Chem. Senses 34: 127–138, 2009
doi:10.1093/chemse/bjn067
Advance Access publication November 12, 2008

Methods of Human Body Odor Sampling: The Effect of Freezing

Pavlina Lenochova[1], S. Craig Roberts[2] and Jan Havlicek[1]

[1]Department of Anthropology, Faculty of Humanities, Charles University, Husnikova 2075, 158 00 Prague 13, Czech Republic and [2]School of Biological Sciences, University of Liverpool, Liverpool, L69 7ZB, UK

Correspondence to be sent to: Jan Havlicek, Department of Anthropology, Faculty of Humanities, Charles University, Husnikova 2075, 158 00 Prague 13, Czech Republic. e-mail: jan.havlicek@fhs.cuni.cz

Abstract

Body odor sampling is an essential tool in human chemical ecology research. However, methodologies of individual studies vary widely in terms of sampling material, length of sampling, and sample processing. Although these differences might have a critical impact on results obtained, almost no studies test validity of current methods. Here, we focused on the effect of freezing samples between collection and use in experiments involving body odor perception. In 2 experiments, we tested whether axillary odors were perceived differently by raters when presented fresh or having been frozen and whether several freeze–thaw cycles affected sample quality. In the first experiment, samples were frozen for 2 weeks, 1 month, or 4 months. We found no differences in ratings of pleasantness, attractiveness, or masculinity between fresh and frozen samples. Similarly, almost no differences between repeatedly thawed and fresh samples were found. We found some variations in intensity; however, this was unrelated to length of storage. The second experiment tested differences between fresh samples and those frozen for 6 months. Again no differences in subjective ratings were observed. These results suggest that freezing has no significant effect on perceived odor hedonicity and that samples can be reliably used after storage for relatively long periods.

Key words: armpit, human, mate choice, MHC, olfaction, smell

Introduction

Human olfaction as a subject of scientific interest had been almost entirely neglected until the late 1970s (Doty 1977). Since then, numerous studies have been published, considerably advancing our understanding of odor-based discrimination of males versus females, kin versus nonkin, and between individuals. Body odor also influences mother–infant interactions and human mate preferences (for reviews, see Schaal and Porter 1991; Lenochova and Havlicek 2008). Yet, experimental research methods in these studies have been highly variable. It is commonly acknowledged that methods used in experimental research may have significant impact on results obtained. Interestingly, there is a lack of studies focusing on the reliability and validity of the methods used in research in human chemical ecology. This also impairs possible comparability between individual studies.

The process of collecting body odor samples commonly includes the following considerations: 1) restrictions on body odor donors, 2) the medium on which odors are collected, 3) length of sampling, and 4) sample storage. Each of these might influence results and therefore should be carefully considered while preparing a study design.

The first step in body odor sampling usually consists of restrictions in activities and in hygienic and dietary habits of the odor donors. In this way, researchers try to minimize the influence of factors that are not of direct interest (i.e., "noise") or might interact with the phenomenon under study. Behavioral restrictions might include exaggerated physical exercise, sexual intercourse, and sleeping with another person or a pet animal in the same bed. Hygienic restrictions mostly proscribe the use of perfumes, deodorants, antiperspirants, and perfumed soaps or shower gels for 1 or 2 days prior to the sampling procedure. Conversely, it is known that some grooming or hygienic habits, such as the shaving of armpit hair, may influence the quality of the armpit odor (Kohoutova et al., forthcoming; Shelley et al. 1953), and yet, most studies have not taken this factor into account.

The majority of studies require at least some restrictions in diet, consumption of alcoholic beverages, and smoking (e.g., Havlicek et al. 2006). However, a list of the foods which should be avoided or the hygienic practices which should be followed varies considerably between studies, and in some studies, the restrictions are limited only to deodorant

application (e.g., Cernoch and Porter 1985). It should be noted that in most cases, the restrictions are based only on anecdotal observations and their actual impact is rather poorly understood. It is, for instance, commonly thought that diet has a major environmental influence on body odor (for recent reviews, see Havlicek and Lenochova 2008; Havlicek and Saxton, forthcoming). However, as far as we know, only the effect of meat consumption on body odor has been experimentally tested (Havlicek and Lenochova 2006).

Researchers mostly agree that, at least in adults, the odor with the highest social impact comes from the armpit (axilla) (Comfort 1971). Therefore, most studies sample odors from this region. In the main, 2 different media are used as substrates into which body odor may become impregnated: T-shirts or cotton pads. The T-shirts are usually worn next to the skin, then removed, and their odor is subsequently rated (e.g., Hold and Schleidt 1977). Although it is presumed that the odor collected is mostly of axillary origin, critiques of the T-shirt method point to the fact that the source of the body odor cannot be specified (i.e., it can include odor from other parts of the torso) (Havlicek et al. 2006). More importantly, the T-shirt method may not sufficiently avoid odor contamination from the environment or from other clothes, which might, for instance, be washed in perfumed washing powder. For these reasons, many researchers prefer cotton pads as sampling media (e.g., Havlicek et al. 2005, 2006; Roberts et al. 2005). The pads are commonly fixed closely to the axilla with surgical adhesive tape or to the clothing adjacent to the axillary region. In this way, the pads are worn in permanent contact with axillary skin.

The length of sampling is a third important variable in human chemical ecology studies. There are huge disparities in sample length across studies, ranging between 30 min (Platek et al. 2001), 1 night (Roberts et al. 2005), 24 h (Havlicek and Lenochova 2006), 2 or 3 consecutive nights (Porter et al. 1986; Wedekind et al. 2007), 4 nights and 1 day (Mallet and Schaal 1998), and 7 nights (Schleidt et al. 1981). These differences in sampling length might not only lead to differences in the strength of the odor, something which is usually negatively related to its pleasantness (Doty et al. 1978, 1982; Havlicek et al. 2005), but also to qualitative variability due to activity of skin bacteria, particularly in the most lengthy sampling procedures (Rennie et al. 1991). On the other hand, a relatively short sampling period may impinge upon the ability of raters to perceive the sample.

Finally, human body odor studies vary in the treatment of samples postcollection and preceding rater presentation. Due to the bacterial activity noted above, odor samples, unlike other stimuli such as photographic images, may change in quality. This fact is acknowledged as a major issue by many researchers. Thus, some studies have used freshly collected samples for each rating session (e.g., Porter et al. 1983; Wedekind et al. 1995; Wedekind and Füri 1997; Havlicek et al. 2005). This approach is not only a highly demanding task for logistical reasons and time coordination but it also restricts the number of raters that may be used and prevents the possibility of fair comparison between samples collected from the same individual at different times. Therefore, it seems to be much easier and simpler to freeze the samples before experimental presentation and many studies employ this approach (e.g., Roberts et al. 2008; Rikowski and Grammer 1999; Singh and Bronstad 2001). Typically, frozen samples are removed from the freezer a few hours before rating begins to allow samples to thaw but otherwise are treated like fresh samples. However, this method is not unquestionably correct: It is assumed that the quality of the odor samples remains stable while in the freezer, but to our knowledge, this assumption has never been systematically tested. Thus, it is a matter of debate whether freezing completely prevents further bacterial action or loss of volatile odorous compounds, each of which might cause changes in the chemical composition of the collected sweat and which might be perceived by human subjects. A shift in odor quality or intensity might not only increase noise in the data collected but could potentially lead to systematic bias in results, especially if time of storage is not balanced for individual samples. Furthermore, some studies (e.g., Wedekind et al. 2007) use odor samples repeatedly over more than 1 day, with additional freeze–thaw cycles, potentially with further unknown effects. We clearly need methodological studies to establish the validity of our research tools. In this study, we carried out 2 experiments that focused on the effect of freezing on the subjective perception of odor samples. In the first experiment, we tested effects of storage duration and freeze–thaw cycles on the quality of body odor samples frozen for a period of 2, 4, or 16 weeks. In the second experiment, we focused on longer term effects (across a 6-month period).

