AUA Working Paper Series No. 2016-4
November 2016

The effect of olfactory sensory cues on economic decision making

Varvara Kechagia
Graduate student, , Department of Agricultural Economics & Rural Development, Agricultural University of Athens
vergeo2007@gmail.com

Andreas C. Drichoutis
Assistant Professor, Department of Agricultural Economics & Rural Development, Agricultural University of Athens
adrichout@aua.gr

Agricultural University of Athens ·
Department of Agricultural Economics
& Rural Development · http://www.aoa.aua.gr

The effect of olfactory sensory cues on economic decision making*

Varvara Kechagia† and Andreas C. Drichoutis‡

Agricultural University of Athens

Abstract: Several studies show that sensory cues influence consumer decision making processes. While scent is a key component of a market's physical environment, it has received far less attention in the academic literature as compared, for example, with visual cues. In addition, most of the studies that examine the effect of ambient scents fail on one or both of these criteria: to properly control the influence of nuisance factors and/or to elicit preferences under real monetary incentives. We collected data from a laboratory experiment where we varied on a between subjects basis the dispersion of a citrus fragrance. We then elicited subjects' willingness to pay for two unbranded products — a mug and a chocolate — by having subjects participate in a 2nd price Vickrey auction. We also elicited subjects' risk preferences using lottery choice tasks. Our results show a statistically and economically significant effect on subjects' willingness to pay: valuations increased between 37% - 43% for subjects who were exposed to a citrus scent as compared to the control group. We do not find a statistically significant effect of the citrus scent on subjects' risk aversion.

Keywords: scent cues, fragrance, olfactory, willingness to pay, risk preferences, risk aversion, laboratory experiment

JEL Classification Numbers: C91, D44, D81, D87.

*We would like to thank Angelos Lagoudakis for excellent research assistance. We'd also like to thank John Hey and Glenn Harrison for providing valuable input with respect to estimation issues.

†Graduate student, Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Iera Odos 75, 11855, Greece, e-mail: vergeo2007@gmail.com.

‡Assistant Professor, Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Iera Odos 75, 11855, Greece, e-mail: adrihout@aua.gr.

"Odors have a power of persuasion stronger than that of words, appearances, emotions, or will. The persuasive power of an odor cannot be fended off, it enters into us like breath into our lungs, it fills us up, imbues us totally. There is no remedy for it."

– Patrick Süskind, Perfume: The Story of a Murderer

1 Introduction

Olfaction is an evolutionarily primitive sense critical for survival across the animal kingdom. Although in humans it is considered less important for survival when compared to other senses like the visual or the auditory sense, the human olfactory repertoire is vast and able to detect millions of airborne odorants at small concentrations (Hoover, 2010). Odorants can exert powerful behavioral effects even at the subconscious level by mediating, for example, the synchronization of menstrual cycles for females (Stern and McClintock, 1998).

There are good reasons to believe that scents may exert a powerful impact on behavior. The fragrance industry exists because of the widespread assumption that pleasant fragrances enhance attractiveness and therefore our social interactions.[1] In the marketing literature a popular quote attributed to Lindstrom (2005) has been used almost like a doctrine whenever it is deemed necessary to highlight the importance of the olfactory sense. The quote can be found in various forms and often reads as '...83% of all commercial communication appeals only to one sense — our eyes. And yet, according to studies, 75% of our day-to-day emotions are influenced by what we smell'.[2] This mismatch between olfactory and visual cues has sparked the development of a 'scent marketing' field. The psychology field has also shown prompt attention in studying the effect of scents on psychology relevant

[1] Sorokowska et al. (2016) have shown that ratings of body odor attractiveness and pleasantness were significantly lower in a natural body odor treatment than in a body odor with fragrance use treatment, which supports the assumption that first impression judgments can be affected by cosmetic use. Demattè et al. (2007) showed that female subjects rated a series of male faces as being significantly less attractive in the presence of an unpleasant odor than in the presence of a pleasant odor. Similarly, Baron's (1981) results indicate that male participants rated as more attractive female confederates in the presence of a perfume.

[2] The quote is likely a compilation of two phrases from Lindstrom's (2005) book, one that appears in the front flap and reads: 'Research shows that a full 75 percent of our emotions are in fact generated by what we smell' and a second phrase which appears in page 83 and reads: '83 percent of the information people retain has been received visually'.

decision making phenomena. Although the effects of olfactory cues on behavior have been predominantly examined in the marketing and psychology fields, the behaviors typically examined (discussed momentarily) are of primary interest to economists as well.

Economists are often worried about the (experimental) methods used in the marketing or psychology field and take such results with a grain of salt. In this paper we bring such prominent results under the scrutiny of the experimental economics lens. We use a very simple experimental design with two treatments: in one of the treatments we use a dispenser to diffuse a scent in the laboratory and the other is a scentless control treatment. More specifically we evaluate the effect of a citrus scent on two economic domains that scents might exert a powerful influence: willingness-to-pay (WTP) and choice under risk.

WTP and other measures of economic value have been a fruitful research area in academia. The appeal of this research agenda is shared by business and corporations which are eager in developing an understanding of factors that affect consumers' WTP that may lead to better pricing decisions. Many companies are now heavily investing in their air design by hiring specialists to develop customized fragrances and by installing complex scent-dispensing systems. One of the implicit assumptions is that by making a store environment distinct (e.g., creating a corporate identity) and pleasant, it will affect consumers' spending by shifting their WTP curve. Therefore, WTP elicitation is a relevant and important domain for examining scent effects.

With respect to risk, our study is motivated by a popular belief that casinos are using scents to get people to gamble more.[3] A paper often cited to back up these claims is an early study by Hirsch (1995) which conducted a field experiment in the casino floor of a large hotel in Las Vegas. Over a weekend, two slot machine areas were scented with different fragrances and higher revenues were observed when compared with weekend days before and after the scent treatment days.

Even if we accept the effect in the Hirsch (1995) study as genuine (in the next section we highlight a few problems with this study) there are questions that remain open about the possible mechanisms that may have driven this particular result. One way by which scents could have affected revenues in the casino, is by attracting a larger group of people in the slot machine area. This explanation would relate to the pleasantness of the encompassing atmosphere. A second explanation is that the scent directly affected individual behavior by making subjects spent more money per spin. This explanation could be rationalized by a direct effect on subjects' risk aversion. Our laboratory experiment rules out the first explanation since the scent is diffused only after subjects have accepted our invitation to

[3]This belief is maintained by blogs or news sites with provocative titles such as 'How Casinos Use Design Psychology to Get You to Gamble More' or 'Casinos Using Scents To Keep People Gambling'.

attend the lab session. We then directly observe whether the scent treatment induces a different risk choice pattern by asking subjects to make choices in lottery choice tasks.

In brief, our results confirm that a citrus scent does exert a statistically and economically significant effect on WTP. We observe some differences between a food and a non-food item which we explain in terms of congruency of the (fruity) citrus scent with the food product. For risk, we find no significant effect of scent on subjects risk aversion. This null result is sensitive to what decision theory and noise story one is ready to accept governing subjects' risk choices, however, it is a null result for our best fitting model.

In what follows we first start with a literature review to set the context of our research questions. In section 3 we describe our experimental design, the scent selection and scent diffusion processes in detail. In section 4 we provide more details with respect to theory and econometrics of risk choice data. We next present results for WTP and choice under risk separately, and conclude in the last section.

2 Literature review

To set the context, we first review the relevant literature in this section. We focus on research that examines scent effects on antecedents of WTP such as attitudes toward products and purchase intentions, on actual WTP or money spent on products and on choice behavior under risk. By design, our literature review only touches upon the aforementioned issues. For more general reviews of the literature on ambient scents with marketing applications see Bradford and Desrochers (2009).

2.1 Attitudes, product evaluations and purchase intention

One strand of the literature that explores the effect of scents on decision making, elicits the effect of scents on antecedents of WTP like attitudes and evaluations and not the effect of scents on WTP *per se*. In one of the first influential studies of olfactory behavioral research, Spangenberg et al. (1996) examined the effect of an ambient scent in a simulated store environment constructed in a consumer behavior laboratory. They used a 2 (scent affect: neutral vs. pleasing) $\times$ 3 (scent intensity: low, medium, high) experimental design with a control (no scent) condition and had significantly pretested a variety of olfactory stimuli that would classify as affectively neutral or affectively pleasing to be used in their treatments. In the simulated store, the product items were selected not to emit any detectable scents: kitchen items, decor items (e.g., nonfloral plants, fans, calendars, framed posters), clothing with the university insignia, books, school supplies, and outdoor athletic gear. The authors measured a variety of outcomes like evaluations of the store, evaluations of the merchandise, intention to visit the store, purchase intention, number of products examined etc. Their results showed that exposure to different pleasant odors (as compared to the no scent condition) led to more positive evaluations of the shop's atmosphere and interior as well as to product evaluations. Subjects also perceived spending less time shopping than subjects in the unscented condition (although they actually spent the same amount of time in the store) and a higher probability of revisiting the store in the future.

Many later studies based their experimental designs on Spangenberg et al. (1996). For example, Morrin and Ratneshwar (2000) examined the effect of a pleasant geranium ambient scent by varying on a within-subjects design brand familiarity (i.e., subjects were shown well known brands and unfamiliar brands). They found that the scent condition improved product evaluations of brands as well as brand recall but more so for the unfamiliar brands. Doucé and Janssens (2013) conducted experiments over two consecutive weeks. In the second week, a pleasant ambient scent (described as a 'slightly minty lemon scent') was diffused in a prestigious clothing store in Belgium while the first week served as the unscented condition. Upon leaving the store, customers were asked to fill out a questionnaire concerning their affective reactions, evaluations, and approach behavior toward the store environment and products. The presence of the ambient scent in the store had a positive effect on all measured outcomes.

de Wijk and Zijlstra (2012) exposed subjects to ambient food-related aromas at identical test rooms at the research facilities of the Restaurant of the Future in Wageningen: one room was scented with a citrus aroma, one with a vanilla aroma and one room was odorless. Among other measures, actual food choice of congruent and non-congruent foods was examined, where the plates consisted of citrus-congruent food (mandarin orange segments and orange juice), vanilla congruent food (vanilla cookies and milk) or neutral in relation to either aroma (cubes of cheese and mineral water). Subjects were told that food in the room was present for their convenience and were free to sample. Consumption of food was measured by tallying food and by weighing drinks present at the beginning and end of the session. Exposure to the ambient citrus aroma increased number of portions of mandarin consumed and reduced selection of cheese.

2.2 Willingness to pay and money spent

Another strand of the literature tries to isolate the effect of scents on consumer spending or WTP. One of the earliest studies that, at the time, received high media attention (Hirsch, 1990 cited in Lindstrom, 2005; copies of the original report can be found in Corbett, 1994, page 97), showed that by placing two identical Nike sneakers in two separate rooms, one room containing a floral scent and one room a neutral scent, had a significant effect on likelihood to purchase the sneakers. Customers in the floral scented room stated they were 84% more likely to purchase the sneakers. In addition, in the original report it is stated that 10% of those effected with the scent, which amounts to three subjects, stated an average WTP of 10.33$ more than participants not exposed to the floral scent. By any kind of standards, descriptive statistics from just three subjects seem absurd, however, this widely publicized result prompted others in pursuing similar research agendas.

Fiore et al. (2000) varied the display of a sleepwear in a room on campus and that of the ambient scent in the room. The display consisted of a female mannequin, a three-fold dressing mirror, two floral pillows, a white textured throw blanket, two candle holders with white candles, a vase with dried flowers, and lighting. Besides a control unscented condition, they varied the scent treatment at two levels. In one scent condition they used a potpouri scent described as 'Lily of the Valley' which was rated by a different group of subjects as appropriate for the sleepwear presentation, whereas in another scent condition they used a potpouri scent described as 'Sea Mist' which was rated as inappropriate for the sleepwear presentation. Among other measures, they asked subjects to indicate (hypothetically) their WTP. The authors found a statistically different WTP between the product on display with an appropriate scent (mean WTP=29.6$) and the product on display with an inappropriate scent (mean WTP=24.8$) while the appropriate fragrance condition did not differ with the control unscented condition (mean WTP=28.4$) when the product was on display.

Mattila and Wirtz (2001) conducted a field experiment where they examined the interaction of scents and music on impulse buying. They adopted a 3 (no scent vs. low arousal (Lavender) scent vs. high arousal (Grapefruit) scent) × 3 (no music vs. low arousal music vs. high arousal music) experimental design inside a gift shop. The study was conducted in three shifts over fourteen consecutive days at the chosen retail and the treatments were randomized across shifts. Impulse buying was measured on a self-reported seven-point Likert scale where subjects had to indicate whether they 'bought more than what had planned to buy'. The congruent condition of a high arousal music and high arousal scent was associated with higher stated impulse buying than other single stimulus or incongruent conditions.

Michon et al. (2006) conducted another field experiment in four consecutive weeks in a mall in Montreal following a 2 (fast 96 bpm vs. slow 60 bpm music tempo of 'light rock' music) × 2 (citrus scent vs. no scent) experimental design.[4] The ambient scent was diffused in the shopping mall's main corridor located between two major retailers. Sampled subjects

[4]The study by Chebat and Michon (2003) seems to be partially reporting two out of four treatments of Michon et al. (2006).

filled-in a self-administered questionnaire where they were asked to indicate how much money they've spent on non-food shopping. Average shopper spending was higher when ambient scent and fast tempo music conditions were combined (57.93$) or when slow tempo music was played with no ambient odor (58.84$).

The congruency-incongruency of music arousal level (slow vs. fast tempo) with scent arousal level (lavender vs. grapefruit) was also examined in Homburg et al. (2012) where they asked subjects to state their willingness to pay for a washing machine and a smartphone. They found a higher WTP for both products when subjects were either in the high music arousal - high scent arousal or in the low music arousal - low scent arousal conditions. That is, congruency of music and scent was a significant factor positively affecting WTP as compared to incongruency.

