in odor identification ability. However, as discussed below, additional sources of variance are most likely obscuring the relationship between measures of olfactory ability. The objective of this study was to better understand how much of the deficit in odor identification associated with aging is directly related to diminished odor sensitivity. It is predicted that a much stronger relationship between odor identification and odor threshold will emerge if confounding variables are taken into account.

Reliability of Odor Sensitivity Estimates

The poor reliability of odor threshold testing is one factor that limits the observed relationship between odor threshold and odor naming tests (Frank, Dulay, & Gesteland, 2003). The work by Punter (1983) speaks to the methodological limits of traditional techniques. He collected repeated threshold data on 11 different odorants. The median test-retest correlation between odors was only 0.40, suggesting that one threshold examination could only explain 16% of the variance in a repeat examination of the same group of participants (Punter, 1983). Similar test-retest correlations have been found by other groups using various techniques (Doty, McKeown, Lee, & Shaman, 1995; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). Even the most favorable studies report test-retest correlations in the 0.80 range, which indicate that threshold tests are associated with a good deal of error (Doty, McKeown, Lee, & Shaman, 1995; Stevens, Cruz, Marks, & Lakatos, 1998).

Given that the correlation between two variables is limited by the stability of the least reliable measure (Cronbach, 1970), it follows that the poor reliability of odor threshold tests reduces the observed correlation between itself and measures of odor naming. This was demonstrated empirically by Frank, Dulay, and Gesteland (2003), who statistically adjusted the correlation between threshold and identification for the measures' test-retest reliabilities and identified an increase in the observed correlation from 0.73 to 0.86.

Punter (1983) concluded that the poor reliability of odor threshold tests was related to the low number of trials employed in the typical assessment. For practical reasons, measures of odor threshold involve fewer trials than are commonly used in other modalities such as vision and audition. The number of trials is limited by the susceptibility of olfactory receptors to stimulus adaptation. In order to avoid adaptation or desensitization to the odor stimuli, there needs to be sufficient time between presentations, which limits the total number of trials that can be administered in a typical testing session. For example, the interstimulus interval used to avoid adaptation when measuring auditory threshold is about 300 to 1400ms (Ashihara, 2007; Lightfoot & Kennedy, 2006). This is dramatically shorter than the 30 second interstimulus interval recommended when assessing odor threshold (Hummel, Knecht, & Kobal, 1996). Test length is especially relevant in older adults, for whom increased test duration has been shown to negatively impact performance on measures of olfaction (Hulshoff Pol, Hijman, Baare, van Eekelen, & van Ree, 2000).

Studies investigating the reliability of olfactory tests have found that 98% of the variance can be accounted for by the length of the test (Doty, McKeown, Lee, & Shaman, 1995). A straightforward approach to increasing the number of trials and improving estimates of olfactory threshold is to increase the number of tests administered. Stevens and Cain (1987) provide data that aptly demonstrated this effect. They measured the odor threshold for three different compounds in a small sample of older adults and compared the threshold scores with the group's performance on an odor identification task. Consistent with other work, there was a moderate to strong correlation between the odor identification test and the individual threshold tests that ranged from 0.51 to 0.63. However, Stevens and Cain were able to improve the between test agreement by calculating a composite threshold composed of the average of the three individual tests. Using the average of the threshold tests improved the correlation to 0.76.

This work was expanded by Stevens and Dadarwala (1993). They had previously observed a surprising degree of overlap between individual threshold scores across groups of older and younger adults. Older adults often performed at a level more consistent with the younger generation and vice versa. This finding is counter-intuitive given the assumption that loss of odor sensitivity is a feature of normal aging. In contrast, when Stevens and Dadarwala (1993) used a composite threshold (averaging across eight separate thresholds tests), they were able to show perfect separation in threshold level between the young and old groups. Age accounted for a significantly greater amount of variance in odor threshold performance when a composite threshold as compared to only one threshold test was used (49% versus 15%). This effect has been replicated in other studies as well (Cain & Gent, 1991; Stevens, Cruz, Marks, & Lakatos, 1998)

Odor Identification Stimulus Intensity

Evidence for an age-related decline in perception of supra-threshold odor stimuli comes from several experiments employing magnitude estimation (Stevens, Bartoshuk, & Cain, 1984; Stevens & Cain, 1985; Stevens, Plantinga, & Cain, 1982). The results of these studies have consistently revealed that older adults judge super threshold odor intensities as weaker compared to similar judgments by younger adults across the dynamic range of odor intensity. If age-related hyposmia relates mostly to a loss of odor sensitivity, one might expect to observe an effect of odor intensity on odor judgments in hyposmic individuals; that is, an increasing of odor intensity may be expected to overcome in part the decrement in sensitivity.

Odor intensity is known to influence odor naming ability. De Wijk and Cain (1994) found that increasing odor concentration improved odor identification performance in older adults to a greater extent than that observed in young adults. Furthermore, Bailie (2006) was able to demonstrate a similar odor intensity effect in older patients with Parkinson's disease.

Despite the probable importance of intensity to performance, the intensities of odors incorporated in the two most common tests of odor identification vary. When the UPSIT was developed, odor intensity was considered in odor selection but, despite an omnibus test revealing a main effect of odor intensity, there was no attempt to match or manipulate intensity (Doty, Shaman, & Dann, 1984). The Sniffin' Sticks battery attempted to match odors for intensity but utilized a fairly liberal criterion, ± 25% mean subjective ratings (Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). The issue of intensity is important when comparing two olfactory measures. Bailie (2006) examined the relationship between odor identification and a novel measure of odor-induced sniff suppression. Previous work had shown that the two tests were correlating at a moderate strength (Frank, Dulay, & Gesteland, 2003) but there were a significant number of participants in which the test results between the two measures were inconsistent. By controlling for the intensity of the odor identification tests, Bailie (2006) found an increase in the agreement between the measures. Taken together, these findings demonstrate that odor intensity is an important variable to consider in the evaluation of olfactory test performance.

Cognitive Impairment

Fluctuations in the cognitive functioning of older adults influences their performance on odor identification tasks (Danthiir, Roberts, Pallier, & Stankov, 2001; Dulay, Gesteland, Shear, Ritchey, & Frank, 2007; Dulay & Murphy, 2002; Larsson, Finkel, & Pedersen, 2000; Larsson, Nilsson, Olofsson, & Nordin, 2004; Morgan, Nordin, & Murphy, 1995; Royall, Chiodo, Polk, & Jaramillo, 2002; Stevens, Cruz, Marks, & Lakatos, 1998; Westervelt, Ruffolo, & Tremont, 2005). Larsson and colleagues investigated the influence of cognition on age-related declines in odor identification in a group of 2,047 adults ranging in age from 45 to 90 (Larsson, Nilsson, Olofsson, & Nordin, 2004). They found that odor identification was moderately correlated to performance on measures of vocabulary and cognitive processing speed. They also found smaller but significant correlations with measures of semantic fluency, phonemic fluency and executive dysfunction. Dulay and colleagues used a more extensive test battery to investigate the same issue in older adults (Dulay, Gesteland, Shear, Ritchey, & Frank, 2007). They reported a significant correlation between odor identification and measures of language, new learning, working memory, and cognitive processing speed. Therefore, in addition to olfactory processes per se, performance on odor identification tasks is affected by the integrity of certain other cognitive abilities.

In contrast to odor identification tests, odor threshold tests are simplistic requiring a person to make a forced choice decision regarding the presence or absence of an odor and, thus, they are less cognitively demanding tasks. However, studies suggest that cognitive functioning influences performance on measures of odor threshold (Dulay & Murphy, 2002; Stevens, Cruz, Marks, & Lakatos, 1998), but the impact appears to be less pronounced than for measures of odor identification (Dulay, Gesteland, Shear, Ritchey, & Frank, 2007). Given this differential impact, it is possible that age-related cognitive decline has a more pronounced confounding effect on measures of odor identification than it does on measures of odor threshold. This differential impact may act to diminish the observed relationship between threshold and identification in that when assessing older adults with cognitive impairments, odor identification performance will be affected both by olfactory sensitivity and cognitive functioning, while the threshold test will be measuring odor sensitivity more purely.

Study Objective

It is clear that many older adults perform poorly on measures of both odor threshold and odor identification. Age-related loss of olfactory sensitivity provides a plausible explanation for both of these findings, but this theory suggests that there should be a strong correlation between measures of odor threshold and odor identification in older adults. The surprisingly low correlation between measures of odor sensitivity and odor naming seems to provide evidence against the age-related hyposmia hypothesis. However, factors such as an individual’s cognitive status and unreliable estimates of odor thresholds may be influencing the relationship between the olfactory measures.

The main goal of the current study was to rigorously evaluate the relationship between odor sensitivity and odor identification performance by controlling several factors that may have affected previous work assessing the interrelationship between these types of olfactory measures. The present study was designed to (1) improve estimates of odor threshold, (2) control for the effect of stimulus intensity on odor identification and (3) assess the impact of age-associated cognitive impairments on the olfactory tasks. This approach was employed in the hope of better understanding age-related olfactory dysfunction and, in particular, the role of sensory deficits in age-related hyposmia.

Adopting the approach of Stevens and Dadarwala (1993), this study attempted to improve the accuracy of odor threshold measurements by administering repeated threshold tests. Odor intensity was manipulated using a technique developed by Bailie (2006) in which odor identification was systematically assessed across three concentrations of odorants. This allowed us to determine the stability of the relationship between odor threshold and odor identification across intensities. Finally, measures of cognition that previously have been shown to be associated with measures of odor threshold and identification (Dulay et al. 2007) were administered to control for the effects of age-related cognitive impairments on the two measures of interest.

Hypotheses

It was predicted that in an older adult sample, the relationship between odor threshold and identification would be substantially improved with more accurate threshold measurement. Specifically, a composite threshold would be a better predictor of odor identification than estimates based on a single test. It was also predicted that the relationship between odor threshold and identification would be stronger at lower odor intensity levels because deficits in olfactory sensitivity would be overcome by stronger odors. Cognitive deficits are believed to mask the relationship between odor threshold and identification by impairing non-olfactory processes involved in odor naming (i.e., memory). Therefore, a significant relationship between neuropsychological tests and odor identification was predicted. Controlling for cognitive functioning was predicted to strengthen the relationship between odor identification and threshold.

