THE STANDARD SNIFF OLFACTOMETER (SO) PROTOCOL AND MORE

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Professional Studies

by

Jiayue Ni

December 2020

ABSTRACT

The Sniff Olfactometer (SO) has emerged as a novel technology to investigate the relationship between odorants and human olfactory perception. However, there is no standard protocol for the SO, which may lead to difficulties for other researchers to follow and verify the methods used. Therefore, research on developing an SO protocol was undertaken based on the feedback of the researchers and the human subjects. There are four parts to the protocol, i.e., subject training-testing, threshold measurements, reproducibility testing, and scale-up. Two patterns of sample presentations were experimented with to test which produced more reproducible data — concentrations with smaller differences in a group (SDG) and concentrations with larger differences of odorants in a group (LDG). Compared with samples in a SDG order, threshold tests with samples in the LDG order led to data with overall lower variance, indicating greater reproducibility.

BIOGRAPHICAL SKETCH

Jiayue Ni was born on June 13th, 1996, in Nanchang, Jiangxi, China. When she was 12, the sensational scandal about infant formula happened in China, which led to infants' illness and deaths. At the time, she decided to focus her efforts on learning about and improving food safety, which also fulfilled her desire to serve the public with what she learned.

In 2018, she was bestowed the Bachelor of Science degree in food science from the University of Massachusetts, Amherst. During her undergraduate period, she won the Chancellor's Scholarship and worked as an undergraduate researcher in Dr. Lili He's laboratory. In 2019, she joined Dr. Terry Acree's laboratory at Cornell University as a Master of Professional Studies student. She researched a standard protocol to use the sniff olfactometer and she also assisted with other projects about odorant perception.

During the summer of 2020, she interned at a startup company and successfully formulated and developed five flavors of natural vegan-friendly protein bars, which are awaiting launch. The experience opened the door to product development for her, encouraging her to delve deeper into food product research and development in the future.

ACKNOWLEDGMENTS

I want to thank my advisors, Dr. Joe Regenstein, and Dr. Terry Acree, for their warm-hearted support for both my academic study and life. Even in the pandemic, they still encouraged me and gave me guidance for my career choice.

I also want to thank the whole research team, Leto, Qi, Kaifeng, and Dave. They are not only strong technical and academic supports, but also sincere and interesting friends.

Finally, I want to thank my family for their never losing faith in me. I would like to appreciate my boyfriend as well, for his support by cheering me up and believing in me during the whole journey. Moreover, without the generous help from the administrative office of the Food Science Department and the College of Agriculture and Life Sciences, I could not have had such a wonderful and unforgettable experience at Cornell.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

2AFC: Two-Alternative Forced Choice

3AFC: Three-Alternative Forced Choice

EOR: Equal Odds Ratio

HEX: Hexanal

LDG: Larger Differences in a Group

PEG: Polyethylene Glycol

SDG: Smaller Differences in a Group

SO: Sniff Olfactometer

TMP: 2,3,5-Tripmethylpyrazine

TOR: Tertiary Odds Ratio

INTRODUCTION

Our living environment is full of various odorants, such as fruity, meaty, bloody, and smokey smells. These smells provide us with cues as to where we should get food, and if we are in danger (Romagny et al., 2016). And olfaction is a pivotal sensory function to help us detect and identify these odorant stimuli from a noisy background composed of millions of other smells (Livermore and Laing, 1998; Bak et al., 2019). To better understand how these stimuli are perceived and analyzed by us, and to investigate how the olfactory system processes and functions are essential (Livermore and Laing, 1998). Prior to all the olfactory investigations, threshold was always determined first (Steinmetz and Cain, 1969). However, according to Wise et al. (2008) and Schmidt et al. (2009), there has been limited previous research about how to measure an odor threshold, and about which is the best choice for olfactory investigations, and it is often unreliable to use the current tools and devices to measure and collect threshold data. Even though at present there are various instruments and methods to measure and analyze olfactory responses, there are unavoidable drawbacks with each, such as a long period of experiments, and over-dependence on subjects, which may lead to subjects' fatigue and bias (Avrunin et al., 2016). For example, when Laing et al. (2003) measured subjects' thresholds, the staircase method was used. Within the staircase method, 11 concentrations of target odorant were prepared. To measure the threshold, 2 samples were presented to the subjects who were asked if they can perceive and identify the samples. If the answer was yes, the researchers moved to a lower concentration and repeated the whole process. The whole procedure was continued until the predetermined criterion (the threshold) or the maximum "number of trials" was reached (Cornsweet, 1962). Cornsweet (1962) also indicated that the method requires the researchers to have a clear experimental design (to prepare proper concentrations) and arbitrary decisions about

when the experiment should be paused and to move to the next step or to redo the experiment with another set of concentrations. The whole procedure is time-consuming, and it is hard for researchers to follow since “arbitrary decisions” are an unclear requirement for researchers. In recent times, Prof. Terry Acree's group (the group) endeavored to contribute to research through development of the olfactometer, the device used to assist olfactory investigations. The sniff olfactometer (SO) was designed as an olfactometer to investigate olfactory discrimination (Rochelle et al., 2018). Humans were assigned as subjects when the SO was used since the human sense of smell is the most valid fashion for odor measurement with olfactometers (Dravnieks et al., 1986). Defined odorant compositions were used. The subjects were exposed to minimal stimuli by puffing the sample bottles for a short amount of time. The puff only lasts for less than 100 ms. In 2018, Rochelle et al. developed and refined the SO used to determine subjects’ thresholds using the equal odds ratio (EOR), which indicates the ratio of the concentrations of a two-odorant mixture that led to identical probability of perceiving either of the components. They also developed the tertiary odds ratio (TOR) as all three components are perceived identically in a mixture of three odorants. On the one hand, the innovative equipment can help researchers solve more problems about olfaction; on the other hand, since the equipment was developed, there are a lack of standard SO protocols, creating difficulties for other researchers to follow and replicate the experiments done with SO. Signals initiated in the olfactory epithelium because of odorants are captured by olfactory bulbs with a 10 to 200 ms latency. Therefore, research has looked at how latencies influence a subject’s ability to identify and distinguish odorants. Ding et al. (2019) tested humans’ binary odor recognition with SO and drew the conclusion that SO is able to measure the effects of latency on subjects’ discrimination including determining the temporal resolution. In their experiments, the EOR of target odorants

was measured. Twelve samples of different concentrations were prepared for each odorant, and each sample was puffed using the SO 36 times, leading to the subject being required to sniff and identify the compound 432 times. When the group replicated the experiment, it was determined that the whole process took 45-80 minutes, and according to the subjects, they lost focus because of fatigue after about 30 minutes, which may lead to human error even if the person would normally be able to do the work properly.

Even though olfactometers have been gaining popularity, few experiments have been successfully done to obtain meaningful threshold measurements (Hayes et al., 2012). Even within the group, different researchers have measured subjects' threshold values in different ways. Rochelle et al. (2018) used an alternated order to organize samples with different concentrations to present to the subjects for evaluation. In the experiment, a high concentration, a low concentration, and a blank (no odorants) sample were grouped in one trial. However, when Ding et al. (2019) measured thresholds, the sequential order was selected as samples with similar concentrations were grouped in each trail, as each group consisted of three high concentrations, or three low concentrations. In early 2020, Tang et al. (unpublished) used both sequential (smaller concentration differences in a group, SDG) and alternated (larger concentration differences in a group, LDG) orders in measuring one subject's threshold with identical odorants and corresponding concentrations presented differently because of the equipment limitations to three samples at a time. According to the subject's feedback, it was easier for her to concentrate and identify the odorants when presented in the alternated (LDG) orders while she felt confused and lost as she smelled "nothing" for an extended time using the sequential orders (SDG) (**Figure 1**).

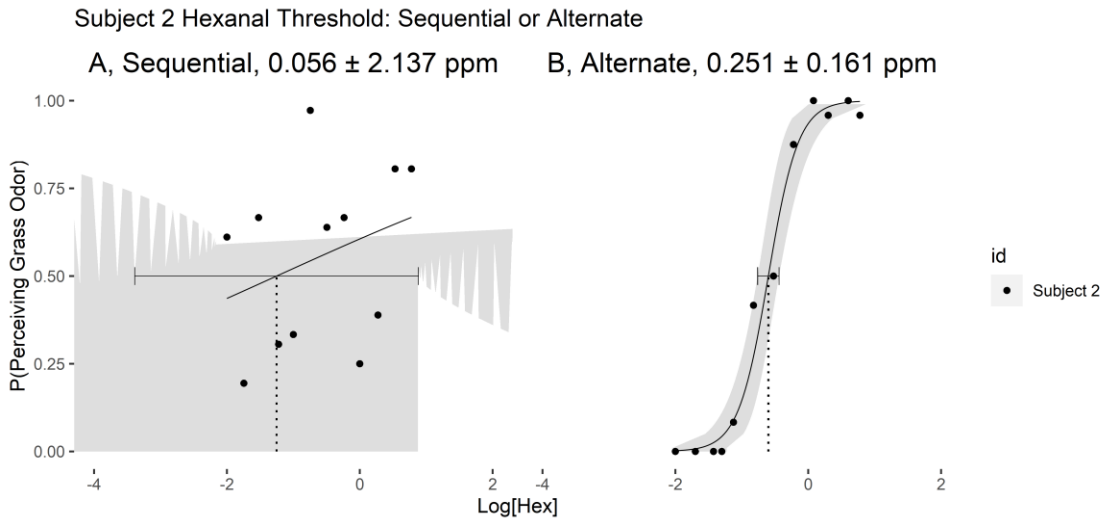


Figure 1. Comparison thresholds of hexanal (smell descriptor: grass) with sequential (SDG) and alternated (LDG) orders of presentation to subject 2. *Twelve different concentrations of hexanal were used in the experiment. Two-alternative forced choice (2AFC) was used in the experiment, the subject was required to choose from "grass" and "blank" when a sample was puffed. The x-axis is the log10 of the sample concentrations (parts/million), the y-axis is the probability that the subject believed they perceived the odorant as "grass", divided by the total number of times they puffed that sample. Each point represents that probability at a given concentration. The concentration at the 0.5 probability was assigned as the subject's threshold. The error bar in the figure of the LDG order was smaller than that of the SDG order.* ***Figure reused with permission from Tang et al. (2020, unpublished).*** *(A more complete explanation of the statistics will be shown in the methods section.)*

In this project, not only was a standard protocol developed for SO, but also the sequential (smaller concentration differences, SDG) and the alternated orders (larger concentration differences, LDG) were analyzed to determine if the latter method is a better choice for researchers to do threshold measurements. Only 6 different concentrations were used in threshold measurement based on the previous studies in the group, and each concentration was puffed 12 times. Subjects took only 15 minutes to finish the measurement, in which they reported feeling more relaxed than being tested 432 times in 45-80 minutes according to the original method. Besides that, it was noticed that if the sample concentrations for threshold

measurement were too high, subjects may find it hard to identify the odorant, compared with using a less concentrated sample. Additionally, the blank sample (no odorant) may cause problems as it may delude the brain to believe something is in the sample although there is nothing presented. This was named by the author as “smell hallucination” in this paper. Different from olfactory hallucination defining the false perception happens on patient with psychiatric disease (Kopala et al., 1994), “smell hallucination” is for healthy subjects with false positive reaction as all subjects used in the experiment claimed that they were healthy. More research is needed to explain why it is difficult for subjects to identity odorants in high concentration and how “smell hallucination” occurs and why subjects have the problem. This study will, hopefully, help SO research obtain more consistent results, which might encourage more studies in this area.

METHODS

Chemicals

The odor chemicals presented to the subjects were 50 mL of hexanal (HEX, the smell being described as "green") (CAS Registry No. 66-25-1, >98%) and 2,3,5-trimethylpyrazine (TMP, the smell being described as "nut") (CAS Registry No. 14667-55-1, >99%) diluted with 10% polyethylene glycol (PEG, smell described as "blank") (CAS Registry No. 9002-88-4) which was prepared with 90% deionized water (the building's carbon filtered deionized water) from the 400 ppm (parts/million) stock solution, made with commercial odorants and PEG. HEX and TMP were purchased from Sigma-Aldrich (Steinheim, Germany) and polyethylene glycol was obtained from Avantor Performance Materials (Radnor, PA, USA). HEX and TMP were selected for their nonidentical structures (**Figure 2**) as well as being easy to be determined by the subjects. And according to Wise and Cain (2000), if the stimuli are too similar subjects may tend to make more discriminative errors. PEG was chosen because of its solubility in water, being a good solvent for odorant chemicals, low volatility, and odorless (Chen et al., 2005). A good solvent for the odorants used for these experiments was required to be less volatile as odorants should be kept in the sample bottle before being puffed; and to be odorless as no other smells could be accepted when testing the target odorants. All solutions used in the solution reproducibility tests were prepared within 4 to 30 hours of the "Day 1" tests while the solutions used in other experiments were freshly prepared within 4 to 12 hours before the measurement time (James et al., 1997). All solutions were stored in amber bottles at room temperature (about 22 °C). Before transferring samples to amber bottles, the latter were ensured to be odor free based on the researcher's decision after manual sniffing. After discarding the solutions, all bottles and caps were cleaned with deionized water and ethanol (70%) for at least 7 times (6

times with water and 1 time with ethanol) until they were odorless, or the bottles and caps would be discarded.

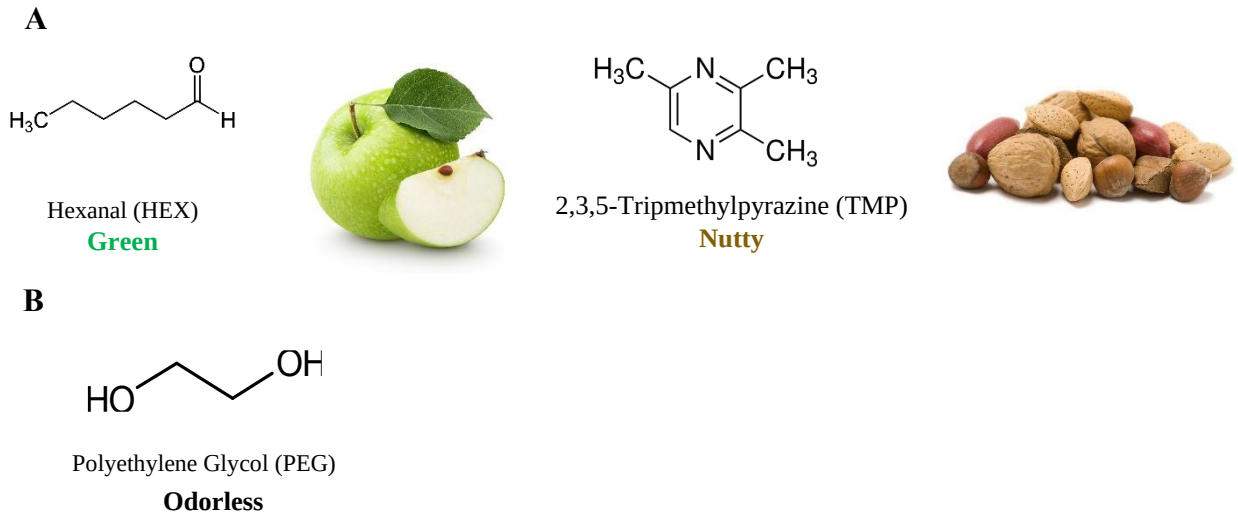


Figure 2. A. Chemical structures of odorants: hexanal (HEX) and 2,3,5-trimethylpyrazine (TMP). B. Chemical structure of solvent (blank sample), polyethylene glycol (PEG)

Subjects

Because of pandemic (COVID-19), only 8 subjects (one failed the testing session, and was not allowed to participate in the threshold measurement experiments) with age range from 22 to 32 were invited to attend the experiments. All subjects, 2 males and 5 females were chosen from graduate students in the Food Science Department at Cornell. Some subjects had SO experience previously, but none of the subjects knew any of the experimental details before the project was completed. The real purpose, developing a SO protocol was not disclosed to the subjects. They were only told to sniff the chemicals and to give the researcher feedback as to how they felt the about the experimental session (if they felt differently among different sessions, if they were tired, etc.), and other comments they wanted to share about the experiment, the equipment, and the researcher. They were also encouraged to ask questions but were told that some questions

would not to be answered by the researcher (asking questions was a way to show the subjects' thought and feelings).

Sniff Olfactometer

The equipment is the only instrument used to present odor puff to the subjects. The SO is from DATU, Inc. (Geneva, NY 14456), and is made up of three 250 mL Teflon™ sniff bottles. These bottles can expel headspace puffs by being easily squeezed by the actuators monitored by a computer program PsychoPy© (PsychoPy3, Pierce et al., 2019), which can be programmed to provide the desired puff timing and presents the puffs in a random order as programmed, as well as the visual instruction shown on the monitor for the subjects (Figure 3).

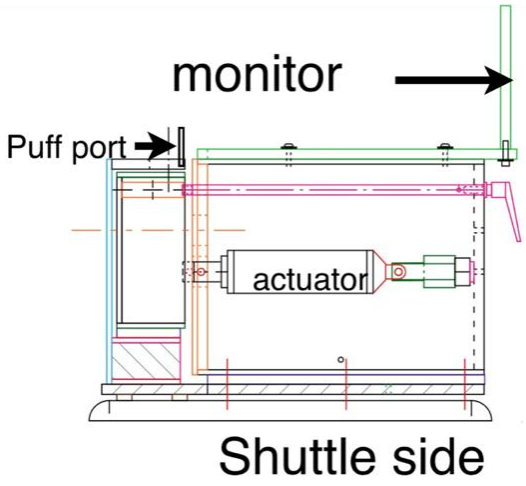


Figure 3. Scheme showing an inside view of the SO. *Scheme reused with permission from Acree et al. (2014)*

Figure 4 shows the side view of the SO. To eliminate any sound disturbances from the surroundings, the subjects were required to put on a headphone or earphones playing the music they liked. All headphones and earphones were provided by the subjects themselves so as to create a more comfortable and familiar situation for them. The subject would put his/her chin on

the chin rest to prevent their head from moving. The monitor installed in front of the subjects would instruct the subjects to follow the experimental process, as it would show "ready to inhale, click when you are ready", "inhale", and "select what you just smelled". All reactions required from subjects were made by their clicking the mouse to move the process along. An aiming system, which is depicted in Figure 5, helped the subject to ensure that his/her head was in the same position all the time, creating a fixed point for sniffing. All the equipment design above was done to eliminate any influence from the environment and to ensure the subject sniffs with the same part of the nostrils throughout the entire experimental session.

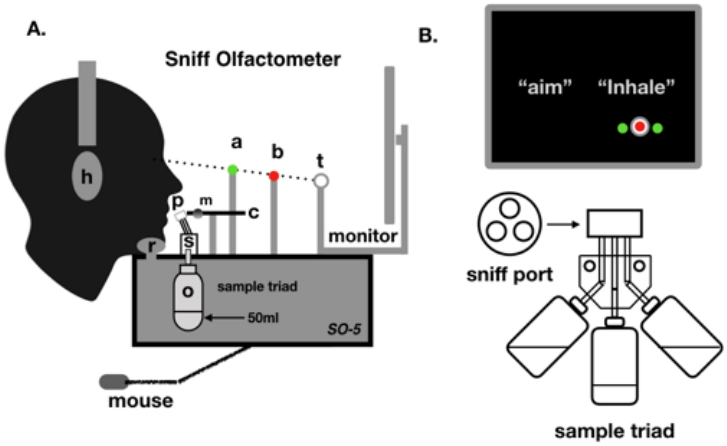


Figure 4. (A) Side view of the SO. headphone h, chin rest r, sniff port p, the slider m, c (to place and fix sample triad), and aiming system a,b, and t. (B)The instructions showed on the monitor. *Figure reused with permission from Ding et al. (2019).*

Figure 5. A picture of the aiming system of the SO. *Picture reused with permission from Tang et al. (2020, unpublished).*

Experimental Preparation

Test Pre-Requisite

All SO analyses were done in a Platinum Lead certified food research laboratory (2012) designed to have an additional 2 room exchanges/ hour in the sensory research wing yielding 6 exchanges/hour. Air in the test room should be filtered to remove undesired odors and be pumped into the room to maintain positive pressure (Delahunty et al., 2006). Neither the subjects nor the researchers are allowed to apply any fragrance products. And to achieve more accurate results and avoid environmental disturbance, the subjects should avoid smoking, drinking, eating and similar activities for at least 1 hour before the experiment (Wise and Cain, 2000; Goyert et al., 2007). To prevent potential bias from the subjects, no obvious difference should be observable when researchers change samples or switch the bottles' position with the sampling triad. The protocol for when to change the bottles will be described below.

Subject Conditioning and Selection

Since most subjects were not familiar with SO nor of testing odorants, all subjects were required to take teaching-testing sessions before the threshold tests (Laing and Francis, 1989). Also, since subjects may have different perceptions and previous experiences (Hettinger and Frank, 2018), it is necessary to train all the subjects before the formal experiments. Before the training session could even take place, the researcher was responsible for explaining the teaching-testing sessions in detail and going through the human-subjects consent form (Appendix) with each person to ensure they understand what their rights and responsibilities are.

During the teaching session, high concentrations, 20-50 ppm of testing odorants were prepared, which were used to help subjects distinguish and become familiar with the odorants. The odorants along with the 10% PEG (blank) were put into the SO equipment and the pre-programed PsychoPy© was run. Instructions were shown on the monitor as described above. The only difference is that when they were asked what they smelled; the correct answer was also shown on the screen along with three choices as "nutty, green, or blank". For example, as shown in Figure 6, if the bottle containing TMP was puffed, the question on the screen would be "What did you smell? Nutty" with the three choices shown at the bottom of the screen to ensure that they put down the right answer. Each bottle would be puffed 5 times and each time the subjects were asked to identify with the correct answers indicated on the screen. The subjects were trained to distinguish among nutty, green, and the blank smell as well as to become familiar with the SO system and question type in the following experimental sessions. Additionally, if the odorant concentrations were not high enough for the subjects, subjects would comment that it was hard for him/her to perceive the smell. The concentration would be increased and the entire teaching session would be done again.