Morrison et al. (2011) varied the presence of a vanilla scent and volume of an upbeat dance music compilation (low vs. high) played on repeat every three hours in a fashion retail store located in a shopping district of a major metropolitan area in Australia. Subjects, when exited the store, where asked to fill in a questionnaire and state, among others, how much money they've spent in the store. They found that the congruency of high volume music and presence of vanilla scent increased pleasure levels, which in turn positively influenced shopping behavior, including time and money spent in the store.

Guéguen and Petr (2006) did a field experiment in a restaurant setting (small pizzeria in Brittany, France) where they administered two scent treatment (lavender vs. lemon) and a no scent treatment over three Saturdays. Subjects in the lavender treatment spent significantly more money (€21.1) than subjects in the lemon and no scent treatment (€18.1 and €17.5, respectively).

Spangenberg et al. (2006) explored the effect of the congruency of gender with gender specific scents that were diffused in a clothing store selling both men's and women's clothing in equivalent quadrate floor spaces. During a two week period, half of the customers were exposed to a masculine scent (rose maroc) and half to a feminine scent (vanilla). This experimental design resulted in congruent and incogruent conditions.[5] Subjects filled in questionnaires that asked them to self-report their spending which was also matched with retailer provided information about the number of individual clothing items purchased and dollars spent by each individual customer. The authors found that subjects in the congruent condition spent more than double the money than subjects in the incogruent condition (55.12$ vs. 23.01$).

[5] The scent was congruent when the scent's gender orientation matched the gender of the products offered (i.e., rose maroc for men's clothing, and vanilla for women's clothing) and incongruent when the scent's gender orientation did not correspond with the product offering (i.e., rose maroc for women's clothing and vanilla for men's clothing).

2.3 Risky decision making

Studies that explore the effect of scents on choice under risk are scarce. Hirsch (1995) conducted the earliest study we are aware of, on a casino floor of a large hotel in Las Vegas. Two different areas of slot machines were scented with two different scents starting on the midnight of a given weekend. The amount of money gambled in the weekend was recorded and compared with the weekends before and after the experiment as well as with a control slot machine area that was not scented. Hirsch (1995) reported that the money gambled increased by 45% in the experimental weekend vs. the weekend before and after the experiment.

Hancock (2009) rightly criticized Hirsch (1995) for not publishing the list of the components of the fragrances that were used, which precluded further testing of Hirsch's (1995) findings. She also points that Hirsch (1995) did not disclose whether the experiment was conducted during a holiday weekend or on a weekend where one or more special events were being held on the casino or in town, which could be a confounding factor of the experimental results. Hancock (2009) improved the experimental design by conducting the experiment in a large United States casino over a period of 20 days. In this period, in five different slot machine locations within the casino, two refreshing and two soothing scents were diffused while the fifth room served as the non-scented control room. The treatments were rotated across rooms in order to randomize possible confounds of popularity of location, ease of access and popularity of machines. Hancock (2009) found that a soothing natural fragrance droved higher and statistically significant coin-in.

More recently, Gagarina and Pikturnienė (2015) manipulated scent type (vanilla vs. peppermint) and intensiveness level (high vs. low concentration) in a laboratory environment and found no effect of any of the treatment variables on risk aversion as measured from hypothetical lottery choice tasks. Admittedly, their sample size per treatment was particularly low (18-19 subjects per treatment).

3 Experimental design

In October 2015 we recruited 160 subjects from the undergraduate population of the university. Subjects participated in sessions of 15 or 10 subjects arranged in the middle of the week. Sessions started from 10 am and concluded by 2 pm. Subjects were split in two treatments: the control treatment and the scent treatment. In the scent treatment subjects were exposed to an olfactory stimuli (described momentarily) that was diffused in the lab room using a dispenser. In a single day only one of the treatments was run to avoid

Table 1: Experimental design

Day	Date	Scent treatment	No scent treatment
Wednesday	21-Oct	-	40 subjects
Thursday	22-Oct	40 subjects	-
Wednesday	04-Nov	40 subjects	-
Thursday	05-Nov	-	40 subjects

any possible contamination between treatments due to fragrance residuals (although the manufacturer reassured us that there will be no residuals left after one hour from turning off the dispenser) and the lab was fully ventilated overnight. Before the first session each morning, the room was 'sniff-tested' by the experimenter and a research assistant and no residual odors were detected. The treatments were counter balanced over weekdays (see Table 1).

Upon arrival, subjects were given a consent form to sign and when all subjects necessary to form an auction group had arrived (subjects participated in auction groups of 5 subjects), each one of them was randomly seated to one of the PC private booths. Printed instructions were given to all subjects and the experimenter read aloud instructions. Subjects were specifically instructed to raise their hand and ask any questions in private and that the experimenter would then share her answer with the group. They received a show-up fee of €4. Subjects could earn or lose money during the experiment (described momentarily), so that average total payouts were €10.8 (S.D.=2.84, min=1.1, max=24). After instructions were read aloud, subjects filled a series of computerized control questions to enhance comprehension of instructions. They were free to advise their printed instructions or ask questions to the experimenter and generally showed a good understanding with an average of 10.5 correct answers out of 12 questions.

The experiment consisted of three stages (experimental instructions are reproduced in English in Appendix C). In Stage 1 subjects went through a typical real effort task where they had to count and report the number of zeros shown in a 5×5 matrix. This task was repeated 10 times (the elements of the matrix where random and changed with each repetition but was the same for all subjects at a given repetition) and subjects could earn €0.5 every time they correctly solved the task within 25 seconds. The task aimed at mitigating house money effects by making subjects earn part of their endowment (e..g., Corgnet et al., 2014; Jacquemet et al., 2009). The zero counting task was purposefully made easy (as evident by the fact that earned real effort money averaged €4.83 with a standard deviation of 0.32 and that 75% and 18.1% of subjects earned exactly €5 and €4.5, respectively), so that subjects would start off in Stage 2 of the experiment with approximately equal endowments.

3.1 The 2^{nd} price auction

In Stage 2 subjects participated in a series of 2^{nd} price Vickrey auctions (Vickrey, 1961) in groups of 5 subjects. Matching in groups was random and remained the same throughout the session. Subjects were unaware of which other subjects in the session composed their group. The group size was decided with three things in mind: a) avoid disengaging off-margin bidders from the auction procedure (Shogren et al., 2001) by having 'too large' groups b) given that price feedback in repeated 2^{nd} price auctions is discouraged (Corrigan et al., 2012), avoid 'too small' groups that would, by design, reveal bidding behavior of other subjects and c) increase the number of independent observations (if we count the auction group as the unit of an independent observation).

The mechanics of the auction were explained in the instructions but were also practiced by allowing subjects to hypothetically bid in three repeated training rounds for two non-focal products: a pack of biscuits and a USB stick (pictures of the products as shown to the subjects can be found in Appendix B; pictures B.1a and B.1b). Bids were entered simultaneously for the two goods. The purpose of the training rounds was to closely mimic the real auctions rounds that followed.

Right after the training rounds, subjects were shown pictures of the real products in their computer screens (shown in Appendix B: pictures B.2a and B.2b) and real products were circulated in the lab for subjects to observe closely if they wished to do so. Both products are not available in the market, were custom made for the experimenters and were purchased at approximately the same price. Subjects were then asked to complete hedonic evaluations of the products (on a scale from 1='dislike very much' to 9='like very much'). Ten repeated rounds of a 2^{nd} price auction followed and subjects were told that only one round and one product would be randomly selected at the end of the session (separately for each auction group) and that the 2^{nd} price would be substracted from the highest bidder's income.

3.2 Risk preference elicitation

In Stage 3, we elicited subjects' risk preferences using the Holt and Laury (2002) task (HL) as well as a modified version which varies the payoff amounts instead of the probabilities (payoff varying - PV). In the HL task individuals are asked to make a series of 10 decisions between two options (see Table 2). In option A, the high payoff amount is fixed at €2 and the low payoff amount is fixed at €1.60 across all 10 decision tasks. In option B, the high payoff amount is fixed at €3.85 and the low payoff amount is fixed at €0.10. The only thing changing across the 10 decisions are the probabilities assigned to the high and low payoffs. Initially the probability of receiving the high payoff is 0.10 but by the tenth decision task, the

Table 2: The Holt and Laury (2002) risk preference task

Lottery A				Lottery B				EVA €	EVB €	EV difference
p	€	p	€	p	€	p	€			
0.1	2	0.9	1.6	0.1	3.85	0.9	0.1	1.640	0.475	1.165
0.2	2	0.8	1.6	0.2	3.85	0.8	0.1	1.680	0.850	0.830
0.3	2	0.7	1.6	0.3	3.85	0.7	0.1	1.720	1.225	0.495
0.4	2	0.6	1.6	0.4	3.85	0.6	0.1	1.760	1.600	0.160
0.5	2	0.5	1.6	0.5	3.85	0.5	0.1	1.800	1.975	-0.175
0.6	2	0.4	1.6	0.6	3.85	0.4	0.1	1.840	2.350	-0.510
0.7	2	0.3	1.6	0.7	3.85	0.3	0.1	1.880	2.725	-0.845
0.8	2	0.2	1.6	0.8	3.85	0.2	0.1	1.920	3.100	-1.180
0.9	2	0.1	1.6	0.9	3.85	0.1	0.1	1.960	3.475	-1.515
1	2	0	1.6	1	3.85	0	0.1	2.000	3.850	-1.850

probability is 1. As shown in Table 2, the expected value of lottery A exceeds the expected value of lottery B for the first four decision tasks. Thus, a risk neutral person should prefer lottery A for the first four decision tasks and then switch to lottery B for the remainder.

Drichoutis and Lusk (2016) argue that the Holt and Laury (2002) task is more accurate at eliciting the shape of the probability weighting function given that it varies probabilities and keeps the monetary amounts constant. They then constructed a task that varies the amounts and keeps probabilities constant at 0.5 for all payoffs. They showed that combining information from the HL and the PV task, greater predictive performance can be achieved. Table 3 shows a payoff varying task that keeps the probabilities constant across the ten decision tasks and changes instead the monetary payoffs down the ten tasks. The monetary payoffs are varied in a way that the pattern of choices for a risk neutral person is similar to the HL task i.e., such a person should prefer lottery A for the first four decision tasks and then switch to lottery B for the remainder.

Instead of providing a table of choices arrayed in an ordered manner all appearing at the same screen as in HL, each choice was presented separately showing probabilities and prizes as in Andersen et al. (2014). The order of appearance of the HL and PV tasks were randomized on a between-subjects basis. An example of one of the decision tasks is shown in Figure B.3. For each subject, one of the choices was randomly chosen and paid out at the end of the session.

3.3 Questionnaire and manipulation check

In Stage 4 subjects went through a short questionnaire that elicited standard demographic characteristics. Subjects were then asked a decoy question of whether they noticed music in

Table 3: The payoff varying risk preference task

Lottery A				Lottery B				EVA €	EVB €	EV difference
p	€	p	€	p	€	p	€			
0.5	1	0.5	1	0.5	1.2	0.5	0.2	1.00	0.70	0.300
0.5	1.2	0.5	1	0.5	1.5	0.5	0.2	1.10	0.85	0.250
0.5	1.4	0.5	1	0.5	1.8	0.5	0.2	1.20	1.00	0.200
0.5	1.6	0.5	1	0.5	2.2	0.5	0.2	1.30	1.20	0.100
0.5	1.8	0.5	1	0.5	2.9	0.5	0.2	1.40	1.55	-0.150
0.5	2.0	0.5	1	0.5	3.5	0.5	0.2	1.50	1.85	-0.350
0.5	2.2	0.5	1	0.5	4.6	0.5	0.2	1.60	2.40	-0.800
0.5	2.4	0.5	1	0.5	6.8	0.5	0.2	1.70	3.50	-1.800
0.5	2.6	0.5	1	0.5	9.2	0.5	0.2	1.80	4.70	-2.900
0.5	2.8	0.5	1	0.5	15	0.5	0.2	1.90	7.60	-5.700

the lab which they could answer with a Yes/No. This question was asked in order to cover up the purpose of the next question which asked subjects whether they noticed a scent in the lab which they could answer with a Yes/No as well.

Given that odors can be either perceived attentively (e.g., 'I smell banana' or 'I smell something') or inattentively (subjects show no evidence of being aware of something in particular), the question about scent perception aimed in classifying subjects according to awareness circumstances. Smeets and Dijksterhuis (2014) have shown that the effects of olfactory stimuli on perceptual and cognitive processing can be conceived of as priming. In this respect, attentive awareness of a scent can be seen as a form of supraliminal priming while inattentive awareness of a scent can be seen as a form of subliminal priming. The differential effects of supraliminal and subliminal scents on consumer behavior have been documented by some studies (e.g., Baron, 1983; Bosmans, 2006; Li et al., 2007). For social psychologists it makes no qualitative difference whether a subject is aware of the stimulus event or not, but whether the individual is aware of the *influence* of the presented stimulus (Bargh, 1992). If subjects are aware of the persuasive power of an olfactory stimuli, they may apply defensive mechanisms towards it to correct for it extraneous effect (see for example Baron, 1983). To account for this fact, subjects were also asked an open ended question about what they think the purpose of the research was. No subject mentioned the word 'scent' or any other synonyms as the topic of exploration of the research project.

An additional set of questions scrutinized subjects for factors that relate to olfactory disorders like antibiotic use, nasal spray use, smoking status as well as direct questions about known taste and smell disorders. A final set of questions asked subjects to evaluate on 7 point Likert scales satisfaction with the lab environment (1='extremely dissatisfied', 7='extremely satisfied'), the lab's ambient conditions (1='very unpleasant', 7='very pleasant'), subjects' feeling during the session (1='extremely relaxed', 7='extremely energetic'), interaction with the experimenters (1='very bad', 7='very good') and overall experience (1='very unpleasant', 7='very pleasant').

3.4 Scent selection considerations

In selecting a scent to use as the olfactory stimuli in the lab, we took into account pleasantness, congruity and memory of scents which are listed as key aspects of scent marketing (Goldkuhl and Styvén, 2007).