Experiment 1

Material and methods

Raters

A group of 28 female students (mean age 23.6, range 20–28 years), using hormonal contraception, rated the odorous samples in 4 sessions over a 4-month period. We chose hormonal contraceptive users to avoid the potential effect of fluctuations in olfactory function during the natural menstrual cycle (Doty et al. 1981; Hummel et al. 1991; Caruso et al. 2001; Navarrete-Palacios et al. 2003). The raters were not paid for their participation, but they were given a perfume tester after each rating session.

Of the 4 experimental sessions, 3, 6, and 7 raters did not attend Sessions 2, 3, and 4, respectively. Therefore, the number of raters used for analysis of individual sessions, as well as for comparison of all 4 sessions, varies. The number of raters used for each analysis is shown in Table 1. After preliminary inspection of the data, we excluded ratings labeled "I cannot smell the sample." The amount of eliminated

Table 1 Experiment 1: Overview on number of raters, samples, total, excluded, and analyzed ratings in analysis aimed to 1) compare fresh samples, 2) test the effect of storing, and 3) the effect of repeated thawing on axillary samples and essences

	Comparison of fresh samples			Effect of storing			Effect of repeated thawing			
	Axillary samples	Essences	Essences	Axillary samples	Axillary samples	Axillary samples	Axillary samples	Essences	Axillary samples	Axillary samples
Ratings from sessions	I (F), IV (F)	I (F), IV (F)	I (F), II (F), III (F), IV (F)	I (F), II (S2), III (S4), IV (S16)	I (F), II (S2)	I (F), III (S4)	I (F), IV (S16)	I (F), II, (T2) III (T4)	II (S2 + T2)	III (S4 + T4)
Raters (N)	21	21	13	14	25	22	21	21	25	22
Samples (N)	8	5	5	8	9	9	8	9	9	9
Total ratings	168 pairs	105 pairs	65 fours	112 fours	225 pairs	198 pairs	168	189 trios	225 pairs	198 pairs
Excluded ratings	22 pairs	5 pairs	1 four	15 fours	35 pairs	28 pairs	14 pairs	33 trios	34 pairs	13 pairs
Analyzed ratings	146 pairs	100 pairs	64 fours	97 fours	190 pairs	170 pairs	154 pairs	156 trios	191 pairs	185 pairs

F, fresh; S, stored; and T, thawed samples. S and T samples are followed by the number of weeks they were treated.

ratings varies among all parts of the analysis according to the number of raters participating (see Table 1).

Axillary samples

Nine healthy male students of Charles University in Prague, Czech Republic, donated odors. The participants were recruited via posters, e-mail advertisements, or personally by the first author. Their mean age was 26.7 years (range 23–31 years), body weight 75.4 kg (minimum 62 kg, maximum 90 kg), and body height 184.1 cm (minimum 175 cm, maximum 197 cm). None smoked, shaved his armpits, or had any serious diseases (2 of them reported small regions of eczema but not in the axillae). As compensation for their time and potential inconvenience, the donors were given CZK 400 (ca., USD 25). One of the 9 men did not take part in the second donation session and was not included in Session 4. Thus, all analyses including Session 4 are based on 8 donors.

Samples of essences

To control for potential changes in raters' hedonic preferences, we used 5 essences: rose (76.3% phenylethylethanol, 11.7% citronellol, 4.1% geraniol, 1.7% geranylacetate, 1.3% citronelylacetate, and several other minor compounds), bergamot (46% limonen, 21% linalool, 17.5% linalylacetate, 10.7% alpha-terpinylacetate, 3.3% gamma-terpinylacetate, 1.0% beta-terpinylacetate, 0.25% beta-myrcene, and 0.25% alpha-terpineol), cinnamon (75% cinnamal, 14% eugenol, and several other minor compounds), geranium (13.5% geraniol, 12.1% citronellol, 5.5% 3,5,5-trimethylhexanol, 5.1% nerol, 5% dipropylene glycol, 3.5% diphenyloxide, 2.8% geranylacetate, 2.4% isoeugenol, 2.2% limonene, 2.2% terpineol, 2.2% citronellylacetate, 1.9% camphene, 1.6% borneol, and several other minor compounds), and castoreum (a musky smell, being a naturally complex compound from the beaver's anal gland, and consisting of a mixture of 33% dipropylene glycol, 16% thujopsene, 10% gurjunene, 9% benzyl benzoate, 8% cederene, and several other minor compounds), all manufactured by AROMA Praha Corp., Zidovice, Czech Republic (http://www.aroma.cz). Two drops of the 100% essence (i.e., ca., 0.1 ml) were applied onto the cotton pad and treated like the human samples.

Schedule and design

The experiment consisted of 4 rating sessions. Session 1 took place in May 2005, followed by subsequent sessions 2 weeks (Session 2), 4 weeks (Session 3), and 16 weeks (Session 4) later. Experimental procedure was based on a within-subject design, that is, the axillary samples used during the whole experiment in a 4-month period were collected from one group of men and were rated each time by the same group of women.

From each donor, we received 4 samples (representing cotton pads from each axilla cut into halves; details below) on the day of Session 1 and 4 more on the day of Session 4 (only one used in this study). Three of the 4 axillary samples from

each donor were frozen immediately upon collection. In addition, we used samples of essences in each session mainly as control for potential changes in preferences across sessions. Thus, during Session 1, raters assessed fresh axillary samples and essences, which were consequently frozen at the end of the session. In Session 2, raters assessed both the axillary samples and essences which they had rated during Session 1, another set of axillary samples collected but not previously rated (i.e., samples stored for a 2-week period), and fresh essences. At the end of the session, the sets already used in Session 1 were returned to the freezer. In Session 3, a twice-thawed set of axillary samples and essences (i.e., rated at Sessions 1 and 2) was assessed again together with nonused samples stored for 1 month and fresh essences. In Session 4, fresh axillary samples collected from the same subjects (i.e., from a second odor donation round), fresh essences, and nonused axillary samples stored for 4 months were rated. The samples repeatedly rated in Sessions 1, 2, and 3 were omitted in Session 4 to avoid odor adaptation effects due to the large number of samples. For the same reason, we also did not use stored essences in Session 2–4.

In summary, a variable number of samples were tested in the course of our study: 9 fresh axillary samples and 5 essences in Session 1; 9 stored and 9 thawed axillary samples and 5 thawed and 5 fresh essences in Sessions 2 and 3; and finally, 8 fresh and 8 stored axillary samples and 5 fresh essences in Session 4.

Odor sampling procedure

Several days before odor sampling (i.e., before Sessions 1 and 4), each participant received instructions and restrictions in written format, along with a sampling set containing a white cotton T-shirt (previously twice washed without washing powder), a block of nonperfumed soap, cotton pads, a surgical tape, and 2 ziplock plastic bags. The donors were required to undergo certain dietary and behavioral restrictions on the day prior to and the day of the sampling. In particular, they were instructed to refrain from 1) using perfumes, deodorants, antiperspirants, aftershaves, perfumed soaps, and shower gels; 2) eating meals containing garlic, onion, chilli, pepper, vinegar, blue cheese, cabbage, radish, fermented milk products, and marinated fish; 3) drinking alcoholic beverages or using other drugs; and 4) smoking. Additionally, they were asked to avoid strenuous physical activities, sexual intercourse, or sleeping in the same bed as their partner on both nights. The night before sampling, the donors were instructed to use a nonperfumed soap and to wear a white cotton T-shirt until the end of the sampling session. The following morning at 7 AM, they taped the cotton pads (100% cotton, elliptical in shape, approximately 7 cm at their longest axis; Ebelin cosmetic pads, DM-drogeriemarkt, Ceske Budejovice, Czech Republic, http://www.dm-drogeriemarkt.cz) with the plaster (3-M Micropore surgical tape) to their armpits and wore them for the following 24 h.

