The influence of ambient color and odor on Stroop task performance and vigilance

An exploratory study

Charley Gielkens Thijs Urlings Frank Baars

Paul Bouchaut Ferdy van Varik Melissa Schoon

Utrecht University

Faculty of Science

Department of Information and Computing Sciences

{cgielken;murlings;fjmbaars;pboucha;fjmvarik;mschoon}@cs.uu.nl

ABSTRACT

Literature has shown a low level link between color and odor. In this paper, we tried to measure the effects of color and odor on user alertness. Secondly, we tried to find out if this effect differs for congruent and incongruent color-odor combinations. Participants performed two different tasks in four different combinations of color and odor. The first task was vigilance task to measure alertness. The second task was a modified Stroop task. The experiments took place in two small rooms with a fixed odor and variable color. Accuracy and reaction times of both tasks were analyzed, but we found neither significant differences nor a Stroop effect. We did find that purple ambient color lowered the reaction time for both tasks. The results are likely influenced by the low number of participants, not perfect consistency in ambient color and odor and the use of a keyboard as input.

1. INTRODUCTION

Hardly ever do we rely upon just one of our five senses. In most cases we use multiple senses to create an interpretation of the world around us, which allows us to interact with it. Even when only one sense is used, interference effects can occur if multiple channels are used or when one channel contains multiple types of information (Hilliset al., 2002). A well known example of this is the Stroop effect. This effect can be shown by asking a subject to read a word, which is a certain color (e.g. red). The word is written in a certain color (e.g. green). The task of the subject is to say what color the word is, so in this case the right answer is green. Their first instinct will be to say the word, rather than the color in which it is displayed (Stroop, 1935; MacLeod, 1991).

Another sense that can cause interference in other perceptions is smell. For example, a white wine colored red artificially will be most likely described with words generally associated with red wine rather than white one (Morrot et al., 2001). In other experiments modified Stroop tasks have been performed to see how smell and other senses interact. Paulli et al. (1999) for example researched the relation between pleasant and unpleasant odors and words. During the experiment they primed subjects with either pleasant or unpleasant odors. The subjects were consecutively shown cards with either pleasant or unpleasant descriptive words. They found that the presence of odor interfered with the performance on odor-congruent Stroop cards.

Another experiment that also used a modified Stroop task is the research of White and Prescott (2007) about the relation between odor and taste. They presented both congruent and incongruent matches between odors and tastes to the subjects. The subjects were asked to identify the tastant as either sweet or sour. They found that participants performed better when the stimuli were congruent (e.g. a sweet sense with the smell of strawberries). In this experiment we also work with congruent and incongruent stimuli, as can be seen in the rest of this paper.

Demattè et al. (2009) researched how visual stimuli influenced odor perception by presenting participants with a patch of color and a short burst of an odorant that was either congruent or incongruent. They found that the visual stimuli always influenced the odor perception, even when participants were instructed to ignore them.

Even though all these kinds of modified Stroop tasks have been used in multisensory perception research, we were not able to find articles that research how ambient odors and colors influence performance on regular Stroop tasks. We expect that with congruent odor and color combinations users will react quicker with a higher success rate to stimuli presented in the same color as the ambient color. Therefore our first proposed hypothesis is:

> **H1:** With congruent ambient odor and color combinations, participants will have a lower reaction time when presented with a stimulus that is congruent in color with their surroundings.

One of the other things this paper tries to prove is the influence of a specific odor on the mental state of a subject. Vernet-Maury et al. (1999) found in a comparative study between five different odorants that lavender causes mainly positive emotional reactions. Other studies also showed that lavender can calm people (Ludvigson & Rottman, 1989; Buchbauer et al. 1991) and improve the mood according to Moss et al. (2003) and Diego et al. (1998). Although lavender improved the mood of the participants in the experiment by Moss et al. (2003), performance in tasks based on memory were significantly lower. Also performance in tasks based on attention or working memory were seriously impaired.

For the orange odor Lehrner et al. (2000) found that it reduced anxiety in females, when being at the dentist's office, but this result was not found with males. Given the previous evidence about lavender odor they looked into the effects of lavender on anxiety at the dentists in their 2005 study. They also used orange odor to compare their previous findings. Their conclusion was that both lavender and orange calmed people down, compared to a control group who were not exposed to any odor.

Another study to different odors is that of Baron & Thomley (1994). They found that lemon and floral odors had a positive effect on performance when performing cognitive tasks. Both lemon and lavender are generally described as pleasant smells, but lavender seems to be more likely to decrease performance while lemon should show a slight increase in performance. (Millot, Brand &Morand, 2002)

These different claims in the literature left us wondering how odors actually influence performance when performing a vigilance task. Based on the literature we found, we assume that orange makes users more vigilant while lavender calms them down. In order to test these assumptions, we have formulated the following hypothesis:

> **H2:** When performing the vigilance task in an incongruent room, performance will be better in the room with the orange odor.

And the following two hypotheses test that if a task is performed in a congruent room this will lead to better results than when it is performed in an incongruent room.