Finally, it was hypothesized that in addition to odor sensitivity (threshold), both odor intensity and cognitive functioning would be significant predictors of identification ability in an older adult sample. This was not predicted to be true in a young adult control group. A young adult control group was included to verify any significant findings for this final hypothesis were age-related changes in performance opposed to unforeseen procedural complications. Further, a young adult group allowed for normative data to be gathered which was important for comparison purposes since that type of information was not available for either of the measures of olfaction.

Chapter 2

Research Design and Methodology

Participants

The study included 45 older adults and 40 younger adults. The older adult sample (OA) ranged in age from 56 to 95 (*M* = 75.76 years, *SD* = 10.30). They were 82.2% female and 93.3% White and had on average 13.79 (*SD* = 2.56) years of education. They were recruited from skilled nursing homes, independent living facilities, and senior citizen support groups in Cincinnati, OH. Participants were compensated $40 for their participation. Any person over the age of fifty-five was eligible to participate in the study. Participants were excluded if they failed to complete less than three of the odor threshold tests or any other measure. Fifty-four older adults were originally recruited. Seven participants were excluded because technical errors resulted in threshold data not being recorded for two of the four odor threshold tests in the study. Two others participants were unable to complete the neuropsychological tests, one because of visual deficits secondary to macular degeneration and the second because of severe deformity of the dominant hand.

The younger adult sample (YA) ranged in age from 18 to 28 years of age (*M* = 19.33 years, *SD* = 1.79) and had on average 12.55 (*SD* = 0.93) years of education. This group was recruited from undergraduate psychology courses at the University of Cincinnati, and they received course credit in exchange for their participation. The sample was 79.5% female and 84.1% White. Any student enrolled in a psychology course at the University of Cincinnati was eligible for participation. As with the older adults, participants were excluded if they failed to complete less than three of the odor threshold tests or any other measure. Originally, 49 participants were recruited into the YA sample but nine participants were excluded. Five of them did not complete the second session of the study. Another's odor threshold data for two tests were not recorded because of technical errors. Three others were excluded because they indicated that English was their second language, which may confound the verbal labeling required for odor identification performance (Frank, Dulay, Niergarth, & Gesteland, 2004).

Written informed consent was obtained from all participants, and all aspects of the study were approved by the Institutional Review Board at the University of Cincinnati.

Olfactory Tests

Odor Threshold: Four odor threshold tests for *n*-butanol were administered to maximize test reliability and reduce error, in accordance with recommendations from Stevens and Dadarwala (1993). Each test used a single-staircase, triple forced-choice procedure. They employed sixteen dilutions that were prepared in a geometric dilution series starting from 4% v/v n-butanol (dilution ratio 1:2; dilutent propylene glycol).

For each trial, three odor samples were presented in a randomized order, with two containing the solute and the third the odorant at a certain dilution. The participants' task was to identify the odor-containing sample. They were provided feedback on the correctness of their response on each trial. When a participant failed to identify a target, a higher concentration of the odorant was administered on the subsequent trial. A reversal of the staircase was triggered when the odor was correctly identified on two successive trials. Threshold was defined as the mean dilution step of the last three of five staircase reversal points. The average of the four threshold tests were used as the composite threshold.

Odors were delivered via an olfactometer. This presentation allowed for the stimuli to be administered with computer controlled timing and consistent airflow. Airflow was provided by a small air pump and regulated by a needle valve. For each trial, the air was filtered through a charcoal tube and then directed to an appropriate odor reservoir. Activation of the selected microvalve channeled airflow through the appropriate odor reservoir and then to the sampling port for 3 sec.

The test was administered with a computer program on a commercial laptop computer. A Visual Basic 6 program provided written and audio test instructions, triggered stimulus release via the olfactometer, and recorded participant responses. The user interacted with the program with a standard mouse. Each testing session included a practice trial, which the participant had to successfully complete prior to beginning the test items. Participants who had difficulty with the technology were provided assistance from a trained research assistant. In most cases, this involved assistance navigating the mouse efficiently.

Odor Identification: A modified version of the Multiple Intensity Odor Identification Test was used to measure odor naming ability (Bailie, 2006). The test uses odorants which are chemicals designed to smell like common odors one might come across in everyday living. These odorants are presented in multiple concentrations levels as a way of manipulating the intensity of the stimuli. Previous work verified that the concentration manipulation resulted in an increase in perceived intensity while maintaining quality necessary for appropriate identification in healthy adults (Bailie, 2006). Five of the original nine odors were selected for incorporation into the present study. The five odors had been observed to have the largest shift in intensity, based on subjective ratings, across changes in concentration. The five odorants were presented to participants at three intensity levels (Low, Medium, and High). Each odor was presented twice for a total of 30 trials (5 odors x 3 intensities x 2 presentations). Odors were presented in a randomized order.

All odorants were diluted in non-odorized solvents (either mineral oil or propylene glycol). The odors included Cherry (Benzaldehyde: 0.0005% v/v, 0.5% v/v, and undiluted); Garbage (Methylthiobutyrate: 0.0001% v/v, 0.001% v/v, and 0.1% v/v); Grass (Cis-3-Hexanol: 0.01% v/v, 0.1% v/v, and 10% v/v); Pineapple (Ethyl 3-Methyl Pentanoate: 0.001% v/v, 1% v/v, and 20% v/v); and Wintergreen (Methyl Salicylate: 0.01% v/v, 1% v/v, and 10% v/v).

Odors were presented in opaque bottles with approximately a 30 sec interstimulus interval. Order of administration was randomized to avoid order effects. The participants were given a sheet that listed each of the five odor names with accompanying pictures. The test was forced choice. After sniffing an odorant, they were asked to select the correct label for the stimulus and then to rate the intensity of the odor on a scale of 0 (no odor detected) to 10 (strongest odor imaginable). The dependent measure was the total number of odors correctly identified (ID Total) as well as the number correct at each intensity level (ID Low, ID Med, and ID High).

Cognitive Tests

The battery of tests used to control for the variability in cognitive functioning was adapted from Dulay and colleagues (2007). Additionally, the Montreal Cognitive Assessment (MoCA) was included as a measure of general cognitive functioning. All measurements reported from the cognitive tests are raw scores. Age-corrected standard scores were not used because the primary hypotheses were focused on how cognitive variables affect the relationship between threshold and identification in older adults; this includes age-related changes in performance found in typical older adults samples in addition to abnormal or pathological changes.

Short form of the Boston Naming Test - 2nd Edition (Kaplan, Goodglass, & Weintraub, 2001). The Boston Naming Test is a 15-item measure of confrontation naming. The participant is asked to identify the name of 15 line drawings of common objects. If the participant does not provide a response within 5 sec, or if their response indicates a misperception of the object, a semantic cue (i.e. "this is type of animal") is provided. The participant's performance is scored based on the total number of correct responses.

Category Fluency (Lezak, Howieson, & Loring, 2004; Rosen, 1980). Category Fluency is a measure of verbal fluency. The participants are given a semantic category and asked to generate as many words as they can within sixty seconds. Three trials with three different categories (animals, fruits, and vegetables) are completed. The total number of words generated across trials is the unit of analysis.

Short form of the California Verbal Learning Test -2nd Edition (Delis, Kramer, Kaplan, & Ober, 2000). The California Verbal Learning Test is a verbal list-learning test that comprises nine items. The list of words is read aloud to the participant. Once completed, the participant is asked to repeat back to the examiner as many words as possible from the list. Four list learning trials are completed consecutively. Performance is measured as the total number of words correctly repeated across trials.

Letter-Number Sequencing, *Digit Symbol-Coding* and *Symbol Search* subtests of the Wechsler Adult Intelligence Scale - 3rd Edition (Wechsler, 1997). The Letter Number Sequencing subtest is a measure of auditory working memory. The task requires the participant to listen to a string of letters and numbers. After the string is read aloud the participant repeats the information, mentally manipulating it so the numbers are repeated first in numeric order and then the letters repeated in alphabetic order. The recorded measure is the total number of strings correctly completed. The Digit Symbol Coding and Symbol Search subtests primarily measure mental and psychomotor processing speed. Both tasks are paper and pencil and incorporate a two-minute time limit. Symbol Search is a visual scanning task on which the participant searches an array of five symbols, looking for one of two target items. During the Digit Symbol Coding subtest, participants are presented with a series of boxes. Each box contains a symbol in the upper section and is blank in the lower section. A key matching symbols to specific numbers is presented at the top of the standard 8.5” x 11” sheet of paper. To complete the test, the participant fills in a specific number for each symbol in the blank spaces provided. The dependent measure for both the Digit Symbol Coding and Symbol Search subtests is the number of successfully completed items.

Trail Making Test, Part A (*Army Individual Test Battery*, 1944). Part A of The Trail Making Test is primarily a measure of mental processing speed. In this task, the participants are presented an arrangement of numbered dots on a sheet of paper. The participant draws a line connecting the dots in order from lowest to highest. The number of seconds to complete the task is recorded as the dependent measure.

Montreal Cognitive Assessment (Nasreddine et al., 2005). The MoCA was included as a supplement to the battery used by Dulay and colleagues (2007). The MoCA is a brief cognitive screener that was developed for detection of mild cognitive impairment in older adults. It consists of 13 brief tasks designed to assess visuospatial functioning, new learning, attention, language, abstraction, and general orientation. The MoCA takes approximately 10 minutes to administer. The test has good test-retest reliability with patient populations (Alzheimer’s disease and Parkinson’s disease) as well as healthy older adults (r = 0.92). The MoCA has been shown to reliably discriminate between normal controls, patients with mild cognitive impairment, and patients with Alzheimer’s disease. Sensitivity for mild cognitive impairment was 90% and Alzheimer’s disease was 100%, far superior to the Mini Mental Status Exam (Nasreddine et al., 2005). These psychometric properties make the MoCA attractive to researchers investigating age-associated hyposmia.

Testing Schedule

The battery of tests was divided into two one-hour sessions in order to avoid test-related fatigue. The sessions were completed approximately one week apart. The first session included the following order of administration: Threshold #1, MoCA, MIOID, and Threshold #2. The second session consisted of Threshold #3, the neuropsychological battery (Boston Naming Test, Category Fluency, California Verbal Learning Test, Letter Number Sequencing, Digit Symbol Coding, Symbol Search, Trail Making Test Part A), and Threshold #4.