After the 24-h sampling, they put the pads into the ziplock plastic bags (each labeled to distinguish samples from the left and right axillae) and brought them back to our laboratory. Using rubber gloves and tweezers, each sample was carefully cut in the middle of the visible sweat mark into 2 similar halves. All samples were labeled and 3 of 4 pieces from each donor were placed into the freezer (freezer brand Candy, temperature –32 °C), whereas one, randomly chosen, was used for the first rating session. We chose the given temperature as this is the maximum freezing temperature obtained in non-deep freezers which are used in most studies. Samples used for each session were randomized between donors according to the side of the axillae. The rating of fresh samples started within an hour of sample collection. The whole sampling process was repeated 4 months after the first sampling.

Donors' conformity with the instructions was checked by questionnaire. No serious violations, particularly on the day of sampling, were found (in May 2005, the day prior to sampling, 2 of the donors reported having a small amount of alcohol (0.5 L beer), 2 others used a deodorant on the morning of the day before sampling (in October 2005, no violation was found).

Odor rating procedure

The ratings were conducted in a quiet, ventilated room between 9 AM and 6 PM. Raters were asked to attend all sessions at approximately the same time to avoid possible temporal changes of rated odors and/or diurnal fluctuation in olfactory abilities. Indoor temperature was 20.5–22 °C (51–58% humidity); 18.5–20.5 °C (64–68% humidity); 22–23 °C (60–65% humidity), and 21–21.5 °C (58–65% humidity) during Sessions 1, 2, 3, and 4, respectively. Frozen samples were removed from the freezer approximately 1.5 h before testing onset. Women rated male axillary samples and samples of essences as a control in each session. All fresh, stored, and thawed samples were enclosed in 500-ml opaque laboratory jars with a glass top and labeled with alphanumeric codes. The codes were changed for each session.

For each session, the samples were randomly divided into 2 sets and assessed by the female raters on 7-point scales for intensity, pleasantness, attractiveness, and masculinity. In the event that raters found any of the samples too weak to assess, they were asked to select "I cannot smell the sample" instead of using the rating scales. The order of the sets and the order of the stimuli within a set were randomized for each rater. Composition of each set (i.e., number of fresh, stored, and thawed samples) was randomly distributed. To avoid odor adaptation, the women were asked to take a break between the 2 rating sessions and were offered some tea, coffee, or mineral water. To check for possible effects of the break on odor ratings, we compared ratings made before and after the break for each session; we did not find significant differences in any of the rated variables. During the break, the raters also filled in additional questionnaires.

Statistical analysis

The within-subject design of our study allowed us to use paired *t*-tests to compare 2 individual sessions and repeated measures analysis of variance (ANOVA) to compare 3 or 4 sessions. We used post hoc tests (Tukey's Honestly Significant Difference test) only if a significant main effect was detected by the ANOVA. Each rater assessed all samples, and therefore these values cannot be judged as independent. To avoid pseudoreplication (Hurlbert 1984), we used mean values as the unit of analysis instead of individual ratings. We performed both types of analysis with either rater or sample (axillary odor or essence) as the unit of analysis. In case of missing values in one of the session, ratings from other sessions were excluded from the analysis to assure equal sampling. For exploratory purposes, we analyzed our data also with individual ratings; however, the results are virtually identical and for above-mentioned reasons are not presented further. All analyses were carried out with the statistical package STATISTICA 7.1.

Results

Comparison of fresh samples

First, we compared the fresh samples collected for Sessions 1 and 4. The analysis is based on 146 ratings of the original 168 rating pairs (21 raters × 8 donors) (Table 1). The rest of the ratings were excluded as at least one sample of the pair was judged to be too weak to be detected by some raters. When a rater was used as the unit of analysis, fresh samples in Session 4 were assessed as more pleasant ($t_{20} = 2.84$; $P = 0.010$) and more attractive ($t_{20} = 3.17$; $P = 0.005$) compared with the fresh body odor samples obtained in Session 1 (Figure 1). Similarly, samples in Session 4 were judged as more attractive ($t_7 = 2.5$; $P = 0.041$) with the donor as the unit of analysis.

In similar fashion, we compared fresh essences in Sessions 1 and 4. The analysis is based on 100 ratings of the original 105 rating pairs (21 raters × 5 essences). We found no significant differences between any rated variables either with the rater (Figure 1) or sample as the unit of analysis.

To test possible changes in female hedonic ratings and perceived intensity across the 4 sessions, we analyzed ratings of essences (the same fresh essences in all sessions). Altogether 64 of 65 hedonic rating tetrads from 13 female raters in Sessions 1–4 were compared by repeated measures ANOVA. With the rater as the unit of analysis, there were no significant differences (Figure 2). However, when using a sample (i.e., essence) as the unit of analysis, we found significant differences in intensity ratings ($F_{3,12} = 6.3$; $P = 0.008$). Essences in Session 1 were rated as stronger than in Session 2 ($P = 0.02$) or Session 4 ($P = 0.01$). No differences in ratings of pleasantness, attractiveness, and masculinity were observed.

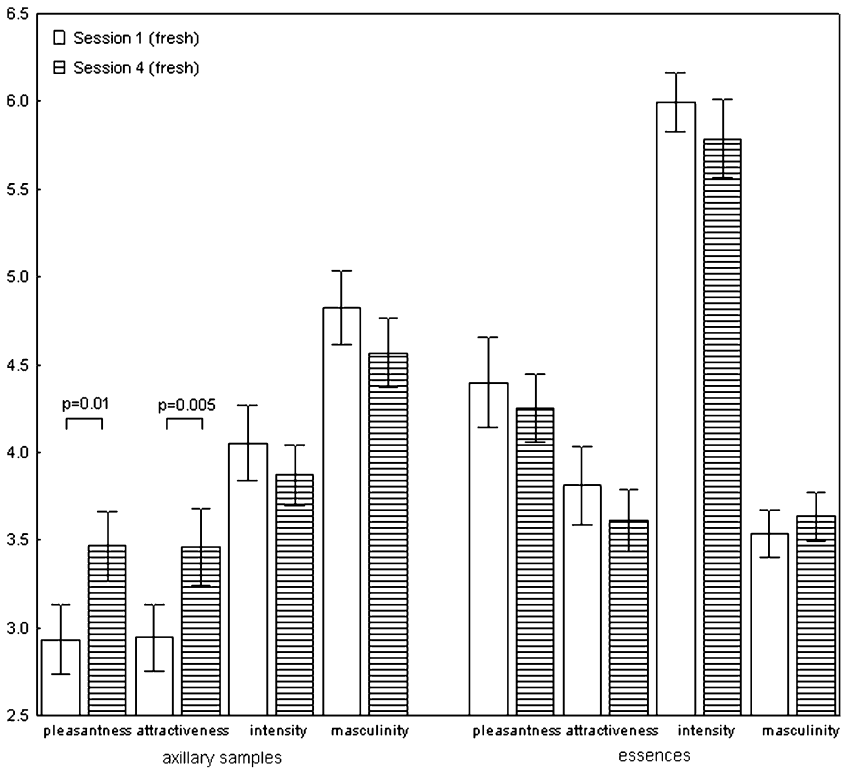

Figure 1 Mean ratings (±standard error) of pleasantness, attractiveness, intensity, and masculinity of fresh axillary samples and essences in Sessions 1 and 4 in Experiment 1 based on 146 (axillary samples) or 100 (essences) individual ratings given by 21 raters.