First we sought in testing the effect of a pleasant stimuli. This is because a pleasant stimuli is likely to be more relevant for marketing applications given the focus of companies in creating a pleasurable shopping experience by modifying the air design aspect of their stores. Although unpleasant stimuli have been explored in the literature (e.g., Grabenhorst et al., 2007; Sutani et al., 2007), these are rather outliers since the vast majority of studies explores pleasant olfactory stimuli.

Second, we opted for an odor mixture rather than an individual odor. This is because perception of an odor can be significantly influenced by a verbal label attributed to the odor (Herz and von Clef, 2001). Once a scent can be verbally labeled, cognitive processing is no longer implicit or automatic (Smeets and Dijksterhuis, 2014) leading to semantic overshadowing (Melcher and Schooler, 1996). Verbalizing an individual odor can be a very difficult task (Cain, 1979), therefore, by selecting an odor mixture we effectively precluded subjects from verbalizing the odor which would interfere with automaticity of cognitive processing.

A third consideration has to do with congruity of the smell and the actual product or service provided. In some studies, congruency has been shown to influence consumers more than incongruent conditions (Mitchell et al., 1995). For example, in Bone and Jantrania (1992) a household cleaner and a sunscreen were more positively evaluated when they were scented with lemon and coconut, respectively, since this is what consumers assume appropriate in these situations. More recently, Olofsson et al. (2012) found that subjects performed faster and more accurately in an object evaluation task when in a congruent scent condition than in a condition which varied valence (pleasantness) of the scent. Parsons (2009) even found a negative effect of an incongruent scent on liking a store and intention to shop at the store. Goldkuhl and Styvén (2007) note the importance of congruent scents for edible products (e.g., the smell of a freshly baked product in a bakery shop) because this allows providers to tangibilise their offerings. The olfactory-visual congruency is not unique to adults but has been shown for infants as well (Wada et al., 2012).

On the other hand, Bosmans (2006) have found that as long as the scent is perceived as pleasant and is not *completely* incongruent, it can still have an effect on product evaluations. Ehrlichman and Halpern (1988) showed that exposure to a pleasant ambient scent led to subjects retrieving a larger number of happy memories than in a non-scented condition because pleasantness can be congruent with the material in long-term memory; thus, they showed that scent congruence with the product is not a necessary condition. Parsons (2009) provides a good overview of the literature on congruent and incongruent scents.

Finally, we opted for a (mixed) citrus scent as citrus scents have been used very often in the literature (e.g., Chebat and Michon, 2003; Chebat et al., 2009; Michon et al., 2005; Liu et al., 2008, for a few examples). This is likely due to an early influential paper by Spangenberg et al. (1996) where they pretested 26 different scents and found that an orange scent scored high in both an affective and activation dimension, and was therefore deemed appropriate to use in an affectively pleasing experimental condition.

After reviewing what is available in the market we finally selected the Airoma® XTREME 'Florida Zest' by Vectair Systems (a picture of the aerosol can is shown in Figure B.4a in Appendix B). The accompanying advertisement — '. . . you can expect to experience a fresh citrus complex made up of orange, grapefruit and mandarin, interlaced with neroli and orange flower' — cleary indicates the mixed citrus nature of the scent although, in practice, it would be very difficult for anyone to distinguish the components of the mix. The advertisement was not communicated to the subjects nor at any point was it made obvious that the room was scented on purpose since the treatment was meant to be kept below subjects' awareness levels.

3.5 Scent diffusion in the lab

To achieve scent diffusion in the lab we used a scent dispenser (shown in Figure B.4b in Appendix B). The scent dispenser was installed at one of the lab walls at a distance of 1.80m from the ground, as suggested by the manufacturer, and was turned on only during the scent treatment days (see Table 1). Scent diffusion started half an hour before the first session in a day and delivered one spray-dose every six minutes in order to maintain continuous scent intensity. The air-conditioning system was set to maintain a constant temperature of 25°C but ventilation was turned off so that the scent would not wear off.

Figure B.5 (see Appendix B) shows the plan room for the laboratory with the position of the dispenser marked in a red circle. Numbers on the vertical and horizontal axis depict the respective distance of any given computer booth from the dispenser in computer units[6].

[6]So, for example, the computer booth corresponding to horizontal=2, vertical=3, denotes a computer that is placed 3 computers away on the vertical axis and 2 computers away in the horizontal axis (either on the left or the right). We make no distinction as per whether a subject was seated on the left or the right of

We use this information later to show that it doesn't matter with respect to scent awareness where a subject was seated in the lab. In any given session, the lab room was either filled from the back to the front or vice versa and the order was counter balanced across sessions.

4 Theory and econometrics of risk preferences

One way to go about estimating treatment effects for risk preferences is to count subjects' number of safe choices (number of times the left lottery is chosen) and then regress this number on the treatment variables. However, Drichoutis and Lusk (2016) present a simple numerical example that demonstrates that if people weigh probabilities non-linearly, then simply observing the switching point in HL types of decision tasks, is insufficient to identify the shape of the utility function and the shape of the probability weighting function. Furthermore, just using the number of times a lottery is chosen in a regression, typically ignores the accumulated literature on stochastic error specifications of risk choice data (e.g., Hey et al., 2010; Hey, 2005; Wilcox, 2008, 2011, 2015). Hey (2014) notes that the stochastic specification is not merely an econometric issue, but also a behavioural one which experts in the field feel that it is the key to understanding behaviour, perhaps even more important than the preference functional.

Therefore, we follow what is considered the gold standard in this literature and employ structural econometric methods (see for example Harrison and Rutström, 2008, for a pedagogical treatise). Let the utility function be the constant relative risk aversion (CRRA) specification[7]:

$$U(M) = \frac{M^{1-r}}{1-r} \tag{1}$$

where r is the relative risk aversion (RRA) coefficient, $r = 0$ denotes risk neutral behavior, $r > 0$ denotes risk aversion behavior and $r < 0$ denotes risk loving behavior. If we assume that Expected Utility Theory (EUT) describes subjects' risk preferences, then the expected utility of lottery i can be written as:

$$EU_i = \sum_{j=1,2} p_i(M_j)U(M_j) \tag{2}$$

where $p(M_j)$ are the probabilities for each outcome M_j that are induced by the experimenter (shown in Tables 2 and 3). Despite the intuitive and conceptual appeal of EUT, a number of experiments suggest that EUT often fails as a descriptive model of individual

the dispenser.

[7]Constant relative risk aversion, rather than increasing or decreasing relative risk aversion, is a realistic assumption given the narrow range of prizes paid out in the lottery choice tasks.

behavior. A popular alternative is Rank Dependent Utility (RDU) developed by Quiggin (1982), which was incorporated into Tversky and Kahneman's (1992) cumulative prospect theory. RDU extends the EUT model by allowing for non-linear probability weighting associated with lottery outcomes.[8] To calculate decision weights under RDU, one replaces expected utility in equation (2) with:

$$RDU_i = \sum_{j=1,2} w_i[p(M_j)]U(M_j) = \sum_{j=1,2} w_{ij}U(M_j) \tag{3}$$

where $w_{i2} = w_i(p_2 + p_1) - w_i(p_1) = 1 - w_i(p_1)$ and $w_{i1} = w_i(p_1)$ with outcomes ranked

[8]As in most experiments of choice under risk, our experiment involved multiple choices over lotteries for which subjects where randomly paid for one of these choices. This payoff mechanism, known as the Random Lottery Incentive Mechanism (RLIM), is under criticism. As first put forward by Holt (1986), given the reduction axiom, RLIM is incentive compatible if and only if the Independence Axiom holds. Given that RDU does not include the independence axiom, then RLIM is inappropriate for non-EUT theories on theoretical grounds. The issue seemed to have been settled for a while perhaps due to open statements from prominent experimentalists. For example, Wakker (2007) argued that the RLIM issue has unduly hindered many papers in the review process and that it is counter-productive to re-hash the issue each and every time and Hey and Lee (2005b) concluded that "...experimenters can continue to use the random lottery incentive mechanism and that this paper can be used as a defense against referees who argue that the procedure is unsafe". However, the issue has been re-opened recently by one group of researchers (Cox et al., 2014; Harrison and Swarthout, 2014) with fairly convincing evidence. Nevertheless, researchers continue to use the RLIM under non-EUT theories as the preferred method of payment (this is true even for researchers that criticized the RLIM for testing non-EUT: Harrison et al., 2015; Harrison and Swarthout, 2016). Using the RLIM under non-EUT specifications either invokes the assumption of the isolation effect i.e., that a subject views each choice in an experiment as independent of other choices in the experiment or assumes two independence axioms as in Harrison and Swarthout (2016): one axiom that applies to the evaluation of a given prospect which is assumed to be violated by non-EUT, and another axiom that applies to the evaluation of the experimental payment protocol. Only the validity of the latter axiom is required to ensure incentive compatibility of the RLIM.

To further defend our choice of the RLIM, we test for the simplest form of contamination that would render isolation invalid and RLIM non-incentive compatible. We test the hypothesis that in answering any question, subjects takes into account the decision made on the immediately preceding question, by hypothesizing that subjects weigh the current decision with ω $(0 \leq \omega \leq 1)$ and the previous decision with $1 - \omega$ (Hey and Zhou, 2014). Other contamination hypotheses have also been considered (Hey and Lee, 2005b,a) which are, admittedly, highly cognitively demanding: 1) in answering each question subjects consider the experiment as a whole; 2) in answering any question subjects take into account their answers to all the preceding questions. We would rationally expect that if a low cognitively demanding contamination hypothesis is rejected, it is unlikely that subjects choose based on more complicated forms of contamination.

The simple contamination form we explore here was first set forth by Hey and Zhou (2014). In notation form, when a subject is facing a decision, she is faced with a choice between the compound lotteries $(d^{n-1}, (1-\omega); A^n, \omega)$ and $(d^{n-1}, (1-\omega); B^n, \omega)$ where A^n, B^n are lotteries A and B, respectively, that subject faces in the n$^{\text{th}}$ decision. d^{n-1} is the lottery chosen in the previous $n-1$ decision, that is, $d^{n-1} = (A_1, p; A_2, 1-p)$ or $d^{n-1} = (B_1, p; B_2, 1-p)$ where A_1, A_2, B_1, B_2, p are the outcomes and probabilities of lotteries shown in Tables 2 and 3. Note that when $\omega = 1$ the subject separates completely and there is no contamination. When we estimate this model for the preferred RDU specification (choosing between alternative probability weighting functions and stochastic error specifications is discussed in the Results section) we estimate $\omega = 0.977$. A Wald test of whether $\omega = 1$ fails to reject the null (p-value=0.779) indicating that isolation of choice tasks is a plausible hypothesis with our data.

from worst to best and $w(\cdot)$ is the probability weighting function.

There are many probability weighting functions that have been used in the literature and here we consider various one and two parameter functions:

1. The power function (Quiggin, 1982): $w(p) = p^{\beta}$
2. Tversky and Kahneman's (1992) (TK) function: $w(p) = \frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{\frac{1}{\gamma}}}$ (if $\gamma = 1$ it collapses to $w(p) = p$)
3. The linear-in-log odds (LinLog) function (Goldstein and Einhorn, 1987; Lattimore et al., 1992; Tversky and Fox, 1995; Gonzalez and Wu, 1999): $w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma}+(1-p)^{\gamma}}$ where $\delta > 0$, $\gamma > 0$ (if $\delta = \gamma = 1$ it collapses to $w(p) = p$; if $\delta = 1$, $\gamma \neq 1$ it collapses to Karmarkar's (1978; 1979) one parameter probability weighting function)
4. Prelec's (Prelec, 1998) one parameter function: $w(p) = exp(-(-lnp)^{a})$ where $0 < a$, $0 < p < 1$ (if $a = 1$ it collapses to $w(p) = p$)
5. General Prelec (two parameter) function (Prelec, 1998)[9]: $w(p) = exp(-\beta(-lnp)^{a})$ where $a > 0, 0 < p < 1, \beta > 0$ (if $a = 1$ it collapses to the power function $w(p) = p^{\beta}$; if $a = \beta = 1$ it collapses to $w(p) = p$)

4.1 Stochastic error specifications

We assume subjects have some latent preferences over risk which are linked to observed choices via a probabilistic model function of the general form:

$$Pr(B) = F\left(\mu\frac{(V_B - V_A)}{D}\right) \tag{4}$$

where $Pr(B)$ is the probability of choosing lottery B (the right lottery), μ is a structural 'noise parameter' (sometimes called a scale or precision parameter) used to allow some errors from the perspective of the deterministic model and V_A, V_B are the decision-theoretic representations of values associated with lotteries A and B i.e., $V_j = EU_j$ for $j = A, B$ if the theory is EU or $V_j = RDU_j$ for $j = A, B$ if the theory is RDU. $F : R \to [0, 1]$ is an increasing function with $F(0) = 0.5$ and $F(x) = 1 - F(-x)$, which is to say that this function takes any argument between $\pm\infty$ and transforms it to a number between 0 and 1 i.e., a probability. The F function comes into two flavors in the respective literature: the

[9] Note, that both Prelec functions are often applied with the constraint $0 < a < 1$ which requires that the probability weighting function exhibits subproportionality (weighting function exhibits an inverse-S shape form). We follow Andersen et al. (2014, 2015); Harrison and Ng (2016) and use the more general specification from Prelec (1998, Proposition 1: (C)), which only requires $a > 0$ and nests the case where $0 < a < 1$.

cumulative standard normal distribution function Φ (the probit link) and the standard logistic distribution function Λ with $\Lambda(\zeta) = 1/(1 + e^{-\zeta})$ (the logit link). D adjusts the scale parameter in heteroskedastic models.