> **H3:** When performing a specific vigilance task in a congruent room participants will perform better than in an incongruent room.

H4: When performing the Stroop task in a congruent room, participants will have a lower reaction in general than in an incongruent room.

In the remainder of this paper the following is presented. In the section 2 we elaborate about the method used to perform the experiment. Next the results will be described in section 3 and after analyzing the results conclusions will be drawn in section 4.

2. METHODOLOGY

2.1. Participants

The participants group consisted of 13people, of which 10 were men and 3 were women. The participants were students of Utrecht University and employees of TNO Defense and Safety. Their mean age was 23.75, with a standard deviation of 2.04. Before starting the experiment participants were tested on colorblindness. One of the participants was found color blind and his experimental data was not taken into account in the data analysis, but he was allowed to participate in our experiment. All the participants reported having normal sense of smell with no history of olfactory dysfunction and normal or corrected – to-normal vision. The experiments lasted for approximately 50 minutes.

2.2. Task description

Participants had to perform two tasks. The first task was vigilance task to measure alertness, for which we used VigTrack. , the second was a modified Stroop task. These will be explained now.

2.2.1. VigTrack

VigTrack is a PDA program developed by TNO. It is used to measure motor accuracy and reaction time. It consists of two tasks, a tracking and a vigilance task,which have to be done simultaneously. One session lasts for exactly 5 minutes. (Valk& Simons, 1998; Simons &Valk, 1998)

Fig. 1 A PDA showing the VigTrack task

The first part of VigTrack is trying to keep a moving ball as close to the target circle as possible. The ball can be manipulated indirectly by pinning the direction in which it has to go. The direction can be pinned by using the stylus and point it in the bull's-eye. The further away from the circle the stylus is pointed, the faster the ball will move in that direction. The ball always reacts with a small delay of around one second. A random force acts on the ball as well, which makes the ball move in random directions and ensures that the participant has to keep paying attention to keep the ball in the middle.

The second part of VigTrack is monitoring the target circle of the same environment. The target circle contained an icon which would change approximately every second. The icons that were shown were a diamond, square and hexagon. Each time a hexagon appeared, the participant had to push a button to indicate that they had seen a hexagon. Multiple hexagons were shown during the session.

For the first task, the difference between the ball and the target circle was measured. For the second task, the reaction time of each button input and correct number of button presses were measured.

The program automatically logged the distance of the ball to the center in both X and Y approximately every 20 ms. Also included in the logfile was the the status of the center cue and the button presses of the participant.

2.2.2. *Stroop task*

The second task we choose for our experiment was a modified Stroop task. The original Stroop task is a well known and popular method in experimental psychology for testing reaction time (Stroop, 1935; MacLeod, 1991).

A custom made Stroop task on the computer was used in our experiment. This Stroop task was programmed in PHP and JavaScript and used a MySQL database to store the results. Due to time and budget limits, the original Stroop task was slightly modified. Instead of measuring verbal input, we used colored keys on the keyboard as input. It is not uncommon to use key input as a substitute of vocal input (MacLeod, 1991;Besner, Stolz&Boutilier, 1997). However, we expect that the Stroop effect will be less when motor reactions are used compared to when participants are required to report verbally. Motivation for this expectation is based on the section *Response Modality: Oral Versus Manual* in MacLeod (1991).

The colors on the keyboard (with a qwerty layout) were purple, orange, blue, red and green, placed on the ‘a’, ‘d’, ‘g’, ‘j’ and ‘l’ key respectively. A TFT screen (1024 x 768, 60Hz) was used to show the program.

We choose to limit the amount of colors to five, because the working memory of humans is limited to four to seven chunks of information (Anderson, 2004). In order to prevent participants from linking colors to fingers and so decreasing their reaction time, they were instructed to only use the index finger of their main hand.

When a participant started with the Stroop task an information screen was shown with instructions about the Stroop task. After this five test stimuli were presented, followed by a message that the real test would begin. The real test consisted of 70 words, given in pseudorandom order. A word would be shown on the screen until the participant responded, followed by a short break of 1.5 second. Color words with the same color were among the set of possible combinations so we could ascertain if the Stroop effect actually occurred.

The language used in the Stroop task was English, because it was not feasible to make a Stroop task in the native languages of all participants due to time constraints and no prior knowledge about the participants. None of the participants had English as their native language, but all were able to read and understand English.

The program logged the participant id, ambient color, ambient scent, reaction time to the cue, what the color of the cue was and which word was used.

2.3. *Materials*

2.3.1. *Rooms*

The experiment took place in two small rooms of around 10 m^2 which were separated by a thin movable wall. The rooms were situated in the cellar section of the TNO building. The equipment used in the rooms for the experiment was identical. In figure 2 the setting of the rooms is presented and figure 3 are photographs of the rooms.

2.3.2. *Olfactory apparatus*

An olfactory apparatus, Xenon of Common Sense Productions (Common Sense Productions, n.d.), was placed in the corner of each room. Once every 5 minutes the olfactory apparatus released odor into the air for a few seconds. The olfactory apparatus was set such that the odor in the room was far above threshold. The two odors used were lavender and orange. Their numbers in the Common Sense Productions Odor list were 210 and 704 respectively. The consumption of the odor dispensers was said to be around 0.5 ml - 3.5 ml of odor liquid on full power.