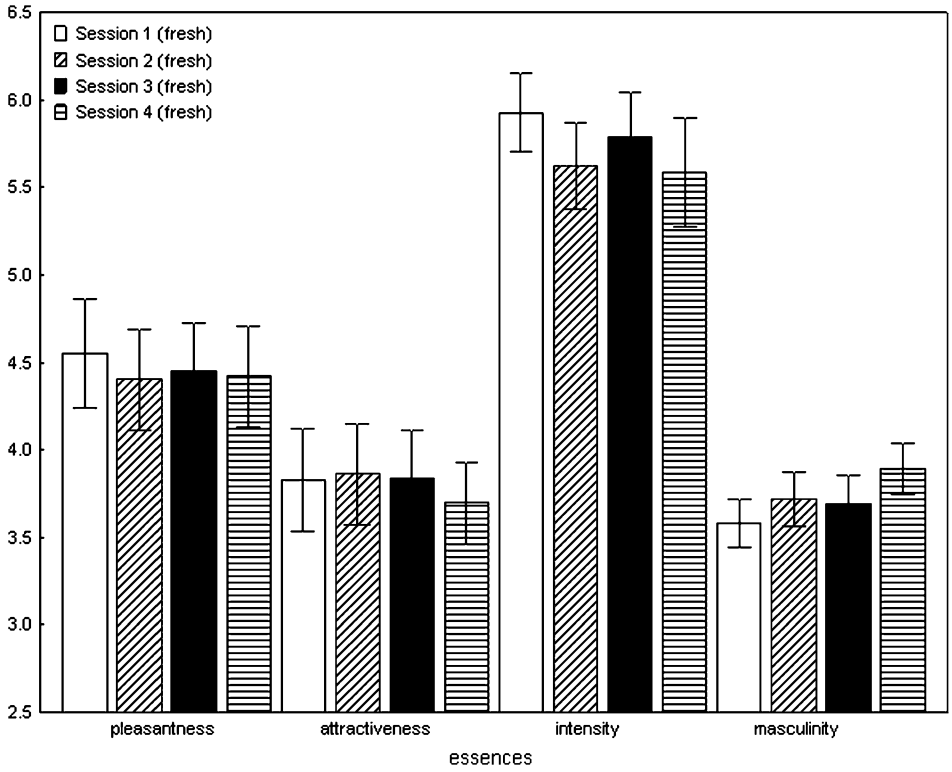

Figure 2 Mean ratings (±standard error) of pleasantness, attractiveness, intensity, and masculinity of fresh essences in Sessions 1–4 in Experiment 1 based on 64 individual ratings given by 13 raters.

Effect of storing

Second, we investigated possible changes in the quality of stored samples. The fresh samples and samples stored for 14 days, 1 month, and 4 months were compared by repeated measures ANOVA. Only the ratings from 14 female raters, available for all 4 sessions, were used. This resulted in a relatively high reduction in the data leaving 97 observations for the analysis. We found no significant changes in the ratings of pleasantness, attractiveness, and masculinity across all sessions. There was, however, a significant effect in the ratings of intensity ($F_{3,39} = 3.0$; $P = 0.041$). Post hoc analysis showed that this was due to nonsignificantly higher intensity ratings in Session 3 compared with Sessions 1 and 4 (in both, $P = 0.06$) (Figure 3). No differences were found when a sample was used as the unit of analysis.

As the analysis across all sessions was based on a relatively small sample, we further analyzed the relation between the fresh samples and the samples stored for 14 days, 1 month, or 4 months using paired *t*-tests. We found no differences between fresh samples and those stored for 14 days or 4 months. The samples stored for 1 month (Session 3) were rated as more intense than the fresh samples ($t_{21} = 3.4$; $P = 0.003$), but there was no difference in ratings of pleasantness, attractiveness, or masculinity. No differences were found when a sample was used as the unit of analysis.

Effect of repeated thawing

Third, we analyzed the potential influence of repeated thawing and storing on the quality of odor samples. In other words, the samples used as fresh during Session 1 were subsequently frozen and used repeatedly during Sessions 2 and 3. The analysis is based on the ratings of 21 female raters and 9 donors (i.e., 189 observations). This resulted in 156 trios of individual ratings after excluding 33 rating trios, which were labeled as too weak. We found no significant changes in pleasantness, attractiveness, and masculinity. There was a significant repeated measures effect of intensity ($F_{2,40} = 4.7$; $P = 0.015$). Post hoc analysis showed that repeatedly thawed samples in Session 3 were rated as significantly more intense compared with the same thawed samples in Session 2 ($P = 0.011$) (Figure 4). No significant differences were found when a sample was used as the unit of analysis. However, similar to the previous type of analysis, there was a nonsignificant trend in intensity ratings ($F_{2,16} = 3.3$; $P = 0.062$) due to higher intensity in Session 3 compared with Session 2 ($P = 0.052$).

Similarly, we analyzed ratings of essences which were treated (i.e., stored and thawed) in the same way as the human samples. The analysis is again based on ratings of 21 female raters of 5 essences; that is, 94 individual ratings after excluding 11 ratings which were labeled as too weak. Mean ratings of all variables on fresh samples and the same samples

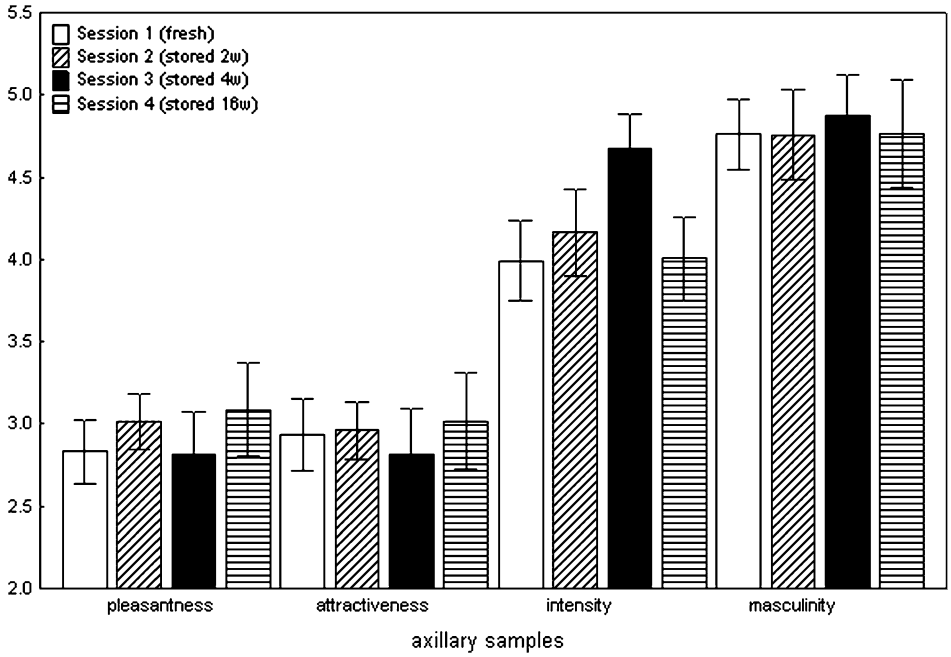

Figure 3 Mean ratings (±standard error) of pleasantness, attractiveness, intensity, and masculinity of fresh (Session 1), stored for 2 weeks (Session 2), 4 weeks (Session 3), and 16 weeks (Session 4) axillary odors in Experiment 1 based on 97 individual ratings given by 14 raters. Note that differences in intensity in Session 3 compared with Session 1 and 4 are approaching formal level of significance (in both, $P = 0.06$).