One popular class of models derives from equation (4) when we restrict $D = 1$. This is a class of homoskedastic latent index models also known as Fechnerian or Strong utility models (see Drichoutis and Lusk, 2014). The model with the logit link is equivalent to $Pr(B) = \Lambda\left(\mu(V_B - V_A)\right) = \frac{exp(\mu V_B)}{exp(\mu V_A)+exp(\mu V_B)}$. Another type of the homoskedastic class of models, called Luce or Strict utility models, uses the logarithm of values in the numerator of equation (4): $Pr(B) = \Lambda\left(\mu(ln[V_B] - ln[V_A])\right)$ which is equivalent to $Pr(B) = \frac{(V_B)^\mu}{(V_A)^\mu+(V_B)^\mu}$.

A second class of models, the heteroskedastic class, derives from equation (4) when $D \neq 1$. Wilcox (2008, 2011) proposed a 'contextual utility' error specification which adjusts the scale parameter by $D = V_{max} - V_{min}$ to account for the range of possible outcome utilities. D is defined as the maximum utility V_{max} over all prizes in a lottery pair minus the minimum utility V_{min} over all prizes in the same lottery pair. It changes from lottery pair to lottery pair, and thus it is said to be contextual. Contextual utility maintains that the error specification is mediated by the range of possible outcome utilities in a pair, so that $Pr(B) = F\left(\mu\frac{(V_B - V_A)}{V_{max}-V_{min}}\right)$.

Another heteroskedastic model which has received some attention in economics lately (Hey et al., 2010; Wilcox, 2015) is prescribed by Decision Field Theory (DFT) (Busemeyer and Townsend, 1992, 1993). DFT allows the decision maker's attention to switch from one event to another across choice pairs. This variability on focus on events is caused by a random difference which Busemeyer and Townsend (1993) name a *valence difference*. The variance of this valence difference in the case of lotteries with just two outcomes is given by $D^2 = w(p_1)(V_{A1} - V_{B1})^2 + (1 - w(p_1))(V_{A2} - V_{B2})^2 - (V_A - V_B)^2$ where V_{A1}, V_{A2}, V_{B1} and V_{B2} are the representations of values associated with the first and second outcome of lottery A and B, respectively. Note that when lotteries are certainties, such as in the last row of the HL task, then $D = 0$ and $Pr(B) = 1$, that is the subject always chooses the dominating lottery.[10]

[10] As a practical note, since $D = 0$ when lotteries are certainties, the last row of the HL task defined over certainties must be excluded from estimation.

4.2 Estimation

After defining the decision theoretical models and error specifications, the log-likelihood function can then be written as:

$$\ln L(y) = \sum_{i=1}^{N} \left[(\ln Z | y_i = 1) + (\ln(1 - Z) | y_i = -1) \right] \tag{5}$$

where $Z = Pr_j$ and j indexes the different error models (j =FP, FL, STRICT, CP, CL, DFTP, DFTL).[11] $y_i = 1$ denotes the choice of lottery B and $y_i = -1$ denotes the choice of the A lottery in the risk preference task i. Subjects were allowed to express indifference between choices and were told that if that choice was selected to be played out, the computer would randomly choose one of the two options for them and that both choices had equal chances of being selected. The likelihood function for indifferent choices is constructed such that it implies a 50/50 mixture of the likelihood of choosing either lottery so that (5) can be rewritten as:

$$\ln L(y) = \sum_{i=1}^{N} \left[(\ln Z | y_i = 1) + (\ln(1 - Z) | y_i = -1) + (\frac{1}{2} \ln Z + \frac{1}{2} \ln(1 - Z) | y_i = 0) \right] \tag{6}$$

Equation (6) is maximized using standard numerical methods. The statistical specification also takes into account the multiple responses given by the same subject and allows for correlation between responses by clustering standard errors i.e., it relaxes the independence assumption and requires only that the observations be independent across the clusters. The robust estimator of variance that relaxes the assumption of independent observations involves a slight modiffication of the robust (or sandwich) estimator of variance which requires independence across all observations (StataCorp, 2013, pp. 312).

5 Results

5.1 Was scent diffusion successful?

Table 4 shows the number of subjects answering with a Yes/No in the scent awareness question. There is a marked shift toward 'Yes' responses in the scent treatment (a χ^2 test rejects the null of no difference between treatments; p-value < 0.001) which is a good

[11] FP and FL stand for the Fechner error with a probit and a logit link, respectively. CP and CL stand for contextual utility with a probit and a logit link, respectively. DFTP and DFTL stand for Decision Field theory with a probit and a logit link respectively. STRICT stands for Luce error or Strict utility

Table 4: Attentive awareness of scent in comparison to the no scent treatment

		Treatment	
		No scent	Scent
Perceived existence of scent	Yes	14	39
	No	66	41

indication that the scent treatment was successful in exogenously varying awareness of the olfactory stimulus.

Table 4 shows that subjects in the scent treatment are about equally split in two groups. We call the group that perceived awareness of the scent as the *supraliminal scent* group and the group that did not perceive the existence of the scent as the *subliminal scent* group. We can explore the factors that contributed to scent awareness by means of a logit regression. Model (1) in Table 5 shows results from a logit regression of scent awareness on the horizontal and vertical distance of a subject's booth from the scent dispenser. As evident none of these variables is statistically significant which is to be interpreted that being close or away from the dispenser was not a factor that determines awareness of the scent.

Model (2) augments the specification by adding variables that aim to capture factors that may affect the sense of smell such as antibiotics and nasal medicine use, smoking status, any known to the subject taste and olfactory dysfunction as well as gender and age. None of the variables is statistically significant. In fact, a χ^2 test of the joint significance of all variables fails to reject the null at conventional statistical significance levels.

These results are reassuring in that they show that perceived awareness of the scent was only determined by subjects' nasal chemosensory performance. Olfactory sensitivity is determined by the odor threshold (i.e., the lowest concentration of a certain odor compound that is perceivable by the human sense of smell) which can vary widely between subjects (e.g., Lawless et al., 1995; Wilby, 1969). This natural variation in odor thresholds reflects the split of subjects into the supraliminal and subliminal scent groups.

5.2 Scent effects on willingness to pay

We can gain some first insights by looking at scatter graphs of bids. Figure 1 shows a scatter plot of bids by treatment where the two axis show bids for the two auctioned products (mug on the vertical axis and chocolate on the horizontal axis). The graph illustrates a larger spread of bids in the scent treatment which implies higher WTP for both products. Figure 2 shows bids for the scent treatment split between the supraliminal and subliminal scent groups. With respect to mug, bids tend to overlap for the two scent groups. For the

Table 5: Logit regressions of supraliminal awareness of scent

	(1)		(2)	
Constant	0.425	(0.792)	-2.655	(3.536)
H1	0.370	(0.632)	0.383	(0.674)
H2	-0.739	(0.628)	-0.616	(0.657)
V2	-0.218	(0.770)	-0.209	(0.834)
V3	-0.672	(0.746)	-0.600	(0.809)
V4	-0.384	(0.690)	-0.249	(0.722)
Male			-0.102	(0.531)
Age			0.177	(0.158)
Olfactory dysfunction: No			0.932	(0.731)
Antibiotics use: No			-0.172	(0.741)
Nasal medicine use: No			-0.489	(1.033)
Smoking: No			-0.414	(0.708)
Smoking: Occasionally			-0.482	(0.972)
Taste dysfunction: No			-0.382	(1.240)
N	80		80	
Log-likelihood	-52.276		-50.611	
χ^2 (p-value)	6.30 (0.278)		9.63 (0.724)	

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

chocolate product bids are spread more to the right of the graph, indicating a higher WTP for the supraliminal scent group but tend to overlap on the vertical axis, indicating a similar WTP for the two scent groups. In both figures, bids are concentrated above the 45°degree line indicating a higher WTP for the mug than the chocolate.

Simple statistical tests support the pattern described above. Table 6 shows mean, median and standard deviation of bids per product, per treatment and per supraliminal/subliminal group. The upper part of the table indicates a higher mean and median WTP for the scent treatment as well as a larger spread. The Kruskal-Wallis test (Kruskal and Wallis, 1952) and the K-sample median test (Mood, 1954) indicate that these differences are statistically significant. The lower part of the table examines differences between the supraliminal and subliminal scent group. As indicated, the null for the chocolate is rejected for both tests. However, both tests fail to reject the null of no difference for the mug at the 5% level.

To check whether the results obtained above hold in the context of conditional analysis as well as to quantify treatment effects, we estimated random effects regression models where the grouping structure of the data consists of three levels of nested groups (i.e., three random effects): the auction group, j, the individual, i, and the auction round, t. The model specification we estimate is of the form:

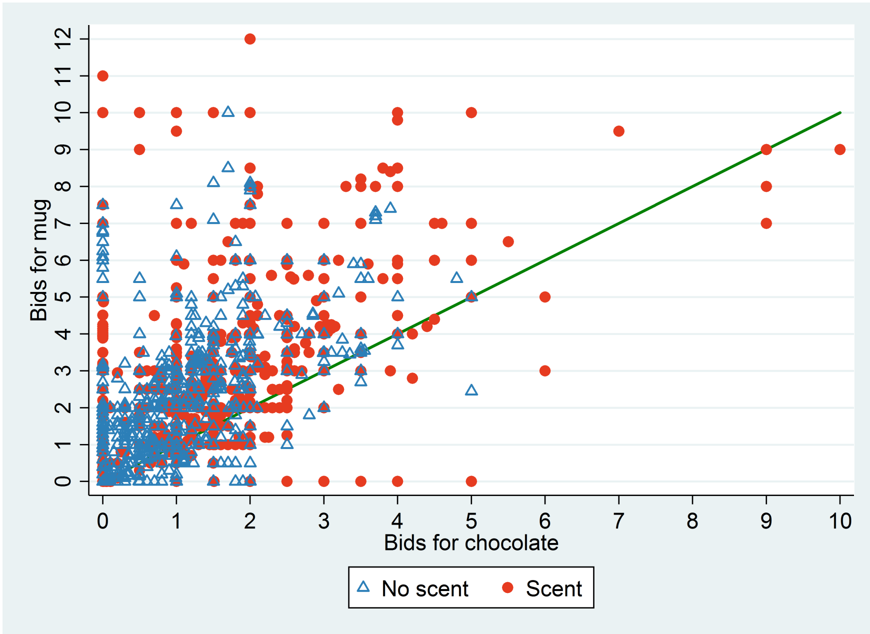


Figure 1: Bids per product and treatment

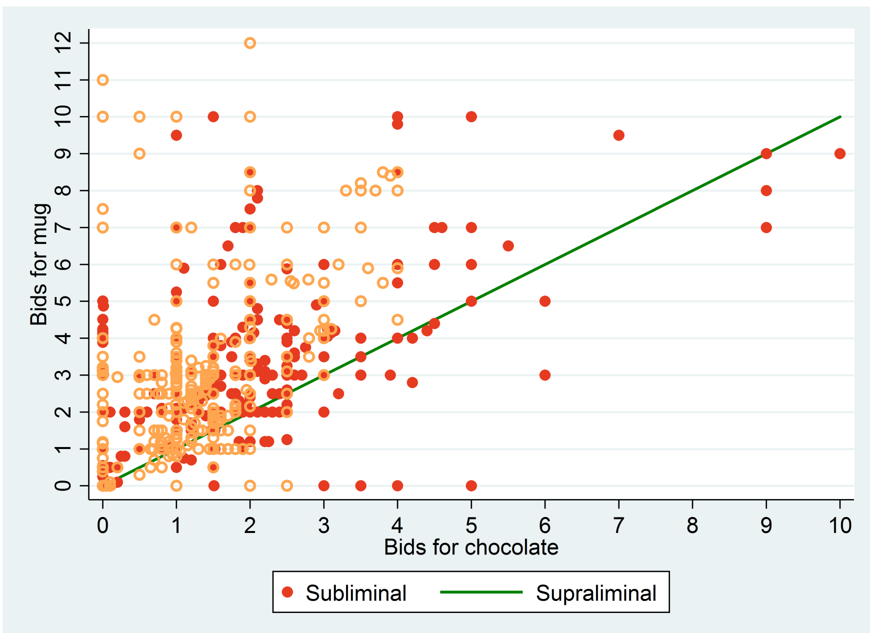


Figure 2: Bids per product for the Scent treatment by supraliminal/subliminal group

Table 6: Descriptive statistics of bids per product and treatment

	Mug			Chocolate		
	Mean	S.D.	Median	Mean	S.D.	Median
No cent	2.15	1.83	1.80	1.10	1.02	0.97
Scent	2.95	2.12	2.60	1.57	1.38	1.32
Kruskal-Wallis test	$\chi^2 = 69.96$, $p < 0.001$			$\chi^2 = 69.94$, $p < 0.001$		
K-sample median test	$\chi^2 = 52.57$, $p < 0.001$			$\chi^2 = 52.67$, $p < 0.001$		
Scent: Supraliminal	3.05	2.13	2.80	1.90	1.66	1.51
Scent: Subliminal	2.85	2.10	2.50	1.25	0.94	1.00
Kruskal-Wallis test	$\chi^2 = 3.36$, $p = 0.067$			$\chi^2 = 42.56$, $p < 0.001$		
K-sample median test	$\chi^2 = 0.176$, $p = 0.676$			$\chi^2 = 32.83$, $p < 0.001$		

$$Bid^*_{jit} = x_{jit}b + u_j + v_{ji} + \varepsilon_{jit} \tag{7}$$

where $j = 1 \ldots J$ indexes the auction groups, $i = i \ldots N$ indexes individuals in an auction group, $t = 1 \ldots T$ indexes auction rounds (in our case $J = 32$, $N = 5$ and $T = 10$) and x is a vector of independent variables. The random effects, u_j, v_{ji} and ε_{jit} are i.i.d. $N(0, \sigma_u^2)$, $N(0, \sigma_v^2)$ and $N(0, \sigma_\varepsilon^2)$, respectively and independently of each other.