Chapter 3

Results

Initial Analysis and Data Consolidation

The data were examined for outliers and missing values, as well as to verify assumptions of the statistical analyses. Explanations about how violations were managed are detailed below.

Composite Threshold: A composite threshold score was created by averaging all available threshold tests. For the majority of participants, this was four tests. However, due to technical errors (the participant completed the test but the data were not recorded by the computer program), two of the YAs and five OAs were missing one of the four threshold tests. In these instances, a composite threshold was calculated by averaging available data points. As mentioned in Chapter 2, participants who were missing more than one threshold test were excluded from the study.

To evaluate the utility of the composite threshold, the data for each of the OAs and YAs who completed all four threshold tests (OA: n = 40; YA: n = 38) were examined. For the following analyses, the first threshold test administered was identified as Threshold #1. Subsequent individual threshold tests were numbered consecutively. Threshold #1 was of particular interest for comparison purposes because it is most similar to the single test threshold estimates typically obtained in research and clinical situations (Cain, Gent, Catalanotto, & Goodspeed, 1983; Cain, Gent, Goodspeed, & Leonard, 1988; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997; Lotsch, Lange, & Hummel, 2004; Seiden, 1997).

The individual threshold tests were strongly correlated with each other, with a median correlation of 0.63 in the OA group (Table 1). The correlation was highest between Thresholds #3 and #4. This may reflect a reduction in test-related error where there is presumably less error in performance as the participants gain more experience with testing procedures; this phenomenon was observed in previous studies adopting a composite threshold (Cain & Gent, 1991; Stevens & Dadarwala, 1993) .

Table 1

Correlations Across Single Threshold Tests in the Older Adult Sample.

	Threshold #1	Threshold #2	Threshold #3	Threshold #4
Threshold #1	1			
Threshold #2	0.61	1		
Threshold #3	0.63	0.63	1	
Threshold #4	0.65	0.59	0.84	1

Notes. All correlations are significant at $p< 0.001$. The analysis only includes participants who completed all four threshold tests.

In the OAs, a four composite threshold (Threshold #1, #2, #3 & #4) resulted in only a slight reduction in the range of participant scores when compared to an individual threshold test. The range in threshold scores for the four threshold composite was 7.25 (*minimum* = 2.25, *maximum* = 9.50). This was somewhat smaller than the range based on a single threshold test estimate (Threshold #1; *range* = 9), a two test composite (Threshold #1 & #2; *range* = 9.83), or a three test composite (Threshold #1, #2, & #3; *range*; 7.44). The effect on the interquartile range was negligible. Figure 1 depicts the change in the median, range of scores, and interquartile range as the number of individual threshold tests used to calculate the composite increases.

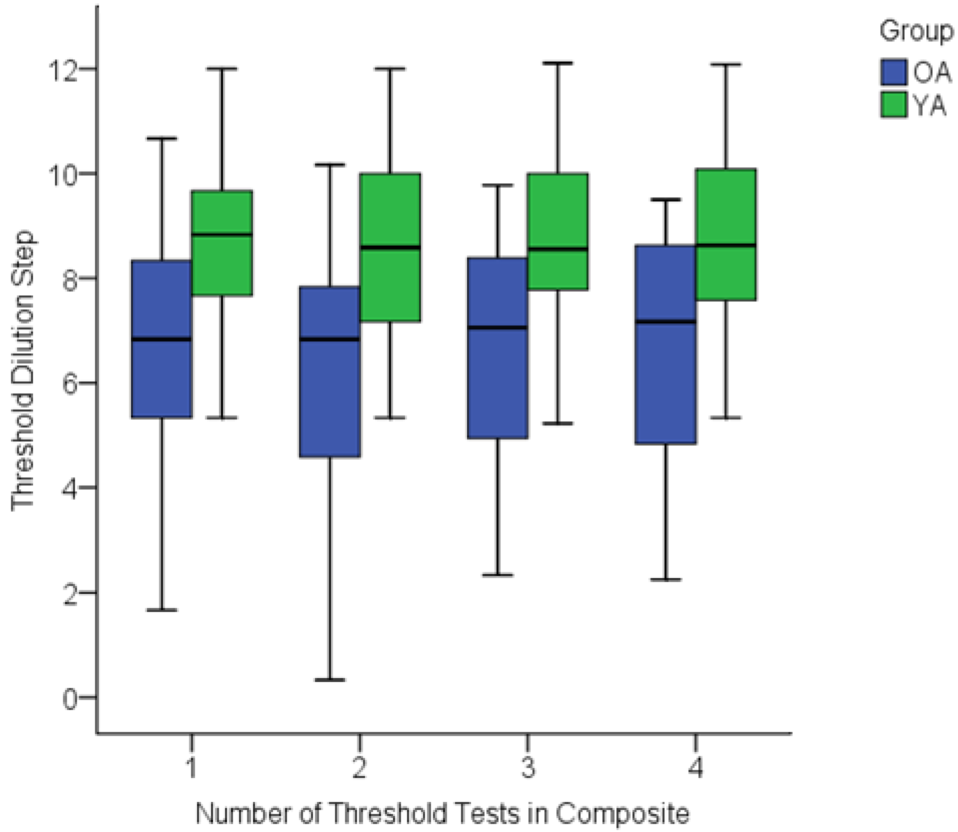

Figure 1. The cumulative effect of the number of tests used to calculate the composite threshold estimate on the spread of threshold scores and group differentiation. Y-axis represents the dilution step in the threshold (0 = 4% v/v n-butanol). The shaded box represents the interquartile range, the black bar is the median, and the whiskers represent the highest and lowest values.

Figure 1 also shows that increasing the number of tests used to calculate the composite threshold only resulted in minimal, if any, improvement in the differentiation between OAs and YAs, which is inconsistent with the findings of Stevens and Dadarwala (1993). On Threshold #1, the OA threshold scores ranged from 1.67 to 10.67 while YAs ranged from 5.33 to 12.00. The four threshold test composite had a similar overlap in performance; YAs ranged from 5.33 to 12.08 and OAs ranged from 2.25 to 9.50. Similarly, the effect size between OAs and YAs across various composites was stable (single threshold: *partial* η^2 = 0.20; two test composite: *partial* η^2 = 0.27; three test composite: *partial* η^2 = 0.26; four test composite: *partial* η^2 = 0.23).

The composite threshold had a moderately strong correlation with Total ID (r = 0.55) in the OAs (Appendix A). However, inconsistent with previous work (Cain & Gent, 1991; Stevens, Cruz, Marks, & Lakatos, 1998; Stevens & Dadarwala, 1993), the change in the strength of the correlation when a single threshold was substituted was not significant; the correlation dropped from 0.55 and 0.49, $t(42)$=0.95, p = 0.17.

Though the data collected do not permit analysis of test-retest reliability of the four test composite threshold, the data do speak to the test-retest stability of a one versus two test composite. The average test-retest reliability of the single test estimates was 0.67. This improved to 0.81 for a two test composite (Table 2).

Table 2

Test-Retest Reliability for Single Threshold Versus Two Threshold Composites in the Older Adult Sample.

	T1	T2	T3	T4	T1_2	T1_3	T1_4	T2_3	T2_4	T3_4
T1	1									
T2	0.61	1								
T3	0.63	0.63	1							
T4	0.65	0.59	0.84	1						
T1_2	0.90	0.90	0.70	0.69	1					
T1_3	0.88	0.68	0.92	0.83	0.87	1				
T1_4	0.90	0.66	0.81	0.92	0.87	0.94	1			
T2_3	0.69	0.88	0.92	0.80	0.87	0.90	0.82	1		
T2_4	0.71	0.88	0.83	0.90	0.88	0.86	0.89	0.94	1	
T3_4	0.67	0.63	0.96	0.96	0.73	0.92	0.90	0.90	0.90	1

Single Threshold: *M*= 0.67

Two Test Composite: *M*= 0.81

Notes. T = threshold; number represents order of threshold test administered; multiple numbers represent a composite threshold composed of the two tests. Note: All correlations are significant at p < o.oo1

Though arguable negligible, the results discussed above offer some support and do not contra-indicate the use of a four threshold composite estimate. Thus for the remainder of the analyses (unless otherwise noted), a four threshold test composite was utilized for both the OAs and YAs.

Intensity Manipulation: Consistent with past work (Bailie, 2006; de Wijk & Cain, 1994), manipulation of odorant concentration resulted in predictable changes in perceived intensity. A one within (odor concentration level), one between (OA and YA) repeated measures analysis of variance revealed a significant main effect of concentration, F (2, 82) = 122.49, $p < 0.001$, *partial* $\eta^2 = 0.75$. As shown in Figure 2, intensity ratings increased linearly with increases in concentration. The OAs rated the odors as significantly weaker then the YAs, F (1, 83) = 46.72, $p = 0.002$, *partial* $\eta^2 = 0.11$. However, this effect was dependent on the concentration of the odorants, F (2, 82) = 7.66, $p = 0.001$, *partial* $\eta^2 = 0.16$. Independent sample t-tests revealed that OAs rated the intensity of the odors as significantly weaker than the YAs for ID Med and ID High, $t(83) = 4.02$, $p < 0.001$ and $t(83) = 4.42$, $p < 0.001$, respectively. However, there was no difference between the groups for ID Low, $t(83) = 0.14$, $p = 0.89$.