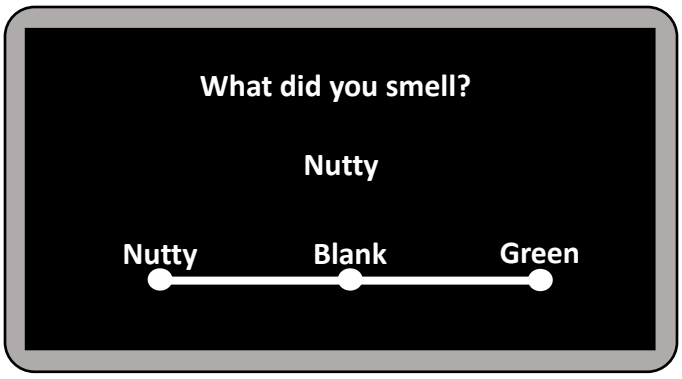


Figure 6. The question shown in training sessions after the TMP bottle was puffed. *Subjects clicked the choice to respond and proceed with the session. This is only a teaching session for the subjects to become familiar with the odorants, the SO system, and question type. If an incorrect choice was selected in this session, the subjects could still proceed the session just by clicking the mouse as data was not collected in this session.*

The teaching session was followed by a testing (selection) session, in which the subjects were tested on their ability to distinguish the two odors. The identical samples (10% PEG, high concentration of HEX and TMP) in the teaching session were used. In the testing session, 5 puffs were done in a random order for each sample (HEX, TMP, and 10% PEG), i.e., 15 puffs in total. Subjects were asked to identify the descriptor of the odorant puffed. Subjects that attained 13 out of 15 correct (13/15) (Appendix) were allowed to move to the next step, whereas the ones that failed this screening were required to take the teaching and testing sessions again until they reached a 13/15 correct score. The subjects who failed to achieve a 13/15 correct score within 5 teaching sessions of the teaching-testing cycle were not used with the project. All subjects were required to take the test again before each formal experiment the following day, i.e., if the experiments were done on separated days, the subjects were required to take the test again before each experiment day.

Threshold Measurements

Threshold measurements are used to determine what the lowest concentrations of a specific odorant subjects are able to identify. This step is important in many research protocols. To measure the threshold, as shown in Figures 7 and 8, a series of various concentrations of odorants need to be presented to the subjects. Two different orders of presenting odorants to the subjects were used in this experiment — SDG and LDG. In the SDG, similar concentrations are put in one group with smaller concentration difference while in the LDG, concentrations in different levels with greater concentration difference are placed in one group.

Six different concentrations of each odorant were chosen to do the threshold measurements, and these concentrations cover levels of high, medium, and low, according to the previous data obtained in the laboratory. The samples were separated into 2 groups, and 3 samples were attached to the sample triad manifold. As subjects can detect which bottle is puffed using the sniff port, each sample's location was switched after each trial (Figure 8). In each trial, every bottle was puffed 4 times randomly, as 12 puffs in total in one trial and 3 trials equally 36 puffs with a set of samples (three different concentrations) was a session. One session would be one of the two sets of SDG or LDG samples and the other session would be the second set of SDG or LDG samples. Two session were considered a completion of a method (SDG or LDG), and subjects were allowed to take a 5-10-minute rest after 2 sessions. Therefore, 4 sessions were needed to do both SDG and LDG samples for one odorant, and 8 sessions were needed to do both HEX and TMP threshold measurement with both SDG and LDG. After all the experiments, all subjects were asked and reported without specific instructions how they felt during sniffing. Subjects were also encouraged to ask questions and comment anytime during the experiments, but they were told in advance that the researcher may not answer all the questions to avoid bias.

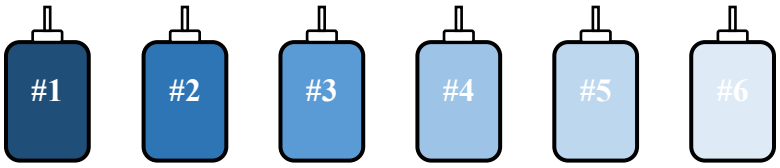


Figure 7. The concentrations in high, medium, and low level. *The darker the color, the more concentrated the sample solution. The concentrations were determined for each experiment.*

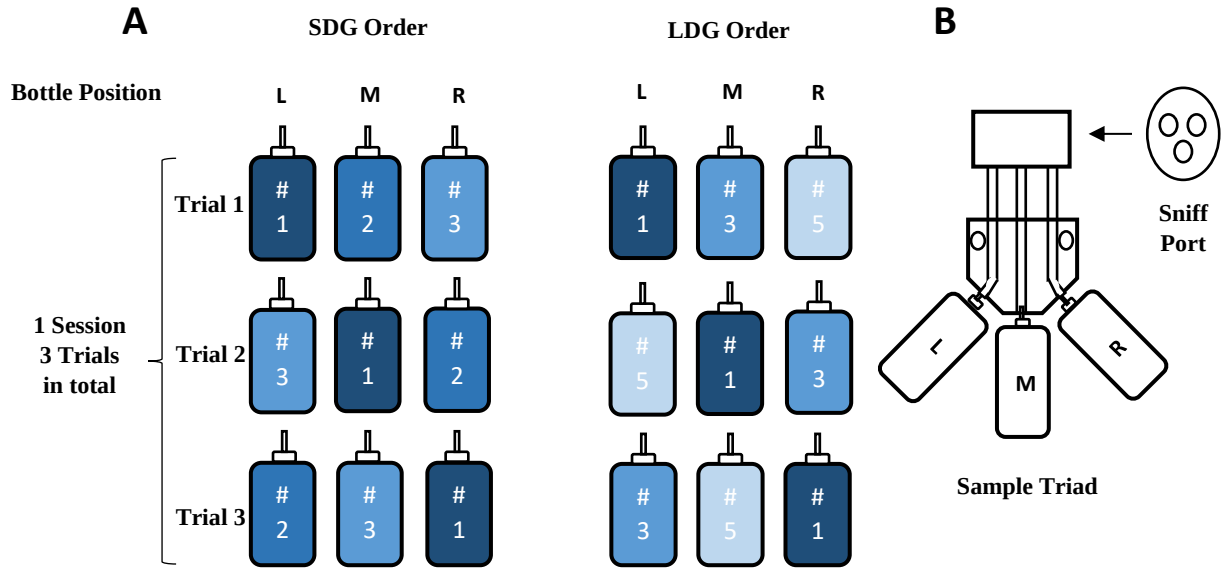


Figure 8. A. Example of SDG and LDG orders of bottles within a trial and the changes when switching bottles as part of a group and B. Bottles in the sample triad. *L (left position), M (middle position), and R (right position). The bottles in the sample triad were switched between each trial.*

Being slightly different from the testing sessions used earlier with each subject, in the threshold measurements, only one odorant with different concentrations were used (Figure 8). Correspondingly, the questions shown on monitor were two-alternative forced choice (2AFC) (Wise and Cain, 2000; Laing et al., 2003) instead of a three-alternative forced choice (3AFC) in the test session (Figure 9).

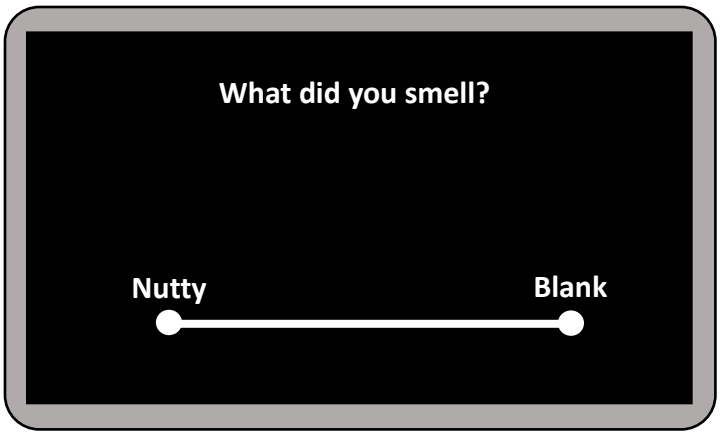


Figure 9*. *The question shown in the threshold measurement of TMP. *Subjects clicked the choice to respond and move the session forward.*

Solution Reproducibility Test

To confirm the results as well as to test if the sample solutions can be recycled, a reproducibility test was undertaken after the threshold measurements with both SDG and LDG orders in both a narrow range and a wider range experiment. Only subject A participated in this test with recycled samples, while the other subjects (B-G) took the reproducibility tests with fresh sample solutions. Subject A was asked to do the whole threshold measurements again with the same samples and the same presentation method used in the previous threshold measurement. The reproducibility tests took place on the second or the third day depending on the subject's schedule (not at the same time) after the first threshold measurement.

Threshold Measurement Design

Narrow Range of Concentrations

For HEX threshold measurement, 15, 10, 5, 1, 0.1, and 0.01 ppm were prepared with 10% PEG solution. For TMP samples, 20, 10, 5, 1, 0.5, and 0.1 ppm were chosen. Subject A and G

participated in this experiment. Both had no previous experience with SO and successfully pass the teaching-testing session.

Wider Range of Concentrations

Wider Range with Higher Concentrations

To get more data points, a wider range of solutions were prepared. For both HEX and TMP threshold measurements, the highest concentration was 100 ppm and lowest one was 0.001 ppm with 10-fold dilutions for the other samples (100, 10, 1, 0.1, 0.01, and 0.001 ppm). Subject A, and subjects D, E, F participated in this experiment. None of these 4 subjects had previous experience with SO, and all of them had a replication test except subject D.

Wider Range with Blank Sample (No Odorant)

A control group was prepared in this experiment to ensure that all subjects and equipment were performing and functioning properly (Steinmetz et al., 1969). PEG (50 mL, 10%) with no odorant chemicals was considered as 0 ppm and the descriptor was set as "blank". For HEX threshold measurements, 20, 1, 0.01, 0.001, 0.0001, and 0 ppm were chosen, whereas for TMP threshold measurements, 30, 1, 0.01, 0.001, 0.0001, and 0 ppm were chosen. Subjects A, B, and C participated in this experiment. Subject B had had experience with SO about 6 months before the experiment.

Statistical Analysis

All data were analyzed with R (Mac/PC 4.0.2. "Taking Off Again"; R Core Team 2020).

All word responses, i.e., "nutty" and "blank" (Figure 9), were converted to binary values, i.e., 1 for the sample being identified by the subject as present, and 0 for the sample not being identified by the subject as present. The number of observed responses at a given concentration

for a single subject was then divided by the total number of puffs of that sample, i.e., 12 in the current studies. And it was then considered as the "probability" of that subject perceiving and identifying (when the subject chose e.g., "nutty" as a response) the presence of the test odorant sample at the given concentration. These were then plotted as the subject's probability of responding e.g., "nutty" at one concentration versus the concentration expressed as Log 10 of the concentrations (in ppm) in the test bottle. In each plot (individual subject for SDG or LDG), there were 6 points (one for each sample), each of which was the average of 12 observations by each subject in each given sample (the given concentration of the odorant), i.e., the completion of two sessions. The points were then used to create dose response curves with standard deviation (confidence interval) using the binomial generalized linear model of the MASS package (Venables and Ripley, 2020) in R. The threshold was set as the concentration corresponding to 50% probability according to this model (Rochelle et al., 2017). Since 2AFC was used in the threshold measurement, the guessing probability, when the subjects could not distinguish "something" from "nothing", but just guessed, was 50% (0.5). As the experiments required the subjects to have an ability to perceive and identify HEX and TMP, the probability of a subject detecting and distinguishing the odorant at a given concentration was required to be higher than 50%, so that, 50% probability was the lowest concentration required (the threshold) for a subject to presume to perceive and identify the given odorants (Lawless, 2010). And since the uncertainty in the concentration of the threshold (50% probability, the independent variable) was measured, the variance was calculated for the concentration (x-axis), and not for the probability (y-axis). All statistics were based on a 95% confidence level.

For the boxplots (all boxplots can be found in the Result and Discussion session), they were generated based on the threshold points and the confidence intervals were extracted from the

curves. Boxplots were generated for each subject and each chemical combination to visualize the comparison between the variance and replicability between threshold measurement with the SDG order and the LDG order.

Comparison of SDG Order and LDG Order

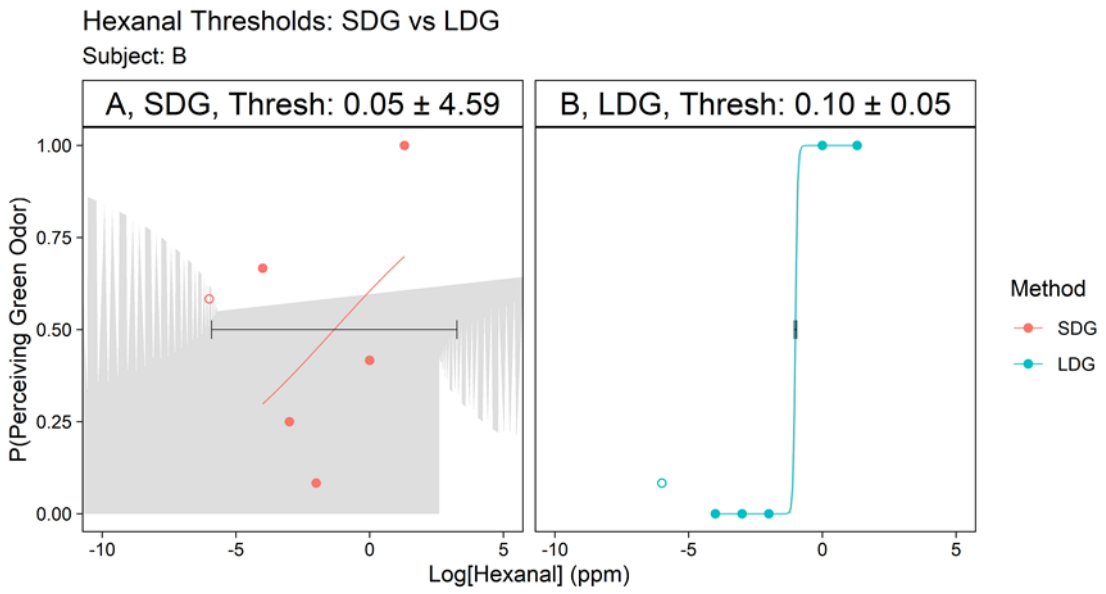


Figure 10. HEX threshold measurement with SDG order and LDG order for subject B, day 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject.*

According to the subjects' feedback, they did not notice any difference between the two methods when they were asked if there was any difference among the different sessions (they did not know two methods were being tested in the experiment). Figure 10 is the threshold curve for subject B in a wider range with blank sample test on day 1. The y-axis shows the probability of the subject's perceiving green odor, which means the subject selected "green" when a sample of HEX was presented to her. When the probability equals 0.5, the corresponding concentration is the subject's threshold for HEX. Since 2AFC was used in the threshold measurement, the guess probability is 50% (0.5), and a higher than guess probability was required, so that the concentration with 50% probability is considered as the lowest concentration (the threshold) for the subjects to have perceived and identified the given odorant (Rochelle et al., 2017). It was also

shown that the data generated with the LDG order had a smaller standard deviation than that of the SDG order. Not only with subject B, the LDG order can give data with smaller standard deviations was found for all subjects (Figures 11 and 12).

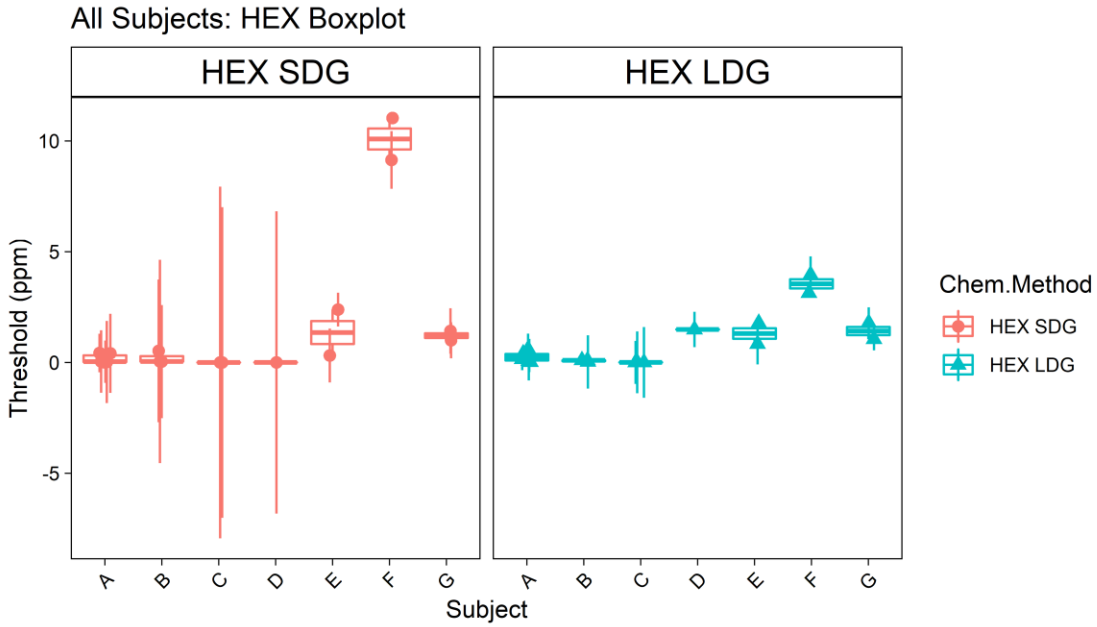


Figure 11. Boxplot of HEX threshold for all the subjects in all experiments. *HEX SDG means HEX threshold measurement with SDG order and HEX LDG means the measurement with LDG order. Each point represents the threshold for the subject on one test day.*

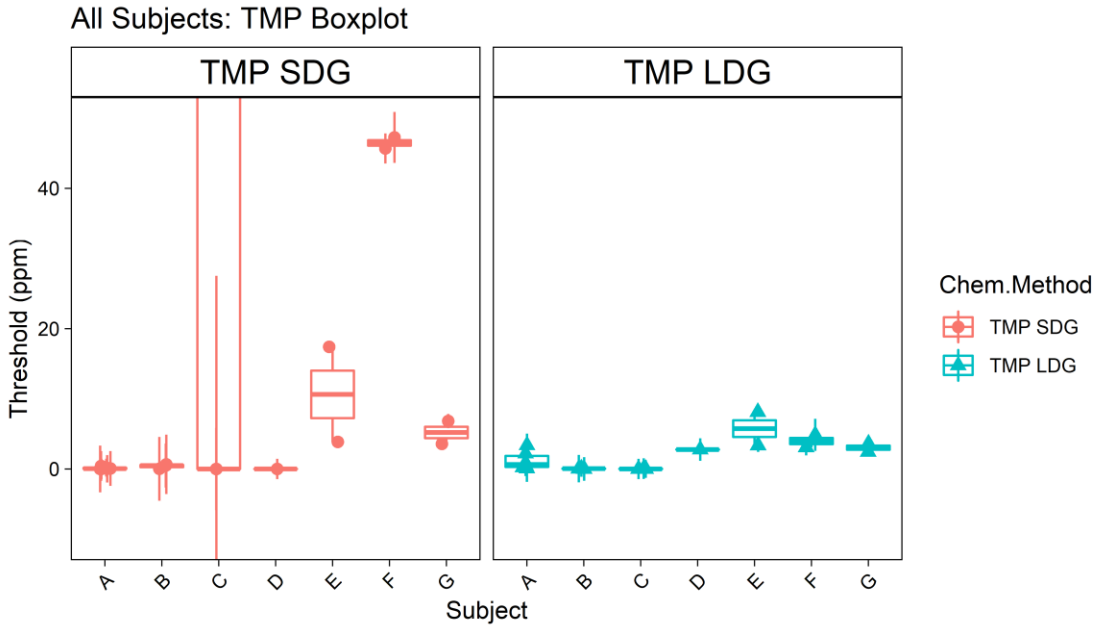


Figure 12.Boxplot of TMP threshold for all the subjects in all experiments. *TMP SDG means TMP threshold measurement with SDG order and TMP LDG means the measurement with LDG order. Each point represents the threshold for the subject on one test day.*

In addition, in the most cases, the curves from the LDG order can be successfully formed while poorer curves were obtained from the SDG order. More data sets can be found in the Appendix. It may be because that the subjects had more examples of "high concentrations" to compare with "low concentrations" in the LDG order, while they had no comparison with the SDG orders. When the high concentration groups were presented to the subjects, they could always smell target odorants and it was easy for them to make decisions. When low concentration groups were shown to the subjects, they got confused as they felt (from their feedback) that they could not smell anything as all samples were less concentrated. It was noticed that more feedback and comments were received from the subjects during the experiments with the SDG order. "That is strange, I can smell everything" (in the high concentration sessions) and "I feel so bad, I cannot smell anything. Is there any blank sample in this group?" (in the low concentration sessions) were common. It may indicate that the subjects began to have doubts and to overthink, which may lead to less consistent threshold data, since a self-generated psychological bias was created.

Rouby and Holley (1993) found that high intensity of an odorant can lead to a masking effect, in which the less intense odorants are suppressed. However, it was not noticed in the LDG order method. As shown in Figures 11 and 12, the thresholds obtained with the LDG order were similar to those attainted with the SDG order.

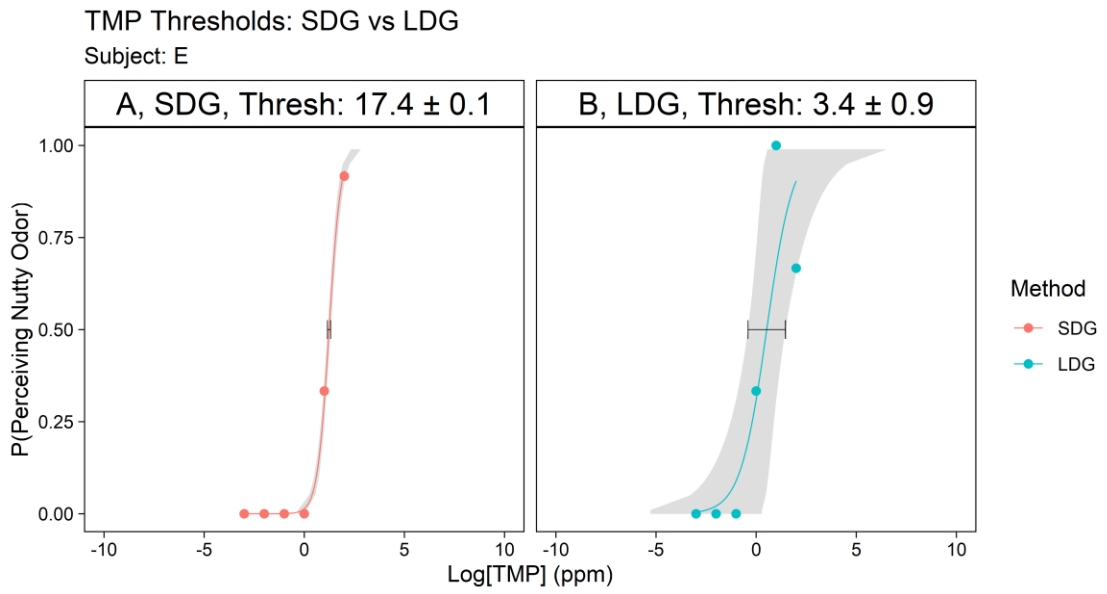


Figure 13. TMP threshold measurement with SDG order and LDG order for subject E on day, replication 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

Although in the most cases, the plot obtained with the LDG order was better than that obtained with the SDG order as it had smaller errors with a few exceptions. As shown in Figure 13, when subject E was doing a TMP threshold measurement (test replication), the error bar from the LDG order was bigger than that from the SDG order. It may, however, be shown that, in most cases, the LDG order can give data with a smaller standard deviation. The difference for a few subjects' perceptions and experiences may lead to exceptions.