2.3.3. Ambient Light

To create an ambient color, a Philips Living Colors Mini LED Lamp (Philips, n.d.) was used. It has a semi-analogue color adjuster with a sensitive touch ring. The colors used in our experiment were purple and orange. Because the color had to be switched between some of the rounds, it was hard to select the precise same color as the round before. Therefore an approximation of the color was used. More about at which rounds the light had to be switched can be found in the procedure of this section.

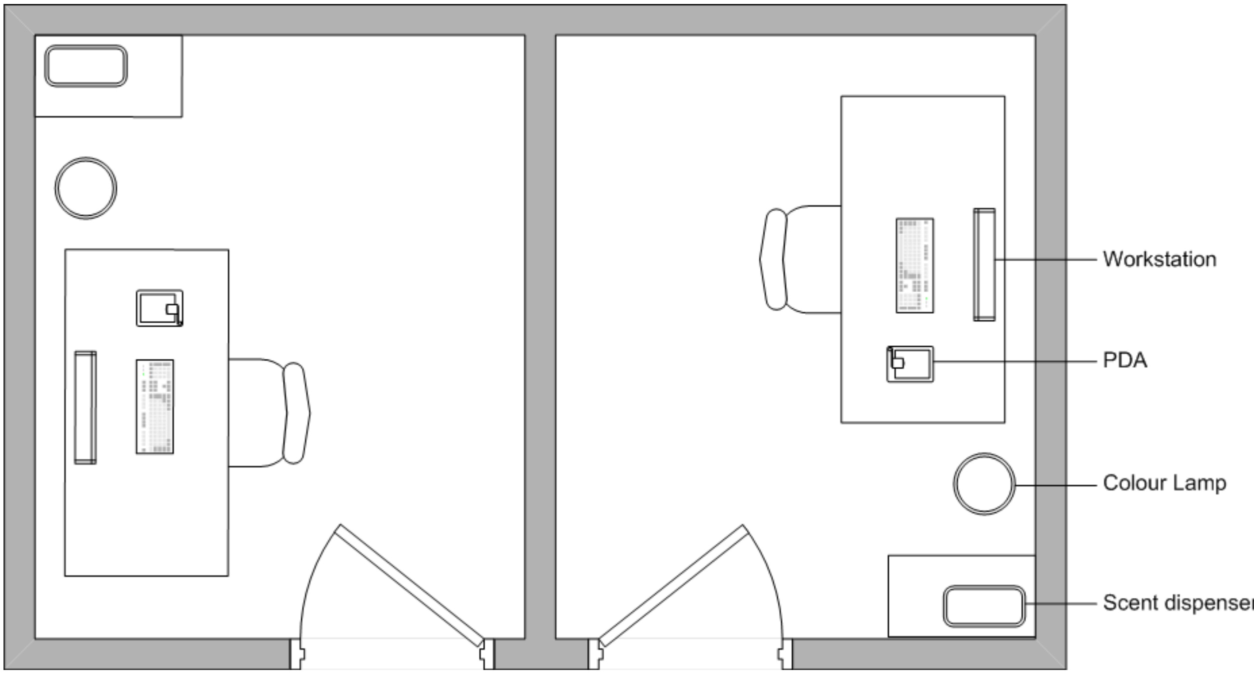

Fig. 2 Layout of the rooms used in the experiment.

Fig. 3 Photographs of the rooms

2.4. Design

The design of this experiment was a 2(odor) x 2 (color) multifactorial within-subjects design that used a number of dependent variables. For the vigilance task these were the reaction time, average deviation and number of hits, while the Stroop task had the dependent variable reaction time. The participant had to perform two tasks in every possible odor/color combinations, which were presented in four rounds, as is described in the procedure. We used eight different sequences in which participants could be tested. Participants were randomly assigned to the sequences so that each sequence was covered by at least one participant.

2.5. Procedure

2.5.1. Preparation

Before testing the participants, the olfactory apparatuses were turned on to ensure a steady smell throughout all the experiments. The rooms were kept closed and were only opened when leading participants in and out. The Living Color lamp was set to either purple or orange color, based on the sequence for the specific participant. The rooms always had the opposite light and smell compared to each other. This was done to lower the preparation time between task sections.

2.5.2. Introduction

The participant was led into the reception room, where a short explanation was given of the experiment they were about to take part in. They then had to sign the consent document. After doing so, the VigTrack task was explained and the participant had to perform a training round to learn how the task worked and to get familiar with it.

After doing the first training round with VigTrack, the participant had to do a test for colorblindness. Data gathered from colorblind participants was not used in the main analysis, but the participants were still allowed to complete the experiment.

When the colorblindness test was filled in, the participant had to perform a second training round. After completing the VigTrack for the second time, the participant was instructed about the Stroop task that they had to perform.