In addition, about 9.3% of all bids for the mug and 13.5% for the chocolate are exactly zero. This calls for the use of a censored regression model to address possible censoring from the left (Tobit model). The Tobit model complicates slightly the analysis since there are four marginal effect that the researcher might be interested in: a) marginal effects on the latent variable, $\frac{\partial E[Bid^*|x]}{\partial x}$ (these are the raw coefficient estimates) b) on the observed variable, $\frac{\partial E[Bid|x]}{\partial x}$ c) on positive bids, $\frac{\partial E[Bid|Bid>0,x]}{\partial x}$ and d) on the probability of being uncensored, $\frac{\partial Pr[Bid>0|x]}{\partial x}$.

Results (raw coefficient estimates) are exhibited in Table 7 (Table A.1 in Appendix A shows results with additional demographic and attitudinal variables added in the model specification). Specifications (1), (3) and (5) show results for the mug, chocolate and a pooled model respectively, where a treatment dummy (Scent) is added in the model specification. The scent treatment dummy is positive and statistically significant which is consistent with a higher WTP under the influence of the olfactory stimuli. Models (2), (4) and (6) replace the Scent treatment dummy with two dummies: one for the supraliminal scent group and one for the subliminal scent group (with the no scent treatment serving as the base category).

Table 7 shows that for the mug product, both the supraliminal and subliminal scent groups exert a positive and statistically significant effect on bids (albeit at the 10% level). For the chocolate product, there is a similar positive effect for the two groups but it is

Table 7: Random effects Tobit models

	Mug		Chocolate		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-1.618	-1.617	-0.061	-0.008	0.214	0.215
	(2.042)	(2.042)	(1.442)	(1.418)	(1.509)	(1.504)
Scent	0.840*		0.500**		0.651**	
	(0.446)		(0.241)		(0.331)	
Scent: Subliminal		0.804*		0.215		0.511
		(0.475)		(0.263)		(0.352)
Scent: Supraliminal		0.877*		0.793***		0.797**
		(0.478)		(0.267)		(0.355)
Round	0.042***	0.042***	-0.006	-0.006	0.017**	0.017**
	(0.007)	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)
Endowment	-0.078	-0.078	-0.165	-0.164	-0.179	-0.179
	(0.393)	(0.393)	(0.281)	(0.277)	(0.304)	(0.303)
Hedonic score	0.573***	0.573***	0.284***	0.275***	0.411***	0.411***
	(0.081)	(0.081)	(0.062)	(0.061)	(0.023)	(0.023)
Chocolate					-1.305***	-1.305***
					(0.037)	(0.037)
σ_u	1.086***	1.085***	0.482***	0.473***	0.786***	0.779***
	(0.185)	(0.185)	(0.129)	(0.127)	(0.141)	(0.140)
σ_v	1.426***	1.426***	1.052***	1.033***	1.104***	1.101***
	(0.099)	(0.099)	(0.074)	(0.072)	(0.075)	(0.075)
σ_ε	0.803***	0.803***	0.568***	0.568***	1.019***	1.019***
	(0.016)	(0.016)	(0.011)	(0.011)	(0.014)	(0.014)
Observations	1600	1600	1600	1600	3200	3200
Log-likelihood	-2081.069	-2081.046	-1557.631	-1554.840	-4560.477	-4559.881

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

statistically significant only for the supraliminal scent group. A Wald test of whether the coefficients of the supraliminal and the subliminal scent groups are equal rejects the null for the chocolate (p-value = 0.017). Taken together, our results imply there is a differential effect of supraliminal and subliminal perception of scent over the food and non-food item. Given the congruency of the chocolate with the citrus scent in the food dimension, we can interpret the effect of the scent on chocolate as the result of conscious awareness of the scent which takes precedence over pleasantness (Olofsson et al., 2012). On the other hand, given the incongruence of the scent with the mug, the effect for the mug comes through the pleasantness of the room which doesn't require a conscious awareness of the scent.

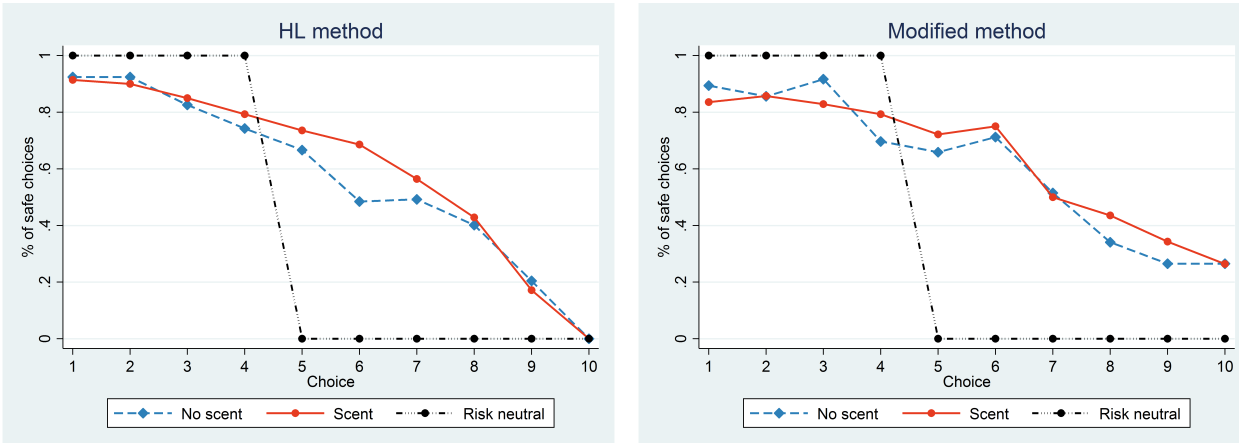
A few other factors that affect bidding behavior are the hedonic score variable which indicates a positive effect on bidding behavior, with higher bids for subjects that liked more the respective product. The pooled model indicates a lower valuation for the chocolate with respect to the mug.

Given that the raw coefficients of Table 7 show the effect on the latent variable, it is more interesting to examine other marginal effects. Table A.2 in Appendix A shows marginal effects for models (2) and (4) of Table 7. For example, for the mug, the supraliminal and the subliminal scent groups have an 8.5% and 8% chance, respectively, of bidding positively as compared to the no scent group (this refers to the $\frac{\partial Pr[Bid>0|x]}{\partial x}$ labeled column). For the chocolate product, the supraliminal group has a 14% chance of bidding positively than the no scent group, while for the subliminal group the effect is not statistically significant.

In addition, in order to get a sense of the economic significance of the estimated marginal effects, one can interpret marginal changes in terms of the predicted WTP. The average predicted WTP for the columns labeled $\frac{\partial E[Bid|x]}{\partial x}$ in Table A.2 is €2.56 and €1.36 for the mug and chocolate, respectively. If we take the estimated marginal effects for the scent groups for the mug and divide over the average prediction, these effects would correspond to a 27.09% (=0.694/2.56) and 29.67%(=0.76/2.56) change for the subliminal and supraliminal group, respectively. For the chocolate, marginal effects as a proportion of predicted WTP are 12.53%(=0.17/1.36) and 49.16%(=0.667/1.36) for the subliminal and supraliminal group, respectively. These are all substantial effects.

5.3 Scent effects and risk preferences

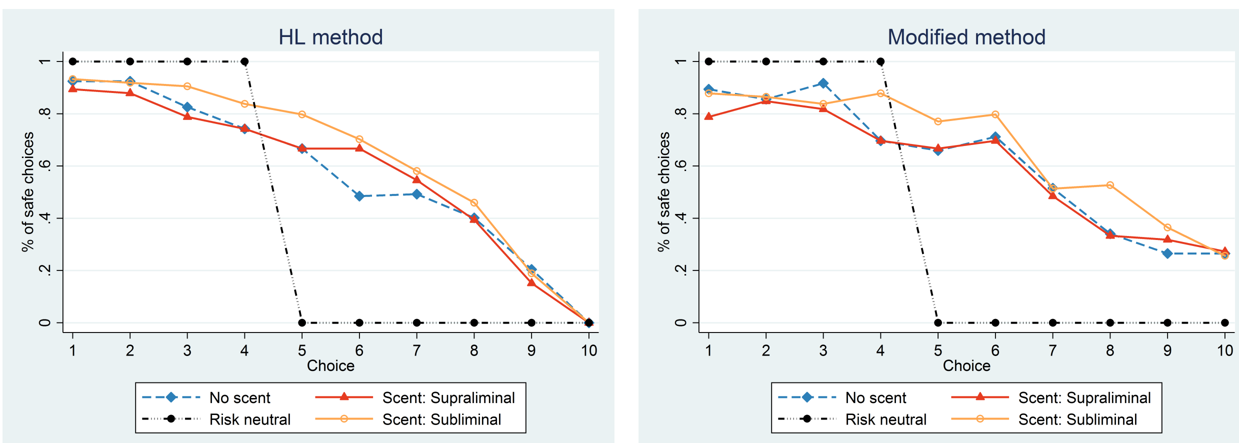
Figures 3 and 4 graph the percent of subjects that chose lottery A at any given choice task. The black dashed line depicts the choices of a risk neutral person assuming a CRRA utility function and EUT. Deviations from the risk neutral line, in the pattern shown in the graph, are taken as indications or risk averse behavior. The differences between the lines are generally small. When we consider the supraliminal and subliminal scent groups separately, Figure 4 shows a slightly more risk averse behavior for the subliminal group but an overlap of lines for the control and the supraliminal group.



(a) Holt and Laury (2002) task

(b) Modified task

Figure 3: Percent of subjects choosing the safe choice per treatment group



(a) Holt and Laury (2002) task

(b) Modified task

Figure 4: Percent of subjects choosing the safe choice per treatment and supraliminal/subliminal groups

However, as mentioned in Section 4, an analysis of risk choice behavior based only on the number of safe choices, ignores a significant strand of the literature concerned with modeling noise in risk choice data. In order to select between the competing stochastic models and probability weighting functions, we first compared models using Akaike's and Bayesian information criteria (AIC and BIC). AIC and BIC do not reveal how well a model fits the data in an absolute sense, i.e., there is no null hypothesis being tested. Nevertheless, these measures offer relative comparisons between models on the basis of information lost from using a model to represent the (unknown) true model.[12]

Given that convergence problems may occur as one tries to add covariates to the basic specification and then end up with specifications with different sets of covariates, we fitted all models at baseline with no covariates and then calculated AIC and BIC. Table A.3 in Appendix A shows AIC and BIC measures for all the combinations of error stories and probability weighting functions. As shown, the Decision Field theory with a logit link shows the best fit with our data for both decision theories (EUT and RDU). Across all model specifications estimated with the DFT with logit link, IC measures show that the one parameter Prelec function should be our choice of a probability weighting function.

Table 8 shows structural estimates where the parameters of interest are modeled with additional treatment covariates (Table A.4 in Appendix A shows results where the specification is augmented with additional demographic and attitudinal variables). Both decision theories are presented (EUT and RDU) for comparison. We note however, that a test of whether RDU collapses to EUT ($\alpha = 1$) is rejected ($p < 0.001$). Nevertheless, we briefly note that the EUT specification echoes the results from the graphs. Results from model (1) which uses only a scent dummy, show no statistical significant association with risk aversion. When we replace the scent dummy with a supraliminal and a subliminal scent dummy (no scent is the base outcome; model (2)), we find a statistically significant effect for the subliminal group but a null effect for the supraliminal group.

Models (3) and (4) show the effect of covariates on r and a, that is the curvature of the utility function and the curvature of the probability weighting function. With respect to the curvatures of the probability weighting function, results unambiguously show no effect of any of the scent dummies. With respect to the curvature of the utility function, when we assume RDU, the effect of the subliminal group lowers slightly in magnitude and is significant only at a higher threshold ($\alpha = 10\%$). Table A.4 shows that when we augment this specification with additional variables, we fail to reject the null of no effect for both the supraliminal and the subliminal scent groups. Thus, we can conclude that we do not observe a significant effect of scent on risk aversion, at least not a robust one.

[12] Drichoutis and Lusk (2016) have shown that AIC and BIC are in agreement with more complex selection criteria such as Vuong's test (Vuong, 1989), Clarke's test (Clarke, 2003) or the out-of-sample log likelihood (OSLLF) criterion (Norwood et al., 2004).

Table 8: Estimates for EUT and RDU given the Decision Field theory stochastic assumption

	EUT		RDU			
	(1)	(2)	(3)		(4)	
	r	r	r	α	r	α
Constant	0.249	0.234	-0.105	-0.394	-0.195	-0.626
	(1.019)	(1.061)	(0.943)	(0.671)	(0.993)	(0.935)
Scent	0.101		0.095	0.006		
	(0.084)		(0.075)	(0.022)		
Scent: Subliminal		0.190**			0.171*	0.006
		(0.093)			(0.098)	(0.026)
Scent: Supraliminal		-0.004			0.011	0.003
		(0.119)			(0.128)	(0.033)
H&L task	-0.033	-0.030	-0.453	-0.257	-0.326	-0.149
	(0.058)	(0.058)	(0.407)	(0.528)	(0.408)	(0.397)
Endowment	0.076	0.078	0.178	0.044***	0.178	0.047**
	(0.209)	(0.217)	(0.177)	(0.013)	(0.188)	(0.019)
μ	2.105***	2.114***	3.022***		2.816***	
	(0.152)	(0.149)	(0.779)		(0.681)	
Observations	2584	2584	2584		2584	
Log-likelihood	-1409.574	-1405.859	-1392.407		-1388.677	

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusions

We examined the effect of a citrus scent on willingness-to-pay and choice under risk on a between subjects basis. Our results generally confirm the large literature from the marketing and psychology fields which indicates that scents may induce consumers in spending more by increasing their valuation for the product.

We also find a differential effect between a food and a non-food item which we attribute to the congruency/incongruency of the scent with the product. More specifically, given the fruity but pleasant nature of the scent which is incongruent with the mug, we find for the mug that it exerts a similar effect in both the supraliminal and the subliminal scent groups. This is because the effect of the scent for the non-food item can be attributed to the general pleasantness of the room despite the incongruence with the product. For the food item, the scent can be considered congruent, thus it is expected to have an effect on WTP only for those subjects that are supraliminally aware of the scent.