"Smell Hallucination" Effect with Blank Sample (No Odorant)

Hallucination is defined as a false perception-sensory experience. It was shown as someone saying that they perceived something in the absence of an adequate stimulus (Greenberg, 1992). Olfactory hallucinations were usually observed in patients with schizophrenia, depression, and eating disorders (Kopala et al., 1994). In the wider range with blank sample experiment, there was a higher probability for all subjects (A, B, and C) to respond as "green" or "nutty" when the

blank sample was presented than the probability they obtained when the lowest concentration with odorants (0.0001 ppm) was presented.

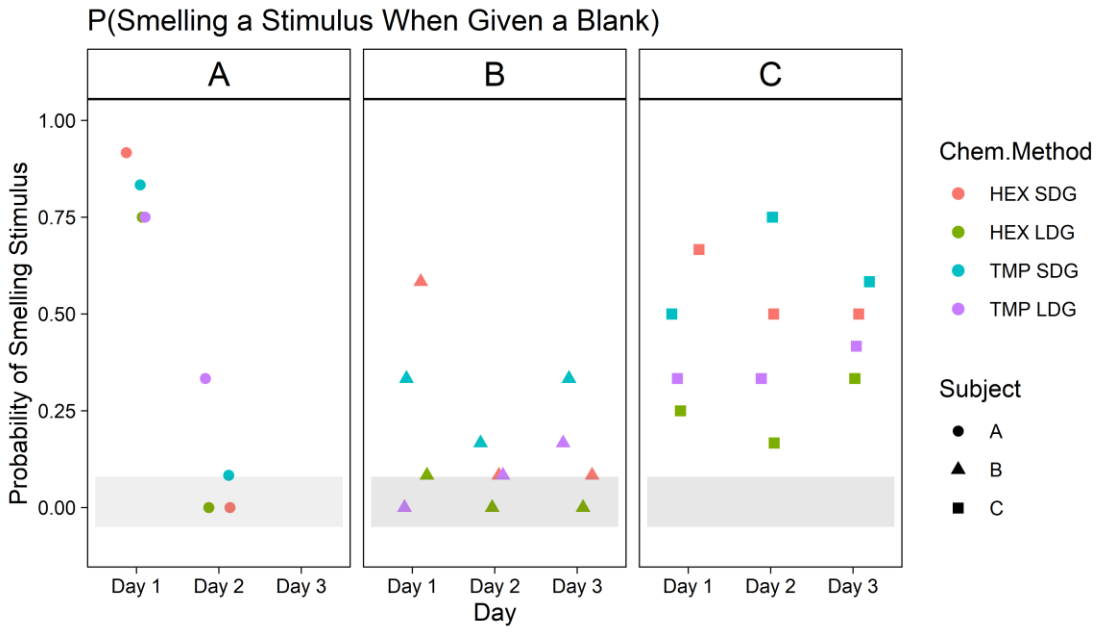


Figure 14. Probability for subjects A, B, and C when the blank sample (no odorant, 0 ppm) was presented*. The grey bar shows the probability for the subjects (A, B, and C) when the lowest concentration with odorants (0.0001 ppm) was presented. The blank and the lowest concentration were tested during the same experiment.*

As shown in Figure 14, in most cases, the subjects responded that they smell something (green or nutty) when a blank sample was presented. And the probability is even higher than when the lowest concentration (0.0001 ppm) was presented to them. This effect is like the olfactory hallucination described above. However, all subjects used in the experiment considered themselves as healthy according to the pre-interview before the teaching-testing session. According to Greenberg (1992), being similar to olfactory hallucination, people with the olfactory pseudo-hallucination can also claim to have a sensory experience without stimulus. The only nuance is that people with pseudo-hallucination are aware that the experience is not real as he/she does not believe that the stimulus was presented to him/her. The olfactory pseudo-

hallucination theory does not fit the situation in the experiments as all subjects were blinded from the experiment design, so that they did not know blank samples were used nor when the blank samples were placed in the trials.

In addition, the LDG order can effectively reduce the probability of false positives (Figure 14), as the LDG order provided subjects with examples of higher concentrations, with which it would be easier for the subjects to distinguish “I smell something” from “I smell nothing”.

Subjects’ Responses to Higher Concentrations (100 ppm)

Delahunty et al. (2006) points out that perception intensity increases when physical concentration of odorants increases. The expectation for the wider range with a high concentration (100 ppm) was that higher odorant concentration would lead to higher probability of subjects’ odorant identification, i.e., the subjects chose “green” or “nutty” in the threshold measurements than that at a lower concentration. However, some subjects had a higher probability of perceiving TMP when a lower concentration (10 ppm) was presented.

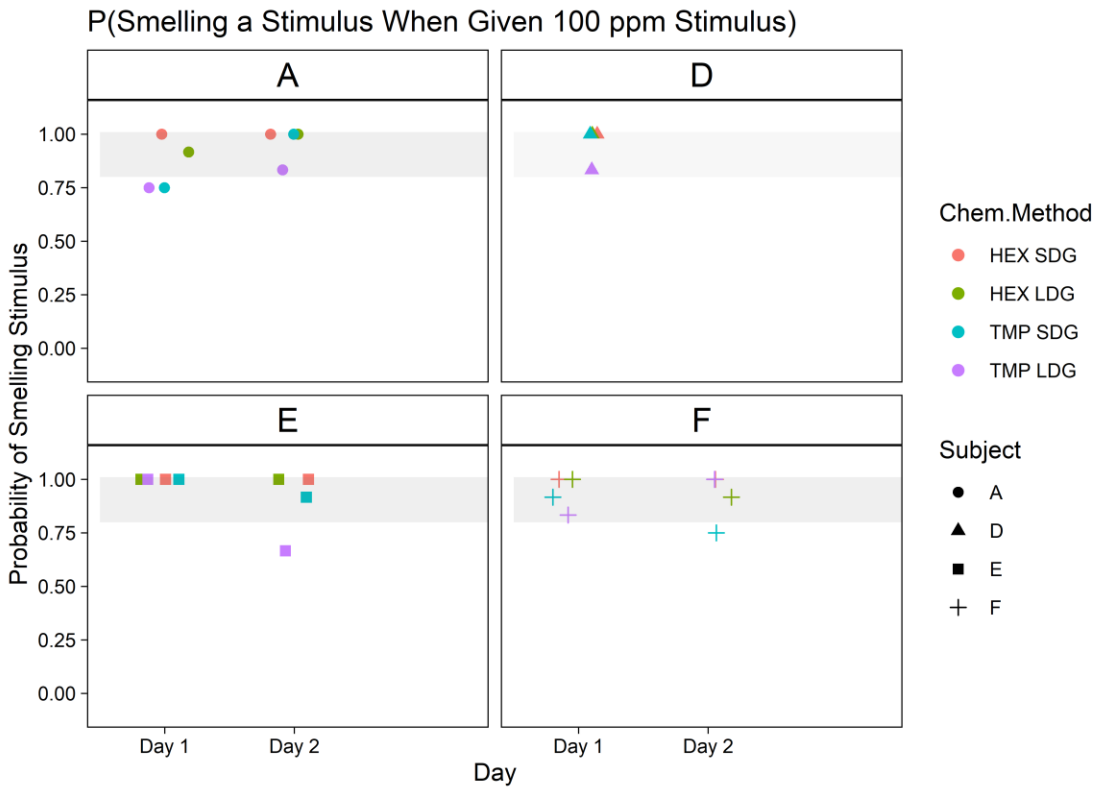


Figure 15. Probability for subjects A, D, E and F when the high concentration sample (100 ppm) was presented*. The grey bar shows the probability for the subjects (A, D, E and F) when a lower concentration (10 ppm) was presented. The 100 ppm and the 10 ppm samples were tested during the same experiment.*

As a part of subject A's feedback, he claimed that he perceived something different from TMP in some trials and that was confirmed as 100 ppm of TMP sample after the experiment was completed. The descriptor of TMP may be changed at high concentrations as there may be a trend to change the odorant descriptor as the concentration increased from low to high (Laing et al., 2003). And the change of odorant (TMP) descriptor led to a lower probability for 100 ppm as subjects can no longer identify the odorant as "nutty".

Solution Reproducibility Test

If the solutions were recycled for the reproducibility test, the thresholds in the replication test were always higher than the ones in the day 1 test. As shown in Figures 16 and 17, subject A's TMP thresholds from day 1 was lower than those from the replication test. However, in Figures 18 and 19, compared with the those collected on the replication test, subject B's TMP thresholds from day 1 were higher. The trend that thresholds collected with the replication test with recycled solutions are higher than the ones collected on day 1 can be observed in the other plots of subject A's result in the Appendix.

There are two possible reasons for that trend – (1) the odorants were gradually evaporated. Since the liquid keeps evaporating as time passes, the longer period between solution preparation and the experiment, the fewer odorant molecules in the bottles. And (2) odorant loss because of puffing. Puffing is an acceleration of odorant evaporation. The number of odorant molecules decreased after being puffed for the day 1 experiment, so that the actual concentrations in the replication test were lower than expected, which led to higher thresholds.

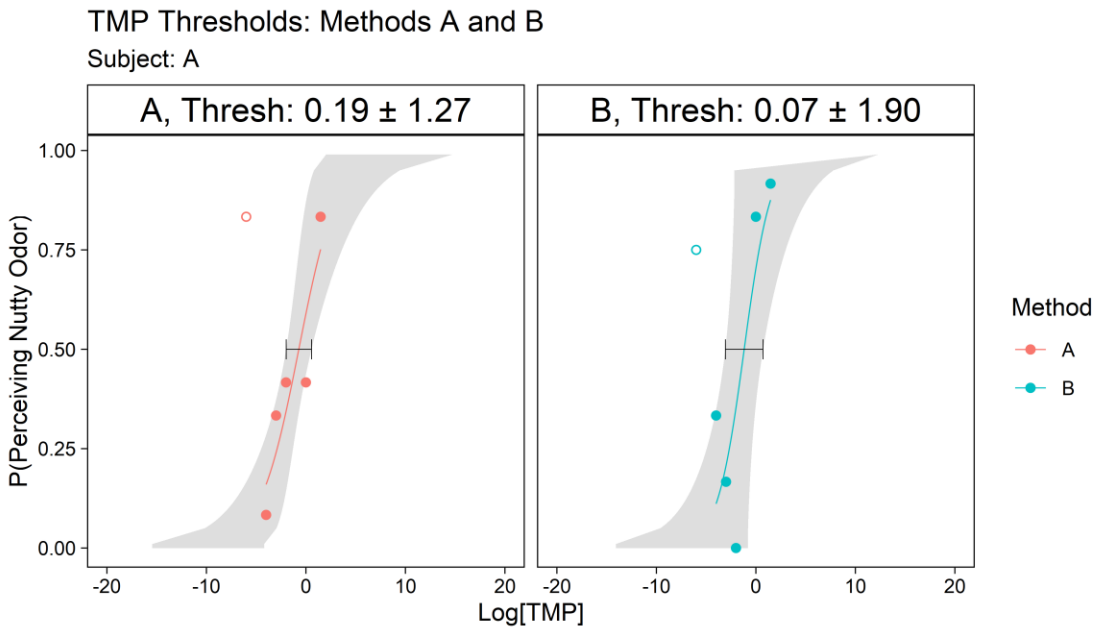


Figure 16. Subject A's TMP threshold measurement with blank sample (no odorant) day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

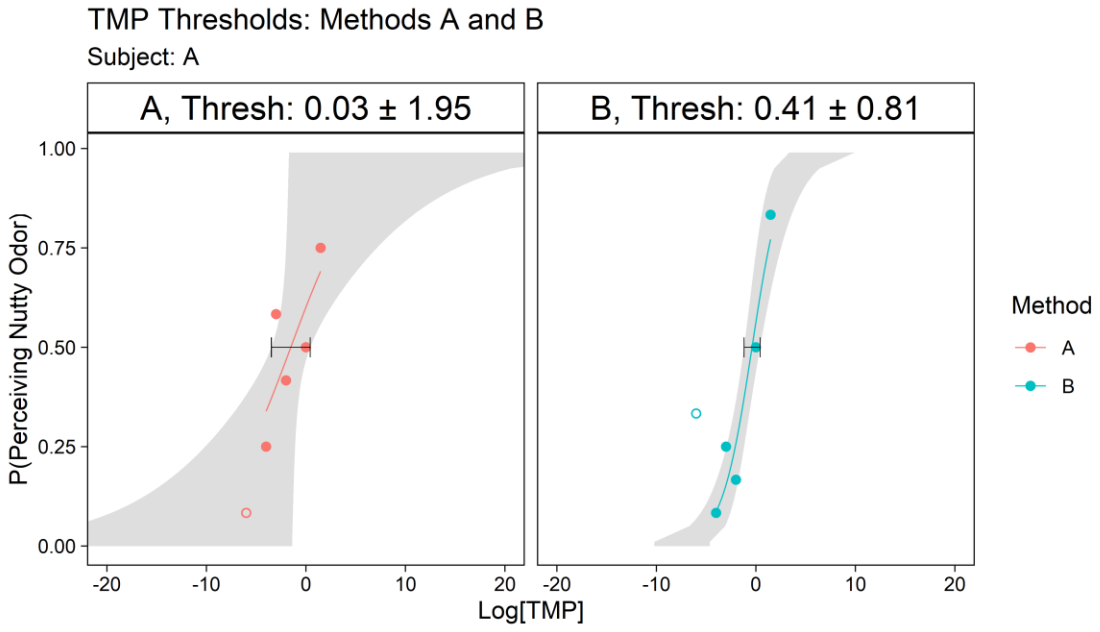


Figure 17. Subject A's TMP threshold measurement with blank sample (no odorant) replication 1. *The solutions were recycled from the experiment on day 1. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

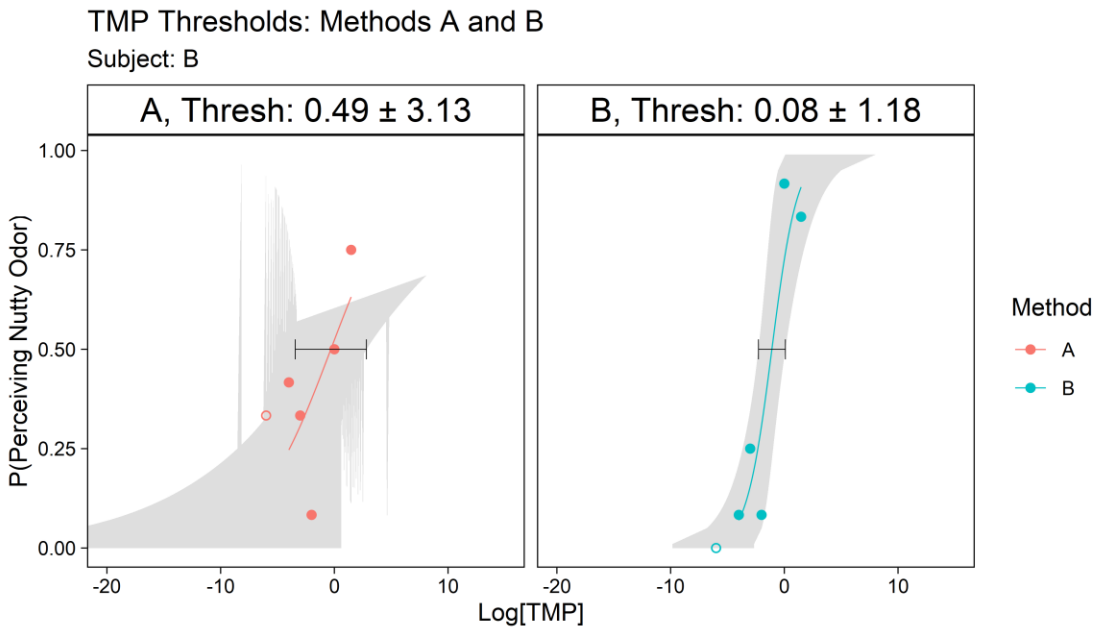


Figure 18. Subject B's TMP threshold measurement with blank sample (no odorant) day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

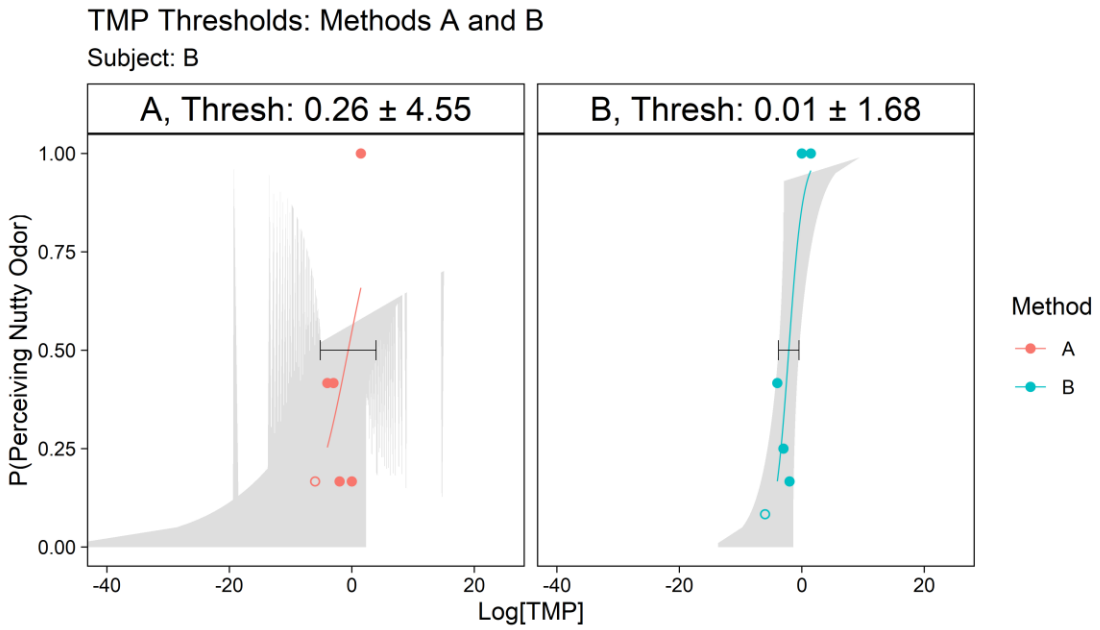


Figure 19. Subject B's TMP threshold measurement with blank sample (no odorant) replication 1*. The solutions were freshly prepared. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

Comparison of Threshold Measurements with a Narrow Range and a Wider Range

Since subjects' thresholds remain unknown until the measurement is taken, a wider range of concentrations should be prepared for the threshold measurement to ensure the thresholds for all the subjects are fully covered. However, how wide the range should be is still unclear. According to Figures 20 and 21, the concentrations failed to include some of the results for subject A as subject A's threshold was higher than the concentrations prepared in the narrow range experiment, while the curves were successfully made for subject G as the concentrations prepared for him fully covered subject G's actual threshold. Since different subjects have different thresholds, and various odorants have various perception intensities, a narrow range may only fit a certain group of subjects, whereas a wider range may serve more subjects. When deciding what range of concentrations should be used in the experiment, researchers should have a pre-test with several subjects to obtain an overview of a possible threshold range of a specific odorant.