2.5.3. Task Procedure

The participant was guided into the experiment room. One round consisted of first doing a VigTrack task and then doing the complete Stroop task. When the participant completed the Stroop task, the participant was instructed to come out of the room again. Participants always had to perform the VigTrack task first, because if the PDA was left unattended too long, it would go into sleeping mode and cause VigTrack to crash after it had awoken again.

Two participants were doing the experiment simultaneously, but in opposite rooms. When both participants had completed the two tasks, their data was stored correctly and new log files were loaded by the instructors. In the second round, the participants were then instructed to go into the opposite room and perform both tasks again in the same sequence. After the second round the ambient colors were switched in both rooms, so from orange to purple and purple to orange. The participants then had to perform two more rounds, performing both tasks again in the same sequence in each room.

When the participants completed all four rounds, they were thanked for their time and participation and would get an explanation about the experiment if they asked for it.

3. RESULTS

All data was analyzed using the statistics package SPSS. The specific analysis used for each case is mentioned with its results.

3.1. Raw data

In order to analyze the VigTrack results we wrote a small script in PHP that analyzed all log files and produced the average distance to the center using simple trigonometry, average reaction time and the number of correct button presses by the participant. This was exported to comma separated files, so that they then could easily be loaded into SPSS.

Data preparation for the Stroop task consisted of writing a SQL query to return the relevant averages per person, per condition. Again data was stored in a comma separated file for easy importability in SPSS

3.2. Stroop effect

To analyze the existence of the Stroop effect, we utilized a two-way ANOVA on the word and color of on-screen stimuli. No interaction is found between these variables ($F(16,1184) = 0.711$, $p = .785$). Also, congruent cue color, ambient odor and ambient color do not enhance reaction times ($F(4,1196) = 0.506$, $p = .731$). An overview of the mean reaction times of the Stroop task can be found in figure 4.

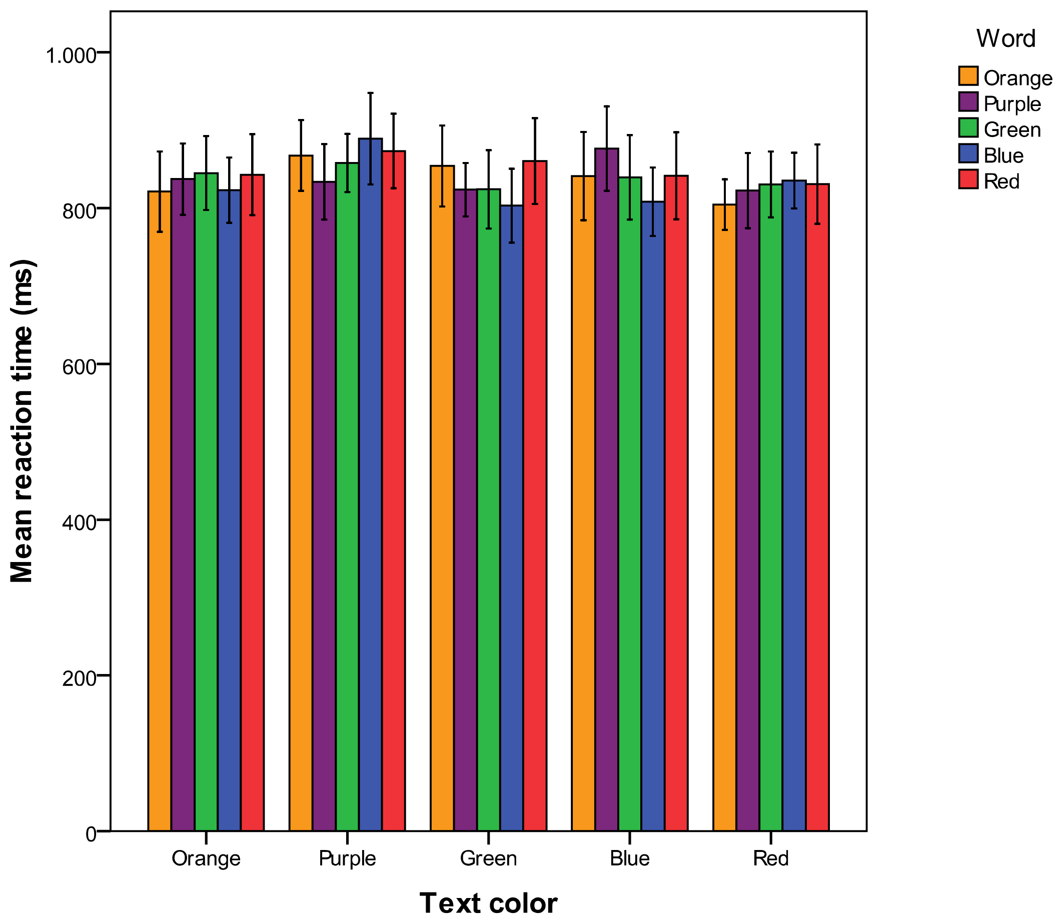

Fig. 4

3.2.1. Cue color influence on reaction time

Using an ANOVA, we found a tendency towards the influence of cue color on average reaction time ($F(4, 1196) = 1.917$, $p = .105$). A Tukey HSD test is used to find the colors that cause this influence. A tendency is found between the purple and red colors (MD = 39.63, p = .077). The text and room color show no interaction effect (F(4,1196) = 0.904, p = .461).