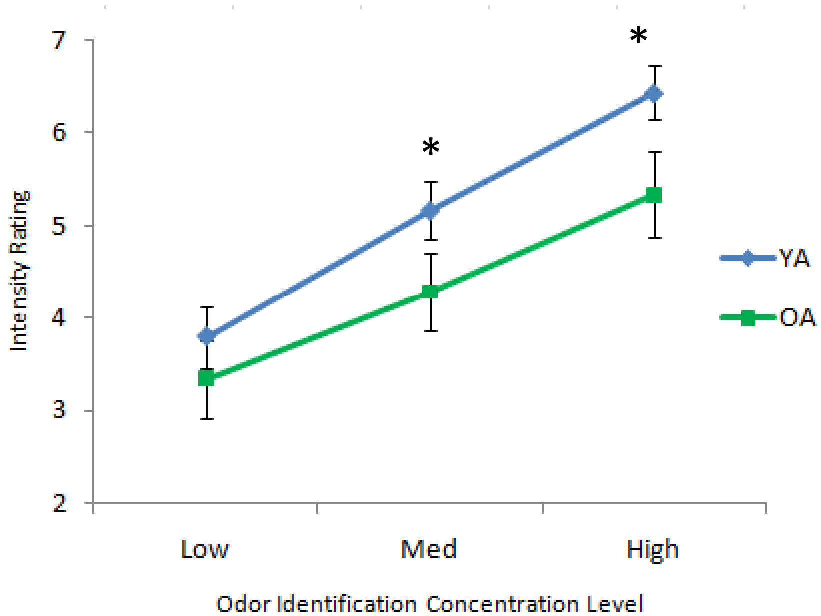

Figure 2. Mean intensity ratings by group across concentration level in the odor identification test. The error bars represent two standard errors of the mean. * represents a significant difference between group ratings, $p < 0.001$.

Cognitive Factors: Dulay and colleagues (2007) used structural equation modeling to identify cognitive factors which influence older adults' performance on measures of olfaction. Using a broad range of neuropsychological measures, they reported three cognitive factors: Verbal Retrieval, Working Memory and Processing Speed. The Verbal Retrieval factor was a summary of performance on the Boston Naming Test (2nd edition), the Category Fluency test, and California Verbal Learning Test (2nd edition). The Processing Speed variable incorporated Symbol Search, Digit Symbol Coding, and the Trail Making Test Part A. Letter Number Sequencing was the sole test in the Working Memory factor. This factor structure was used to guide factor analysis in the present study.

In the present study, the same cognitive battery was administered, and the test scores were consolidated according to the results of the Dulay et al. (2007) study. Both the OAs and

YAs were included in construction of the factors. Two of the factors (Verbal Retrieval and Processing Speed) were constructed using confirmatory factor analysis. Each factor was calculated separately, and the analysis was restricted to one factor using a principal component extraction. The Verbal Retrieval factor had an initial eigenvalue of 1.69 that explained 59.20% of the variance in scores on the Boston Naming Test (2nd edition), California Verbal Learning Test (2nd edition), and the Category Fluency test. The Processing Speed factor had an initial eigenvalue of 2.47, explaining 82.16% of the variance in scores on the Digit Symbol Coding, the Symbol Search, and the Trail Making Test Part A.

The third factor identified in the model by Dulay and colleagues (2007), Working Memory, was composed of only one variable, Letter Number Sequencing. So for the purpose of this study, performance on the Letter Number Sequencing task was transformed into a z-score across all participants. This allowed for ease of comparison with the other factors.

Group Differences

Descriptive statistics comparing OAs and YAs are presented in Table 3. As would be expected, a multivariate analysis of variance examining summary measures (composite threshold, ID Total, MoCA, Verbal Retrieval, Verbal Retrieval, Working Memory, and Processing Speed) revealed that older adults performed significantly worse than their younger counterparts across measures, F (6, 78) = 32.20, $p < 0.001$, *partial* $\eta^2 = 0.71$. The Processing Speed factor yielded the largest effect between groups. Measures of olfaction (threshold and ID Total), Working Memory, and general cognitive functioning (MoCA) had similar effect sizes between groups. A smaller but significant effect was found on the Verbal Retrieval factor. Group differences across the neuropsychological tests are presented in Appendix B.

Table 3

Performance by Group Across Olfactory and Neuropsychological Measures. M(SD)

	OA	YA	*F*	*p-value*	*partial η^2*
Threshold	6.17 (2.43)	8.80 (1.76)	31.98	<0.001	0.28
ID Total	18.69 (5.83)	23.48 (3.90)	30.28	<0.001	0.27
MoCA	23.73 (3.68)	28.23 (1.39)	52.94	<0.001	0.39
Verbal Retrieval	-0.36 (1.15)	0.41 (0.59)	14.77	<0.001	0.15
Working Memory	-0.55 (0.74)	0.89 (0.73)	44.60	<0.001	0.35
Processing Speed	-0.72 (0.73)	0.81 (0.53)	120.38	<0.001	0.59

Primary Hypotheses

A generalized linear model was used to address the main hypothesis which predicted that threshold, intensity, and cognitive factors were all significant predictors of the OAs' odor identification performance. For the purpose of this analysis, odor concentration was treated as a categorical variable with three levels (Low, Med, and High) rather than as a continuous variable. Since each participant was evaluated at each level of concentration, each participant generated three observations. The dependent variable was the number of odors correctly identified at the specific concentration level of the dummy variable. The distribution of the dependent variable was normal for the OA sample and an identity link function was used with a model-based covariance matrix. For the YA sample, identification was negatively skewed, so in the generalized linear model a gamma distribution was indicated with a log link function and a model-based covariance matrix. However, the distribution did not normalize with log transformation. Threshold and the cognitive factors (MoCA, Verbal Retrieval, Working Memory, and Processing Speed) were put into the model as covariates. The interaction between threshold and concentration level was also entered. Group (OA and YA) was not placed into a single model since the hypotheses focused on performance only in the OAs. The generalized linear model was conducted separately for the OA and YA samples. The YAs served primarily as a reference group to help determine that the findings were unique to age-related hyposmia. The model for the OAs was significant, $\chi^2(9) = 248.11$, $p < 0.001$. Consistent with predictions, the model for the YAs was not significant, $\chi^2(9) = 0.91$, $p = 1.00$. The results from the generalized linear models are presented in Table 4 and 5.

Table 4

Results of the Generalized Linear Model Predicting Odor Identification for the Older Adults

	Model Effect			Parameter Estimates		
	$\chi 2$	*df*	*p-value*	*b*	95% *CI*	*p-value*
Omnibus Test	248.11	9	<0.001			
(intercept)	5239.12	1	<0.001	4.51	2.47 to 6.55	<0.001
Threshold	149.91	1	<0.001	0.11	-0.02 to 0.24	0.089
Intensity	5.53	2	0.0631			
Low				-1.20	-2.34 to 0.06	0.039
Med				-1.06	-2.20 to 0.08	0.069
High				0[a]		
Threshold*Intensity	9.16	2	0.01025			
Low*Threshold				0.21	0.04 to 0.38	0.018
Med* Threshold				0.25	0.08 to 0.42	0.005
High*Threshold				0[a]		
MoCA	46.29	1	<0.001	0.08	0.02 to 0.14	0.013
Verbal Retrieval	18.61	1	<0.001	0.24	-0.02 to 0.50	0.067
Working Memory	0.01	1	0.909	-0.05	-0.14 to 0.04	0.269
Processing Speed	18.60	1	<0.001	0.74	0.40 to 1.07	<0.001

a. High intensity is the reference group

Table 5

Results of the Generalized Linear Model Predicting Odor Identification for the Younger Adults

	Model Effect			Parameter Estimates		
	χ2	*df*	*p-value*	*b* [a]	95% *CI*	*p-value*
Omnibus Test	0.91	9	1			
(intercept)	0.52	1	0.471	1.38	-2.84 to 5.60	0.52
Threshold	0.33	1	0.556	0.04	-0.14 to 0.22	0.68
Intensity	0.04	2	0.98			
Low				0.23	-2.10 to 2.55	0.85
Med				0.05	-2.27 to 2.37	0.97
High				0[b]		
Threshold*Intensity	0.05	2	0.98			
Low*Threshold				-0.02	-0.28 to 0.24	0.85
Med* Threshold				0	-0.26 to 0.26	0.99
High*Threshold				0[b]		
MoCA	0.02	1	0.88	0.1	-0.12 to 0.14	0.88
Verbal Retrieval	0.25	1	0.61	0.09	-0.25 to 0.43	0.61
Working Memory	0.04	1	0.85	0.01	-0.06 to 0.07	0.85
Processing Speed	0.32	1	0.57	-0.11	-0.50 to 0.28	0.58

a. Identification scores for the YAs were transformed thus *b*-weights should not be compared with the OAs.

b. High intensity is the reference group

Hypothesis 1: Threshold Effect. As expected, the generalized linear model revealed that threshold performance, as measured by the four threshold composite, was a significant predictor of odor identification in the OA sample, $\chi^2(1) = 149.91$, $p < 0.001$. However, the corresponding change was dependent on intensity level, which is discussed below.

The above model was also run using Threshold #1 in place of the composite threshold. Inconsistent with initial predictions, the inclusion of the composite threshold appears to have made little difference in the analysis. A single threshold measure (Threshold #1) was also a significant predictor when substituted into the OA model in place of the composite measure, $\chi^2(1) = 116.98$, $p < 0.001$. This is consistent with the insignificant change this same substitution had on the correlation between threshold and ID Total, which was discussed above.

Hypothesis 2: Intensity Effect. The main effect of intensity (odor concentration level) approached significance as a predictor of odor identification in the OAs, $\chi^2(1) = 5.53$, $p = 0.06$, and there was a significant interaction between threshold and intensity level, $\chi^2(2) = 9.16$, $p = 0.01$. This interaction was explored through examination of the correlations between the identification and threshold measures (See Appendix A). The strength of the correlation between identification and threshold was significantly weaker for ID High ($r = 0.36$) than ID Low ($r = 0.52$), $t(42) = 1.79$, $p = 0.04$. The correlation between ID Med ($r = 0.54$) and threshold was also significantly stronger than High, $t(42) = 2.05$, $p = 0.02$. The difference in correlation between threshold-ID Low and threshold-ID Med was not significant, $t(42) = 0.18$, $p = 0.43$. This pattern indicates that performances on the identification and threshold tests were most similar at low odorant concentrations.

The nature of the interaction between intensity and threshold is also depicted in Figure 3, which compares odor identification ability across intensity levels for the top and bottom 15th percentile of the OA sample on threshold. The bottom 15th percentile of the sample in terms of threshold (*n* = 7) had improved odor naming ability when intensity was increased. A pair-wise t-test comparing ID Low to ID High revealed that this improvement was statistically significant, $t(6) = -2.52$, $p = 0.045$. This was not the case for the top 15th percentile, whose change in performance was not statistically significant and more similar to that of the YAs.