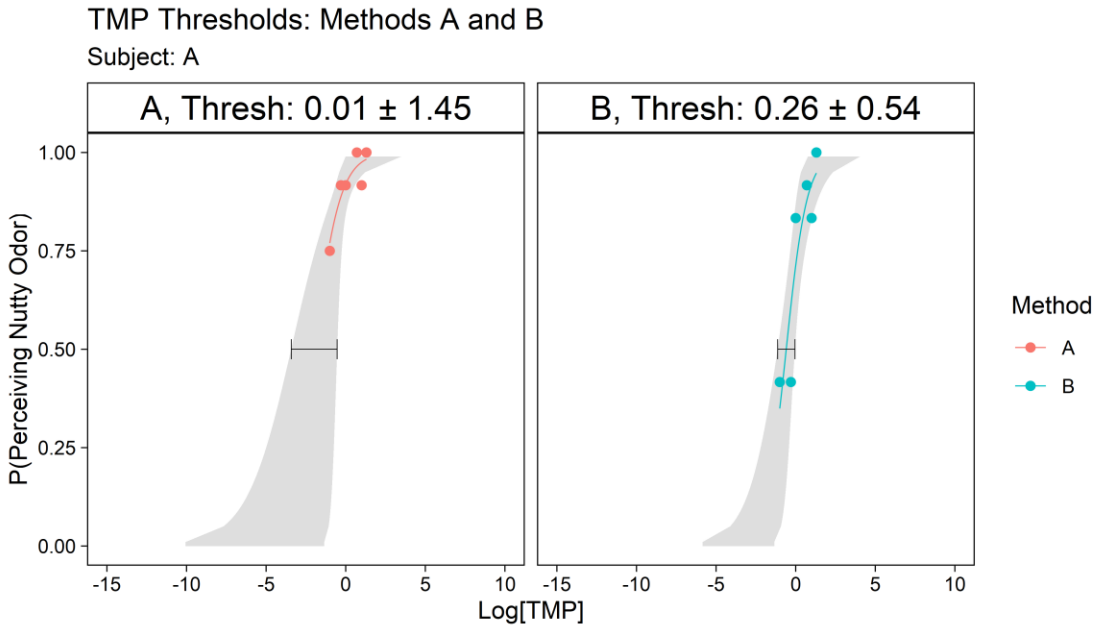


Figure 20. Subject A's TMP threshold measurement with the narrow range day 1, *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

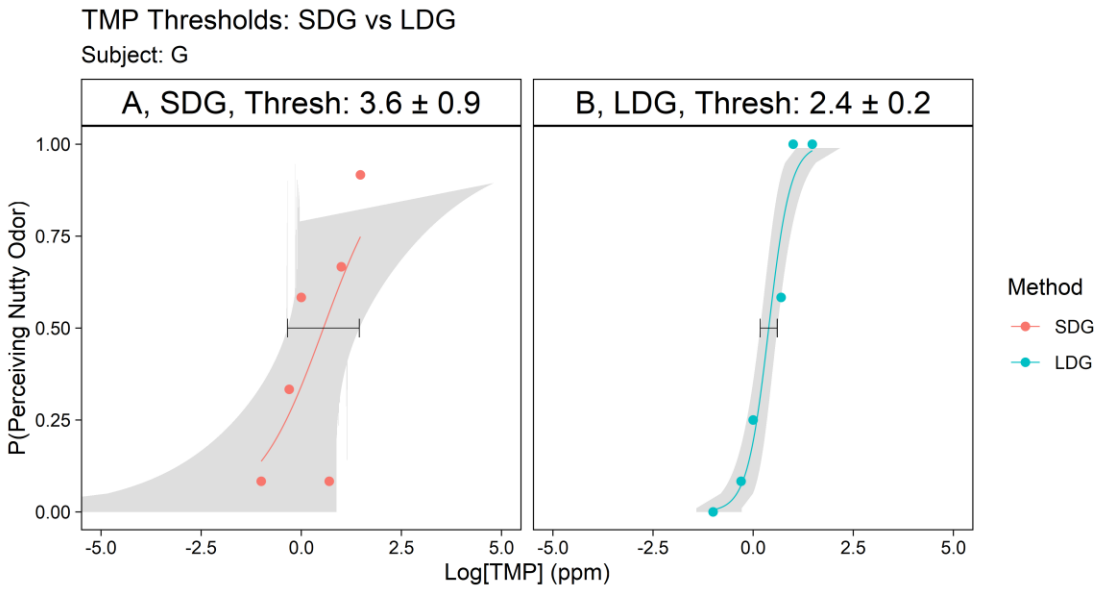


Figure 21. Subject G's TMP threshold measurement with the narrow range day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

Experimental Limitations

Because of the pandemic (COVID-19), sample size (the number of subjects) was limited. To produce higher quality data, the number of subjects need to be increased. According to Richardson and Zucco (1989), human beings have an outstanding ability to perceive and identify some odorants, but they still may have difficulties in discriminating other odorants. During the experiment, some subjects complained that it was hard for them to distinguish TMP from blank (PEG, 10%), while other subjects claimed that it was easy for them to do. Further research with more subjects and more varieties of odorants are needed to confirm the current conclusions.

Besides recruiting more subjects, the sex of subjects should also be balanced. In the experiment, 5 females and 2 males were chosen. Most studies have suggested that compared with men,

women outperform on the odorant identification and discrimination test for some odorants (Bengtsson et al., 2001; Doty and Cameron, 2009). Since more females were recruited in the experiment, there may be some bias although there seems to be limited information about whether men and women have consistently different thresholds.

Since a vigorous sniff of an odorant at a low concentration may transfer a similar quantity of the odorant molecules to the olfactory receptors as from a weak sniff of the odorant at a high concentration (Mainland and Sobel, 2005), it is hard to control the number of odorant molecules transferred into subjects' receptor. Even though with the SO, identical amount of odorant from the headspace are supposed to be delivered with each puff, to measure how vigorously the subjects sniff would be difficult and even harder to keep consistent, which will in turn affect whether they identify the sample as "nutty" (in the TMP threshold test), "green" (in the HEX threshold test) or "blank" when the sample concentration is near their personal threshold level and those small difference may be significant. The SO equipment may need to be improved to ensure each subject receive the same amount of odorant molecule in each puff.

CONCLUSIONS

When measuring thresholds with human subjects with the sniff olfactometer (SO), researchers should group the samples with a larger concentration difference in an LDG order at different levels of concentrations (high, medium, and low). This seems to provide the subjects with an easier task in finding their thresholds. If the samples are grouped in a SDG order, it would be harder for the subjects to distinguish “something” from “nothing”, as all samples smell similar in one trial. Although in most psychophysical experiment, blank samples are necessary to correct for false positives, blank samples may lead to problems in threshold measurement with the SO. Some subjects may have “smell hallucination” when a blank sample is presented as they smell “something” even there is “nothing”. More research is needed to better understand what causes the hallucination. A wider range of concentration may be helpful in threshold measurements, but it mainly depends on the human subjects. To choose a proper range, researchers should avoid selecting concentrations that are too high, as the descriptor may be changed, and the subjects can no longer identify the odorant. Further research is necessary to confirm this point.

Because of the pandemic, the researcher had a limited number of subjects. To achieve a more convincing conclusion, the whole study should include more subjects and the sex of the subjects should also be more balanced.

REFERENCES

1. Avrunin, O., Shushlyapina, N., Nosova, Y., & Bogdan, O. (2016). Olfactometry diagnostic at the modern stage. *Bulletin of the National Technical University «Khpi» Series: New Solutions in Modern Technologies*, 12 (1184), 95–95. https://doi.org/10.20998/2413-4295.2016.12.13
2. Bak, J. H., Jang, S. J., & Hyeon, C. (2019). Modular structure of human olfactory receptor codes reflects the bases of odor perception. *Biorxiv, Jan. 20, 2019*. https://doi.org/10.1101/525287
3. Bengtsson, S., Berglund, H., Gulyas, B., Cohen, E., & Savic, I. (2001). Brain activation during odor perception in males and females. *Neuroreport*, *12*(9), 2027-2033.
4. Chen, J., Huddleston, J. G., Rogers, R. D., & Spear, S. K. (2005). Polyethylene glycol and solutions of polyethylene glycol as green reaction media. *Green Chemistry*, 7(2), 64–82. https://doi.org/10.1039/B413546F
5. Cornsweet, T. N. (1962). The staircase-method in psychophysics. *American Journal of Psychology, 75*(3), 485-491.
6. Delahunty, C. M., Eyres, G., & Dufour, J. P. (2006). Gas chromatography-olfactometry. *Journal of Separation Science*, *29*(14), 2107-2125. https://doi.org/10.1002/jssc.200500509
7. Ding, K., Wang, X., Rinberg, D., & Acree, T. (2019). Sniff Olfactometry: Temporal effects on odorant mixture perception in humans. *Chemrxiv, Nov. 12, 2019.* https://doi.org/10.26434/chemrxiv.10293395.v1

8. Doty, R. L., & Cameron, E. L. (2009). Sex differences and reproductive hormone influences on human odor perception. *Physiology & Behavior, 97*(2), 213-228. https://doi.org/10.1016/j.physbeh.2009.02.032
9. Dravnieks, A., Schmidtsdorff, W., & Meilgaard, M. (1986). Odor thresholds by forced-choice dynamic triangle olfactometry: reproducibility and methods of calculation. *Journal of the Air Pollution Control Association*, *36*(8), 900-905. https://doi.org/10.1080/00022470.1986.10466128
10. Goyert, H. F., Frank, M. E., Gent, J. F., & Hettinger, T. P. (2007). Characteristic component odors emerge from mixtures after selective adaptation. *Brain Research Bulletin*, *72*(1), 1-9. https://doi.org/10.1016/j.brainresbull.2006.12.010
11. Greenberg, M. S. (1992). Olfactory hallucinations. In *Science of Olfaction* (pp. 467-499). Springer, New York, NY, USA. https://doi.org/10.1007/978-1-4612-2836-3_17
12. Hayes, J. E., Jinks, A. L., & Stevenson, R. J. (2013). A comparison of sniff bottle staircase and olfactometer-based threshold tests. *Behavior Research Methods*, *45*(1), 178-182. https://doi.org/10.3758/s13428-012-0220-2
13. Hettinger, T. P., & Frank, M. E. (2018). Stochastic and temporal models of olfactory perception. *Chemosensors*, *6*(4), 44. https://doi.org/10.3390/chemosensors6040044
14. James, C. E., Laing, D. G., & Oram, N. (1997). A comparison of the ability of 8–9-year-old children and adults to detect taste stimuli. *Physiology & Behavior*, *62*(1), 193-197. https://doi.org/10.1016/S0031-9384(97)00030-9
15. Kopala, L. C., Good, K. P., & Honer, W. G. (1994). Olfactory hallucinations and olfactory identification ability in patients with schizophrenia and other psychiatric

disorders. *Schizophrenia Research, 12*(3), 205-211. https://doi.org/10.1016/0920-9964(94)90030-2

16. Laing, D. G., & Francis, G. W. (1989). The capacity of humans to identify odors in mixtures. *Physiology & Behavior*, *46*(5), 809-814. https://doi.org/10.1016/0031-9384(89)90041-3
17. Laing, D. G., Legha, P. K., Jinks, A. L., & Hutchinson, I. (2003). Relationship between molecular structure, concentration and odor qualities of oxygenated aliphatic molecules. *Chemical Senses, 28*(1), 57-69. https://doi.org/10.1093/chemse/28.1.57
18. Livermore, A., & Laing, D. G. (1998). The influence of chemical complexity on the perception of multicomponent odor mixtures. *Perception & Psychophysics*, 60(4), 650-661. https://doi.org/10.3758/BF03206052
19. Lawless, H. T. (2010). A simple alternative analysis for threshold data determined by ascending forced-choice methods of limits. *Journal of Sensory Studies*, 25(3), 332–346. https://doi.org/10.1111/j.1745-459X.2009.00262.x
20. Mainland, J., & Sobel, N. (2006). The sniff is part of the olfactory percept. *Chemical Senses, 31*(2), 181-196. https://doi.org/10.1093/chemse/bjj012
21. Peirce, J. W. (2007). PsychoPy — Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13. https://doi.org/10.1016/j.jneumeth.2006.11.017
22. Richardson, J. T., & Zucco, G. M. (1989). Cognition and olfaction: A review. *Psychological Bulletin, 105*(3), 352. https://doi.org/10.1037/0033-2909.105.3.352
23. Rochelle, M. M., Prévost, G. J., & Acree, T. E. (2017). Computing odor images. *Journal of Agricultural and Food Chemistry, 66*(10), 2219-2225. https://doi.org/10.1021/acs.jafc.6b05573

24. Romagny, S., Coureaud, G., & Thomas-Danguin, T. (2018). Key odorants or key associations? Insights into elemental and configural odour processing. *Flavour and Fragrance Journal*, *33*(1), 97-105. https://doi.org/10.1002/ffj.3429

25. Rouby, C., & Holley, A. (1995). Temporal competition between odorants: Effect of different time intervals on the perception of monorhinic and dichorhinic binary mixtures. *Perception*, 24(9), 1083-1097. https://doi.org/10.1068/p241083

26. Schmidt, R., & Cain, W. S. (2010). Making scents: Dynamic olfactometry for threshold measurement. *Chemical Senses*, *35*(2), 109-120. https://doi.org/10.1093/chemse/bjp088

27. Steinmetz, G., Pryor, G. T., & Stone, H. (1969). Effect of blank samples on absolute odor threshold determinations'. *Perception & Psychophysics*, *6*(3), 142-144. https://doi.org/10.3758/BF03210080

28. Tang, Q., & Acree T. E. (2020). The temporal suppression effect of binary mixture on human. (Unpublished)

29. Team, R. C. (2013). R: A language and environment for statistical computing. http://cran.univ-paris1.fr/web/packages/dplR/vignettes/intro-dplR.pdf

30. Venables, W. N. & Ripley, B. D. (2002) *Modern Applied Statistics with S. Fourth Edition*. Springer, New York, USA. ISBN 0-387-95457-0

31. Wise, P. M., & Cain, W. S. (2000). Latency and accuracy of discriminations of odor quality between binary mixtures and their components. *Chemical Senses*, *25*(3), 247-265. https://doi.org/10.1093/chemse/25.3.247

32. Wise, P. M., Bien, N., & Wysocki, C. J. (2008). Two rapid odor threshold methods compared to a modified method of constant stimuli. *Chemosensory Perception*, *1*(1), 16-23. https://doi.org/10.1007/s12078-008-9010-8

APPENDICES

Testing Design (Explanation for the 13/15 Bar)

Since there were 3 samples used in a testing session—nutty, green, and blank, so that the guessing probability is 1/3. To ensure the subjects can distinguish these smells, a correct score above 1/3 was required, which was assigned as a 3/5 correct score for each odorant. On the other hand, for each odorant (nutty, green, and blank), up to 2 errors were accepted, if a 13/15 correct score were achieved by the subject.

If one of the errors was actually "green" (the correct answer) but "nutty" was chosen by the subject, it would indicate that in that specific puff, the subject was confused for both "green" and "nutty" (one error for the green, and one error for the nutty).

If the other one error was actually "blank" (the correct answer) as "nutty" was chosen by the subject, as above, the subject actually made one error for the blank and one error for the nutty.

In total, there would be one error for green, one error for blank, and 2 errors for nutty, which approached the bar being set for nutty (error of 2/5 for each odorant). From the direct result of the testing, only 2 errors were observed as the subject made mistakes on a green and a blank. But, one mistake led to confusions (errors) for 2 odorants—the correct one and the chosen one. Since the following experiments needed subjects to have an almost perfect ability to distinguish among all three odorants (nutty, green, and blank), a bar of 13/15 correct score was set for the testing sessions.

Consent Form

Research Participant Information and Consent Form

Description of the research

You are invited to participate in a research study about [Insert the name of the project]. The participation is voluntary. The purpose of this study is to [insert object of study].

Confidentiality

This study is anonymous. We will not be collecting or retaining any information about your identity. The record of this study will be kept strictly confidential. Research records will be kept in a locked file, and all electronic information will be coded and secured using a password-protected file. Your identity will be disclosed in the material that is published.

What will my participation involve?

In the research, you will first be trained to be familiar with aromas of [names of aroma], and all chemicals used in the study are food-grade. After the training, you will be tested on the identification of the same aromatic substances. If 87% correct score is achieved, you will move to the 2nd part of the study, or you have to retake the training session. The repetition will not exceed 5 times.

Before the 2nd part, you are allowed to take 5-10 minutes of rest. In the 2nd part, you will [insert the details of the experiment].

The experiment is a forced-choice experiment, so you will have to make a choice even if you are not sure about your answer. The total running time per session is about [insert time], and you are allowed to take 10-20 minutes of rest between any 2 sessions.

Responsibilities

If you decide to participate in this study, you will be asked to follow these items--

1. You are responsible *NOT* to apply any fragrance products (e.g., perfume).
2. You are responsible *NOT* to smoke, drink, or eat at least 1 hour before the study.
3. You are responsible for informing the researcher about your nose condition (e.g., stuffed nose).
4. You are responsible for wearing headphones/earphones throughout the entire experiment to eliminate experimental error (headphones and earphones can be prepared by yourself).
5. You are encouraged to ask questions and comment anytime during the experiment, but *NOT* all the questions will be answered.
6. You *MAY* be asked about your opinion about the entire experimental procedure.

Compensation

Upon proper completion of the study, you will receive [insert award detail].

Who to contact

You may present your questions to the experiment administrator [insert name] at any time. The researcher conducting this study is [Insert researcher's name]. Please feel free to contact [Insert researcher's name] at [Insert researcher's email address]. Your participation in this study is completely voluntary, and should you feel it necessary at any time to withdraw, alert the administrator. If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Cornell Institutional Review Board (IRB) at 607-255-5138 or access

their website at http://www.irb.cornell.edu. You may also report your concerns or complaints anonymously through Ethicpoints(www.hotline.cornell.edu) or by calling at 1-866-293-3077. Ethicpoints is an independent organization that serves as a liaison between the University and the person bringing any complaints so that anonymity can be ensured.

By signing the consent form, I understand the purpose of this study and what I will be asked to do. I know the confidentiality will be ensured and I will receive proper compensation. I also am aware of my responsibilities and rights.

Name: _____ Date: _____

(Prepared originally and modified by the author with permission from Terry Acree)

The Standard Protocol for Using the Sniff Olfactometer

Part 1: Teaching and Testing

1. Talk before the Teaching

   A talk between the researchers and the subject before any session is necessary. In the talk, researchers are obligated to tell the subjects what will happen in the experiments, including all steps and what should be noticed, i.e., They have to prepare to inhale when they click "yes" to "Ready to inhale?". It is unnecessary for the subjects to know the real purpose of the project to avoid bias.

2. Teaching Session

   A teaching session is designed for the subject to become familiar with the odor chemicals used in the formal test, the process, and the operation of the SO.

   In the teaching session, one blank (10% PEG) and odor chemicals in a high concentration (20-50 ppm) are used to ensure the subject can distinguish all substances. If the concentrations prepared for the subjects are too low for them to perceive, please increase the concentrations. To get the subject familiar with the PsychoPy© system, a two-step teaching program is shown in the flowchart. The puffing time and order can be programed with PsychoPy© based on the experiment design.

   a) "You will smell Aroma A."

      Then the A chemical is puffed from a Teflon™ sniff bottle. The subject inhales to learn and remember the aroma.

   b) "Identify the aroma you just smelled."

      Multiple choices are given to the subject, and he/she should choose "Aroma A" as the correct answer (Figure I).

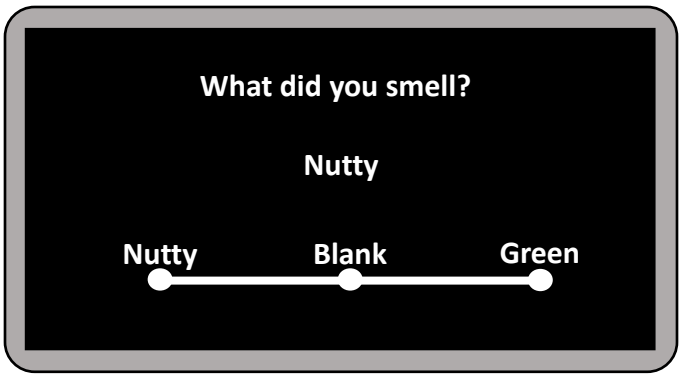


Figure I. The question shown in a training session after "nutty" bottle was puffed. Subjects clicked the choice to respond and proceed the session.

All samples should be puffed 5 times in random order.

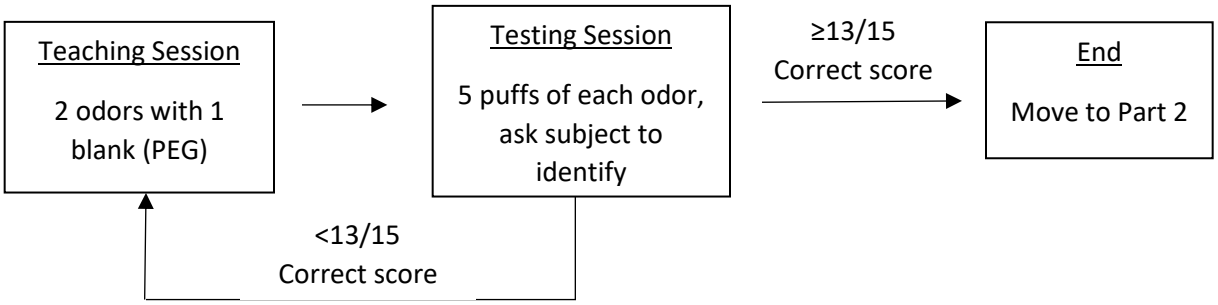
Preparation	→	Teaching
1-2 different odors with 1 blank (PEG). A high concentration (20-50 ppm) should be prepared. *If the concentration is too low, please increase the concentration		1. "You will smell Aroma A". 2. Puff the bottle containing Aroma A. 3. "Identify the aroma" from choices A, B, and C Each odor will be puffed 5 times randomly.

3. Testing Session

To verify how well the subject learned in the teaching session, a testing session is created.

In the testing session, each odor chemical used in the teaching session will be puffed 5 times, and the subject is asked to identify what odor he/she is smelling. The question form and the choices are the same as in the teaching session. If the subject can obtain a 13/15 correct score, the session ends, or otherwise he/she has to go back to the teaching session. If a subject cannot complete the testing session within 5 repetitions of the "teach-

test" cycle, the experiment should be terminated, and a new subject should be selected and tested on.



*If the person cannot achieve 13/15 correct score within 5 repetitions of the "Teach-Test" cycle, terminate the session, and a new subject should be tested on.

Part 2: Threshold Measurements

To determine the threshold of the specific odorant, 6 different concentrations should be prepared. The 6 samples are divided into 2 groups with the LDG order (each group has one sample from high concentration, one from medium concentration, and one from low concentration; the high, medium, and low are relative levels in the 6 samples). There are 3 trials for each group since the position of each sample bottle is switch after each trial (each bottle will be place in all three positions once as left, middle, and right). In each trial, each bottle is puffed 4 times, and subjects are required to identify the sample with either its descriptor or as blank (Figure II).