3.2.2. Ambient odor influence on Stroop task reaction time

With a two-way ANOVA, we find no influence of ambient odor on cue reaction time in the Stroop task (F(1,1199) = 0.000, p = .985). However, an influence of the ambient colors is found (F(1,1199) = 4.718, p = .030). The corresponding graph can be found in figure 5.

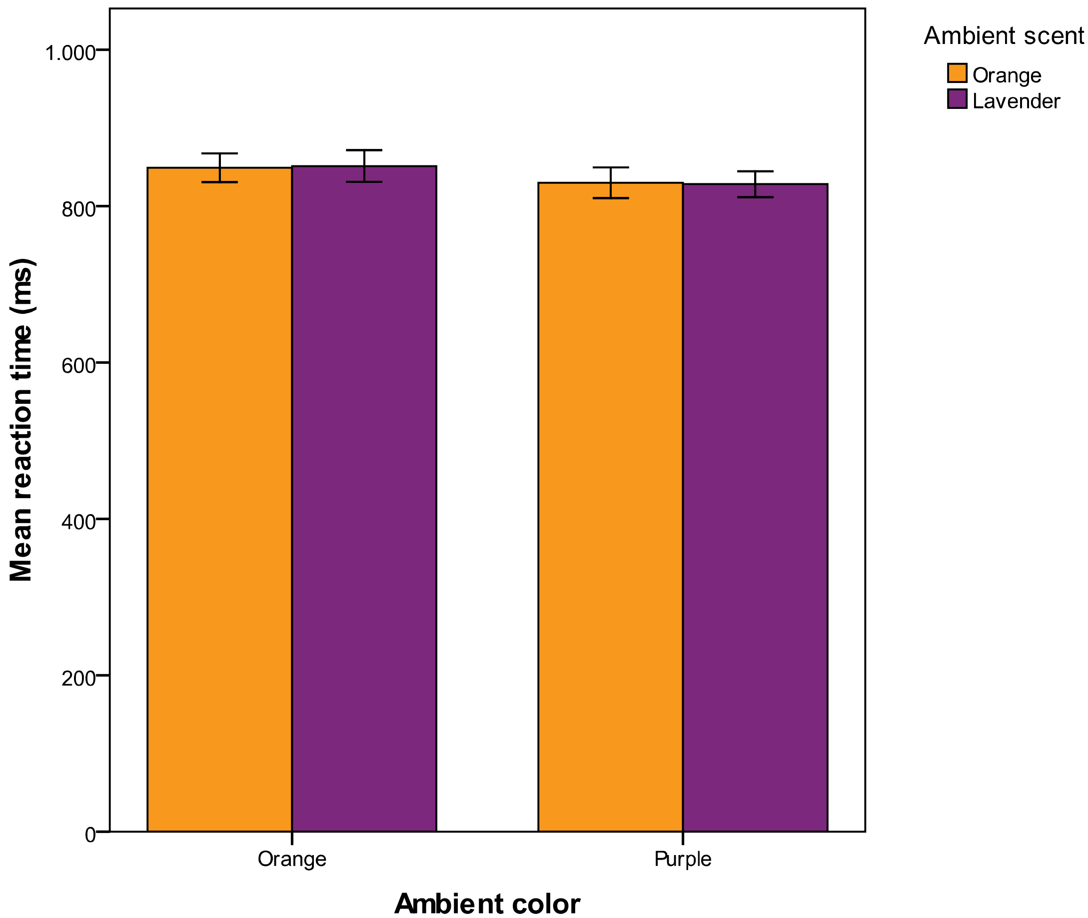

Fig. 5

3.3. VigTrack

3.3.1. Number of hits

The number of hit hexagonal cues in the VigTrack task in each of the ambient settings has been analyzed using a ANOVA-test. Ambient color and odor do not influence the number of hits in this task (F(1,48) = 0.001, p = .997 and F(1,48) = 0.129, p = .722 respectively). See figure 6.