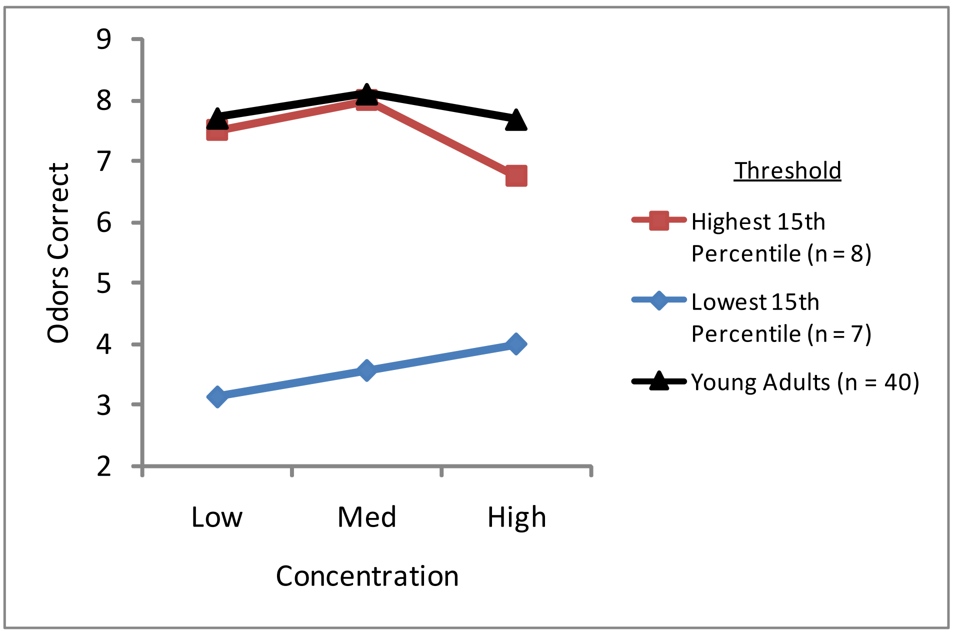

Figure 3. Comparison of top and bottom 15th percentiles of the OA sample on threshold for odor identification performance across intensity levels. Young adults are included for comparison purposes.

Hypothesis 3: Cognition. General cognitive functioning (as measured by the MoCA), Verbal Retrieval, and Processing Speed all emerged as significant covariates of odor identification in the OAs (Table 4 above). Working Memory was not significant, $\chi^2(1) = 0.43$, $p = 0.51$.

General cognitive functioning, verbal retrieval, and processing speed were all significantly correlated with Total ID in the OA sample (Appendix A). The strength of these effects varied between moderate and strong and was generally consistent across intensity levels.

Inconsistent with the initial hypothesis, there was a substantial drop in the strength of the correlation between threshold and Total ID when performance on the cognitive measures was controlled (Table 6). When controlling for the cognitive factors, the relationships between ID High and threshold was no longer significant. This drop may in part be related to the cognitive factors relationship with age. The cognitive variables and the olfactory variables are all related to age in the OAs (Appendix A). This may suggest that the controlling for the cognitive factors may have served to reduce shared age-related changes in the correlation between threshold and identification. This at best is only a small source of the change observed in Table 6 given a partial correlation between identification and threshold controlling for age resulted in only a modest reduction in the strength of the correlations (ID Low: $r = 0.47$, $p = 0.001$; ID Med: $r = 0.48$, $p < 0.001$; ID High: $r = 0.29$, $p = 0.03$) compared to what was observed when controlling for the cognitive factors (Appendix C).

Table 6

Partial Correlations Between Composite Threshold and Identification Low, Med, and High in the Older Adults

Variable Partialed	ID Low	ID Med	ID High
None	0.52**	0.54**	0.36**
MoCA	0.42**	0.45**	0.22
VR	0.34*	0.37**	0.16
PS	0.42**	0.44**	0.21
MoCA, VR, & PS	0.33*	0.36**	0.14

Notes. VR= Verbal Retrieval; PS= Processing Speed; * $p < 0.05$ (1-tailed); ** $p < 0.01$ (1-tailed).

Chapter 4

Discussion

Research investigating age-related hyposmia indicates that there is a primary loss of olfactory sensitivity associated with the aging process (Bhatnagar, Kennedy, Baron, & Greenberg, 1987; Chen, Getchell, Sparks, & Getchell, 1993; Frank, Dulay, Niergarth, & Gesteland, 2004; Gilbert, Pirogovsky, Ferdon, & Murphy, 2006; Kareken, Mosnik, Doty, Dzemidzic, & Hutchins, 2003; Murphy, 1995; Paik, Lehman, Seiden, Duncan, & Smith, 1992; Schiffman, 1997; Stevens & Cain, 1987). Decreased olfactory sensitivity is believed to be the primary factor that contributes to the impaired performance of older adults on measures of odor identification and odor threshold (Doty, Smith, McKeown, & Raj, 1994). However, the strength of the correlation between measures of identification and threshold across a variety of studies indicates that other factors may be affecting performance (Doty, Smith, McKeown, & Raj, 1994; Dulay, Gesteland, Shear, Ritchey, & Frank, 2007; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). The specific objective of this study was to evaluate the relationship between odor sensitivity (as measured by absolute thresholds) and odor identification in older adults, while controlling for the limited reliability of odor threshold estimates and the impact of age-associated cognitive impairments. Further, it was predicted that the strong relationship between odor sensitivity and identification would be observed when weaker odors were used in the identification test. The relationship would weaken when stronger odors were used in the identification test, because age-related sensitivity deficits would be overcome by stronger odors.

Reliability of Odor Sensitivity Estimates

Though odor threshold was a significant predictor of odor identification ability, the anticipated benefit of using a composite threshold to increase the reliability of threshold estimates was minimal. Based on previous work (Cain & Gent, 1991; Stevens & Cain, 1987; Stevens, Cruz, Marks, & Lakatos, 1998; Stevens & Dadarwala, 1993) it was predicted that using a composite threshold estimate in substitution of a single threshold test would reduce measurement error and increase the accuracy of threshold estimates. This would strengthen the observed correlation between threshold and the measure of odor identification. Inclusion of a four test composite threshold resulted in only a modest decrease in the variability of threshold scores in the OA sample. The range of the threshold estimates decreased by approximately one dilution step which is considerably lower than the five dilution step reduction found by Stevens and Dadarwala (1993). More germane to the primary focus of the present study, use of the composite threshold did not significantly change the strength of the relationship between threshold and identification nor did it substantially affect the value of threshold as a predictor of identification in a generalized linear model.

The inconsistency between this study, which revealed little or no benefit from a composite threshold, and past work documenting a dramatic effect (Cain & Gent, 1991; Stevens & Cain, 1987; Stevens, Cruz, Marks, & Lakatos, 1998; Stevens & Dadarwala, 1993) raises questions about differences in methodology. Past studies have utilized various approaches to the assessment of threshold and have shown the effect of the composite is consistent across odorants as well as threshold calculation and administration procedures (e.g., method of ascending limits versus staircase procedure). However, an important difference between the present study and past work lies is the odor delivery system. In all of the previous studies (Cain & Gent, 1991;

Stevens & Cain, 1987; Stevens, Cruz, Marks, & Lakatos, 1998; Stevens & Dadarwala, 1993), the odorant was placed in plastic “shampoo” bottles and released by the participant squeezing. This technique does not control for variability in pressure (i.e., how hard a person squeezes on a given trial) or the duration of the odorized air being released. The use of an olfactometer in the present study allowed for air pressure and duration of stimulus release to be standardized across trials as well as across participants. Using olfactometers in the assessment of odor sensitivity has been identified as an important variable that may reduce the error associated with olfactory threshold estimates (Dravnieks & Jarke, 1980; Punter, 1983).

This study also used a three alternative procedure (two blanks and one odorant) while the previous studies (Cain & Gent, 1991; Stevens & Cain, 1987; Stevens, Cruz, Marks, & Lakatos, 1998; Stevens & Dadarwala, 1993) used only two alternatives (one blank and one odorant). A three-alternative procedure is more consistent with what has been referred to as the “gold standard” of threshold assessment (Lotsch, Lange, & Hummel, 2004). A three alternative design reduces the probability of a participant correctly responding simply by chance and in so doing is assumed to improve measurement accuracy.

It is possible that the use of an olfactometer and a three-alternative procedure reduced measurement error to a sufficient level thereby diminishing the anticipated benefit of incorporating a composite threshold. The test-retest correlation in the OA sample ranged from 0.59 to 0.84. Though generally consistent with the literature (Doty, Smith, McKeown, & Raj, 1994; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997), the test-retest correlation is notably higher than the same correlation of “about r = .4” reported by Stevens and Dadarwala (1993). As those authors noted, “the reduction in variability that was effected by averaging the tests…could occur only because the test-retest correlation between any two tests is low" (Stevens and Dadarwala, 1993, p. 300).

The moderate correlation between Total ID and threshold that was found in the present study ($r = 0.55$) is comparable to the correlations noted throughout the literature (Doty, Smith, McKeown, & Raj, 1994; Dulay, Gesteland, Shear, Ritchey, & Frank, 2007; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997; Lotsch, Reichmann, & Hummel, 2008); though not as high as the correlation reported by Doty, Shaman, and Dann (1984). Despite techniques that should have improved threshold estimate accuracy (e.g., the use of a composite threshold, an olfactometer, and a three-alternative testing procedure), there was little impact on the relationship between the measure of identification and threshold.

Odor Identification Stimulus Intensity

As hypothesized, the generalized linear model revealed a significant interaction between odor intensity and threshold. As the intensity of the stimuli in the odor identification test was increased, threshold became significantly less useful in predicting performance. This is most likely a reflection of the identified pattern in which those OAs with impaired threshold scores benefitted from the high intensity odors, whereas those with normal detection scores did not change across intensities (and actually trended towards worse performance at the high intensity). This pattern is consistent with past work investigating the effect of odor intensity manipulation and identification performance (Bailie, 2006; de Wijk & Cain, 1994). As shown in Figure 3, the difference in identification performance between the OAs who had the most sensitive odor threshold scores and those with the lowest threshold scores (least sensitive) was greatest at the low and medium intensity levels and least for the high intensity odors. This reduction in the range of threshold scores is probably why the strength of the correlation between threshold and identification diminished with increasing odor intensity.