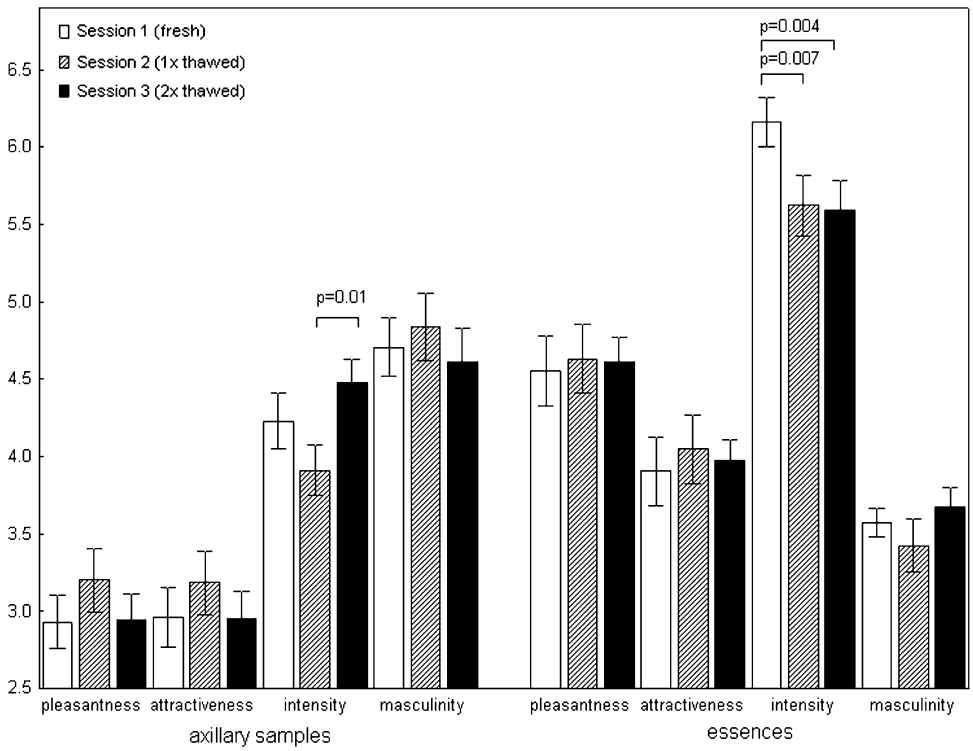

Figure 4 Mean ratings (±standard error) of pleasantness, attractiveness, intensity, and masculinity of fresh (Session 1), once-thawed (Session 2), and twice-thawed (Session 3) axillary odors and essences in Experiment 1 based on 156 (axillary samples) or 94 (essences) individual ratings given by 21 raters.

thawed after 14 days and 1 month were tested using repeated measures ANOVA. We found an effect only for intensity ($F_{2,40} = 7.3$; $P = 0.002$). Post hoc analysis showed that fresh samples (Session 1) were rated significantly more intense compared with repeatedly thawed samples in Session 2 ($P = 0.007$) and Session 3 ($P = 0.004$) (Figure 4). Using sample as the unit of analysis, we similarly found an effect only for intensity ratings ($F_{2,8} = 10.6$; $P = 0.006$), which was caused by higher intensity in Session 1 compared with Sessions 2 and 3 (both $P = 0.010$).

Repeated manipulation with the samples might influence the odor differently than simple storing. To test this, we further compared ratings of stored with repeatedly thawed axillary samples within individual sessions (analyses are based on 191 of 225 and 185 of 198 rating pairs in Sessions 2 and 3, respectively). We found lower intensity ($t_{24} = 3.8$; $P = 0.001$) in repeatedly thawed samples compared with stored ones in Session 2. No differences were found between repeatedly thawed samples (i.e., previously used in Sessions 1 and 2) and stored samples in Session 3. Using sample as the unit of analysis, we similarly found lower intensity ($t_8 = 3.6$; $P = 0.007$) in repeatedly thawed samples compared with stored ones in Session 2. However, no other significant results were observed by this type of the analysis.

Experiment 2

To test the effect of longer storage duration, we designed a new experiment in which the samples were stored without thawing for 6 months. Consequently, their quality was compared with fresh samples from the same donors.

Methods and materials

Raters

As in the previous experiment, a group of 27 women (mean age 24.3 years, range 19–32 years), using hormonal contraception, judged the odor samples. All of them but 2 were students of various universities in Prague. According to the raters' own time schedules, they were assigned either to Sessions A or B. The women were not paid for their participation but they received a perfume tester and a 150-g chocolate bar.

Altogether, we obtained 291 ratings (9 donors × 15 raters in Session A + 13 donors × 12 raters in Session B). Preliminary data inspection revealed that at least one of the samples in the pair was labeled by the rater as "I cannot smell the sample" for 35 pairs of ratings, and consequently, these ratings were excluded from the analysis.

Axillary samples

Thirty young men, mostly Charles University students, participated in this experiment as odor donors. Twenty-five of them completed the whole experimental procedure. Their mean age was 24.1 years (range 20–32 years), body weight 73 kg (minimum 62 kg, maximum 91 kg), and body height 180.6 cm (minimum 170 cm, maximum 200 cm). As in the previous experiment, none smoked, shaved his armpits, or had serious diseases (2 men reported small regions of eczema but not in the axillae; 3 donors used antiallergic pills). In compensation for their time and potential inconvenience, the donors were given CZK 500 (i.e., ca., USD 33).

Schedule and design

The first sampling was carried out in April 2006. Immediately after delivery, these samples were put into the freezer and stored for 6 months. In October 2006, the second sampling of fresh axillary odors took place. In all, 25 of the original 30 men participated at this stage.

To avoid olfactory adaptation and exhaustion effects associated with rating 50 samples (i.e., the whole odor set; 25 stored and 25 fresh samples), we randomly assigned donors into groups and carried out 2 rating sessions (A and B) on 2 adjacent days. In each session, both fresh and stored samples of the particular donor group were assessed. In Session A, samples obtained from 11 men were assessed by 15 women; in Session B, samples from 14 men were assessed by 12 women. For all analyses, we pooled the data obtained in Session A and B.

After preliminary inspection of the data, we found that 3 odor samples from different men were described as smelling of deodorant by more than half of the women. For this reason, we excluded these donors from the analysis and the number of donors decreased to 22 (9 used in Session A and 13 in Session B).

Odor sampling procedure

Sampling was organized in the same way for both sessions. Approximately, 1 week before sampling, each participant received an experimental pack including all instructions and a sampling set. All restrictions in diet, hygienic practices, activity, as well as sampling procedure were identical to the previous experiment.

Immediately after the first sampling, each cotton pad was cut into halves, placed into the labeled ziplock plastic bags, and put into the freezer. As the donors wore pads in both axillae, halving of the pads resulted in 4 samples from each donor. From these 4 samples only 1, randomly assigned, was used for the rating session 6 months later. In the second sample session, after 6 months, again only one-half of 1 pad (taken from the same armpit as the first sample pad) was used for the ratings. Other samples were retained for additional research.