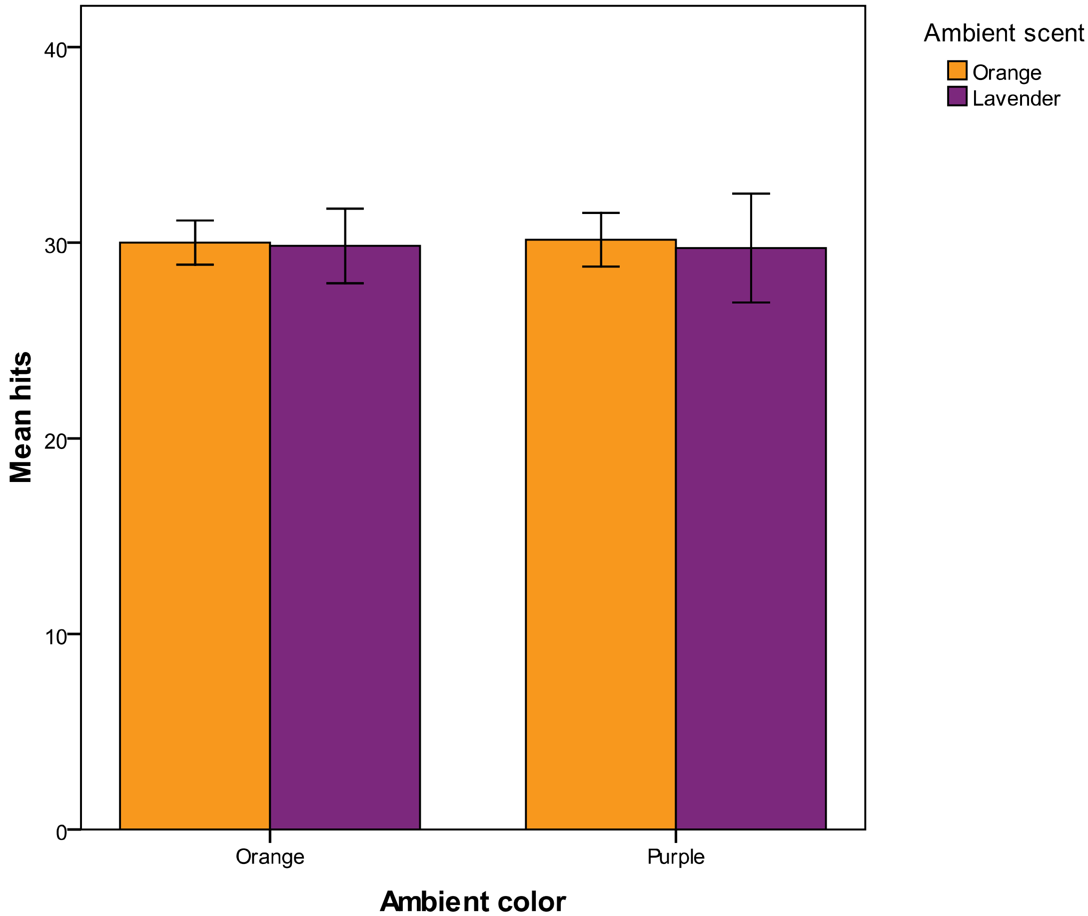

Fig. 6

3.3.2. Distance to center

The distance of the red sphere to the center of the screen is not significantly influenced by ambient odor or ambient color, according to the performed ANOVA-test ($F(1,48) = 1.334$, $p = .254$ and $F(1,48) = 0.548$, $p = .463$ respectively). However, when looking at figure 7, a tendency towards the influence of ambient odor on the mean distance to the center can be seen.