The interaction between intensity and threshold supports the conclusion that diminished sensitivity is an important factor in age-related hyposmia. It would follow that the insufficient standardization of odor intensity in traditional odor identification tests may have ramifications for the interpretation of studies investigating age-related hyposmia. For example, the moderate to strong correlation between threshold and identification documented in the literature may be an underestimation because the high intensity supra-threshold odors common in identification tasks decreases the tests ability to detect subtle variations in the olfactory sensitivity. Threshold tests are designed to estimate a specific transition point at which a person can reliably detect the weakest odor possible. This approach may allow for a detailed assessment of normosmic patients who are able to differentiate between weak and ephemeral odors. On the other hand, odor identification tests have been shown to have a ceiling effect (Minor, Wright, & Park, 2004). Odor identification tests are sensitive to olfactory deficits, especially in the elderly (Frank et al., 2006). However, odor identification tests have difficulty discriminating average from superior olfactory ability because most normosmic patients get the majority of the items correct (Minor, Wright, & Park, 2004).

Cognitive Impairment

The strength of the correlation between threshold and identification was predicted to increase if age-associated cognitive variables were controlled. Though general cognitive functioning (MoCA), Processing Speed and Verbal Retrieval were all significant predictors of odor identification ability, subsequent analysis revealed that controlling for cognitive functioning resulted in a decrease in the observed correlation between identification and threshold. This effect was consistent across intensities of odor identification stimuli and not entirely related to shared variance with age.

The initial hypothesis of this study was based on the assumption that cognitive impairments often found in older adults confound performance on odor identification tests to a greater extent than odor threshold tests. This was based on findings from Dulay and colleagues (2007). This group had compared performance on a threshold test and an odor identification test to a broad battery of neuropsychological measures in a sample of older adults. They found that the measure of odor threshold was not correlated with cognitive measures except for performance on the Trail Making Test Part A, a measure of mental processing speed (r = -0.26, p = <0.001) with some additional evidence which suggested a weak correlation with Letter Number Sequencing, a measure of working memory (r = 0.16, p = 0.06). In contrast, they found the odor identification test (UPSIT) was significantly correlated to performance on all of the neuropsychological measures.

Inconsistent with Dulay and colleagues, in the present study performance on the threshold task was significantly correlated with all of the neuropsychological measures with the exception of one (Boston Naming Test, a measure of language functioning). The strength of these correlations was generally comparable to the strength of the correlations between the odor identification test and the neuropsychological tests (see Appendix A). This is not an aberrant finding as other studies have reported that odor threshold performance is moderately correlated to cognitive functioning (Dulay & Murphy, 2002; Stevens, Cruz, Marks, & Lakatos, 1998). In the present study, it appears that both threshold and identification measured a common source of variance with the cognitive factors. When the cognitive factors were partialled out, that shared variance was removed which reduced the strength of the correlation between measures of identification and threshold.

There are two primary differences in the threshold techniques employed in this study and the methods used by Dulay and colleagues (2007) which may account for the increased relationship between threshold and the cognitive factors. First, Dulay and colleagues used a two alternative threshold procedure while this study incorporated three alternatives. Two and three alternative threshold procedures are both commonly used in studies of olfactory functioning (Cain, Gent, Goodspeed, & Leonard, 1988; Doty, 2000; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). However, it is possible that inclusion of the additional alternative contributed to an increase in the cognitive demand of the threshold test. Consistent with this hypothesis, threshold performance was significantly correlated to the Working Memory factor (Letter Number Sequencing), $r = 0.33$, which is larger than the same correlation reported by Dulay and colleagues (2007), $r = 0.16$.

Though the difference in alternatives may have had some influence on the increased cognitive demand associated with the threshold test, it seems more likely that the utilization of a computer interface was a critical difference. As mentioned in the Research Design and Methodology section, some of the participants in the older adult sample had difficulty operating the computer. Complications associated with the use of the computer had been anticipated. Both audio and written test instructions were provided to increase comprehension and the test was simplified so participants merely had to operate the mouse in a “point and click” fashion. Further, a practice trial was included and assistance was provided from a trained research assistant when needed. Despite these provisions, some older adults had increased difficulty navigating the mouse pointer and keeping pace with the program (e.g., they had trouble getting their nose to the odor sampling port at the appropriate time).

Researchers working in the area of computers and human behavior argue that psychometric equivalence should not be assumed when a test is converted from a manual to a computer format (Schulenberg & Yutrzenka, 2004; Williams & McCord, 2006). A person's attitude towards computers as well as their computer experience can influence success and motivation in patient-computer interactions (Schulenberg & Yutrzenka, 2004; Weber et al., 2003). For example, studies investigating the validity of computerized versions of neuropsychological tests have found poorer performance on the computerized versions compared to paper and pencil counterparts (Thompson, Ennis, Coffin, & Farman, 2007; Williams & McCord, 2006). This deficit is associated with decreased ability to concentrate during the computer test as well as the ease and comfort participants attribute to traditional formats (Thompson, Ennis, Coffin, & Farman, 2007). Successful patient-computer interactions have been associated with performance on measures of attention and visual pursuit tracking (Spinhoven, Labbe, & Rombouts, 1993; Weber, Fritze, Schneider, Kuhner, & Maurer, 2002). Furthermore, older adults have been shown to have increased computer anxiety that was associated with increased decision making time (Laguna & Babcock, 1997). In the present study, these factors may have resulted in the cognitively impaired participants doing worse on the threshold tests. This may account for the unexpected strength of the correlation between the threshold test and the neuropsychological tests.

Since the failure to observe a discrepancy between the relationship of the cognitive factors and the two olfactory tasks may be an anomaly associated with the current study's methodology, it is difficult to make any firm conclusions regarding the effect cognitive impairment may have on the relationship between threshold and identification tests in older adults. However, the results do bring to light issues that will need to be addressed as computer technology is incorporated into future approaches to chemosensory assessment.

Conclusions and Future Directions

The generalized linear model revealed that threshold and cognitive status were significant predictors of odor identification. The model revealed that the relationship between threshold and identification was contingent on the intensity of the odors in the odor identification test. Odor naming ability for low and medium intensity odors was more strongly related to a participant's odor sensitivity than odor naming for high intensity odors. This interaction between threshold and odor intensity suggests that the failure to take into account odor intensity in odor identification tests affects the assessment of age-associated hyposmia. The impact of measurement error in traditional threshold estimates and the influence of cognitive impairments were less clear in part because of methodological limitations.

It is noteworthy that, despite efforts to control confounding variables, threshold at best accounted for only 27% of the variance in odor identification performance. This is consistent with past work investigating this area as discussed above. The magnitude of this relationship is similar to what was noted in a recent study that examined sensory threshold changes associated with aging in other sensory modalities (Humes, Busey, Craig, & Kewley-Port, 2009). Hume and colleagues attempted to examine the relationship between sensory thresholds in older adults (n = 137; age ranged from 60 – 88) across stimulus frequencies, perceptual tasks (traditional threshold and gap detection), and modality (hearing, vision, touch). They instituted very strict methodological criteria. The results of their study indicated that correlation strength within a modality and within a task ranged from 0.4 to 0.7. Within modality but across tasks, which would be most akin to the relationship between threshold and identification in the present study, correlations were significantly weaker; the correlation strength ranged from -0.04 to 0.36. In light of these findings, the correlation strength between threshold and identification found in this study may represent a ceiling imposed by the two methodologies. The ceiling might be related to differences between the psychometrics of the two tests. For example, as mentioned above odor identification tests seem to have limitations when assessing average to above average olfactory abilities while threshold tests have limitations related to reliability. Alternatively, the ceiling may suggest that the tests are sensitive to different aspects of age-related brain changes. For example, a recent study showed that age related changes to the olfactory bulb differentially effected performance on measures of threshold and identification (Buschhuter et al., 2008).

Future studies of age-associated hyposmia will certainly attempt to improve existing methodologies by eliminating confounding variables. Other studies have attempted to improve existing techniques so that they are less influenced by confounding variables. Instruments such as the cross-cultural UPSIT (Doty, Marcus, & Lee, 1996) or the adaptive maximum likelihood procedure for olfactory threshold measurement (Linschoten & Harvey, 2004; Linschoten, Harvey, Eller, & Jafek, 2001) are good examples of these approaches. Others have pursued novel approaches to olfactory assessment such as odor-induced sniff suppression (Frank, Dulay, & Gesteland, 2003) or examining patterns of complex odor ratings (Hudry, Saoud, D'Amato, Dalery, & Royet, 2002; Royet et al., 2001). However, this study highlights the potential hurdles associated with the incorporation of new methodological approaches in psychophysical research (e.g., computers and olfactometers).

Adaptation of new techniques is necessary but must be done with caution. Appropriate studies must be conducted so psychometric properties are delineated and there is a clear understanding of how the new approaches compare to traditional measures. These studies must also address limitations specific to various patient populations such as older adults, cognitively impaired individuals, or people from different cultures.

This study highlights four specific areas of inquire. First, what effect does incorporation of an olfactometer have on the reliability of olfactory measures? Historically, technical hurdles have limited the use of an automated olfactometer but advancements are underway in hopes of increasing its applicability (e.g., see Osmic Enterprises at www.osmicenterpirses.com). It is assumed that adoption of more airflow and air pressure control results in improved accuracy. However, the similarity between the test-retest reliability in this study and that of past work calls this prediction into question. Further, work has been done demonstrating air flow and pressure are not necessarily important factors in odor identification (Laing, 1986). More specific empirical inquiry is needed to provide clear answers in regards to the advantage of this technology in psychophysical testing.

In a similar vein, what is the effect on the cognitive demands of the task when a test is converted into a computer format? Computerization improves standardization, eases test administration burden, and decreases cost (Weber et al., 2003). However, there is evidence that computerized testing may result in decreased test performance (Thompson, Ennis, Coffin, & Farman, 2007; Williams & McCord, 2006), a possible indication that computers increase the cognitive load of tests in certain populations. This issue must be addressed empirically as computers are used to “improve” tests of olfactory functioning and the study of age-associated hyposmia.