All donors filled in questionnaires concerning conformity with the restrictions during the sampling. We found no serious violations, particularly, during the day of sampling. In April, 3 of the donors reported having a small amount of alcohol (2 dL wine or 0.5 L beer) on the day before sampling and 2 others ate a little onion. In October, 4 men admitted having a small amount of onion or garlic and 3 of them also a glass of wine or beer, another donor reported having 2 glasses of beer, all of them on the day before sampling.

Odor rating procedure

All samples were enclosed in 500-ml opaque jars and labeled with a 1-letter code. Rating sessions took place in a quiet, ventilated room with a relatively constant temperature and humidity (21–21.5 °C, 60–64%). The women rated all samples for their intensity, pleasantness, attractiveness, and masculinity on 7-point scales as in the previous experiment. After assessing half of the samples, raters were asked to have a break, lasting at least 10 min, for refreshment and for the completion of an additional questionnaire. Stimuli and set orders were randomized for each rater. All other details were the same as in Experiment 1.

Statistical analysis

The design of our study was within subjects; that is, the stored and the fresh samples from the same men were assessed by the same raters. The data were analyzed, using paired samples *t*-tests, in 2 different ways: 1) using mean odor donor ratings and 2) using mean odor raters ratings as unit of analysis. We compared the mean ratings of each dependent variable (e.g., pleasantness) obtained from both fresh and stored samples. As in the previous experiment, we do not present analyses on individual ratings as they are subject to pseudoreplication and hence inflation of the degrees of freedom. However, our exploratory analysis using individual ratings as the unit of analysis showed virtually identical results. The statistical package STATISTICA 7.1 was used for all analyses.

Results

First, we analyzed the data using rater as the unit of analysis; there were no significant differences, indicating that the mean ratings for fresh samples, as calculated for each of 27 raters, were not different from the ratings of the samples stored for 6 months (Figure 5). The analysis is based on 256 ratings of the original 291 rating pairs (15 raters × 9 donors in Session A + 12 raters × 13 donors in Session B).

Second, we used individual odor donor as the unit of analysis. We did not find any difference between the rated quality characteristics of the fresh and the stored body odor samples of the 22 men.

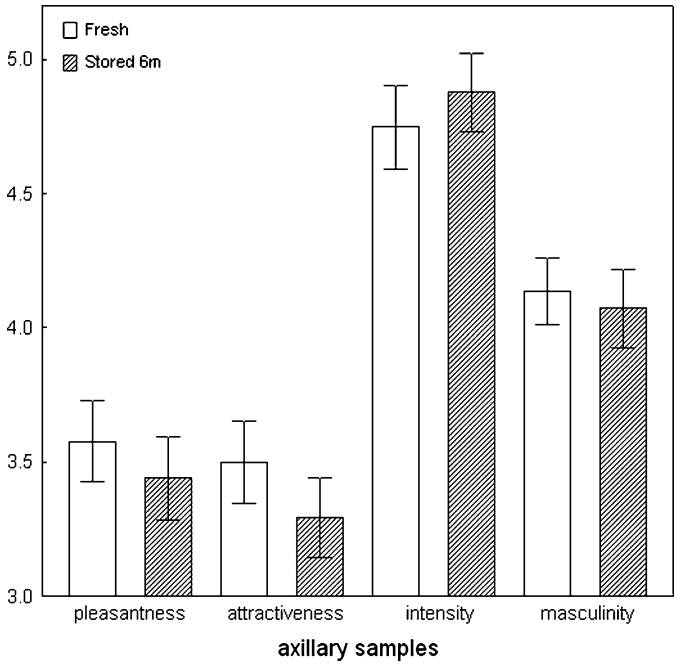

Figure 5 Mean ratings (±standard error) of pleasantness, attractiveness, intensity, and masculinity of fresh (open bars) and stored for 6 months (shaded bars) axillary odors in Experiment 2 based on 256 individual ratings given by 27 raters.

Discussion

In this study, we aimed to test the effect of freezing on subjectively perceived quality of odor samples. In general, we found no significant differences in odor pleasantness, attractiveness, and masculinity between fresh samples and those kept in a freezer for periods between 2 weeks and 4 months. These results were subsequently confirmed by a follow-up experiment where no significant effect of freezing over 6 months was found.

In experiment 1, the analysis across all rating sessions resulted in relatively high drop out rates for participants. We therefore additionally compared samples with different lengths of storage with the fresh samples. Again, we found no significant difference between fresh and stored samples in most rated variables. The only exception in both analyses was significantly higher intensity in the samples stored for 1 month (Session 3) compared with the fresh samples (Session 1). This unexpected finding might be a result of slightly higher indoor temperature during Session 3 (22–23 °C) compared with Session 1 (20.5–22 °C). Environmental temperature, together with humidity, may affect volatility of the odor samples and consequently their perceived intensity. However, we did not find a similar effect in ratings of the essences. It is therefore possible that various chemicals are affected by the environmental temperature to a different degree. This issue should certainly be explored in future studies.

Several previous studies have used frozen body odor samples instead of fresh ones. Its main advantage is that stimulus collection and ratings can be separated, and this makes the whole procedure less logistically demanding. More importantly, freezing allows experimenters to obtain larger rater samples as the rating session is not restricted to the day of sample collection. This is an important issue as most body odor studies suffer from relatively small sample sizes, reducing their reliability and perhaps their validity as well. Sources and media of stored bodily samples have included T-shirts (Fleming et al. 1995, 1997; Rikowski and Grammer 1999; Singh and Bronstad 2001; Jacob et al. 2002; Wedekind et al. 2007; Roberts et al., 2008), cotton pads worn in axillae

(Roberts et al. 2005), and bottled amniotic fluid (Schaal et al. 1995, 1998; Soussignan et al. 1997; Marlier et al. 1998a, 1998b; Schaal and Marlier 1998). Some have stored the samples in a regular freezer (ca., –20 to –32 °C) (Fleming et al. 1995, 1997; Schaal et al. 1995; Soussignan et al. 1997; Schaal and Marlier 1998; Singh and Bronstad 2001), whereas others (Rikowski and Grammer 1999; Jacob et al. 2002; Roberts et al. 2005; Wedekind et al. 2007) have used a deep freezer (–85 °C) instead. In the current study, we did not manipulate freezing temperature. The reason why we chose –32 °C was that it is the lowest possible temperature obtained in a regular freezer. At this point, it is not clear how storage temperature influences perceptive qualities of the human samples. It may affect both further bacterial action via their slower metabolic rates and also various oxidative processes. To our knowledge, the only studies to test various variables during freezing process on odors is in the field of food science. It was, for instance, shown in carrots that refrigeration (1 °C) compared with freezing (–24 °C) affects some but not other chemicals constituting the aroma of this vegetable (Kjeldsen et al. 2003). This makes direct comparison highly difficult as it involves structurally different chemicals. However, a study on fish–meat freezing suggests that freezing at –20 °C (but not at –80 °C) does not prevent further lipid and fatty acid conversions (Baron et al. 2007).

In virtually all studies on newborns, the samples (T-shirts or bottled amniotic fluid) were kept in the freezer for only 1–4 days. On the other hand, the length of storage in most studies on adults is often not specified. The aims of experiments using frozen samples have varied from individual odor recognition (infant by mother) (Fleming et al. 1995, 1997), matching odors of twins (Roberts et al. 2005) and hedonic ratings related to newborn's own amniotic fluid (Schaal et al. 1995; Soussignan et al. 1997; Schaal and Marlier 1998), the genetic underpinning of mate preferences (Jacob et al. 2002; Roberts et al. 2008), facial symmetry (Rikowski and Grammer 1999), and ovulation (Singh and Bronstad 2001). At this point, it is not clear whether various bodily odors are susceptible to degradation in different degree. Our results show that a regular freezer might be sufficient for human bodily odor studies. Similarly, freezing for a period up to 6 months should not create significant effects on perceived hedonic judgments.