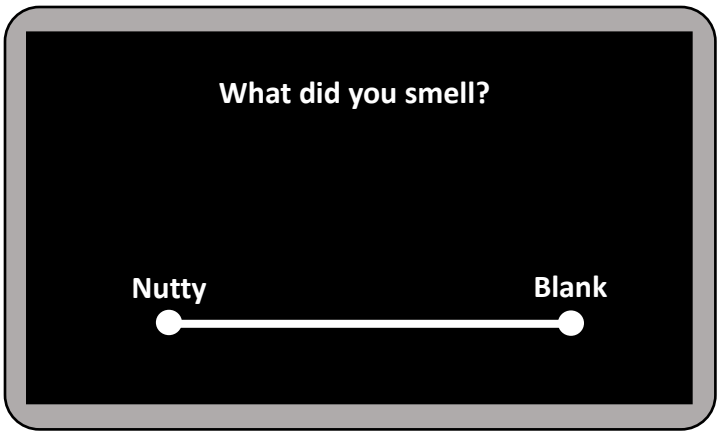
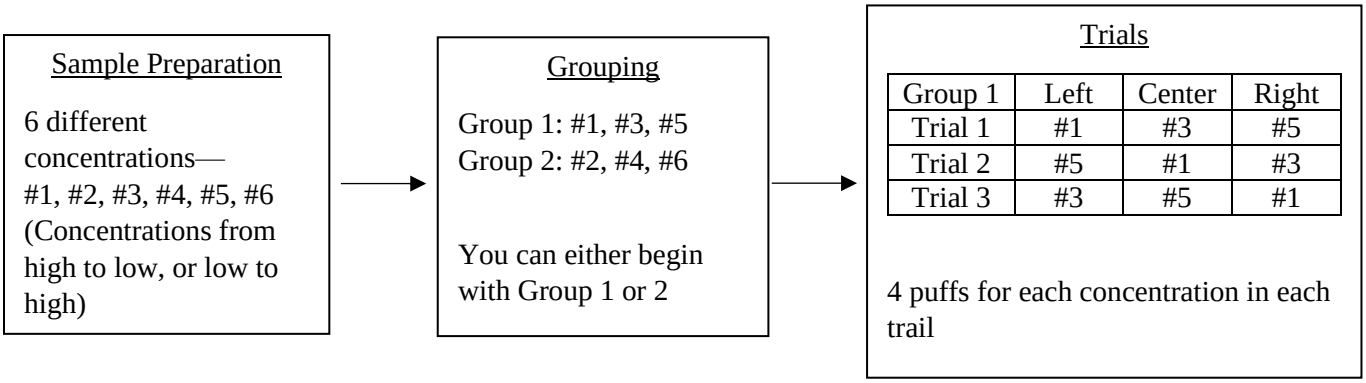


Figure II. The question shown in the threshold measurement for the "nutty" odorant.

Subjects clicked the choice to respond and proceed the session.



Part 3: Formal Test and Reproducibility Test

The formal test could be done according to the project design, and the test should be repeated with the same fresh samples and same method after the first test. The replication test could be taken any time after the first day.

Part 4: Use of More Subjects

More subjects should be recruited for the experiment to attain more convincing data.

Notice:

Apart from the 4 parts mentioned above, there are several issues that needed to be addressed:

1. The test room should remain tidy and ventilated.
2. The working condition of the subject should be recorded in detail (e.g., stuffed nose because of illness).
3. All the aromatic solutions should be prepared on the same day of the experiment.
4. Neither the subjects nor the researchers can apply any fragrance products.
5. To achieve the most accurate results, the subjects should avoid smoking, drinking, and eating for at least 1 hour before the experiment.
6. No obvious difference should be noticed between any two trials (researchers should spend a similar length of time to change bottles among trials).

Threshold Curves for All Subjects and All Experiments

1. Solution Reproducibility Test (Recycled Solution)

1.1 Narrow Range of Concentration

Day 1

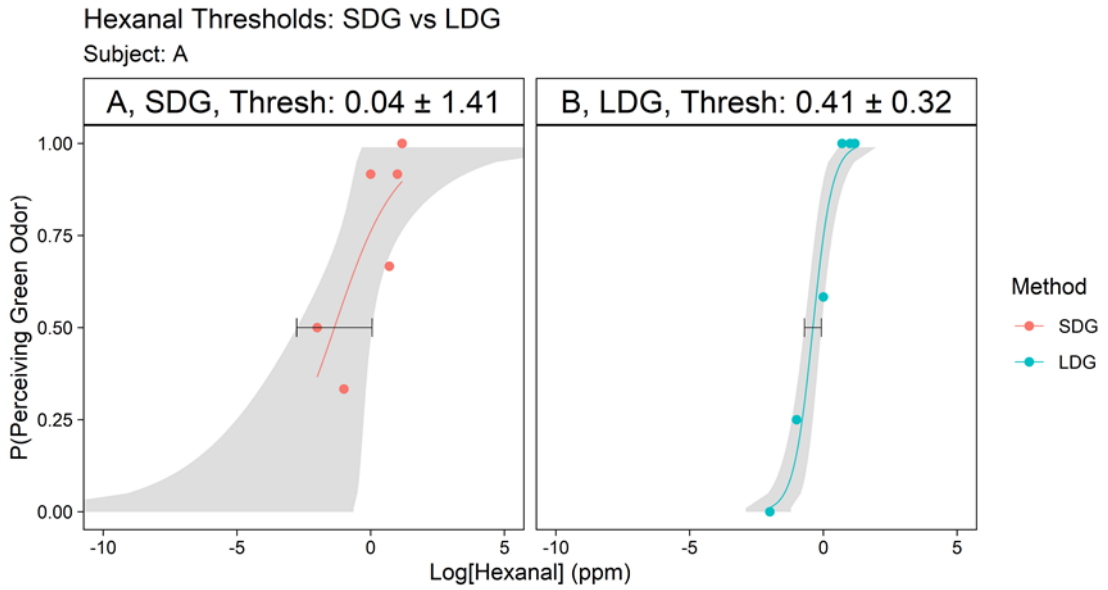


Figure 1.1.1. Solution reproducibility test -- HEX threshold measurement with a narrow range, subject A, day 1. *Solutions were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "green".*

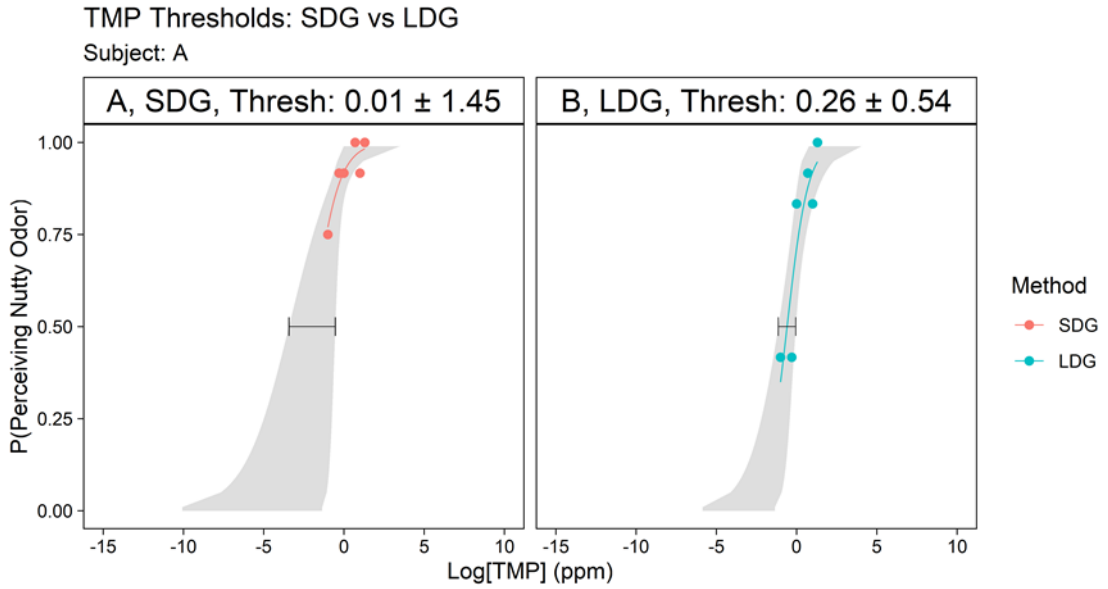


Figure 1.1.2. Solution reproducibility test -- TMP threshold measurement with a narrow range, subject A, day 1. *Solutions were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

Replication

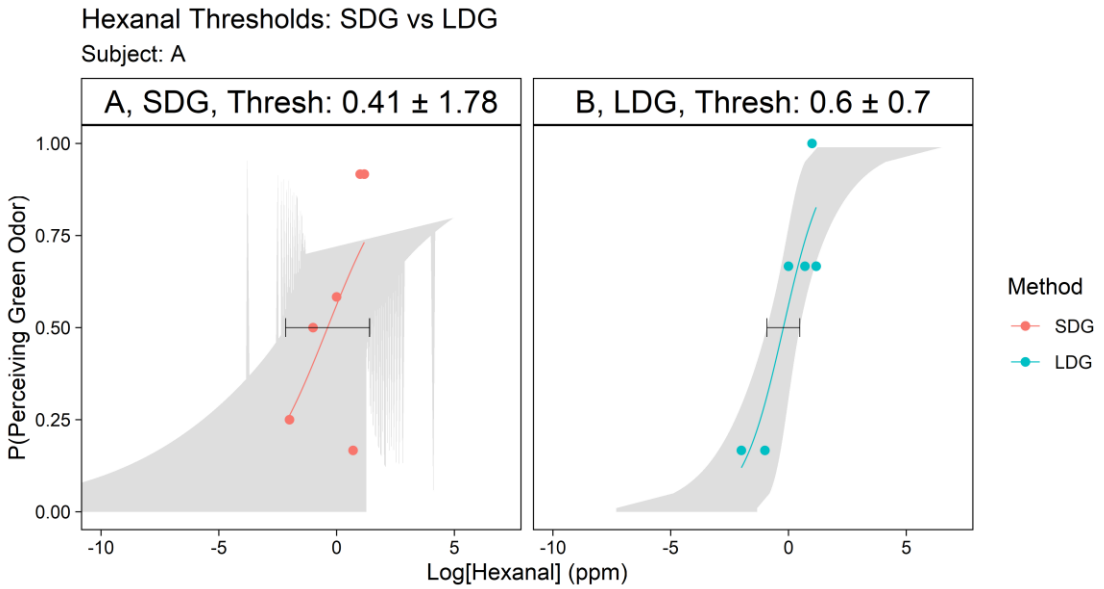


Figure 1.1.3. Solution reproducibility test -- HEX threshold measurement with a narrow range, subject A, replication. *Solutions were recycled from day 1. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

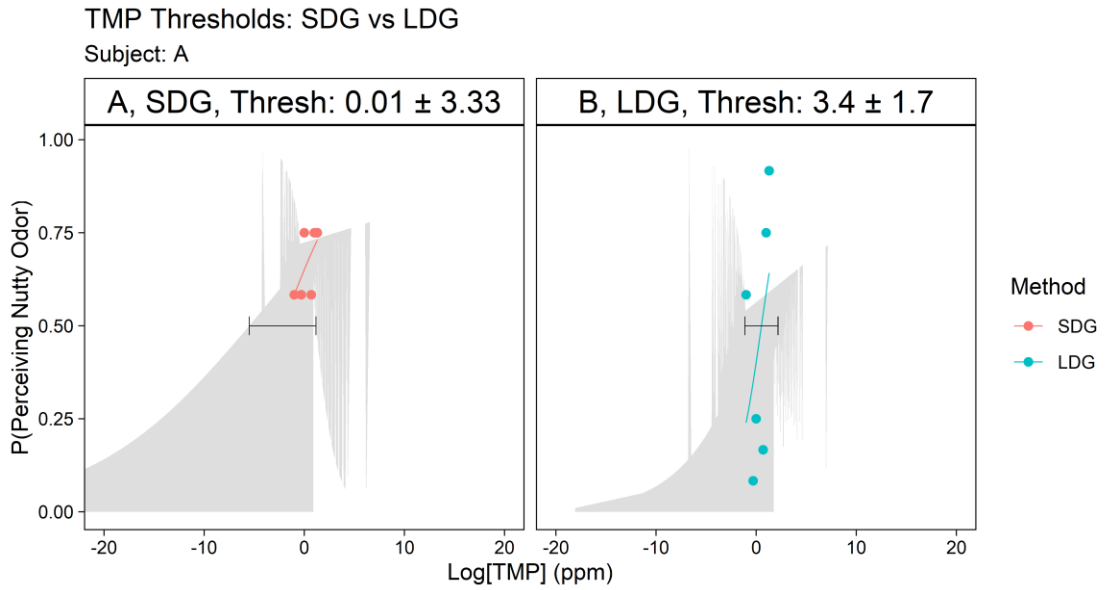


Figure 1.1.4. Solution reproducibility test -- TMP threshold measurement with a narrow range, subject A, replication. *Solutions were recycled from day 1. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

1.2 *Wider Range with Higher Concentrations (100 ppm)*

Day 1

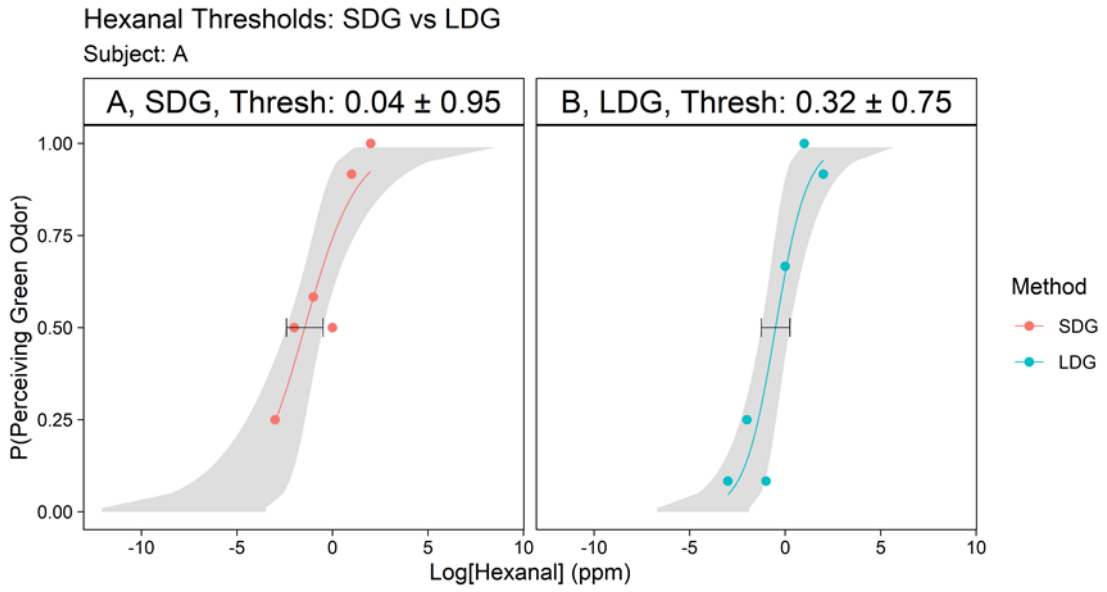


Figure 1.2.1. Solution reproducibility test -- HEX threshold measurement with a high concentration (100 ppm), subject A, day 1. *Solutions were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

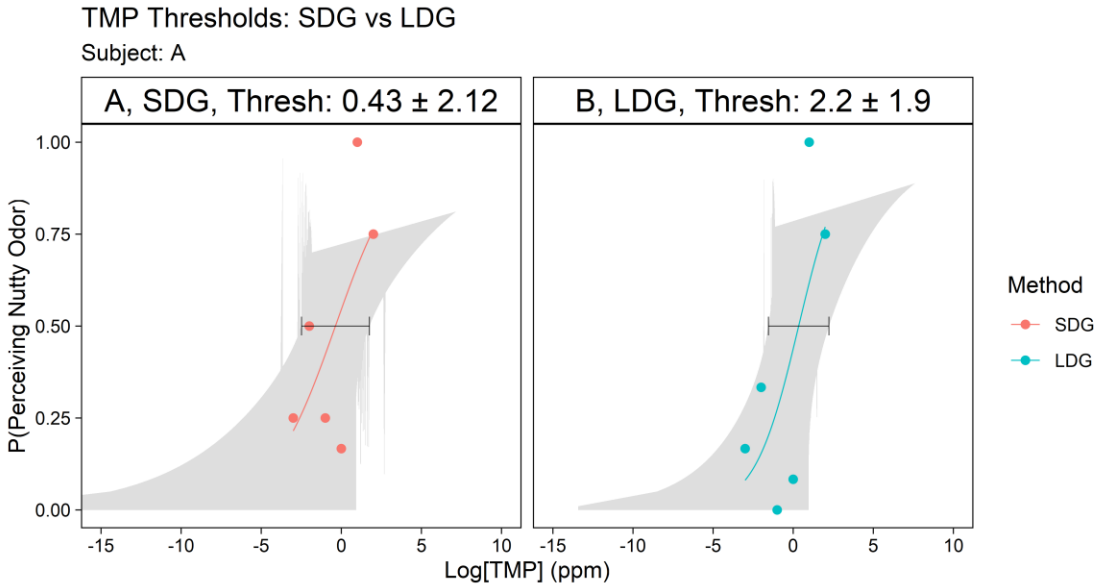


Figure 1.2.2. Solution reproducibility test -- TMP threshold measurement with a high concentration (100 ppm), subject A, day 1. *Solutions were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

Replication

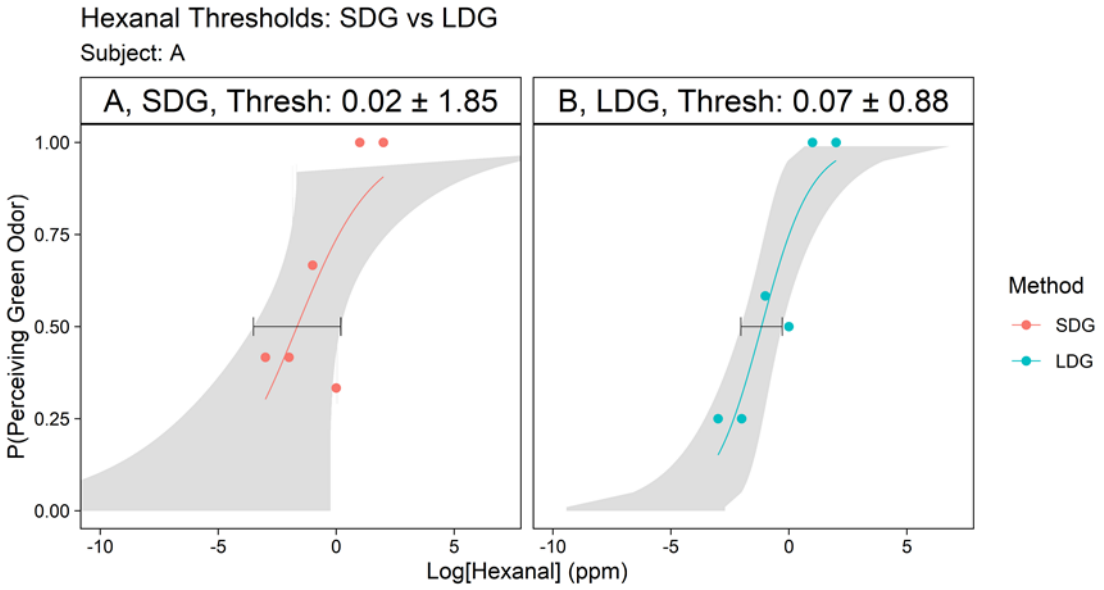


Figure 1.2.3. Solution reproducibility test -- HEX threshold measurement with a high concentration (100 ppm), subject A, replication. *Solutions were recycled from day 1. The x-axis*

is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".

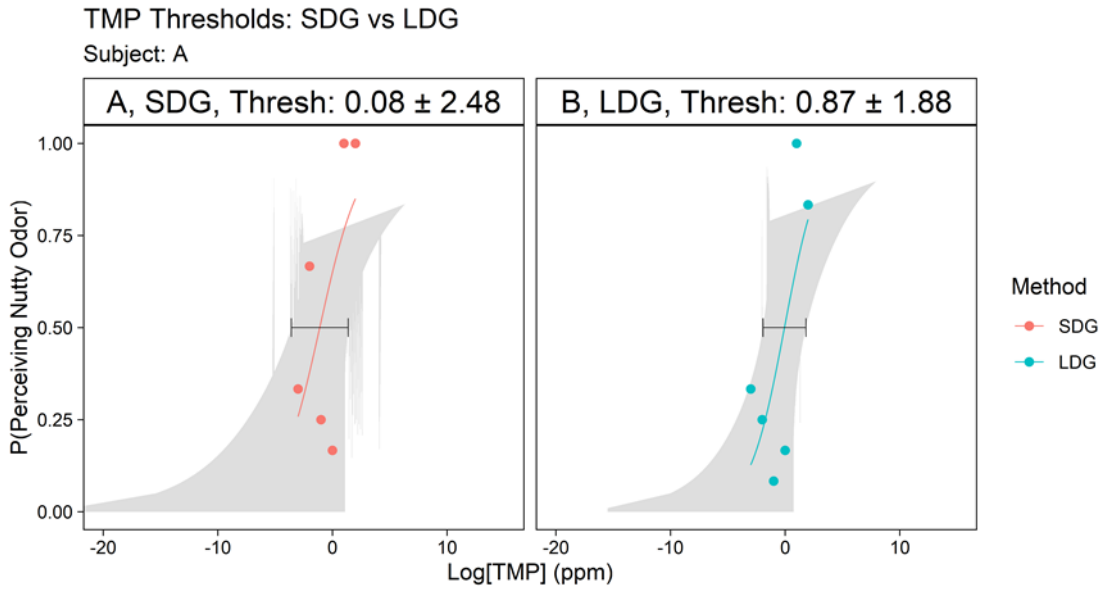


Figure 1.2.4. Solution reproducibility test -- TMP threshold measurement with a high concentration (100 ppm), subject A, replication. *Solutions were recycled from day 1. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

1.3 Wider Range with a Blank Sample (No Odorant)

Day 1

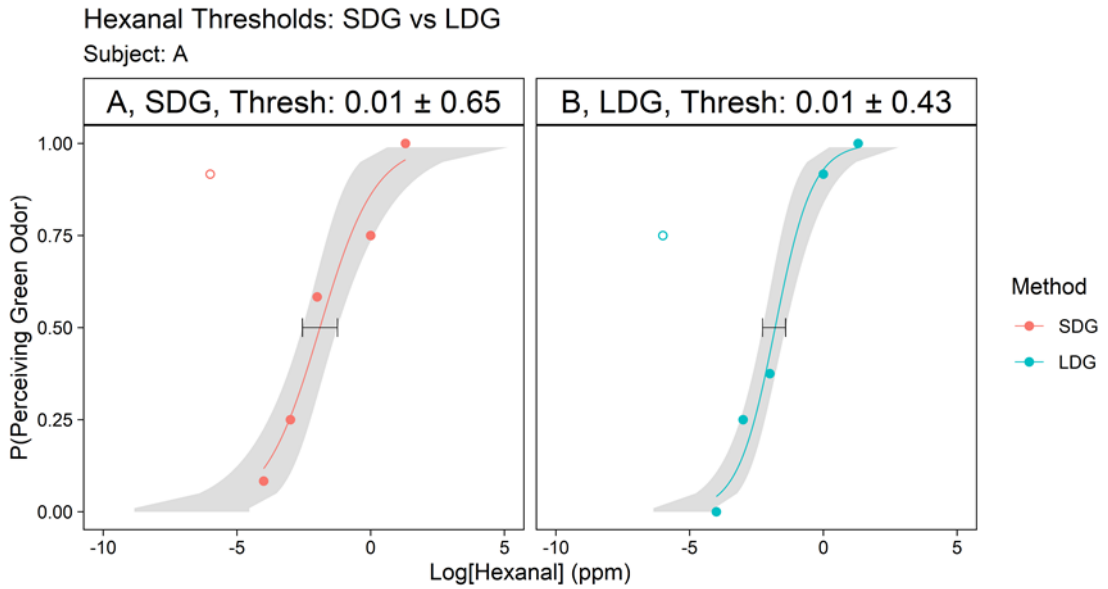


Figure 1.3.1. Solution reproducibility test -- HEX threshold measurement with a blank sample (no odorant), subject A, day 1. *Solutions were freshly made. The hollow point is when the blank sample (0 ppm) was exposed to the subject. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "green".*