Third, future studies in age-related hyposmia may benefit from use of the MoCA to control for cognitive impairments in the elderly. As can be seen in Appendix A, the MoCA was strongly correlated with the cognitive factors previously identified as influential to measures of threshold and identification (Dulay, Gesteland, Shear, Ritchey, & Frank, 2007). Given its high sensitivity and short administration time, the MoCA may be a useful screening tool in future studies. This may be especially salient in large scale epidemiological studies where efficient use of time and resources is paramount.

Finally, future investigations into the primary cause of olfactory dysfunction in older adults may benefit from the use of a battery of olfactory measures. This study focused on two of the most common measures. However, there may be a limit to the ability of these tests to characterize the diverse nature of age-related hyposmia. This study emphasizes the importance of using multiple measures. The inclusion of both tests allowed for a much broader understanding of the limitations of each test individually. This is essential if the goal is to make sense of age-related changes to olfactory functioning. The use of multiple tests, as in an olfactory battery, has been shown to improve diagnostic sensitivity and there is evidence that different tests may measure distinct aspects of the olfactory system (Lotsch, Reichmann, & Hummel, 2008). The utility of a battery of tests may be expanded further through incorporation of novel approaches to the assessment of olfactory functioning (Frank, Dulay, & Gesteland, 2003; Hudry, Saoud, D'Amato, Dalery, & Royet, 2002; Royet et al., 2001). Studies that incorporate a broad range of olfactory tasks may be the most effective way of addressing barriers to the understanding of age-related olfactory dysfunction.

References

Army Individual Test Battery (1944). Washington, DC: War Department, Adjutant General's Office.

Ashihara, K. (2007). Hearing thresholds for pure tones above 16 kHz. *J Acoust Soc Am, 122*(3), EL52.

Bailie, J. M. (2006). *Clarifying the olfactory impairment in patients with Parkinson's disease.* Unpublished Master's, University of Cincinnati, Cincinnati.

Berendse, H. W., Booij, J., Francot, C. M., Bergmans, P. L., Hijman, R., Stoof, J. C., et al. (2001). Subclinical dopaminergic dysfunction in asymptomatic Parkinson's disease patients' relatives with a decreased sense of smell. *Ann Neurol, 50*(1), 34-41.

Bhatnagar, K. P., Kennedy, R. C., Baron, G., & Greenberg, R. A. (1987). Number of mitral cells and the bulb volume in the aging human olfactory bulb: a quantitative morphological study. *Anat Rec, 218*(1), 73-87.

Bramerson, A., Johansson, L., Ek, L., Nordin, S., & Bende, M. (2004). Prevalence of olfactory dysfunction: the skovde population-based study. *Laryngoscope, 114*(4), 733-737.

Buschhuter, D., Smitka, M., Puschmann, S., Gerber, J. C., Witt, M., Abolmaali, N. D., et al. (2008). Correlation between olfactory bulb volume and olfactory function. *Neuroimage, 42*(2), 498-502.

Cain, W. S., Gent, J., Catalanotto, F. A., & Goodspeed, R. B. (1983). Clinical evaluation of olfaction. *Am J Otolaryngol, 4*(4), 252-256.

Cain, W. S., & Gent, J. F. (1991). Olfactory sensitivity: reliability, generality, and association with aging. *J Exp Psychol Hum Percept Perform, 17*(2), 382-391.

Cain, W. S., Gent, J. F., Goodspeed, R. B., & Leonard, G. (1988). Evaluation of olfactory dysfunction in the Connecticut Chemosensory Clinical Research Center. *Laryngoscope, 98*(1), 83-88.

Cain, W. S., Stevens, J. C., Nickou, C. M., Giles, A., Johnston, I., & Garcia-Medina, M. R. (1995). Life-span development of odor identification, learning, and olfactory sensitivity. *Perception, 24*(12), 1457-1472.

Chen, Y., Getchell, T. V., Sparks, D. L., & Getchell, M. L. (1993). Patterns of adrenergic and peptidergic innervation in human olfactory mucosa: age-related trends. *J Comp Neurol, 334*(1), 104-116.

Cronbach, L. J. (1970). *Essentials of Psychological Testing* (3rd Ed.). NY: HarperCollins Publishers.

Danthiir, D., Roberts, R. D., Pallier, G., & Stankov, L. (2001). What the nose knows. Olfaction and cognitive abilities. *Intelligence, 29*, 337-361.

de Wijk, R. A., & Cain, W. S. (1994). Odor identification by name and by edibility: Life-span development and safety. *Human Factors, 36*, 182-187.

Delis, D., C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test, Second Edition, Adult Version*. San Antonio, TX: The Psychological Corporation.

Devanand, D. P., Michaels-Marston, K. S., Liu, X., Pelton, G. H., Padilla, M., Marder, K., et al. (2000). Olfactory deficits in patients with mild cognitive impairment predict Alzheimer's disease at follow-up. *Am J Psychiatry, 157*(9), 1399-1405.

Doty, R. L. (2000). *Odor Threshold Test Administration Manual*. Haddon Hts., NJ: Sensonics, Inc.

Doty, R. L., Marcus, A., & Lee, W. W. (1996). Development of the 12-item Cross-Cultural Smell Identification Test (CC-SIT). *Laryngoscope, 106*(3 Pt 1), 353-356.

Doty, R. L., McKeown, D. A., Lee, W. W., & Shaman, P. (1995). A study of the test-retest reliability of ten olfactory tests. *Chem Senses, 20*(6), 645-656.

Doty, R. L., Shaman, P., & Dann, M. (1984). Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. *Physiol Behav, 32*(3), 489-502.

Doty, R. L., Smith, R., McKeown, D. A., & Raj, J. (1994). Tests of human olfactory function: principal components analysis suggests that most measure a common source of variance. *Percept Psychophys*, *56*(6), 701-707.

Dravnieks, A., & Jarke, F. (1980). Odor Threshold Measurement by Dynamic Olfactometry: Significant Operational Variables. *Journal of the Air Pollution Control Association, 30*(12), 1284-1289.

Dulay, M. F., Gesteland, R. C., Shear, P. K., Ritchey, P. N., & Frank, R. A. (2007). Assessment of the influence of cognition and cognitive processing speed on three tests of olfaction. *J Clin Exp Neuropsychol*, 1-11.

Dulay, M. F., & Murphy, C. (2002). Olfactory acuity and cognitive function converge in older adulthood: support for the common cause hypothesis. *Psychol Aging, 17*(3), 392-404.

Frank, R. A., Dulay, M. F., & Gesteland, R. C. (2003). Assessment of the Sniff Magnitude Test as a clinical test of olfactory function. *Physiol Behav, 78*(2), 195-204.

Frank, R. A., Dulay, M. F., Niergarth, K. A., & Gesteland, R. C. (2004). A comparison of the sniff magnitude test and the University of Pennsylvania Smell Identification Test in children and nonnative English speakers. *Physiol Behav, 81*(3), 475-480.

Frank, R. A., Gesteland, R. C., Bailie, J., Rybalsky, K., Seiden, A., & Dulay, M. F. (2006). Characterization of the sniff magnitude test. *Arch Otolaryngol Head Neck Surg, 132*(5), 532-536.

Gilbert, P. E., Pirogovsky, E., Ferdon, S., & Murphy, C. (2006). The effects of normal aging on source memory for odors. *J Gerontol B Psychol Sci Soc Sci, 61*(1), P58-60.

Hofer, S. M., Berg, S., & Era, P. (2003). Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance, and cognitive functioning. *Psychol Aging, 18*(2), 285-305.

Hofer, S. M., Flaherty, B. P., & Hoffman, L. (2006). Cross-sectional analysis of time-dependent data: Problems of mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohorts. *Multivariate Behavioral Research, 41*(165-187).

Hudry, J., Saoud, M., D'Amato, T., Dalery, J., & Royet, J. P. (2002). Ratings of different olfactory judgments in schizophrenia. *Chem Senses, 27*(5), 407-416.

Hulshoff Pol, H. E., Hijman, R., Baare, W. F., van Eekelen, S., & van Ree, J. M. (2000). Odor discrimination and task duration in young and older adults. *Chem Senses, 25*(4), 461-464.

Humes, L. E., Busey, T. A., Craig, J. C., & Kewley-Port, D. (2009). The effects of age on sensory thresholds and temporal gap detection in hearing, vision, and touch. *Atten Percept Psychophys, 71*(4), 860-871.

Hummel, T., Knecht, M., & Kobal, G. (1996). Peripherally obtained electrophysiological responses to olfactory stimulation in man: electro-olfactograms exhibit a smaller degree of desensitization compared with subjective intensity estimates. *Brain Res, 717*(1-2), 160-164.

Hummel, T., Sekinger, B., Wolf, S. R., Pauli, E., & Kobal, G. (1997). 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem Senses, 22*(1), 39-52.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test, Second Edition.* Philadelphia, PA: Lippincott Williams & Wilkins.

Kareken, D. A., Mosnik, D. M., Doty, R. L., Dzemidzic, M., & Hutchins, G. D. (2003). Functional anatomy of human odor sensation, discrimination, and identification in health and aging. *Neuropsychology, 17*(3), 482-495.

Kovacs, T., Cairns, N. J., & Lantos, P. L. (1999). Beta-amyloid deposition and neurofibrillary tangle formation in the olfactory bulb in ageing and Alzheimer's disease. *Neuropathol Appl Neurobiol, 25*(6), 481-491.

Laguna, K., & Babcock, R. L. (1997). Computer anxiety in young and older adults: Implications for human-computer interactions in older populations. *Computers in Human Behavior, 13*(3), 317-326.

Laing, D. G. (1986). Identification of single dissimilar odors is achieved by humans with a single sniff. *Physiol Behav, 37*(1), 163-170.

Larsson, M., Finkel, D., & Pedersen, N. L. (2000). Odor identification: influences of age, gender, cognition, and personality. *J Gerontol B Psychol Sci Soc Sci, 55*(5), P304-310.

Larsson, M., Nilsson, L. G., Olofsson, J. K., & Nordin, S. (2004). Demographic and cognitive predictors of cued odor identification: evidence from a population-based study. *Chem Senses, 29*(6), 547-554.

Lezak, M., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment* (4th Ed.). New York: Oxford University Press.