For the risk choice tasks, we find that the effect of scents on risk aversion is sensitive to the decision theory one assumes. Under EUT we find a significant effect of the scent on the curvature of the utility function while statistical significance vanishes under RDU. We find no effect of any of the scent dummy variables on the curvature of the probability weighting function.

Coming back to the casino studies (Hirsch, 1995; Hancock, 2009), based on our null result it would be tempting to rule out any effect of scents on risk and conclude that increased revenues in those studies could be attributed to scents altering the pleasantness of rooms and thus attracting larger groups of people in the slot machine areas. We need a larger pool of studies to allow for more definite conclusions, so we urge researchers to embark on a research agenda that will evaluate sensory experiences on economic decision making using rigorous experimental economics methods.

References

Andersen, S., G. W. Harrison, M. Lau, and E. E. Rutström (2015). Multiattribute utility theory, intertemporal utility and correlation aversion. *Center for the Economic Analysis of Risk, Working Paper 2011-03*.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2014). Discounting behavior: A reconsideration. *European Economic Review 71*, 15–33.

Bargh, J. A. (1992). Does subliminality matter to social psychology? awareness of the stimulus versus awareness of its influence. In R. F. Bornstein and T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives*, pp. 236–255.

Baron, R. A. (1981). Olfaction and human social behavior: Effects of a pleasant scent on attraction and social perception. *Personality and Social Psychology Bulletin 7*(4), 611–616.

Baron, R. A. (1983). "Sweet smell of success"? The impact of pleasant artificial scents on evaluations of job applicants. *Journal of Applied Psychology 68*(4), 709–713.

Bone, P. F. and S. Jantrania (1992). Olfaction as a cue for product quality. *Marketing Letters 3*(3), 289–296.

Bosmans, A. (2006). Scents and sensibility: When do (in)congruent ambient scents influence product evaluations? *Journal of Marketing 70*(3), 32–43.

Bradford, K. D. and D. M. Desrochers (2009). The use of scents to influence consumers: The sense of using scents to make cents. *Journal of Business Ethics 90*(2), 141–153.

Busemeyer, J. R. and J. T. Townsend (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences 23*(3), 255–282.

Busemeyer, J. R. and J. T. Townsend (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review 100*(3), 432–459.

Cain, W. (1979). To know with the nose: keys to odor identification. *Science 203*(4379), 467–470.

Chebat, J.-C. and R. Michon (2003). Impact of ambient odors on mall shoppers' emotions, cognition, and spending: A test of competitive causal theories. *Journal of Business Research 56*(7), 529–539.

Chebat, J.-C., M. Morrin, and D.-R. Chebat (2009). Does age attenuate the impact of pleasant ambient scent on consumer response? *Environment and Behavior 41*(2), 258–267.

Clarke, K. A. (2003). Nonparametric model discrimination in international relations. *Journal of Conflict Resolution 47*(1), 72–93.

Corbett, J. (1994). *Extended Play: Sounding Off from John Cage to Dr. Funkenstein.* Durham, London: Duke University Press.

Corgnet, B., R. Hernn-Gonzlez, P. Kujal, and D. Porter (2014). The effect of earned versus house money on price bubble formation in experimental asset markets. *Review of Finance 19*(4), 1455–1488.

Corrigan, J. R., A. C. Drichoutis, J. L. Lusk, J. R.M. Nayga, and M. C. Rousu (2012). Repeated rounds with price feedback in experimental auction valuation: An adversarial collaboration. *American Journal of Agricultural Economics 94*(1), 97–115.

Cox, J., V. Sadiraj, and U. Schmidt (2014). Paradoxes and mechanisms for choice under risk. *Experimental Economics*, 1–36.

de Wijk, R. A. and S. M. Zijlstra (2012). Differential effects of exposure to ambient vanilla and citrus aromas on mood, arousal and food choice. *Flavour 1*(1), 1–7.

Demattè, M. L., R. Österbauer, and C. Spence (2007). Olfactory cues modulate facial attractiveness. *Chemical Senses 32*(6), 603–610.

Doucé, L. and W. Janssens (2013). The presence of a pleasant ambient scent in a fashion store: The moderating role of shopping motivation and affect intensity. *Environment and Behavior 45*(2), 215–238.

Drichoutis, A. C. and J. L. Lusk (2014). Judging statistical models of individual decision making under risk using in- and out-of-sample criteria. *PLoS ONE 9*(7), e102269.

Drichoutis, A. C. and J. L. Lusk (2016). What can multiple price lists really tell us about risk preferences? *Journal of Risk and Uncertainty 53*(2).

Ehrlichman, H. and J. N. Halpern (1988). Affect and memory: Effects of pleasant and unpleasant odors on retrieval of happy and unhappy memories. *Journal of Personality and Social Psychology 55*(5), 769–779.

Fiore, A. M., X. Yah, and E. Yoh (2000). Effects of a product display and environmental fragrancing on approach responses and pleasurable experiences. *Psychology and Marketing 17*(1), 27–54.

Gagarina, A. and I. Pikturnienė (2015). The effect of ambient scent type and intensiveness on decision making heuristics. *Procedia - Social and Behavioral Sciences 213*, 605–609.

Goldkuhl, L. and M. Styvén (2007). Sensing the scent of service success. *European Journal of Marketing 41*(11/12), 1297–1305.

Goldstein, W. M. and H. J. Einhorn (1987). Expression theory and the preference reversal phenomena. *Psychological Review 94*(2), 236–254.

Gonzalez, R. and G. Wu (1999). On the shape of the probability weighting function. *Cognitive Psychology 38*(1), 129–166.

Grabenhorst, F., E. T. Rolls, C. Margot, M. A. A. P. da Silva, and M. I. Velazco (2007). How pleasant and unpleasant stimuli combine in different brain regions: Odor mixtures. *The Journal of Neuroscience 27*(49), 13532–13540.

Guéguen, N. and C. Petr (2006). Odors and consumer behavior in a restaurant. *International Journal of Hospitality Management 25*(2), 335–339.

Hancock, G. D. (2009). The efficacy of fragrance use for enhancing the slot machine gaming experience of casino patrons. *Doctoral dissertation, William F. Harrah College of Hotel Administration, University of Nevada, Las Vegas*.

Harrison, G. W., A. Hofmeyr, D. Ross, and J. T. Swarthout (2015). Risk preferences, time preferences and smoking behaviour. *Center for the Economic Analysis of Risk, Working Paper 2015-11*.

Harrison, G. W. and J. M. Ng (2016). Evaluating the expected welfare gain from insurance. *Journal of Risk and Insurance 83*(1), 91–120.

Harrison, G. W. and E. E. Rutström (2008). Risk aversion in the laboratory. In J. C. Cox and G. W. Harrison (Eds.), *Research in Experimental Economics Vol 12: Risk Aversion in Experiments*, Volume 12, pp. 41–196. Bingley, UK: Emerald Group Publishing Limited.

Harrison, G. W. and J. T. Swarthout (2014). Experimental payment protocols and the bipolar behaviorist. *Theory and Decision 77*(3), 423–438.

Harrison, G. W. and J. T. Swarthout (2016). Cumulative prospect theory in the laboratory: A reconsideration. *Center for the Economic Analysis of Risk, Working Paper 2016-02*.

Herz, R. S. and J. von Clef (2001). The influence of verbal labeling on the perception of odors: Evidence for olfactory illusions? *Perception 30*(3), 381–391.

Hey, J. (2005). Why we should not be silent about noise. *Experimental Economics 8*(4), 325–345.

Hey, J. (2014). My experimental meanderings. *Theory and Decision 77*(3), 291–295.

Hey, J. and W. Zhou (2014). Do past decisions influence future decisions? *Applied Economics Letters 21*(3), 152–157.

Hey, J. D. and J. Lee (2005a). Do subjects remember the past? *Applied Economics 37*(1), 9–18.

Hey, J. D. and J. Lee (2005b). Do subjects separate (or are they sophisticated)? *Experimental Economics 8*(3), 233–265.

Hey, J. D., G. Lotito, and A. Maffioletti (2010). The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. *Journal of Risk and Uncertainty 41*(2), 81–111.

Hirsch, A. R. (1990). Preliminary results of olfaction nike study. *Smell and Taste Treatment and Research Foundation Ltd. Chicago*.

Hirsch, A. R. (1995). Effects of ambient odors on slot-machine usage in a las vegas casino. *Psychology and Marketing 12*(7), 585–594.

Holt, C. A. (1986). Preference reversals and the independence axiom. *The American Economic Review 76*(3), 508–515.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *The American Economic Review 92*(5), 1644–1655.

Homburg, C., M. Imschloß, and C. Kühnl (2012). Of dollars and senses - Does multisensory marketing pay off? *Research Insights Paper RI009 of the Institute for Market-Oriented Management, University of Mannheim.*

Hoover, K. C. (2010). Smell with inspiration: The evolutionary significance of olfaction. *American Journal of Physical Anthropology 143*(S51), 63–74.

Jacquemet, N., R.-V. Joule, S. Luchini, and J. F. Shogren (2009). Earned wealth, engaged bidders? Evidence from a second-price auction. *Economics Letters 105*(1), 36–38.

Karmarkar, U. S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance 21*(1), 61–72.

Karmarkar, U. S. (1979). Subjectively weighted utility and the allais paradox. *Organizational Behavior and Human Performance 24*(1), 67–72.

Kruskal, W. H. and W. A. Wallis (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association 47*(260), 583–621.

Lattimore, P. K., J. R. Baker, and A. D. Witte (1992). The influence of probability on risky choice. *Journal of Economic Behavior & Organization 17*(3), 377–400.

Lawless, H. T., C. J. C. Thomas, and M. Johnston (1995). Variation in odor thresholds for l-carvone and cineole and correlations with suprathreshold intensity ratings. *Chemical Senses 20*(1), 9–17.

Li, W., I. Moallem, K. A. Paller, and J. A. Gottfried (2007). Subliminal smells can guide social preferences. *Psychological Science 18*(12), 1044–1049.

Lindstrom, M. (2005). *Brand sense: Build powerful brands through touch, taste, smell, sight, and sound*. New York: Free Press.

Liu, Y., F. Tovia, K. Balasubramian, J. D. Pierce, and J. Dugan (2008). Scent infused textiles to enhance consumer experiences. *Journal of Industrial Textiles 37*(3), 263–274.

Mattila, A. S. and J. Wirtz (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of Retailing 77*(2), 273–289.

Melcher, J. M. and J. W. Schooler (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language 35*(2), 231–245.

Michon, R., J. C. Chebat, and R. Michon (2006). The interaction effect of music and odour on shopper spending. *Doctoral dissertation, School of Retail Management, Ryerson University, Toronto, Canada.*

Michon, R., J.-C. Chebat, and L. W. Turley (2005). Mall atmospherics: the interaction effects of the mall environment on shopping behavior. *Journal of Business Research 58*(5), 576–583.

Mitchell, D. J., B. E. Kahn, and S. C. Knasko (1995). There's something in the air: Effects of congruent or incongruent ambient odor on consumer decision making. *Journal of Consumer Research 22*(2), 229–238.

Mood, A. M. (1954). On the asymptotic efficiency of certain nonparametric two-sample tests. *The Annals of Mathematical Statistics 25*(3), 514–522.

Morrin, M. and S. Ratneshwar (2000). The impact of ambient scent on evaluation, attention, and memory for familiar and unfamiliar brands. *Journal of Business Research 49*(2), 157–165.

Morrison, M., S. Gan, C. Dubelaar, and H. Oppewal (2011). In-store music and aroma influences on shopper behavior and satisfaction. *Journal of Business Research 64*(6), 558–564.

Norwood, B. F., M. C. Roberts, and J. L. Lusk (2004). Ranking crop yield models using out-of-sample likelihood functions. *American Journal of Agricultural Economics 86*(4), 1032–1043.

Olofsson, J. K., N. E. Bowman, K. Khatibi, and J. A. Gottfried (2012). A time-based account of the perception of odor objects and valences. *Psychological Science 23*(10), 1224–1232.

Parsons, A. G. (2009). Use of scent in a naturally odourless store. *International Journal of Retail & Distribution Management 37*(5), 440–452.

Prelec, D. (1998). The probability weighting function. *Econometrica 66*(3), 497–528.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization 3*(4), 323–343.

Shogren, J. F., M. Margolis, C. Koo, and J. A. List (2001). A random nth-price auction. *Journal of Economic Behavior and Organization 46*(4), 409–421.

Smeets, M. A. and G. B. Dijksterhuis (2014). Smelly primes - when olfactory primes do or do not work. *Frontiers in Psychology 5*.

Sorokowska, A., P. Sorokowski, and J. Havlicek (2016). Body odor based personality judgments: The effect of fragranced cosmetics. *Frontiers in Psychology 7*.

Spangenberg, E. R., A. E. Crowley, and P. W. Henderson (1996). Improving the store environment: Do olfactory cues affect evaluations and behaviors? *Journal of Marketing 60*(2), 67–80.

Spangenberg, E. R., D. E. Sprott, B. Grohmann, and D. L. Tracy (2006). Gender-congruent ambient scent influences on approach and avoidance behaviors in a retail store. *Journal of Business Research 59*(12), 1281–1287.

StataCorp (2013). *Stata user's guide release 13*. College Station, Texas, USA: Stata Press.

Stern, K. and M. K. McClintock (1998). Regulation of ovulation by human pheromones. *Nature 392*(6672), 177–179. 10.1038/32408.

Sutani, K., S. Iwaki, M. Yamaguchi, K. Uchida, and M. Tonoike (2007). Neuromagnetic brain responses evoked by pleasant and unpleasant olfactory stimuli. *International Congress Series 1300*, 391–394.

Tversky, A. and C. R. Fox (1995). Weighing risk and uncertainty. *Psychological Review 102*(2), 269–283.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance 16*(1), 8–37.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica 57*(2), 307–333.