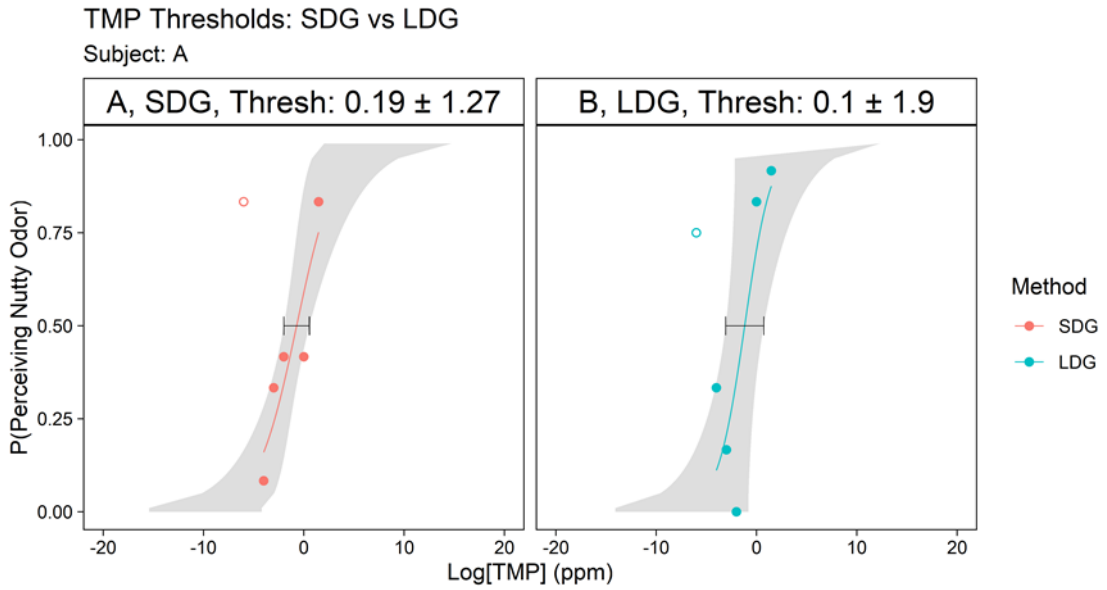


Figure 1.3.2. Solution reproducibility test -- TMP threshold measurement with blank sample (no odorants), subject A, day 1. *Solutions were freshly made. The hollow point is when the blank sample (0 ppm) was exposed to the subject. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

Replication

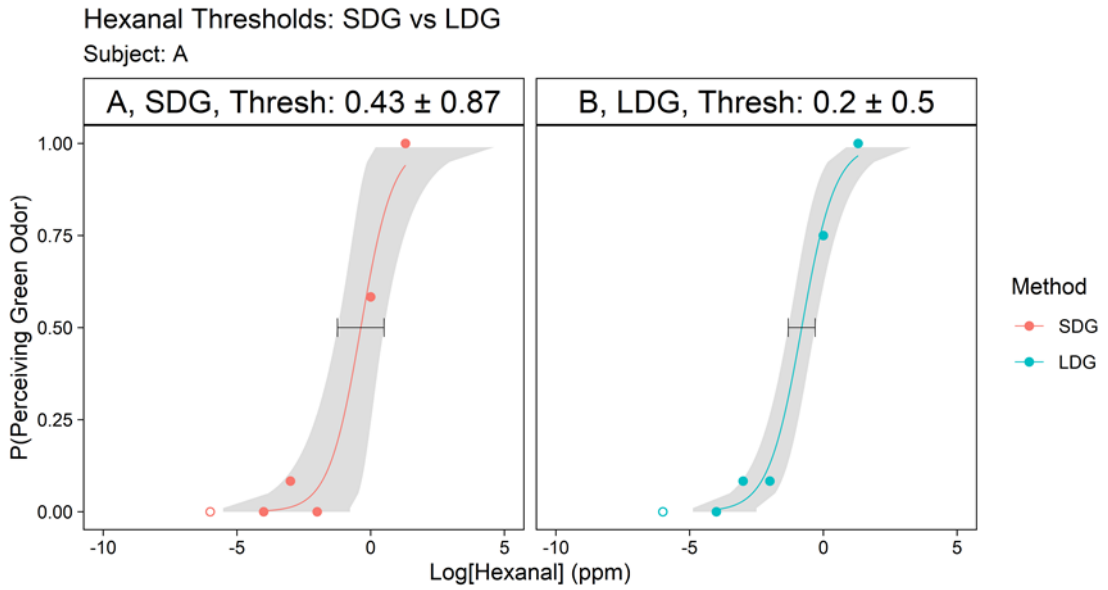


Figure 1.3.3. Solution reproducibility test -- HEX threshold measurement with blank sample (no odorants), subject A, replication. *Solutions were recycled from day 1. The hollow point is when the blank sample (0 ppm) was exposed to the subject. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "green".*

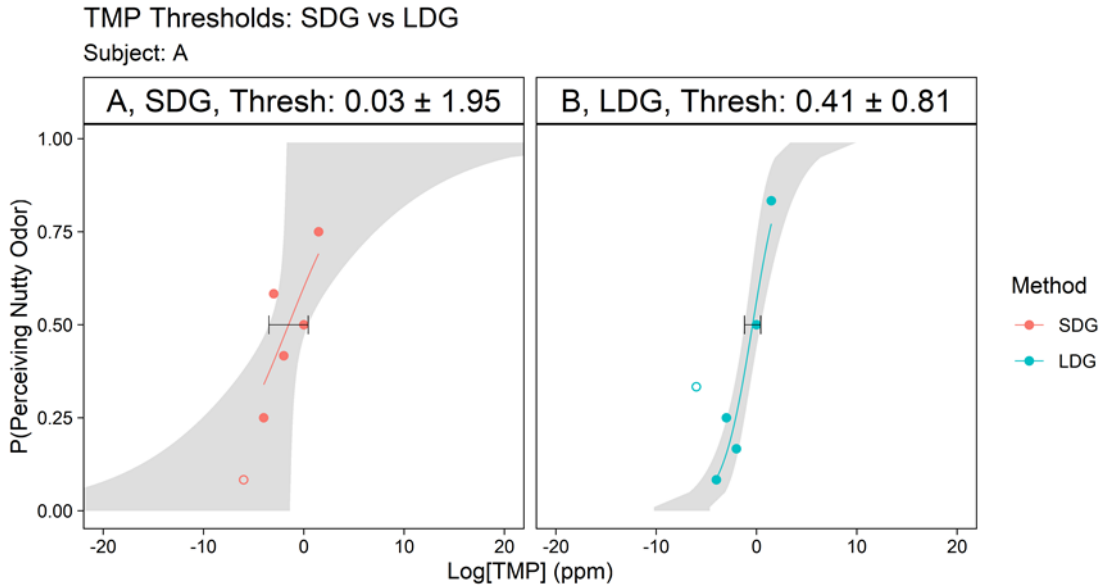


Figure 1.3.4. Solution reproducibility test -- TMP threshold measurement with blank sample (no odorants), subject A, replication. *Solutions were recycled from day 1. The hollow point is when the blank sample (0 ppm) was exposed to the subject. The x-axis is the log10 of the sample*

concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".

2. Narrow Range of Concentrations

2.1 Day 1

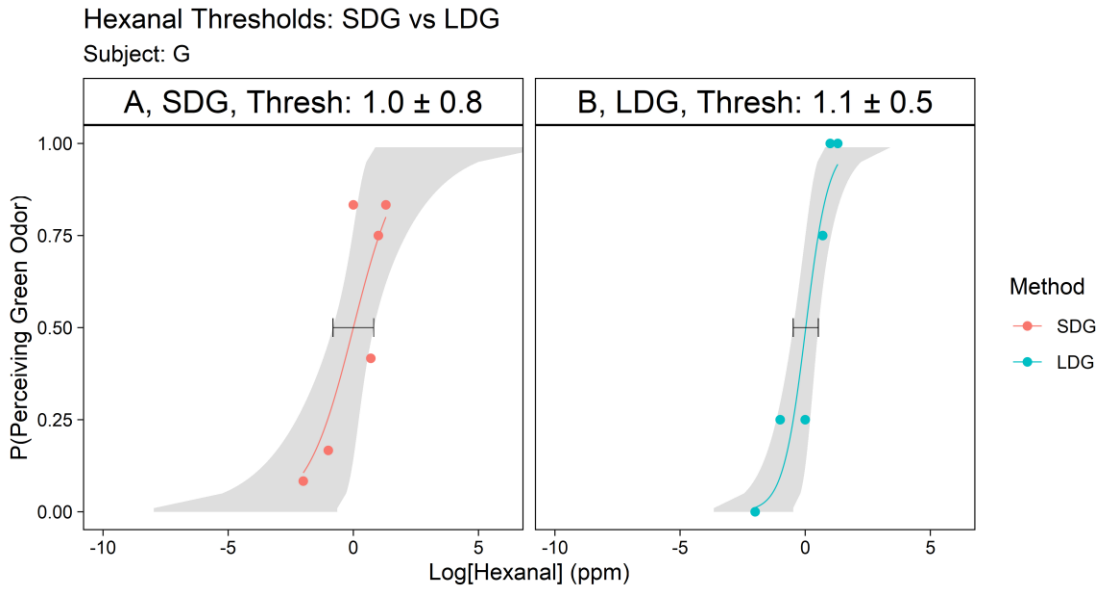


Figure 2.1.1. HEX threshold measurement with narrow range, subject G, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "green".*

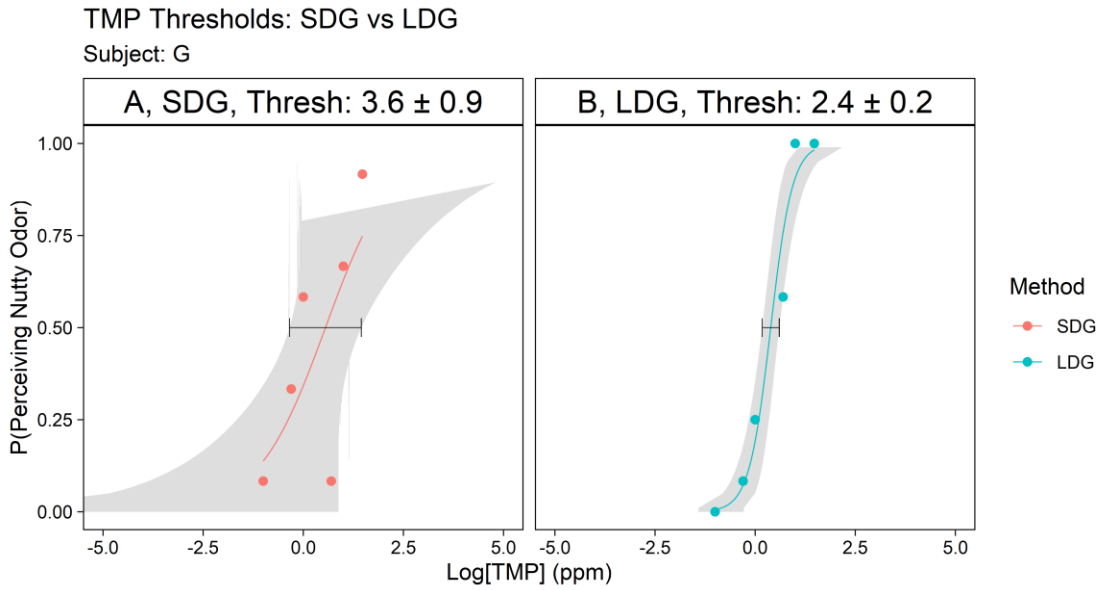


Figure 2.1.2. TMP threshold measurement with narrow range, subject G, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

2.2 Replication

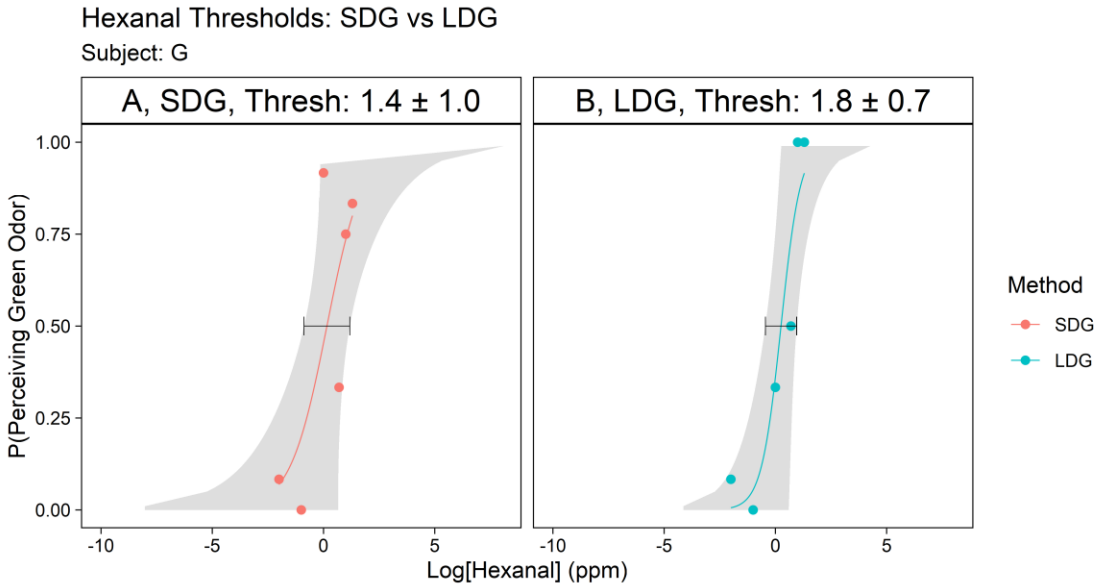


Figure 2.2.1. HEX threshold measurement with narrow range, subject G, replication. *The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, as he identified the sample as "green".*

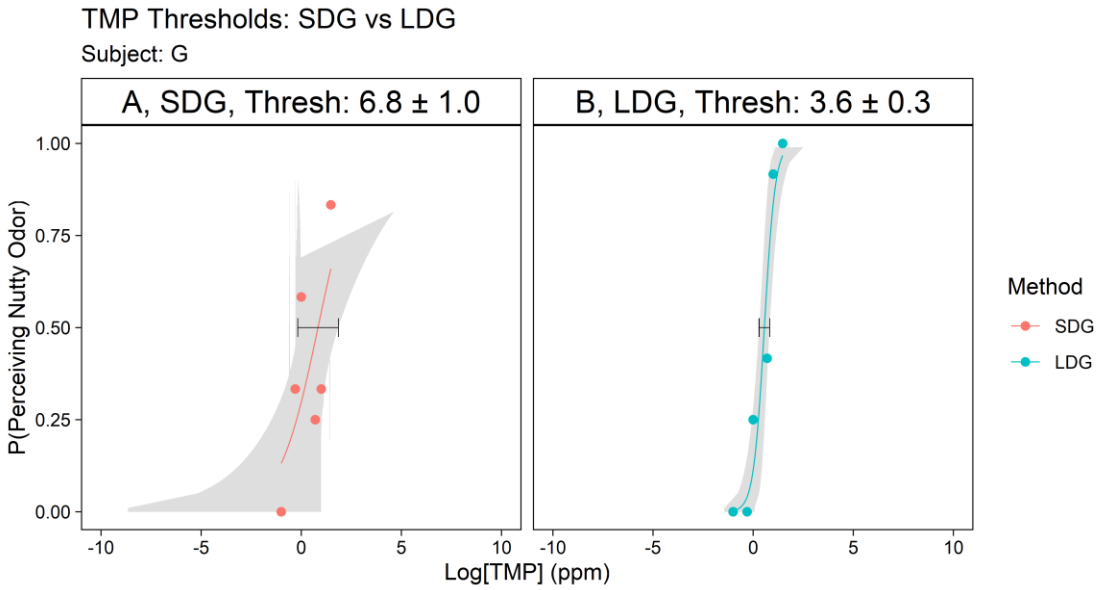


Figure 2.2.2. TMP threshold measurement with narrow range, subject G, replication. *The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, he identified the sample as "nutty".*

3. Wider Range with Higher Concentrations (100 ppm)

3.1 Day 1

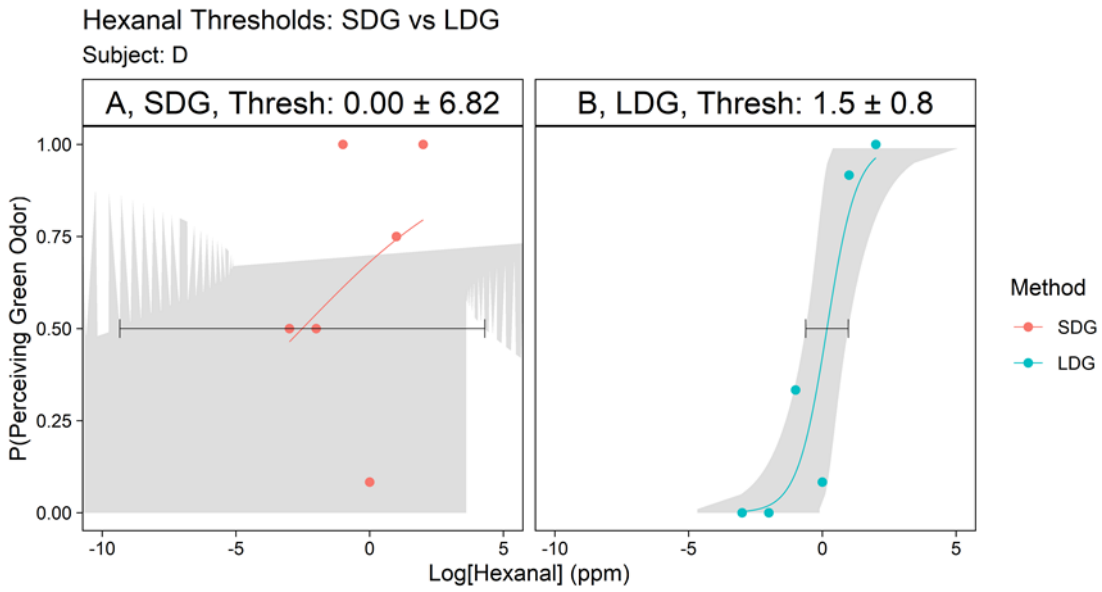


Figure 3.1.1. HEX threshold measurement with a higher concentration (100 ppm), subject D, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

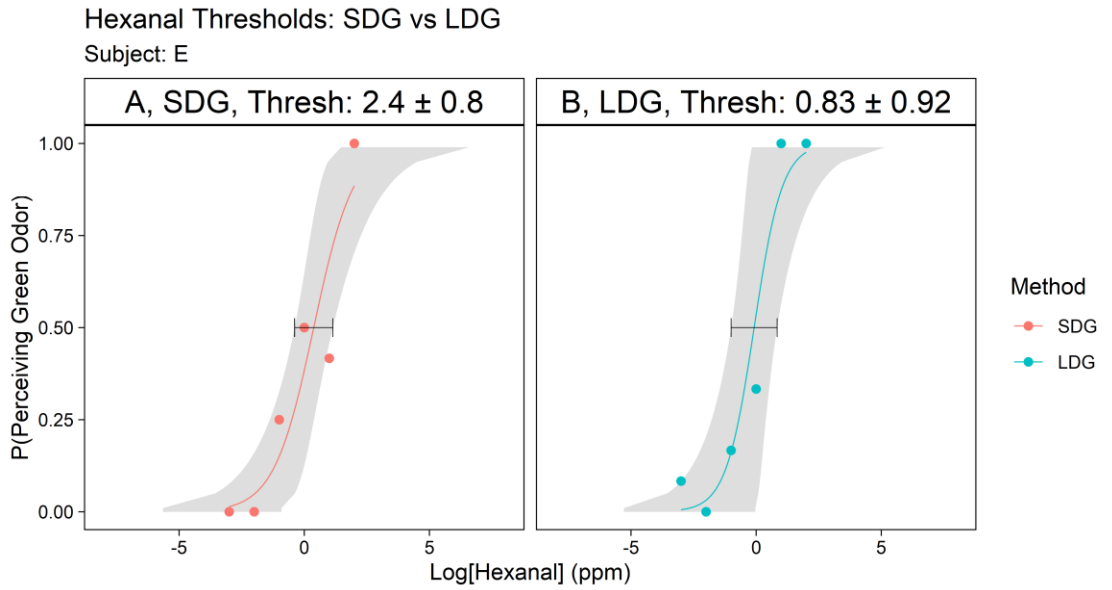


Figure 3.1.2. HEX threshold measurement with a higher concentration (100 ppm), subject E, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