The only other study to test potential effects of freezing samples is that of Roberts et al. (2008), whose main aim was testing the effect of major histocompatibility complex (MHC) on body odor preferences. As these authors used frozen samples, they also tested reliability of their method by comparing hedonic ratings of fresh samples and those which were kept in freezer for a period of 1, 2, or 3 months, all samples being collected from the same individuals. Consistent with this study, they found no significant effect of the length of frozen storage, although the authors noted a nonsignificant decrease in pleasantness and desirability between fresh and frozen samples.

Another aim of our study was to test the effect of freeze–thaw cycles on the odor samples. Axillary odor develops from odorless substances produced mainly by apocrine glands that are metabolized by axillary microflora (e.g., Shelley et al. 1953; Leyden et al. 1981). Thus, a possible difference between simple freezing of the samples and repeatedly thawing them is that in the latter case, there could potentially be further bacterial action during the manipulation with the samples which might change their odor quality. The only ratings that were affected by repeated freezing were those of odor intensity. This applies to both the body odor samples and the essences. As might be expected, the fresh essence samples were rated more intense than the same samples used for the second time after 14 days (Session 2) and once again 1 month later (Session 3). Therefore, it seems likely that the decrease in intensity was related to the repeated manipulation of the samples. However, axillary samples that had been twice frozen and thawed (Session 3) were rated as more intense compared with the ratings of the same samples in Session 2. This might suggest that further microbial action had occurred between Sessions 2 and 3. Alternatively, the conditions in Session 3 may have been in some way unusual because a similar, but unexpected, finding was also found in the Session 3 ratings of axillary samples that had been frozen and thawed only once.

In contrast to intensity ratings, we found no significant changes in hedonic ratings in the course of repeated freezing and thawing. This is an interesting finding as a relatively high correlation between odor intensity and pleasantness/attractiveness is a common finding in body odor studies (e.g., Havlicek et al. 2006). Moreover, this relationship is not restricted to bodily odors and seems to be more a general phenomenon of odor perception. There is currently lively debate about whether odor intensity and pleasantness represents one perceptive dimension or consists of various cognitive processes (Rouby and Bensafi 2002). Our results might support the latter.

Repeatedly thawed body odor samples (T-shirts) have been used in several previous studies. For instance, Wedekind et al. (2007) tested whether verbal descriptions of body odor are correlated with one's MHC. Although the number of freeze–thaw cycles that samples were exposed to were not described, they state that the T-shirts were stored in freezer in between the days of evaluation. Rikowski and Grammer (1999), in their study on correlations between odor and facial attractiveness, also used their frozen samples twice. Although the authors do not give any further details on this procedure, the fact that they were able to detect the hypothesized effect (i.e., positive correlation between facial and odor attractiveness) indicates that more than one freeze–thaw cycle had no systematic effect on their results (Rikowski and Grammer 1999).

We further tested long-term stability of odor quality of freshly collected odor samples; that is, fresh samples collected from the same men at 2 different times. This is an important issue in studies performing long-term or repeated sampling designs, which are based on the assumption that

body odor of their subjects remains unchanged. For this purpose, we compared the samples collected from the same individuals at the beginning with end of a 4-month period (Experiment 1, Session 1 and 4). Unexpectedly, we found that pleasantness and attractiveness of the samples were rated significantly higher in Session 4. Such a result could be explained either by systematic change of the stimuli (e.g., the donor's odor itself or differences in sampling or testing conditions across sessions) or changes in raters at a perceptional level. For instance, it has been shown repeatedly that the quality of individual body odor can fluctuate, according to donor's health status, emotional and reproductive state, and diet (for a review, see Havlicek and Lenochova 2008). Here, the quality of samples could be systematically affected by seasonal differences in donors' diet (Havlicek and Lenochova 2006) as the first sampling session was carried in May and the other one in October. Another reason could be differences in ambient temperature between sampling sessions, affecting the intensity of perspiration and thus the humidity in the axillary area. Increasing humidity can stimulate rates of growth (Hartmann 1983) or colonization of skin microflora, leading to changes in the intensity of axillary odor (Hopwood et al. 2005).

Differences between the sessions might also be attributed to systematic changes in raters' perception. One may, for instance, argue that seasonal variations in mood might have influenced raters' judgments. However, we believe that the most likely explanation is that due to repeated exposure to body odors during the 4-month period, raters may have become familiar with it. For most of our raters, it was the first time they had taken part in this kind of experiment and they may possibly have rated samples more negatively in their first exposure, a tendency to rate unfamiliar odors rather negatively is a well-known phenomenon which is not restricted to bodily odors (Royet et al. 1999). This is supported by the data for essences. The reason why we included essences in our testing sessions was to test for possible changes in raters' olfactory perception, providing a control comparison against rating of the body odor samples. Ratings of essences did not vary in most recorded variables across the sessions. The only difference we observed was in intensity. In particular, ratings in Session 1 were rated as stronger compared with Session 2 or 4 (sample as the unit of analysis). This supports the suggestion that a repeat exposure effect was responsible for the differences. Therefore, we suggest that future studies using repeated measures designs should carry out pilot ratings to avoid bias in collected data.

In all data analyses, we used 2 approaches: 1) with rater and 2) sample as the unit of analysis. In general, analyses based on the rater as the unit stress changes in perception more and analyses based on samples stress changes in their chemistry more. However, in our study design, where the same raters repeatedly rated samples obtained from the same individuals, results are expected to converge. In all analyses, both approaches showed the same trends, although not always significant by both corresponding analyses. When such discrepancy occurred, it was mostly an analysis with the rater as unit which appeared to be significant, perhaps due to a higher number of units used and consequently also degrees of freedom.

In conclusion, our 2 experiments showed that freezing of bodily odor samples for at least a 6-month period does not significantly affect their perceived quality. Although we found some fluctuations in intensity ratings, these were not directly linked to the length of freezing. Similarly, ratings of repeatedly frozen and thawed samples were not significantly different from fresh samples obtained from the same individual and there was again no systematic link between the number of thawing cycles and their perceptual quality. These results suggest that this approach could be used in studies on the social impact of human odors. One should note, however, that our study aimed to investigate the potential effects of freezing only on ratings of hedonicity and intensity. Whether our results can be generalized to other olfaction-related cognitive tasks, identification or recognition, for example, should be determined in future studies.

Funding

Josef, Marie and Zdenka Hlavka Foundation (6/2/19/2006 to P.L.), Grant Agency of Charles University (393/2005 to J.H.); Grant Agency of Czech Republic (406/06/P377 to J.H.); AROMA Praha Corp. (provided essences); Optimum Distribution CZ&SK (provided perfume testers).

Acknowledgements

We would like to thank all the volunteers for their participation in our studies, Tamsin Saxton, Nassima Boulkroune, Jitka Lindova, and Tom Hadrava and 2 anonymous reviewers for many helpful comments and language corrections on previous versions of the manuscript.

References

Baron CP, Kjaersgard IVH, Jessen F, Jacobsen C. 2007. Protein and lipid oxidation during frozen storage of rainbow trout (*Oncorhynchus mykiss*). J Agric Food Chem. 55:8118–8125.