Wada, Y., Y. Inada, J. Yang, S. Kunieda, T. Masuda, A. Kimura, S. Kanazawa, and M. K. Yamaguchi (2012). Infant visual preference for fruit enhanced by congruent in-season odor. *Appetite 58*(3), 1070–1075.

Wakker, P. P. (2007). Message to referees who want to embark on yet another discussion of the random-lottery incentive system for individual choice. Available at http://people.few.eur.nl/wakker/miscella/debates/randomlinc.htm.

Wilby, F. V. (1969). Variation in recognition odor threshold of a panel. *Journal of the Air Pollution Control Association 19*(2), 96–100.

Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox and G. W. Harrison (Eds.), *Research in Experimental Economics Vol 12: Risk Aversion in Experiments*, pp. 197–292. Bingley, UK: Emerald Group Publishing Limited.

Wilcox, N. (2015). Error and generalization in discrete choice under risk. *Economic Science Institute (Chapman University) Working paper*.

Wilcox, N. T. (2011). 'Stochastically more risk averse:' A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics 162*(1), 89–104.

A Appendix: Additional tables/figures

Table A.1: Random effects Tobit models (with additional controls)

	Mug		Chocolate		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-3.486	-3.481	-0.978	-0.739	-1.125	-1.053
	(2.319)	(2.321)	(1.656)	(1.632)	(1.707)	(1.705)
Scent	0.818*		0.397*		0.568*	
	(0.443)		(0.241)		(0.324)	
Scent: Subliminal		0.811*		0.163		0.475
		(0.465)		(0.260)		(0.342)
Scent: Supraliminal		0.828*		0.702**		0.694*
		(0.484)		(0.275)		(0.356)
Round	0.042***	0.042***	-0.006	-0.006	0.017**	0.017**
	(0.007)	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)
Endowment	-0.277	-0.276	-0.430	-0.403	-0.443	-0.433
	(0.398)	(0.399)	(0.281)	(0.276)	(0.304)	(0.303)
Hedonic score	0.586***	0.586***	0.296***	0.288***	0.413***	0.413***
	(0.079)	(0.079)	(0.061)	(0.060)	(0.023)	(0.023)
Chocolate					-1.305***	-1.305***
					(0.037)	(0.037)
Male	-0.215	-0.214	-0.203	-0.192	-0.161	-0.155
	(0.249)	(0.249)	(0.178)	(0.175)	(0.190)	(0.190)
Household size	0.025	0.024	-0.041	-0.046	0.000	-0.002
	(0.136)	(0.136)	(0.095)	(0.093)	(0.104)	(0.103)
Education: 3rd semester student	0.327	0.326	-0.236	-0.263	0.057	0.048
	(0.384)	(0.384)	(0.259)	(0.255)	(0.291)	(0.291)
Education: 5th semester student	0.638	0.640	0.001	0.050	0.352	0.380
	(0.447)	(0.449)	(0.301)	(0.297)	(0.339)	(0.340)
Education: 7th semester student	0.902*	0.899*	0.281	0.185	0.715*	0.676*
	(0.538)	(0.542)	(0.369)	(0.365)	(0.409)	(0.411)
Education: 9th semester student	-0.033	-0.037	0.021	-0.082	0.066	0.018
	(0.627)	(0.632)	(0.427)	(0.424)	(0.477)	(0.479)
Education: >9th semester student	0.703	0.702	0.591	0.552	0.760*	0.746*
	(0.528)	(0.528)	(0.371)	(0.365)	(0.403)	(0.402)
Income: Above average	0.727*	0.728*	0.396	0.422	0.581*	0.593*

	(0.411)	(0.412)	(0.289)	(0.284)	(0.312)	(0.312)
Income: Average	0.439	0.440	0.509**	0.542**	0.478*	0.495*
	(0.364)	(0.365)	(0.257)	(0.253)	(0.276)	(0.276)
Income: Below average	-0.334	-0.333	0.201	0.253	-0.085	-0.064
	(0.429)	(0.431)	(0.303)	(0.299)	(0.327)	(0.327)
Income: Bad or Very bad	-0.389	-0.388	0.011	0.055	-0.255	-0.233
	(0.452)	(0.453)	(0.324)	(0.318)	(0.344)	(0.344)
Body Mass Index	0.112***	0.112***	0.058**	0.054*	0.087***	0.086***
	(0.040)	(0.040)	(0.028)	(0.028)	(0.030)	(0.030)
Smoking: No	0.421	0.422	0.405	0.419	0.430	0.434
	(0.375)	(0.375)	(0.265)	(0.260)	(0.288)	(0.287)
Smoking: Occasionally	0.420	0.420	0.228	0.238	0.406	0.410
	(0.506)	(0.506)	(0.357)	(0.351)	(0.386)	(0.385)
Auction comprehension	-0.091	-0.092	0.054	0.025	-0.015	-0.027
	(0.102)	(0.104)	(0.071)	(0.071)	(0.077)	(0.078)
σ_u	1.065***	1.065***	0.483***	0.484***	0.759***	0.758***
	(0.186)	(0.186)	(0.125)	(0.123)	(0.138)	(0.138)
σ_v	1.340***	1.340***	0.975***	0.956***	1.023***	1.020***
	(0.094)	(0.094)	(0.069)	(0.068)	(0.071)	(0.071)
σ_ε	0.804***	0.804***	0.568***	0.568***	1.019***	1.019***
	(0.016)	(0.016)	(0.011)	(0.011)	(0.014)	(0.014)
Observations	1600	1600	1600	1600	3200	3200
Log-likelihood	-2072.354	-2072.353	-1547.532	-1545.013	-4549.733	-4549.375

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base categories are: No Scent, Education: 1st semester student, Income: Good or Very good, Smoking: Yes.

Table A.2: Marginal effects for random effects Tobit models (2) and (4) of Table 7

	Mug			Chocolate		
	$\frac{\partial E[Bid\|x]}{\partial x}$	$\frac{\partial E[Bid\|Bid>0,x]}{\partial x}$	$\frac{\partial Pr[Bid>0\|x]}{\partial x}$	$\frac{\partial E[Bid\|x]}{\partial x}$	$\frac{\partial E[Bid\|Bid>0,x]}{\partial x}$	$\frac{\partial Pr[Bid>0\|x]}{\partial x}$
Scent: Subliminal	0.694*	0.553*	0.080*	0.170	0.126	0.046
	(0.411)	(0.331)	(0.047)	(0.209)	(0.156)	(0.056)
Scent: Supraliminal	0.760*	0.608*	0.085*	0.667***	0.516***	0.140***
	(0.416)	(0.336)	(0.047)	(0.227)	(0.179)	(0.046)
Round	0.036***	0.029***	0.004***	-0.005	-0.004	-0.001
	(0.006)	(0.005)	(0.001)	(0.004)	(0.003)	(0.001)
Endowment	-0.067	-0.054	-0.008	-0.133	-0.101	-0.032
	(0.338)	(0.271)	(0.038)	(0.224)	(0.171)	(0.054)
Hedonic score	0.493***	0.396***	0.056***	0.223***	0.170***	0.054***
	(0.067)	(0.057)	(0.010)	(0.047)	(0.037)	(0.012)

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Akaike and Bayesian Information criteria by error story and model specification

			EUT	RDU				
				Power	TK	LinLog	Prelec	General Prelec
AIC	Fechner (Strong utility)	Probit	2945.30	2916.24	2919.10	2916.54	2931.95	2918.23
		Logit	2931.01	2913.61	2917.50	2914.01	2922.72	2915.59
	Luce (Strict utility)		2956.76	2931.74	-	-	2925.77	2927.50
	Contextual utility	Probit	2865.49	2848.33	2850.89	2843.67	2848.22	2843.93
		Logit	2860.20	2846.89	2861.65	2841.71	2846.68	2841.88
	Decision Field theory	Probit	2845.84	2837.98	2816.14	2817.51	2810.14	2811.93
		Logit	2827.86†	2825.54†	2810.06†	2811.04†	2805.90†	2807.86†
BIC	Fechner (Strong utility)	Probit	2957.12	2933.97	2936.83	2940.17	2949.68	2941.86
		Logit	2942.83	2931.33	2935.23	2937.64	2940.45	2939.23
	Luce (Strict utility)		2968.57	2949.47	-	-	2943.50	2951.18
	Contextual utility	Probit	2877.31	2866.10	2868.6	2867.31	2865.95	2867.57
		Logit	2872.01	2864.62	2879.37	2865.35	2864.40	2865.52
	Decision Field theory	Probit	2857.56	2855.55	2833.76	2840.97	2827.71	2835.36
		Logit	2839.58†	2843.19†	2827.63†	2834.47†	2823.47†	2831.29†

AIC: Akaike information criterion. BIC: Bayesian information criterion. The dagger (†) indicates column-wise for each information criterion the model which minimizes information lost.

Table A.4: Estimates for EUT and RDU given the Decision Field theory stochastic assumption (with additional controls)

	EUT		RDU			
	(1)	(2)	(3)		(4)	
	r	r	r	α	r	α
Constant	0.640	0.657	0.356	-2.806	0.444	-2.693
	(1.018)	(1.047)	(1.065)	(11.621)	(1.155)	(11.629)
Scent	0.132		0.104	-0.065		
	(0.090)		(0.091)	(0.456)		
Scent: Subliminal		0.189*			0.168	0.072
		(0.103)			(0.108)	(0.466)
Scent: Supraliminal		0.060			0.015	-0.485
		(0.120)			(0.133)	(2.037)
H&L task	-0.030	-0.029	-0.302	-1.340	-0.306	-2.162
	(0.059)	(0.059)	(0.521)	(4.088)	(0.421)	(7.052)
Endowment	0.167	0.157	0.230	0.450	0.212	0.641
	(0.206)	(0.213)	(0.190)	(1.001)	(0.201)	(1.416)
Male	0.110	0.105	0.065	-0.053	0.056	-0.126
	(0.093)	(0.091)	(0.116)	(0.284)	(0.104)	(0.536)
Household size	0.003	0.003	0.011	0.024	0.009	0.020
	(0.055)	(0.054)	(0.059)	(0.158)	(0.056)	(0.236)
Education: 3rd semester student	-0.310**	-0.302**	-0.266*	-0.130	-0.251*	-0.140
	(0.126)	(0.128)	(0.153)	(0.240)	(0.430)	(0.574)
Education: 5th semester student	-0.350**	-0.345**	-0.295**	-0.114	-0.284**	-0.109
	(0.146)	(0.148)	(0.144)	(0.226)	(0.298)	(0.475)
Education: 7th semester student	-0.395**	-0.363**	-0.244	0.314	-0.171	0.767
	(0.160)	(0.171)	(0.217)	(0.383)	(1.052)	(2.504)
Education: 9th semester student	-0.366**	-0.347**	-0.099	0.612	-0.070	1.064
	(0.172)	(0.172)	(0.327)	(0.513)	(1.711)	(3.126)
Education: >9th semester student	-0.460**	-0.438**	-0.374*	-0.050	-0.338	0.082
	(0.219)	(0.210)	(0.218)	(0.397)	(0.541)	(1.137)
Income: Above average	-0.063	-0.069	0.062	0.395	0.063	0.670
	(0.136)	(0.133)	(0.154)	(1.096)	(0.146)	(1.898)
Income: Average	0.074	0.070	0.091	0.143	0.089	0.249
	(0.119)	(0.118)	(0.112)	(0.563)	(0.112)	(0.997)
Income: Below average	0.056	0.042	0.279**	0.746	0.272**	1.192

	(0.129)	(0.128)	(0.132)	(1.959)	(0.128)	(3.129)
Income: Bad or very bad	0.072	0.065	0.084	0.063	0.070	-0.000
	(0.197)	(0.194)	(0.215)	(0.826)	(0.204)	(1.130)
Body Mass Index	-0.028*	-0.026*	-0.033*	-0.053	-0.033*	-0.093
	(0.016)	(0.015)	(0.018)	(0.160)	(0.019)	(0.305)
Smoking: No	-0.024	-0.030	0.003	0.080	-0.005	-0.181
	(0.104)	(0.105)	(0.122)	(0.658)	(0.127)	(1.116)
Smoking: Occasionally	0.019	0.025	0.047	0.152	0.048	0.246
	(0.168)	(0.183)	(0.184)	(0.807)	(0.206)	(1.209)
μ	2.152***	2.154***	2.841***		2.854***	
	(0.152)	(0.151)	(0.878)		(0.719)	
Observations	2584	2584	2584		2584	
Log-likelihood	-1388.589	-1387.081	-1363.766		-1362.273	

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base categories are: No Scent, Education: 1st semester student, Income: Good or Very good, Smoking: Yes.