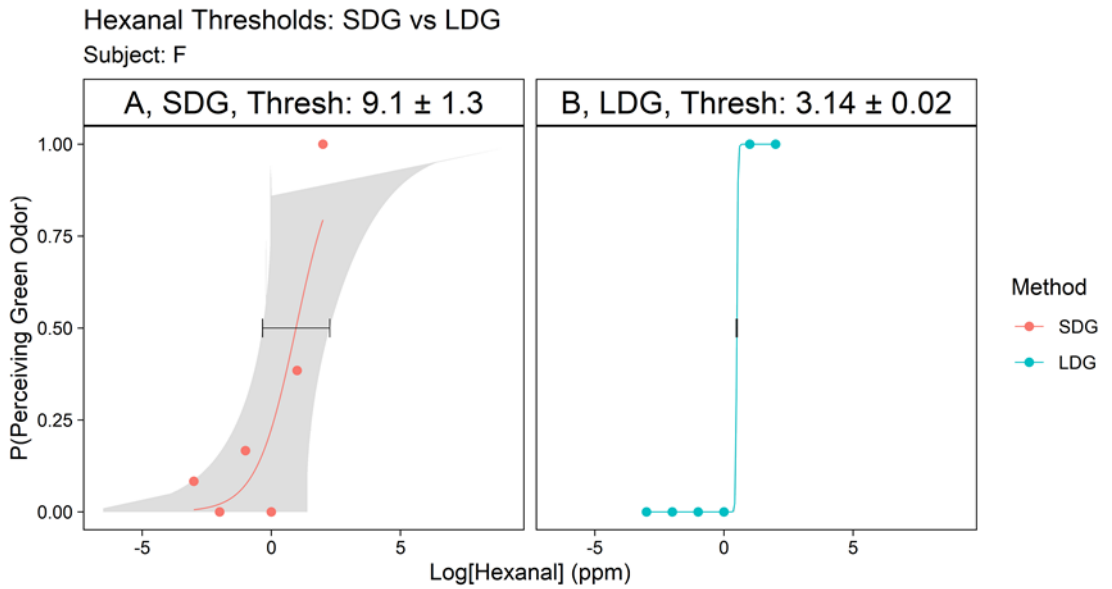


Figure 3.1.3. HEX threshold measurement with a higher concentration (100 ppm), subject F, day 1. *x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

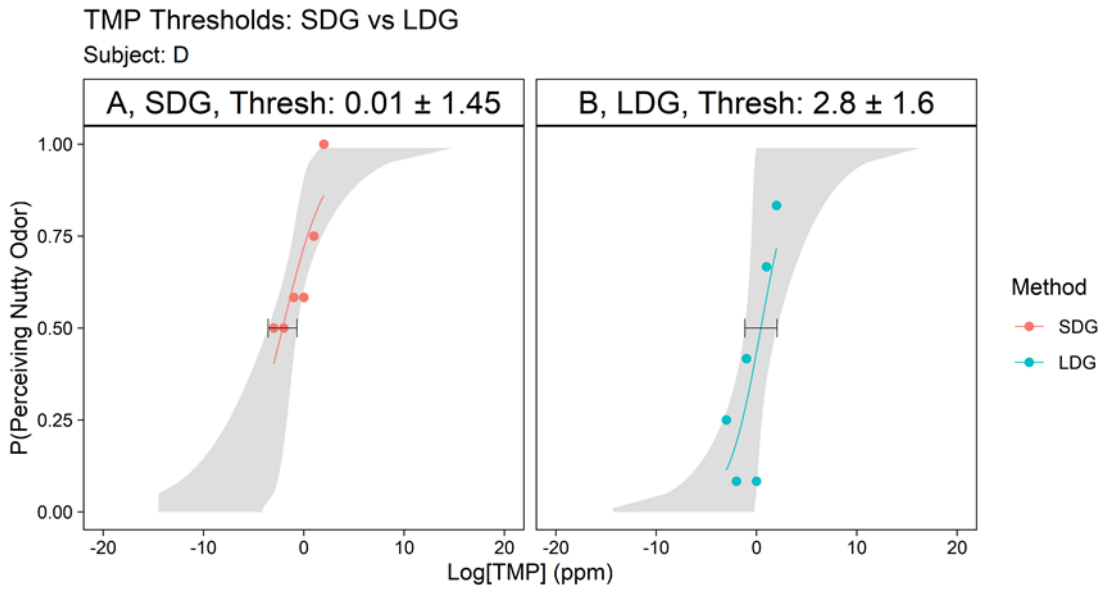


Figure 3.1.4. TMP threshold measurement with a higher concentration (100 ppm), subject D, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

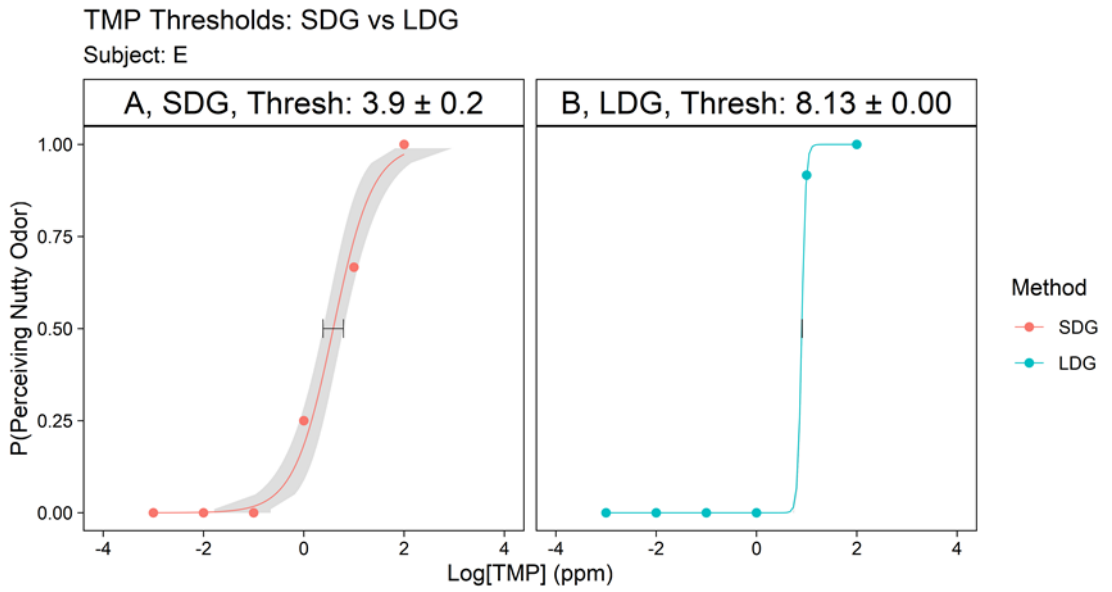


Figure 3.1.5. TMP threshold measurement with a higher concentration (100 ppm), subject E, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

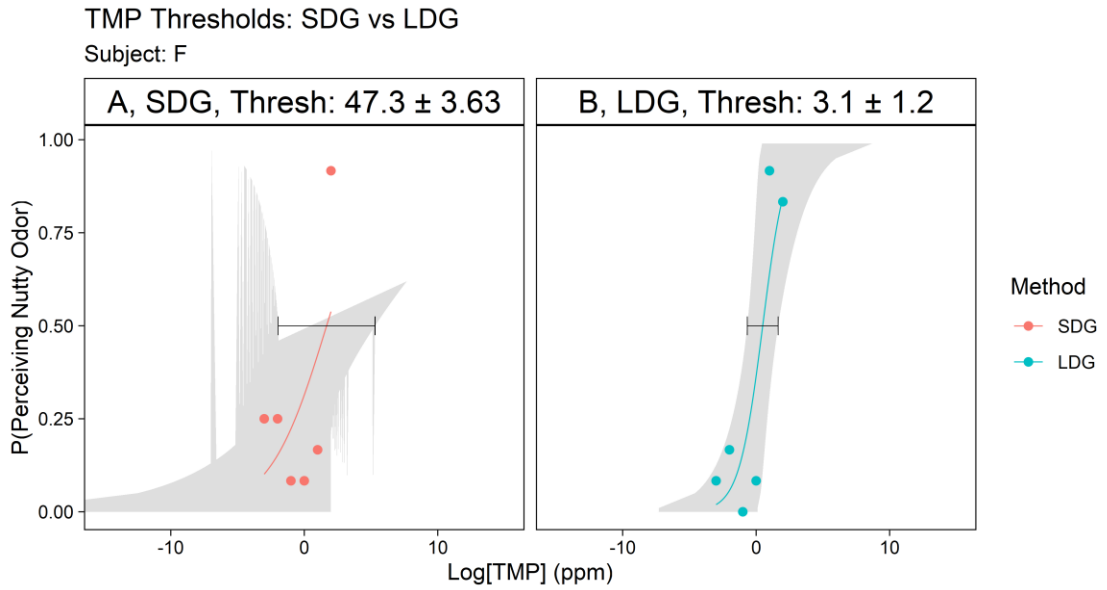


Figure 3.1.6. TMP threshold measurement with a higher concentration (100 ppm), subject F, day 1. *The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

3.2 Replication

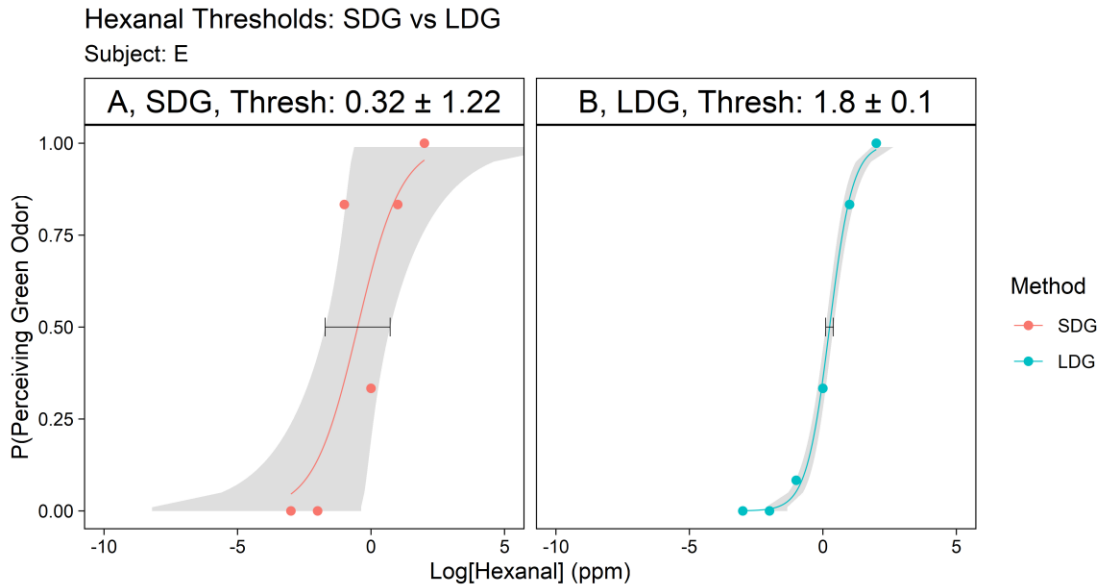


Figure 3.2.1. HEX threshold measurement with a higher concentration (100 ppm), subject E, replication. *The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

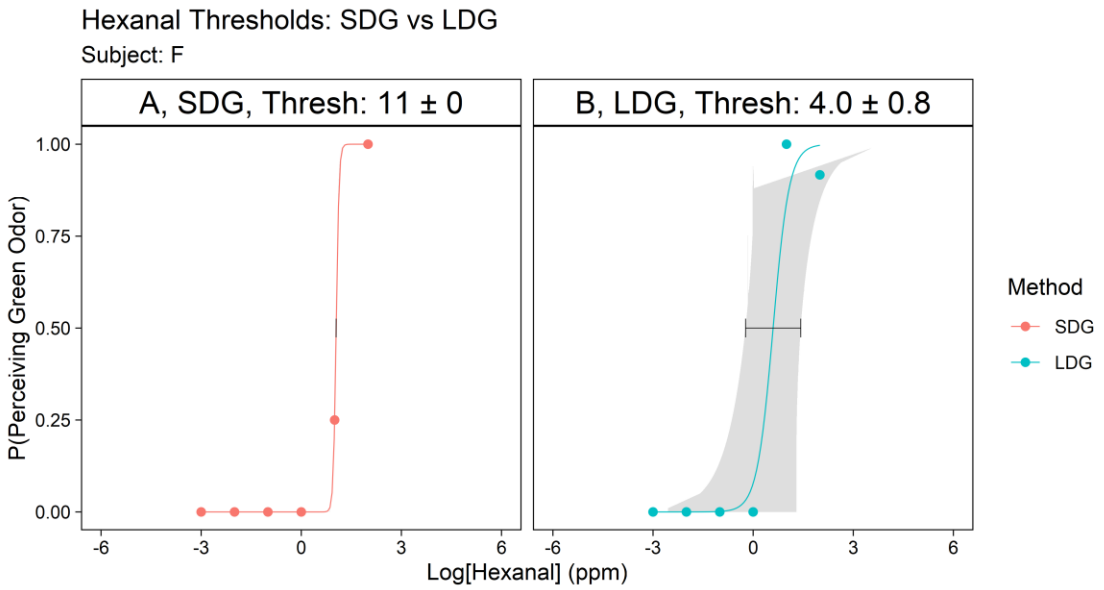


Figure 3.2.2. HEX threshold measurement with dense concentrations, subject F, replication. *The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

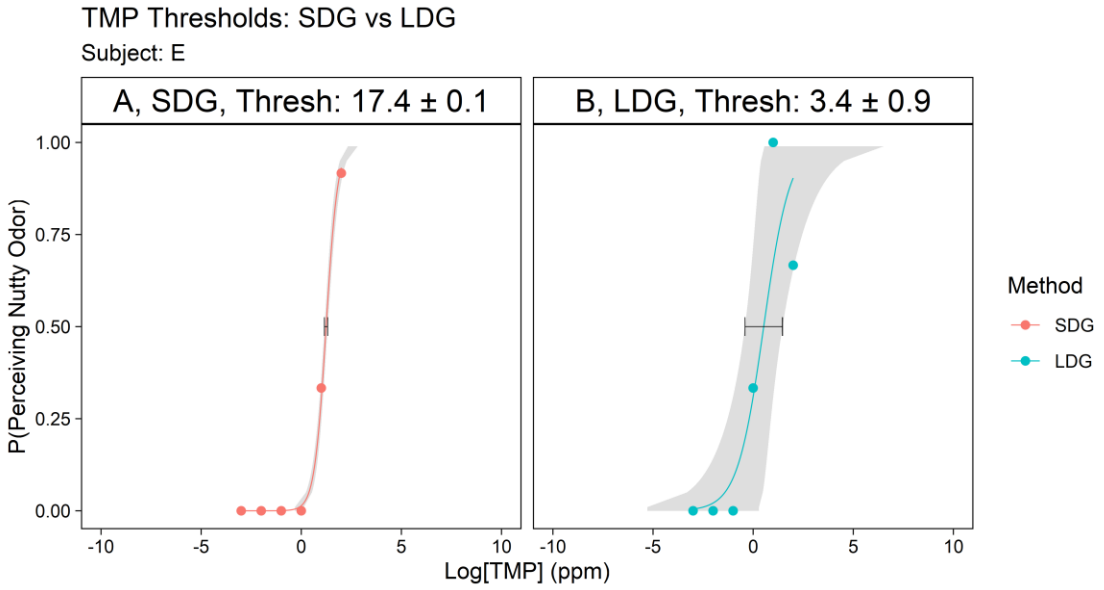


Figure 3.2.3. TMP threshold measurement with dense concentrations, subject E, replication. *The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

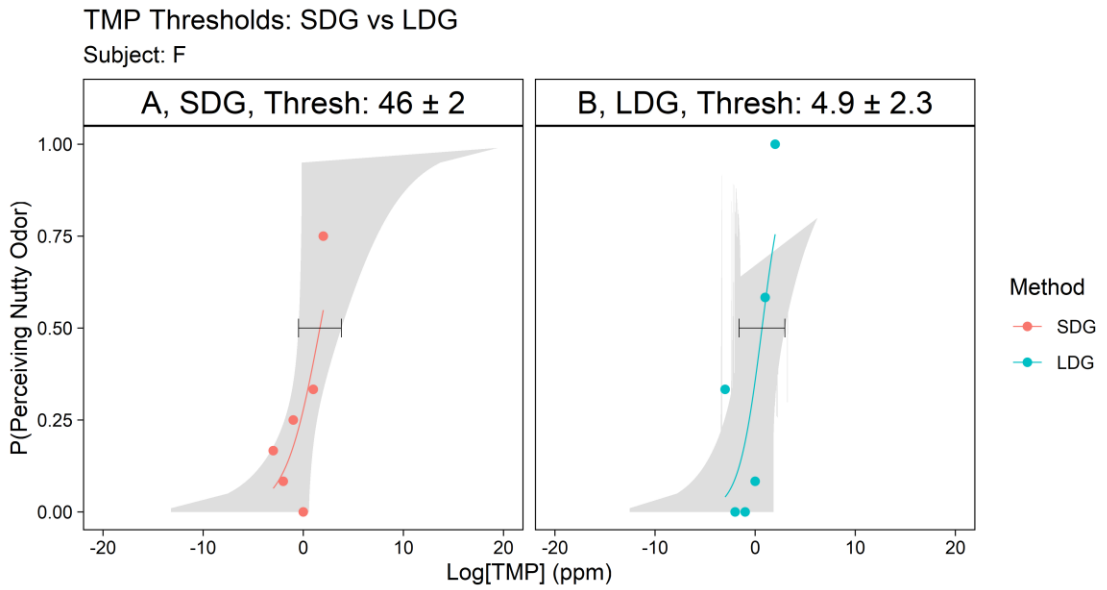


Figure 3.2.4. TMP threshold measurement with a higher concentration (100 ppm), subject F, replication. *The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

4. Wider Range with Blank Sample (No Odorant)

4.1 Day 1

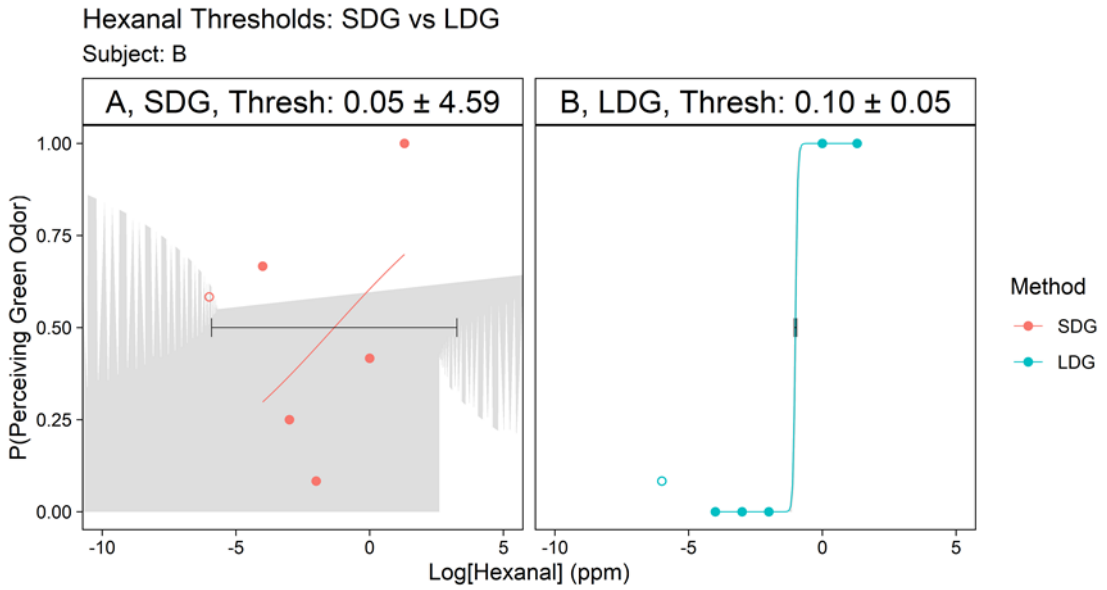


Figure 4.1.1. HEX threshold measurement with a blank sample (no odorant), subject B, day 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

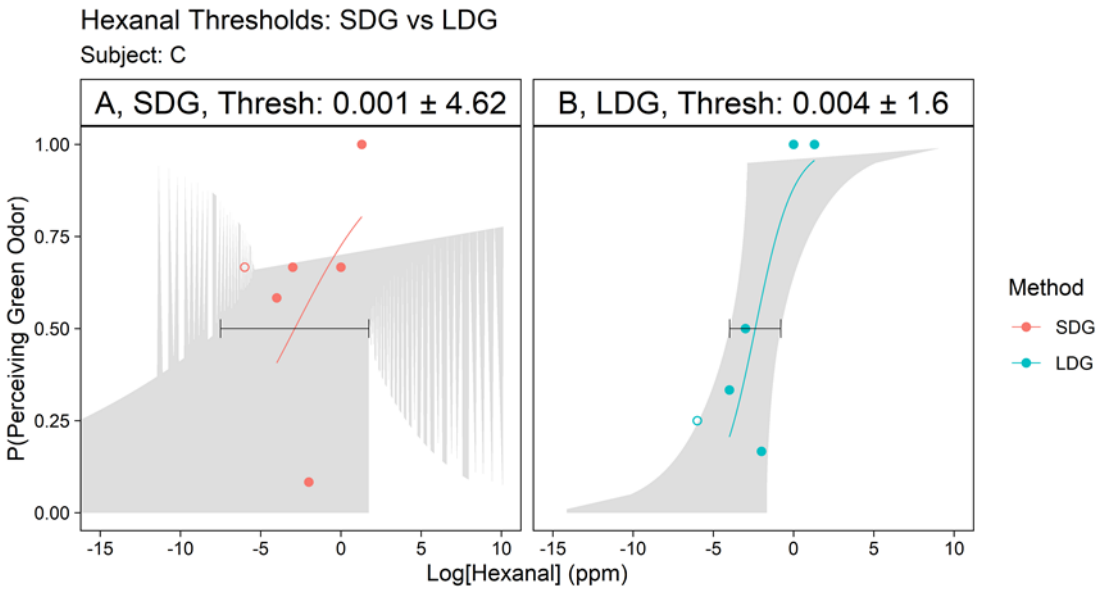


Figure 4.1.2. HEX threshold measurement with a blank sample (no odorant), subject C, day 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