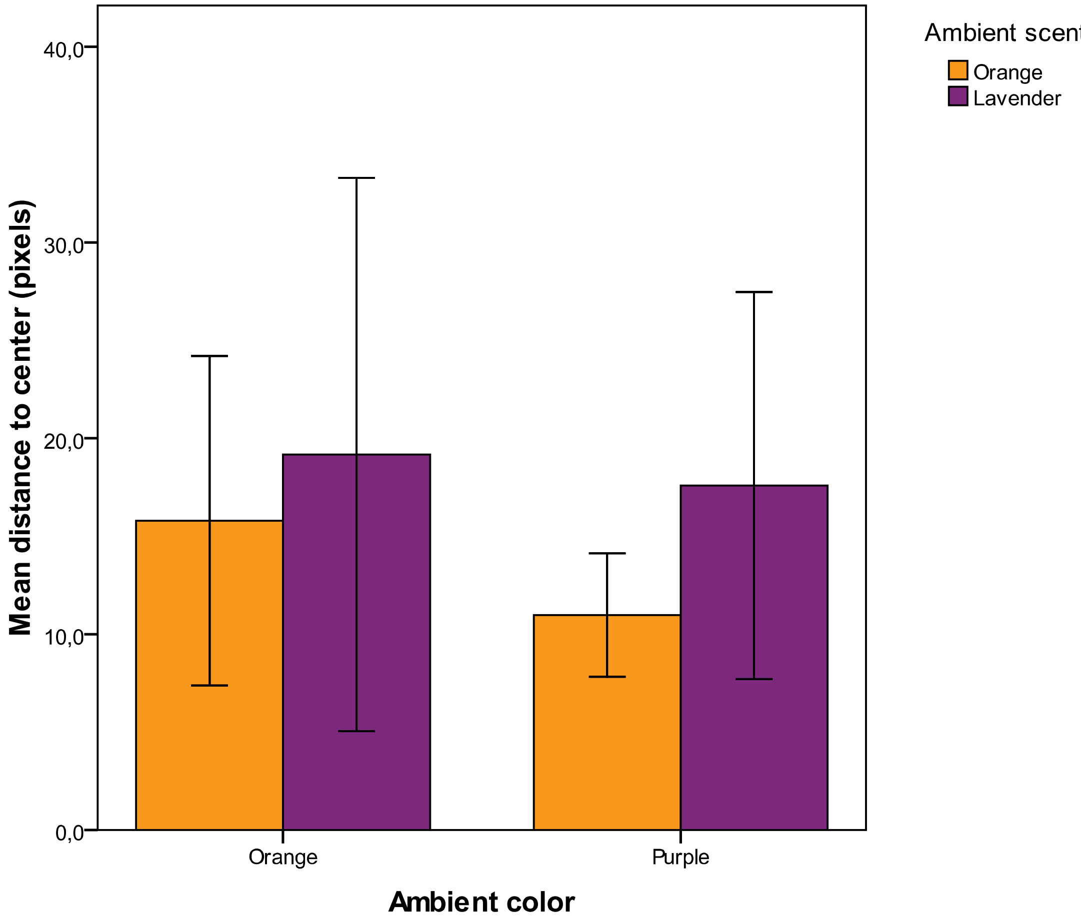

Fig. 7

3.3.3. Reaction time

The ANOVA-test showed no influences of ambient color or ambient odor on VigTrack cue reaction time ($F(1,48) = 0.634$, $p = .439$ and $F(1,48) = 0.088$, $p = .768$ respectively). See figure 8.

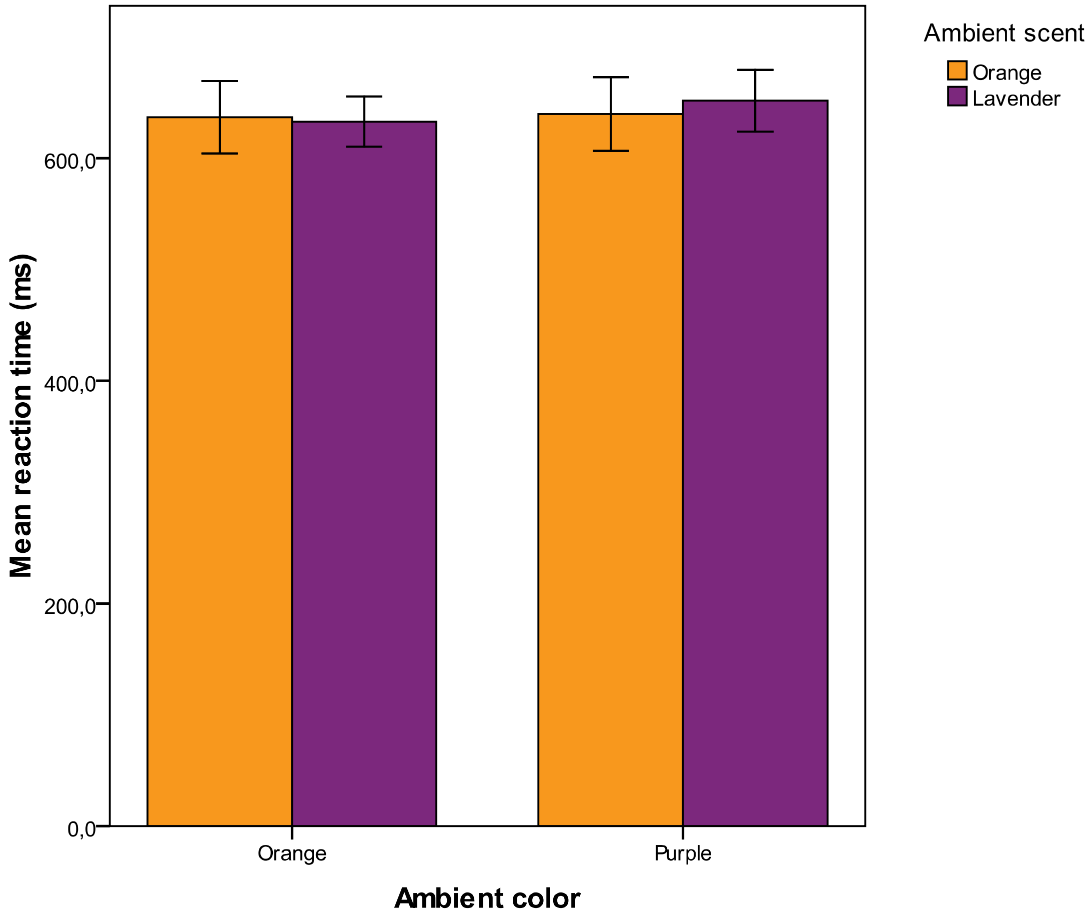

Fig. 8

4. CONCLUSION AND FUTURE WORK

A broad range of literature has shown a low-level link between odor and color. In this experiment however we hoped to find a significant combined relation for color and odor on the performance of our subjects. We did this by letting our subjects perform two different tasks in different conditions varying in color and odor combination.