B Appendix: Pictures and other experimental stimuli

(a) Biscuits

(b) USB stick

Figure B.1: Picture stimuli used in the training auction rounds

(a) Chocolate with university logo

(b) Mug with university logo

Figure B.2: Picture stimuli used in the real auction rounds

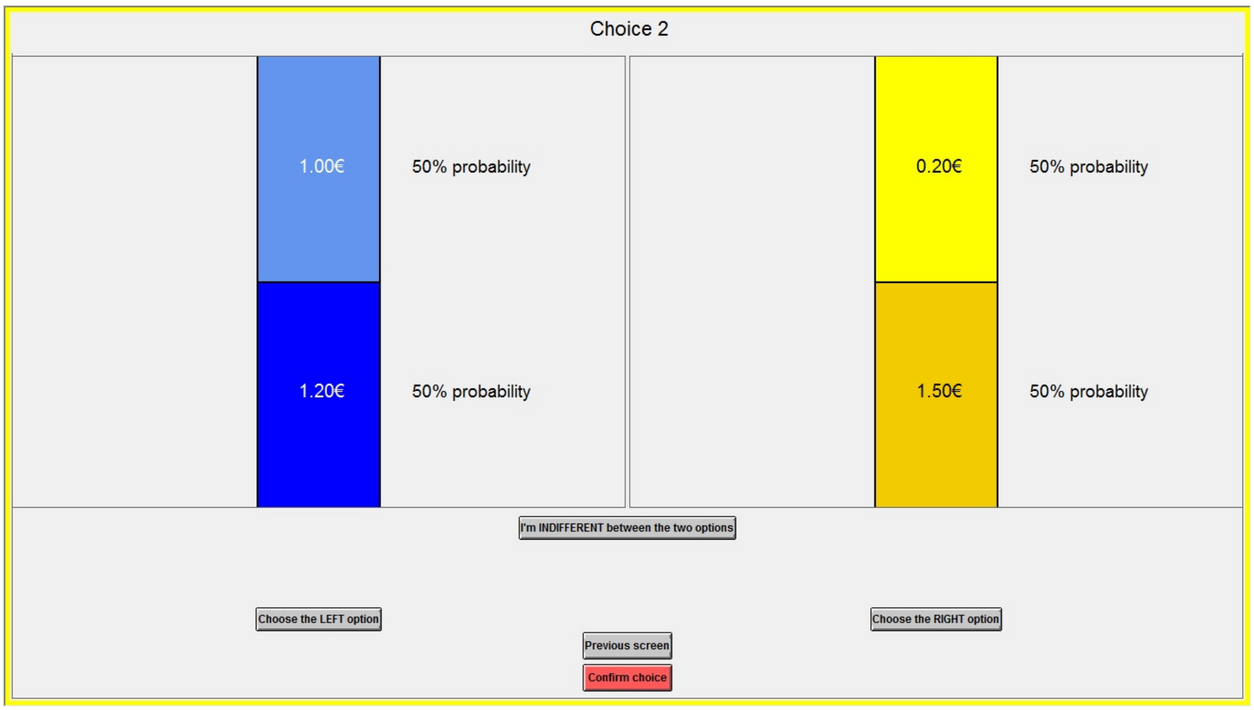


Figure B.3: Risk preference choice task example screen

(a) Airoma® XTREME 'Florida Zest' by Vectair Systems

(b) Scent dispenser

Figure B.4: Olfactory stimuli and scent dispenser

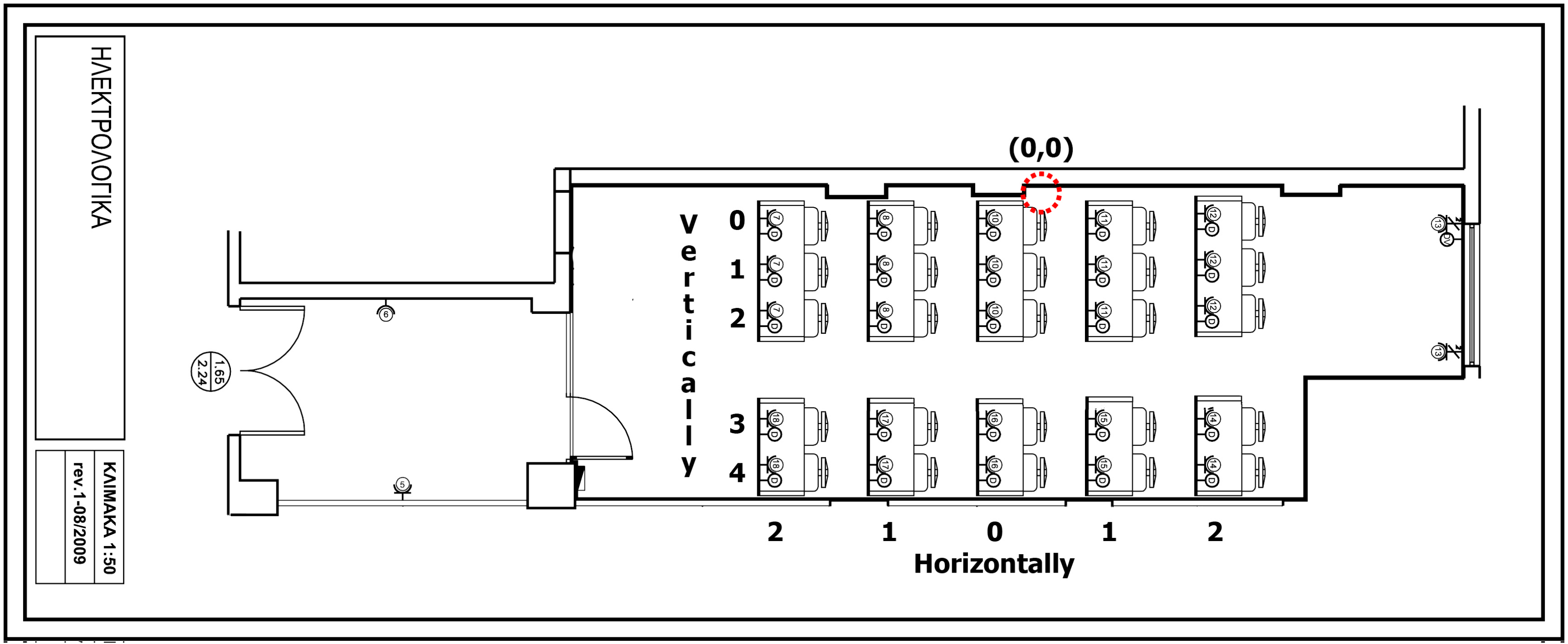


Figure B.5: Laboratory room plan with scent dispenser position marked in red circle (position: 0,0) and horizontal/vertical distance of computer cubicles from dispenser position

C Appendix: Experimental Instructions

Instructions

[This is a translation of the original instructions written in Greek]

Welcome to our survey. First of all, I would like to thank you for your presence here, today. Before we begin I would like to ask you to turn off or mute you mobile phones so that the survey process is not interrupted. Furthermore, you should not communicate with other participants until the survey is over, in order to prevent your responses being influenced by others. Any attempt to communicate with other participants will result in the failure of the survey. During this survey you are going to interact with other participants but you will not know with who you are interacting with, so that anonymity is preserved.

If you have any questions during the survey, you can raise your hand **and address the researcher in private, not in public**. We will answer all questions except questions concerning the way you should behave during the survey. The reason is that no one, including us, can tell you how you should behave since if we knew that, we wouldn't need to conduct this survey today.

You should know that there are no "right" or "wrong" decisions, nonetheless, choices made by you and other participants will affect your final earnings, therefore we advise you to pay attention to the instructions you will receive from now on. Every participant will receive 4€ as a show-up fee and an additional amount of money during the survey; this is money that you will receive at the end of the session.

Initially I will ask each one of you to pick a three-digit number from a jar. This number is strictly personal and unique to each one of you and we will use it for your payment.

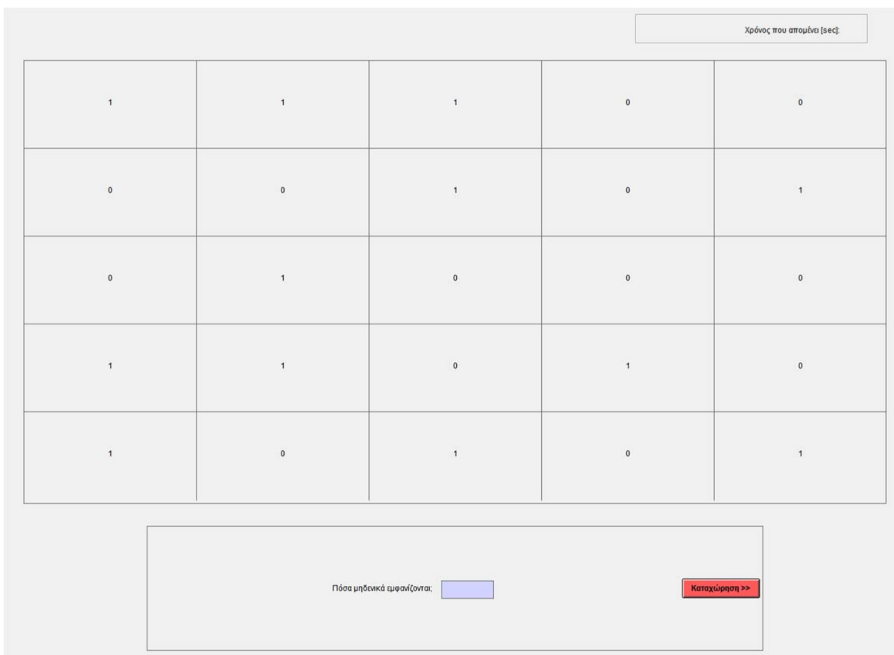
The survey

This survey consists of three different stages. In stages one and three you will be given the chance to increase your income through short tasks which will be assigned to you, while in the second stage you will take part in a series of auctions. After the end of the third stage you will be asked to answer a short questionnaire and after filling it out you will receive your payment and will be free to leave the lab. More detailed instructions for the three stages follow.

~1~

Stage 1

In this stage you'll have to complete a task in which you'll determine the right number of zeros that will appear in a matrix with dimensions 5x5, like the one shown below. The table will contain the numbers 0 and 1 and you'll have 25 seconds to count the exact number **of zeros** that you'll see on your screen and type the number in the box below the matrix. This process **will be repeated for 10 times** and each time the matrix will change. For every **correct answer** you give, you **earn 0.50€ on top** of your final income. If you give a **wrong answer** or if you **do not respond within the time limit**, you do not win or lose money. Therefore, if you answer correctly all the times, you can earn up to **5€, on top** of your participation fees.



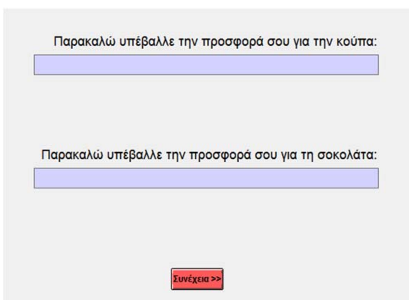
Stage 2

In this stage you will be asked to take part in a type of auction known as the 2nd price auction. In an auction procedure, several participants submit a bid for a good they wish to purchase. The person that provides the highest bid, buys the item. The difference is that in a 2nd price auction the person submitting the highest bid **purchases the product at the second highest bid** submitted by another participant.

Something similar to the above, will take place in Stage 2. The computer will randomly split you in groups of 5 people but you will not know who will be the members of your group. Afterwards, we will provide you with information and we'll show you one chocolate and one mug for which you are going to submit separate bids to purchase them and depending on whether your bid is the highest in your group, you will eventually buy or not one of the two products. The process includes the following steps:

1. Description and presentation of products
2. Bidding for each product separately by all participants
3. Computer ranks bids from highest to lowest (separately for each group of 5 persons)
4. Provide feedback to participants of whether they were the highest bidders or not, for each product.

ATTENTION!! If you bid the highest bid and you are the highest bidder you will actually have to BUY the product and the (second largest) bid will be subtracted from your final payment. For this reason you should be honest and during the auction offer the **highest** amount of money that each product is worth to you. **NOT** the price you believe that it costs **but** the **maximum** price at which you would be willing to buy it if it was available in a store.



If you are not willing to pay any amount in order to purchase any of these products, you should offer a zero bid. The screen you'll see when you submit your bids, looks like the one shown to the left.

~3 ~

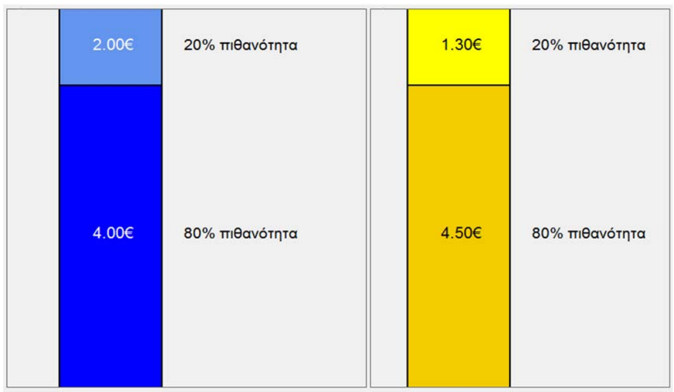
Before the auction, **five practice rounds** will be performed for some products that are not available in the lab and that you are not going to buy. These rounds are made solely for familiarizing yourself with the process and will not contribute to your final earnings, that is, it is not possible to gain or lose money.

In the actual auction, **ten auction rounds** will take place for both products simultaneously.

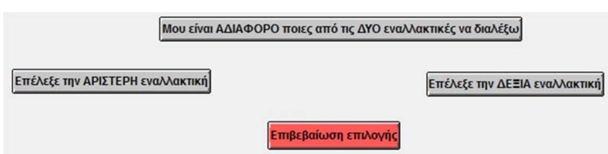
Depending on yours and other participants' bids, you could be or not the highest bidder of your group, in one or more rounds, for one or both products. For this reason, you'll receive feedback from the computer. At the end of the session, the computer will randomly choose one of the two products (mug or chocolate) and one of the ten auction rounds as binding. This means that today **one and only product will be given to the highest bidder** of each group.

Stage 3

In this final stage, 20 different screens will be shown. Each screen will give you the opportunity to choose between two options. These options are lotteries and will be shown in a screen similar to the image below:



Each column represents a lottery where two different monetary amounts are displayed and right next to each amount, the respective probability to win the amount is displayed. You will have to choose one of the two alternatives shown in the left and in the right of your screen by clicking on the corresponding button. You will also have the option to indicate indifference between the two options, as shown below. If you state indifference for the lotteries, that is, if you state that both lotteries are equally likeable, then the computer will randomly select one of the lotteries for you.



The screen will be shown 20 times with different monetary amounts and different probabilities assigned to each amount and **you will choose between the two different lotteries 20 times.**

At the end of Stage 3, each one of you will carry out a random draw in which you will randomly pick a number from 1 to 20 and one number from 1 to 100. These numbers will determine the binding round in the third stage (one of the 20 screens) and the probability which will determine the amount paid by the lottery you selected in the round that was drawn. The additional amount of money from this stage will be added to your income and the result will appear on your screen. You can win only one of the two amounts that appear on each lottery and only in one of the 20 lottery choices; in the one that will be randomly drawn. For this reason you should be very careful and make every choice as if it is the one that will be randomly drawn.