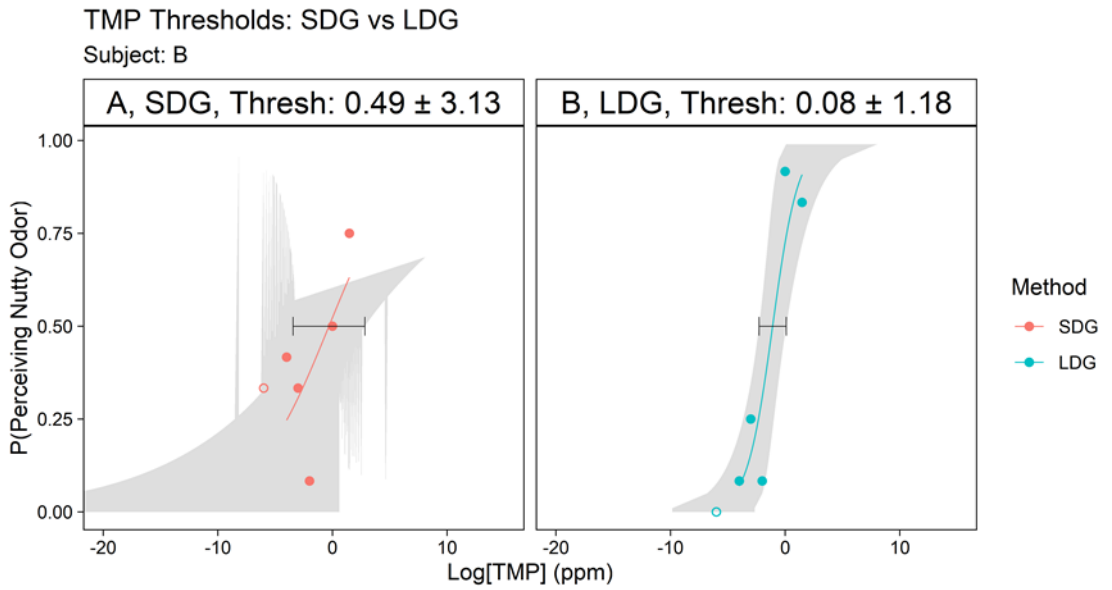


Figure 4.1.3. TMP threshold measurement with a blank sample (no odorant), subject B, day 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

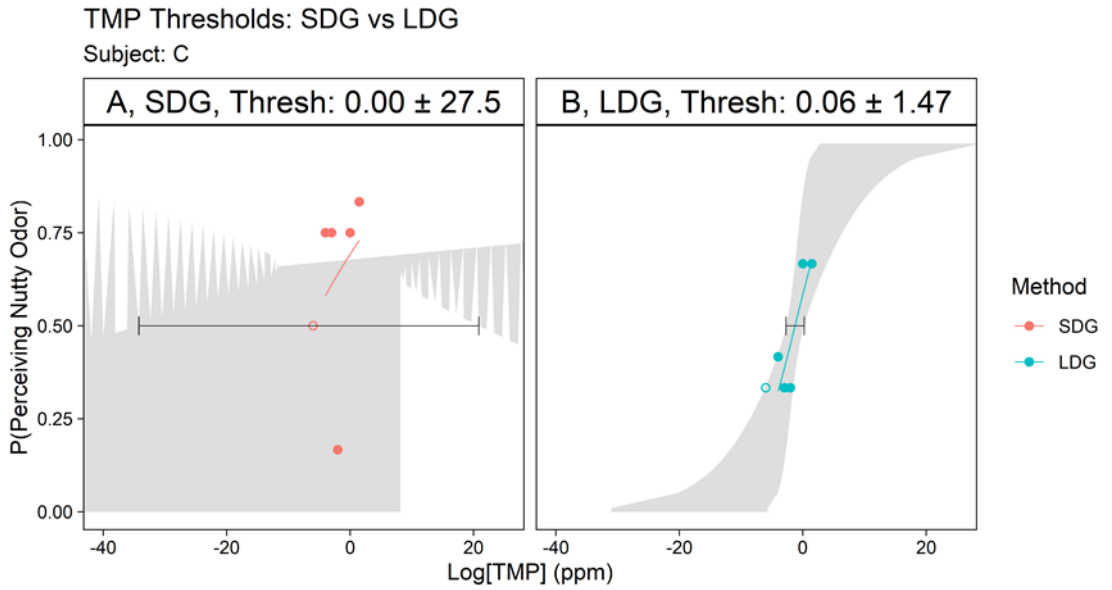


Figure 4.1.4. TMP threshold measurement with a blank sample (no odorant), subject C, day 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

4.2 Replication 1

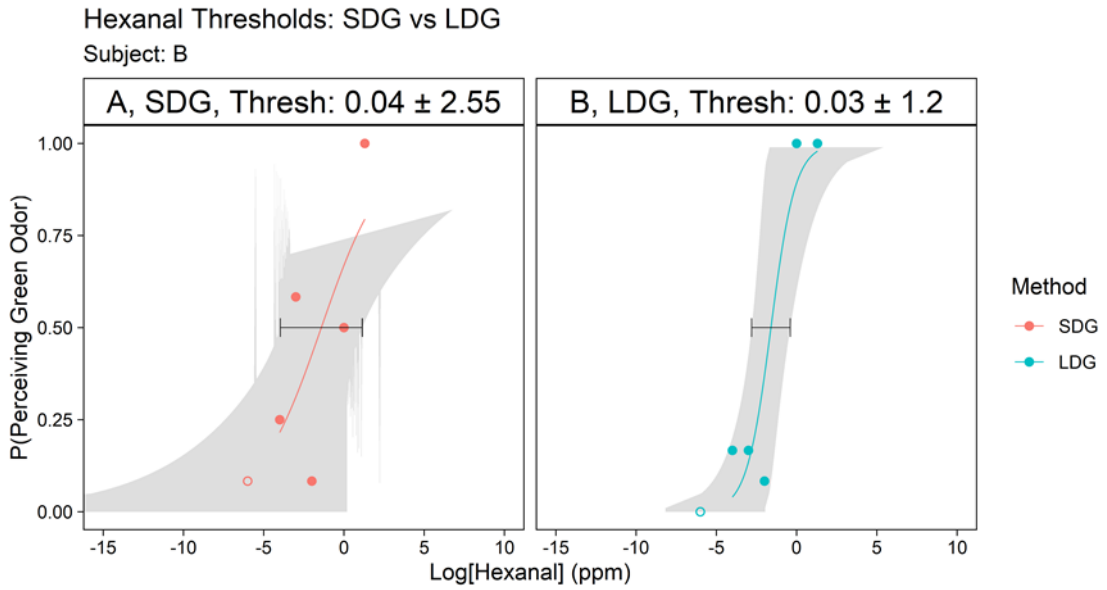


Figure 4.2.1. HEX threshold measurement with a blank sample (no odorant), subject B, replication 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

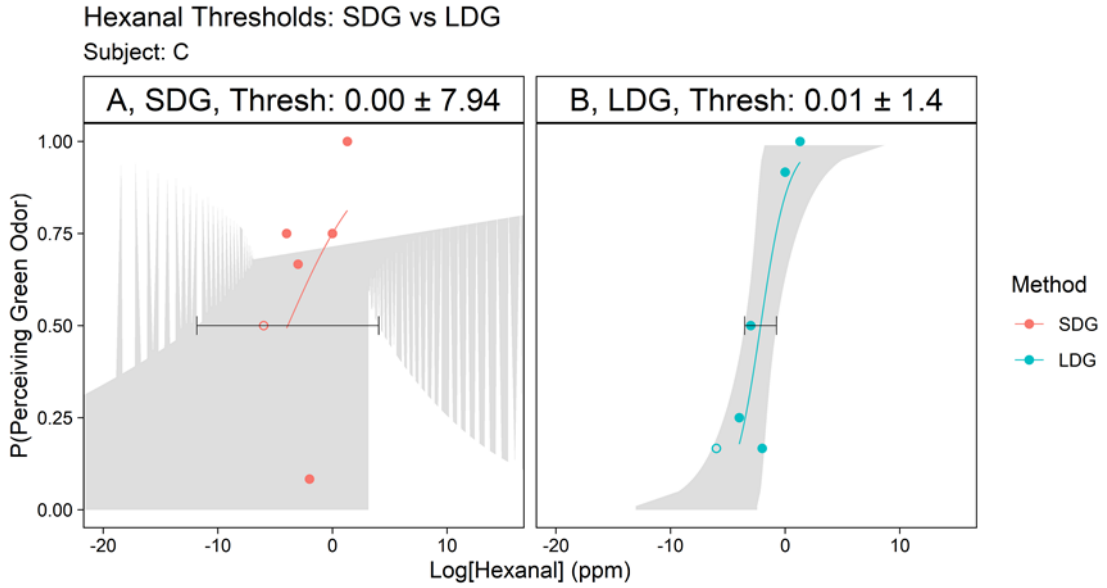


Figure 4.2.2. HEX threshold measurement with a blank sample (no odorant), subject C, replication 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

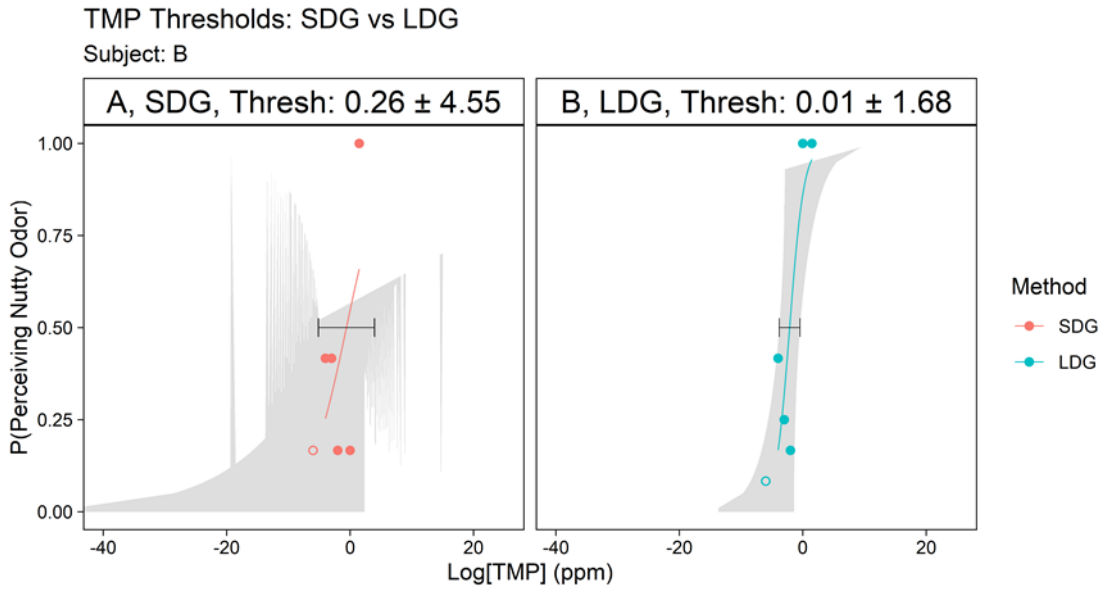


Figure 4.2.3. TMP threshold measurement with a blank sample (no odorant), subject B, replication 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

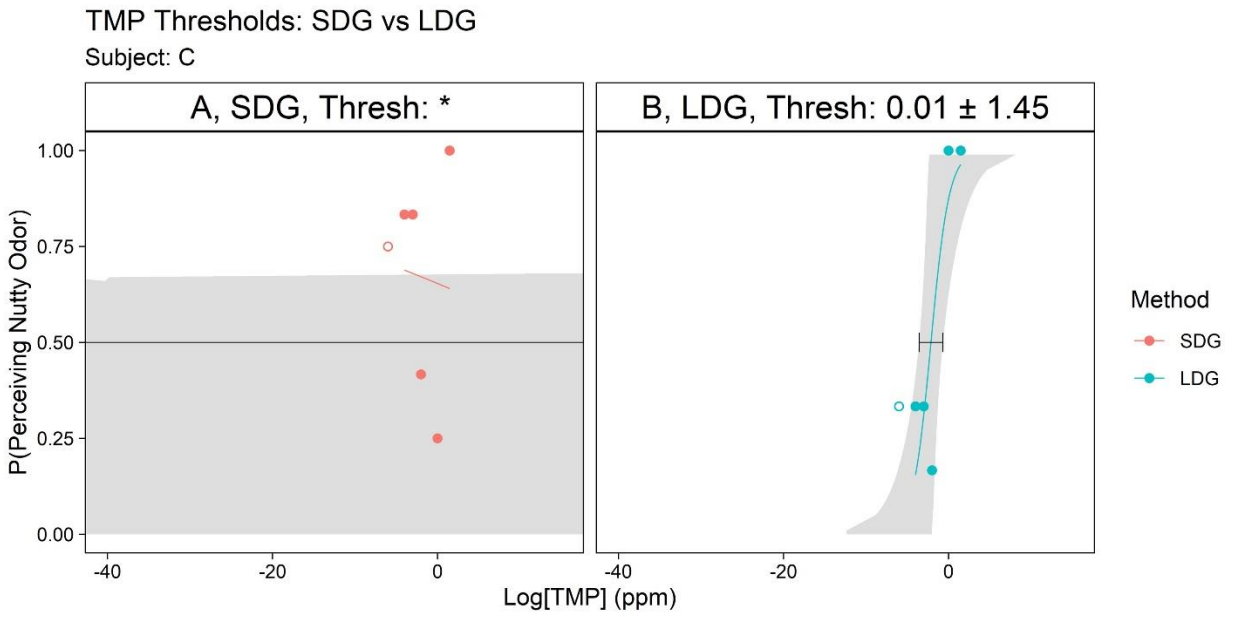


Figure 4.2.4. TMP threshold measurement with a blank sample (no odorant), subject C, replication 1. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty". * In this figure the calculated threshold was* 1.3×10^{16} *ppm, which is an unrealistic value. The possible explanation for that was that the subject was not able to systematically determine the concentrations in approximately their order of concentration with the SDG order.*

4.3 Replication 2

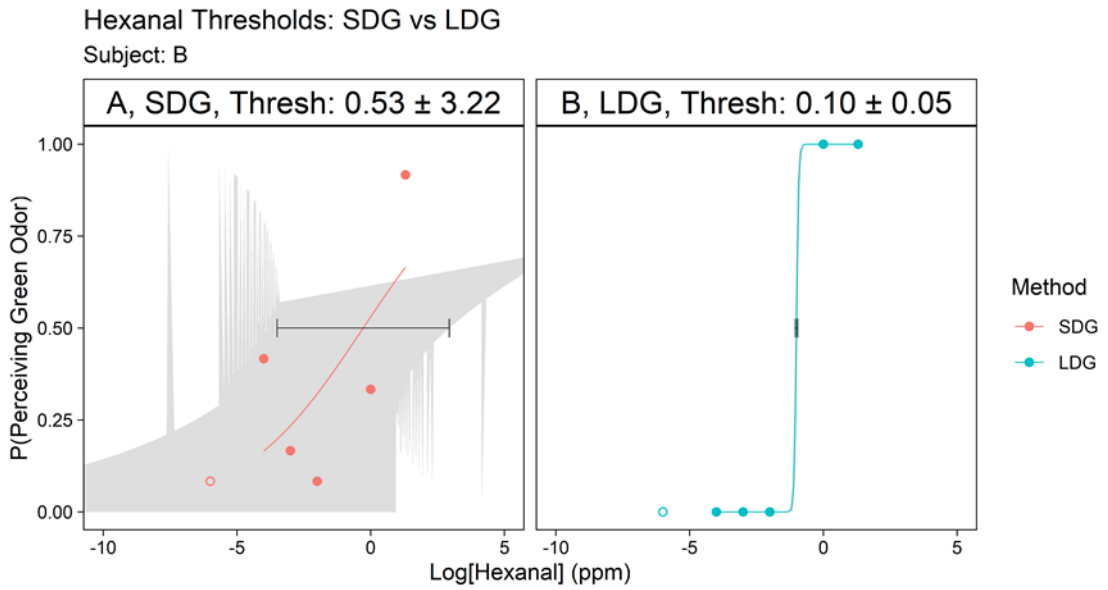


Figure 4.3.1. HEX threshold measurement with a blank sample (no odorant), subject B, replication 2. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

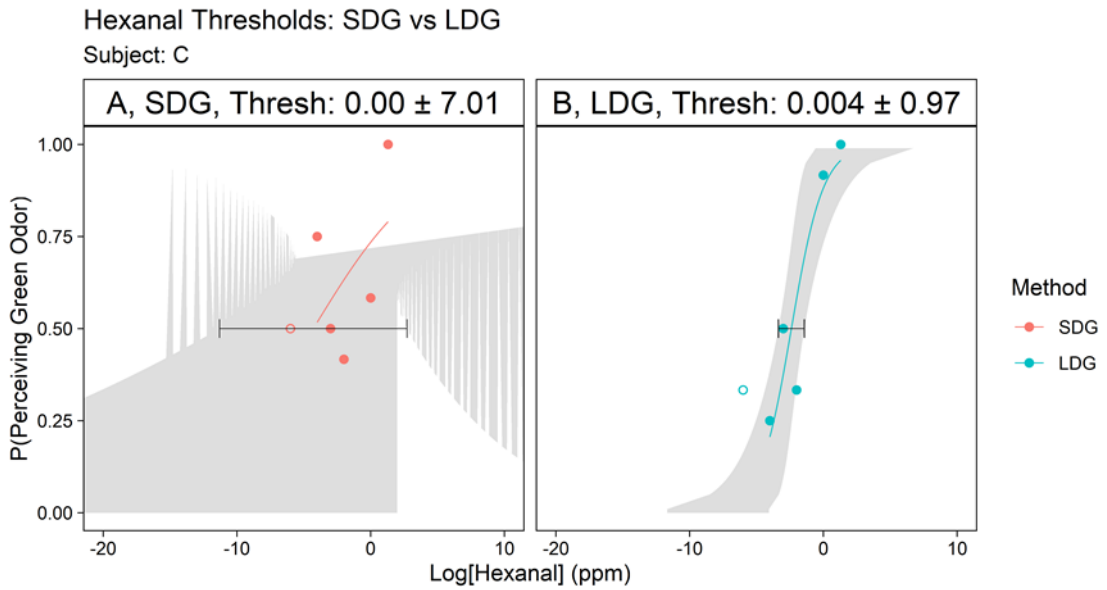


Figure 4.3.2. HEX threshold measurement with a blank sample (no odorant), subject C, replication 2. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "green".*

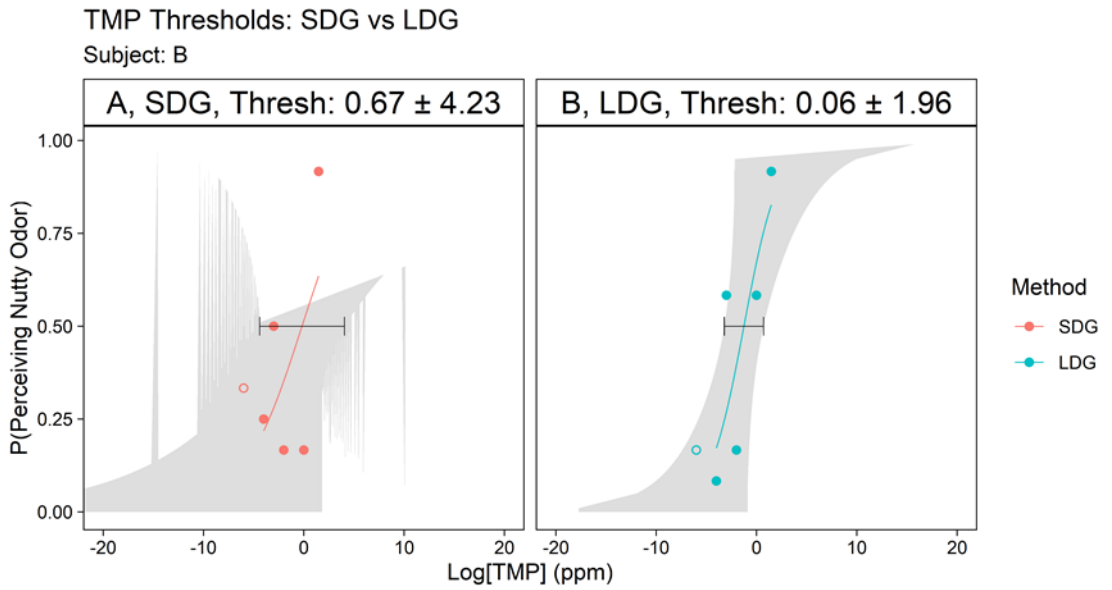


Figure 4.3.3. TMP threshold measurement with a blank sample (no odorant), subject B, replication 2. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*

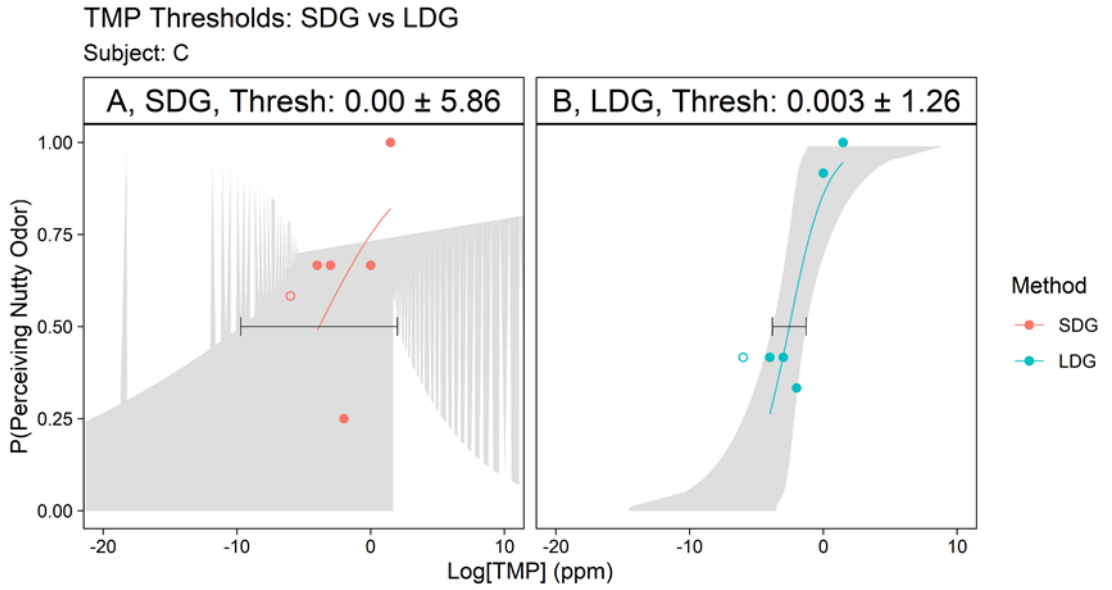


Figure 4.3.4. TMP threshold measurement with a blank sample (no odorant), subject C, replication 2. *The hollow point is when the blank sample (0 ppm) was exposed to the subject. The samples were freshly made. The x-axis is the log10 of the sample concentrations (in ppm), and the y-axis is the probability when the subject perceived a sample, she identified the sample as "nutty".*