Lightfoot, G., & Kennedy, V. (2006). Cortical electric response audiometry hearing threshold estimation: accuracy, speed, and the effects of stimulus presentation features. *Ear Hear, 27*(5), 443-456.

Linschoten, M. R., & Harvey, L. O., Jr. (2004). Detecting malingerers by means of response-sequence analysis. *Percept Psychophys, 66*(7), 1190-1201.

Linschoten, M. R., Harvey, L. O., Jr., Eller, P. M., & Jafek, B. W. (2001). Fast and accurate measurement of taste and smell thresholds using a maximum-likelihood adaptive staircase procedure. *Percept Psychophys, 63*(8), 1330-1347.

Lotsch, J., Lange, C., & Hummel, T. (2004). A simple and reliable method for clinical assessment of odor thresholds. *Chem Senses, 29*(4), 311-317.

Lotsch, J., Reichmann, H., & Hummel, T. (2008). Different odor tests contribute differently to the evaluation of olfactory loss. *Chem Senses, 33*(1), 17-21.

Mesholam, R. I., Moberg, P. J., Mahr, R. N., & Doty, R. L. (1998). Olfaction in neurodegenerative disease: a meta-analysis of olfactory functioning in Alzheimer's and Parkinson's diseases. *Arch Neurol, 55*(1), 84-90.

Minor, K. L., Wright, B. D., & Park, S. (2004). The Smell Identification Test as a measure of olfactory identification ability in schizophrenia and healthy populations: a Rasch psychometric study. *J Abnorm Psychol, 113*(2), 207-216.

Morgan, C. D., Nordin, S., & Murphy, C. (1995). Odor identification as an early marker for Alzheimer's disease: impact of lexical functioning and detection sensitivity. *J Clin Exp Neuropsychol, 17*(5), 793-803.

Murphy, C. (1995). Age-Associated Differences in Memory for Odors. In F. R. Schab & R. G. Crowder (Eds.), *Memory for Odors* (pp. 109-132). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Murphy, C., Schubert, C. R., Cruickshanks, K. J., Klein, B. E., Klein, R., & Nondahl, D. M. (2002). Prevalence of olfactory impairment in older adults. *Jama, 288*(18), 2307-2312.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc, 53*(4), 695-699.

Paik, S. I., Lehman, M. N., Seiden, A. M., Duncan, H. J., & Smith, D. V. (1992). Human olfactory biopsy. The influence of age and receptor distribution. *Arch Otolaryngol Head Neck Surg, 118*(7), 731-738.

Ponsen, M. M., Stoffers, D., Booij, J., van Eck-Smit, B. L., Wolters, E., & Berendse, H. W. (2004). Idiopathic hyposmia as a preclinical sign of Parkinson's disease. *Ann Neurol, 56*(2), 173-181.

Punter, P. H. (1983). Measurement of human olfactory thresholds for several groups of structurally related compounds. *Chemical Senses, 7*(3/4), 215-233.

Rosen, W. G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology, 2*, 135-146.

Royall, D. R., Chiodo, L. K., Polk, M. S., & Jaramillo, C. J. (2002). Severe dysosmia is specifically associated with Alzheimer-like memory deficits in nondemented elderly retirees. *Neuroepidemiology, 21*(2), 68-73.

Royet, J. P., Croisile, B., Williamson-Vasta, R., Hibert, O., Serclerat, D., & Guerin, J. (2001). Rating of different olfactory judgments in Alzheimer's disease. *Chem Senses, 26*(4), 409-417.

Schiffman, S. S. (1997). Taste and smell losses in normal aging and disease. *Jama, 278*(16), 1357-1362.

Schulenberg, S. E., & Yutrzenka, B. A. (2004). Ethical issues in the use of computerized assessment. *Computers in Human Behavior, 20*, 477-490.

Seiden, A. M. (1997). The initial assessment of patients with taste and smell disorders. In A. M. Seiden (Ed.), *Taste and Smell Disorders*. New York: Thieme.

Serby, M., Larson, P., & Kalkstein, D. (1991). The nature and course of olfactory deficits in Alzheimer's disease. *Am J Psychiatry, 148*(3), 357-360.

Spinhoven, P., Labbe, M. R., & Rombouts, R. (1993). Feasibility of computerized psychological testing with psychiatric outpatients. *J Clin Psychol, 49*(3), 440-447.

Stevens, J. C., Bartoshuk, L. M., & Cain, W. S. (1984). Chemical senses and aging: Taste versus smell. *Chem Senses, 9*, 167-179.

Stevens, J. C., & Cain, W. S. (1985). Age-related deficiency in perceived strength of six odors. *Chem Senses, 10*, 517-529.

Stevens, J. C., & Cain, W. S. (1987). Old-age deficits in the sense of smell as gauged by thresholds, magnitude matching, and odor identification. *Psychol Aging, 2*(1), 36-42.

Stevens, J. C., Cruz, L. A., Marks, L. E., & Lakatos, S. (1998). A multimodal assessment of sensory thresholds in aging. *J Gerontol B Psychol Sci Soc Sci, 53*(4), P263-272.

Stevens, J. C., & Dadarwala, A. D. (1993). Variability of olfactory threshold and its role in assessment of aging. *Percept Psychophys, 54*(3), 296-302.

Stevens, J. C., Plantinga, A., & Cain, W. S. (1982). Reduction of odor and nasal pungency associated with aging. *Neurobiol Aging, 3*(2), 125-132.

Thompson, S. B. N., Ennis, E., Coffin, T., & Farman, S. (2007). Design and evaluation of a computerized version of the Benton visual retention test. *Computers in Human Behavior, 23*, 2383-2393.

Weber, B., Fritze, J., Schneider, B., Kuhner, T., & Maurer, K. (2002). Bias in computerized neuropsychological assessment of depressive disorders caused by computer attitude. *Acta Psychiatrica Scandinavica, 105*, 126-130.

Weber, B., Schneider, B., Fritze, J., Gille, B., Hornung, S., Kuhner, T., et al. (2003). Acceptance of computerized compared to paper-and-pencil assessment in psychiatric inpatients. *Computers in Human Behavior, 19*(81-93).

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Westervelt, H. J., Ruffolo, J. S., & Tremont, G. (2005). Assessing olfaction in the neuropsychological exam: the relationship between odor identification and cognition in older adults. *Arch Clin Neuropsychol, 20*(6), 761-769.

Williams, J. E., & McCord, D. M. (2006). Equivalence of standard and computerized versions of the Raven Progressive Matrices Test. *Computers in Human Behavior, 22*(5), 791-800.

Appendix A

	Age	TMT	DSC	SS	Processing	Working M	CVLT	CF	BNT	Verbal Ret	MoCA	ID Hi	ID Med	ID Low	Total ID	Threshold
Age	1	0.23	-0.51**	-0.49**	-0.44**	-0.35**	-0.39**	-0.25	-0.18	-0.36**	-0.20	-0.42**	-0.44**	-0.37**	-0.46**	-0.28*
TMT		1	-0.59**	-0.59**	-0.88**	-0.38**	-0.45**	-0.43**	-0.50**	-0.60**	-0.41**	-0.31*	-0.40**	-0.35**	-0.41**	-0.35*
DSC			1	0.74**	0.87**	0.54**	0.45**	0.44**	0.41**	0.57**	0.49**	0.59**	0.38**	0.43**	0.52**	0.27*
SS				1	0.85**	0.52**	0.53**	0.43**	0.50**	0.64**	0.46**	0.54**	0.46**	0.43**	0.54**	0.39**
Processing					1	0.53**	0.54**	0.50**	0.54**	0.69**	0.51**	0.52**	0.47**	0.46**	0.55**	0.39**
Working M						1	0.48**	0.56**	0.33*	0.60**	0.28*	0.38**	0.10	0.38**	0.32*	0.33*
CVLT							1	0.47**	0.31*	0.77**	0.51**	0.28*	0.30*	0.46**	0.40**	0.54**
CF								1	0.34*	0.79**	0.53**	0.55**	0.59**	0.62**	0.67**	0.46**
BNT									1	0.74**	0.43**	0.24	0.29*	0.12	0.25*	0.20
Verbal Ret										1	0.64**	0.47**	0.51**	0.52**	0.57**	0.52**
MoCA											1	0.47**	0.41**	0.44**	0.36**	0.40**
ID Hi												1	0.63**	0.77**	0.89**	0.36**
ID Med													1	0.58**	0.85**	0.54**
ID Low														1	0.89**	0.52**
Total ID															1	0.55**
Threshold																1

0.01 (1-tailed) p-value <

0.05 (1-tailed) p-value

Appendix B

Performance by Group Across Neuropsychological Measures. M(SD)

	OA	YA	*F*	*p-value*	*partial* η^2
MoCA	23.73 (3.68)	28.23 (1.39)	52.94	<0.001	0.39
Boston Naming Test, 2nd Ed.	12.40 (2.98)	13.40 (1.22)	3.92	0.051	0.05
Category Fluency	40.80 (10.14)	45.95 (7.95)	6.67	0.012	0.07
California Verbal Learning Test, 2nd Ed.	24.73 (4.74)	28.13 (3.27)	14.39	<0.001	0.15
Letter Number Sequencing	8.16 (2.39)	11.98 (2.88)	44.60	<0.001	0.35
Symbol Search	21.91 (6.57)	41.08 (6.82)	173.88	<0.001	0.68
Digit Symbol Coding	48.31 (15.38)	80.75 (16.09)	90.23	<0.001	0.52
Trail Making Test, Part A	52.56 (20.70)	32.08 (9.44)	33.02	<0.001	0.29

Appendix C

Partial Correlations Between Composite Threshold and Identification Low, Med, and High in the Older Adults controlling for Age and Cognitive Functioning

Variable Partialed	ID Low	ID Med	ID High
None	0.52**	0.54**	0.36**
Age	0.47**	0.48**	0.29*
Age, MoCA	0.38**	0.41**	0.15
Age, VR	0.33*	0.36**	0.13
Age, PS	0.41**	0.42**	0.18
Age, MoCA, VR, & PS	0.32*	0.35**	0.11

Notes. VR= Verbal Retrieval; PS= Processing Speed; * $p < 0.05$ (1-tailed); ** $p < 0.01$ (1-tailed).