After analyzing the data of the Stroop task, we did not find a Stroop effect. Also ambient odor and color stimuli had no effect on the performance of participants in the Stroop task. This disproves our first and fourth hypotheses. A possible explanation might be that the ambient color and odor negated the effect, but to definitively say something about this a control group would have need to been used. Another explanation might be that none of the participants performed the task in their native language or that the use of a keyboard for user response contributed to this.

Due to the fact that no Stroop effect was found, it we can only look into the differences between colors.

The second and third hypotheses could not be proven, but the results show that the purple ambient color has caused a lower reaction time while performing the Stroop task, but not so for VigTrack. Although the color purple has an influence on the reaction time, we did not find similar evidence for the ambient odors orange and lavender.

The ambient color and odor do not influence the number of hits in the vigilance task, nor the mean distance to the center.

Although literature showed that orange and lemon odors improve performance, this seemed not to be the case for us. A possible explanation might be that the rooms were not sealed air tight from each other, which might have resulted in odors leaking from one room into the other thus obfuscating our results. This might be a point of attention for future work.

For setting the ambient colors a more consistent way of determining the setting should also be found, just as with the strength of the odor. More test-subjects could also show that the differences are significant.

5. ACKNOWLEDGEMENTS

We would like to thank Sander Janssen for guiding us and TNO for giving us the opportunity to perform the experiments at their Soesterberg facility as well as providing the hardware needed for the experiment.

6. REFERENCES

Anderson, J.R. (2004) *Cognitive Psychology and its Implications.* Worth Publishers.

Baron, R.A. and Thomley, J., (1994) Positive affect as a potential mediator of the effects of pleasant fragrances on task performance and helping, Environ. Behav., 26 766–784.

Besner D, Stolz JA, Boutilier C. (1997), The Stroop Effect And The Myth Of Automaticity, *Psychonomic Bulletin & Review* , 4(2), 221-225.

Buchbauer, G., Jirovetz, L., Jäger, W., Dietrich, H., Plank, C., & Karamat, E. (1991). Aromatherapy: evidence for sedative effect of essential oil of lavender after inhalation. *Z. Naturforsch , 46c*, 1067-1072.

Common Sense Productions. (n.d.). Geurmachines. Retrieved June 26, 2009, from http://www.commonsenseproductions.nl/pdf/Geurmachines-csp-a3.pdf

Demattè, M. L., Sanabria, D., & Spence, C. (2006). Cross-Modal Associations Between Odors and Colors. *Chemical Senses , 31*, 531-538.

Demattè, M. L., Sanabria, D., & Spence, C. (2009). Olfactory Discrimination: When Vision Matters? *Chemical Senses , 34*, 103-109.

Diego, M. A., Jones, N. A., Field, F., Hernandez-Reif, M., Schanberg, S., Kuhn, C., et al. (1998). Aromatherapy Positively Affects Mood, Eeg Patterns of Alertness and Math Computations. *International Journal of Neuroscience , 96* (3), 217 – 224.

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining Sensory Information: Mandatory Fusion Within, but Not Between, Senses. *Science , 298*, 1627-1630.

Lehrner, J., Eckersberger, C., Walla, P., Pötsch, G., & Deecke, L. (2000). Ambient odor of orange in a dental office reduces anxiety and improves mood in female patients. *Physiology & Behavior , 71*, 83-86.

Lehrner, J., Marwinski, G., Lehr, S., Johren, P., & Deecke, L. (2005). Ambient odors or orange and lavender reduce anxiety and improve mood in a dental office. *Physiology & Behavior , 86*, 92-95.

Ludvigson, H. W., & Rottman, T. R. (1989). Effects of ambient odors of lavender and cloves on cognition, memory, affect and mood. *Chem Senses , 14*, 525-536.

MacLeod CM. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin* , Vol. 109, No. 2, 163-203.

Millot, J-L., Brand, G., Morand, M. (2002), Effects of ambient odors on reaction time in humans, *Neuroscience Letters*, 322, 79-82.

Morrot, G., Frédéric, B., & Dubourdieu, D. (2001). The Color of Odor. *Brain and Language , 79* (2), 309-320.

Moss, M., Cook, J., Wesnes, K., & Duckett, P. (2003). Aromas of rosemary and lavender essential oils differentially affect cognition and mood in healthy adults. *International Journal of Neuroscience , 113*, 15-38.

Pauli, P., Bourne, J. L., Diekmann, H., & Birbaumer, N. (1999). Cross-modality priming between odors and odor-congruent words. *American Journal of Psychology , 112* (2), 175-186.

Philips. (n.d.). Living Colors Mini. Retrieved June 26, 2009., website: http://www.lighting.philips.com/microsite/living_colors/?lang=nl

Simons, M., Valk, P.J.L. (1998). Early starts: effects on sleep, alertness and vigilance. NASA no. 19990025676, csa.com

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* (18), 643-662.

Valk, P.J.L., Simons, M. (1998) Pros and cons of strategic napping on long haul flights, NASA no. 19990025675, csa.com

White TL, Prescott J. 2007. Chemosensory Cross-Modal Stroop Effects: Congruent Odors Facilitate Taste Identification. *Chem. Senses* 32: 337-341, 2007.