Caruso S, Grillo C, Agnello C, Maiolino L, Intelisano G, Serra A. 2001. A prospective study evidencing rhinomanometric and olfactometric outcomes in women taking oral contraceptives. Hum Reprod. 16:2288–2294.

Cernoch JM, Porter RH. 1985. Recognition of maternal axillary odors by infants. Child Dev. 56:1593–1598.

Comfort A. 1971. Likelihood of human pheromones. Nature. 230:432–433.

Doty RL. 1977. A review of recent psychophysical studies examining the possibility of chemical communication of sex and reproductive state in humans. In: Doty RL, Müller-Schwarze D, Mozell MM, editors. Chemical signals in vertebrates. New York: Plenum. p. 273–286.

Doty RL, Green PA, Ram C, Yankell SL. 1982. Communication of gender from human breath odors: relationship to perceived intensity and pleasantness. Horm Behav. 16:13–22.

Doty RL, Kligman A, Leyden J, Orndorff MM. 1978. Communication of gender from human axillary odors: relationship to perceived intensity and hedonicity. Behav Biol. 23:373–380.

Doty RL, Snyder PJ, Huggins GR, Lowry LD. 1981. Endocrine, cardiovascular, and psychological correlates of olfactory sensitivity changes during menstrual-cycle. J Comp Physiol Psychol. 98:45–60.

Fleming A, Corter C, Surbey M, Franks P, Steiner M. 1995. Postpartum factors related to mother's recognition of newborn infant odours. J Reprod Infant Psychol. 13:197–210.

Fleming AS, Steiner M, Corter C. 1997. Cortisol, hedonics, and maternal responsiveness in human mothers. Horm Behav. 32:85–98.

Hartmann AA. 1983. Effect of occlusion on resident flora, skin-moisture and skin-pH. Arch Dermatol Res. 275:251–254.

Havlicek J, Dvorakova R, Bartos L, Flegr J. 2006. Non-advertized does not mean concealed. Body odor changes across the human menstrual cycle. Ethology. 111:81–90.

Havlicek J, Lenochova P. 2006. The effect of meat consumption on body odor attractiveness. Chem Senses. 31:753–759.

Havlicek J, Lenochova P. 2008. Environmental effects on human body odour. In: Hurst JL, Beynon RJ, Roberts SC, Wyatt TD, editors. Chemical signals in vertebrates XI. New York: Springer, p. 199–212.

Havlicek J, Roberts SC, Flegr J. 2005. Women's preference for dominant male odour: effects of menstrual cycle and relationship status. Biol Lett. 1:256–259.

Havlicek J, Saxton T. Forthcoming. The effect of diet on human bodily odors. In: Hasegawa K, Takahashi H, editors. New research on food habits. New York: Nova Science Publishing.

Hold B, Schleidt M. 1977. The importance of human odor in nonverbal communication. Z Tierpsychol. 43:225–238.

Hopwood D, Farrar MD, Bojar RA, Holland KT. 2005. Microbial colonization dynamics of the axillae of an individual over an extended period. Acta Derm Venereol. 85:363–364.

Hummel T, Gollisch R, Wildt G, Kobal G. 1991. Changes in olfactory perception during the menstrual cycle. Experientia. 47:712–715.

Hurlbert SH. 1984. Pseudoreplication and the design of ecological field experiments. Ecol Monogr. 54:187–211.

Jacob S, McClintock M, Zelano B, Ober C. 2002. Paternally inherited HLA alleles are associated with women's choice of male odour. Nat Genet. 30:175–179.

Kjeldsen F, Christensen LP, Edelenbos M. 2003. Changes in volatile compounds of carrots (*Daucus carota L.*) during refrigerated and frozen storage. J Agric Food Chem. 51:5400–5407.

Kohoutova D, Rubesova A, Havlicek J. Forthcoming. The effect of shaving on axillary odor pleasantness. J Chem Ecol.

Lenochova P, Havlicek J. 2008. Human body odour individuality. In: Hurst JL, Beynon RJ, Roberts SC, Wyatt TD, editors. Chemical signals in vertebrates XI. New York: Springer, p. 189–198.

Leyden JJ, Mc Ginley KJ, Holzle E, Labows JN, Kligman AM. 1981. The microbiology of the human axilla and its relationship to axillary odor. J Invest Dermatol. 77:413–416.

Mallet P, Schaal B. 1998. Rating and recognition of peers' personal odors by 9-year-old children: an exploratory study. J Gen Psychol. 125:47–64.

Marlier L, Schaal B, Soussignan R. 1998a. Bottle-fed neonates prefer an odor experienced in utero to an odor experienced postnatally in the feeding context. Dev Psychobiol. 33:133–145.

Marlier L, Schaal B, Soussignan R. 1998b. Neonatal responsiveness to the odor of amniotic and lacteal fluids: a test of perinatal chemosensory continuity. Child Dev. 69:611–623.

Navarrete-Palacios E, Hudson R, Reyes-Guerrero G, Guevara-Guzman R. 2003. Lower olfactory threshold during the ovulatory phase of the menstrual cycle. Biol Psychol. 63:269–279.

Platek SM, Burch RL, Gallup GG. 2001. Sex differences in olfactory self-recognition. Physiol Behav. 73:635–640.

Porter RH, Balogh RD, Cernoch JM, Franchi C. 1986. Recognition of kin through characteristic body odors. Chem Senses. 11:389–395.

Porter RH, Cernoch JM, McLaughlin FJ. 1983. Maternal recognition of neonates through olfactory cues. Physiol Behav. 30:151–154.

Rennie PJ, Gower DB, Holland KT. 1991. In-vitro and in-vivo studies of human axillary odor and the cutaneous microflora. Br J Dermatol. 124:596–602.

Rikowski A, Grammer K. 1999. Human body odour, symmetry and attractiveness. Proc R Soc Lond B Biol Sci. 266:869–874.

Roberts SC, Gosling LM, Carter V, Petrie M. 2008. MHC-correlated odour preferences in humans and the use of oral contraceptives. Proc R Soc Lond B Biol Sci. 275:2715–2722.

Roberts SC, Gosling LM, Spector TD, Miller P, Penn DJ, Petrie M. 2005. Body odor similarity in noncohabiting twins. Chem Senses. 30:651–656.

Rouby C, Bensafi M. 2002. Is there a hedonic dimension to odors? In: Rouby C, Schaal B, Dubois D, Gervais R, Holley A, editors. Olfaction, taste and cognition. Cambridge: Cambridge University Press. p. 140–159.

Royet JP, Koenig O, Gregoire MC, Cinotti L, Lavenne F, Le Bars D, Costes N, Vigouroux M, Farget V, Sicard G, et al. 1999. Functional anatomy of perceptual and semantic processing for odors. J Cogn Neurosci. 11: 94–109.

Schaal B, Marlier L. 1998. Maternal and paternal perception of individual odor signatures in human amniotic fluid—potential role in early bonding? Biol Neonate. 74:266–273.

Schaal B, Marlier L, Soussignan R. 1995. Responsiveness to the odor of amniotic fluid in the human neonate. Biol Neonate. 67:397–406.

Schaal B, Marlier L, Soussignan R. 1998. Olfactory function in the human fetus: evidence from selective neonatal responsiveness to the odor of amniotic fluid. Behav Neurosci. 112:1438–1449.

Schaal B, Porter RH. 1991. Microsmatic humans revisited—the generation and perception of chemical signals. Adv Study Behav. 20:135–199.

Schleidt M, Hold B, Attili G. 1981. A cross-cultural study on the attitude towards personal odors. J Chem Ecol. 7:19–31.

Shelley WB, Hurley HJ Jr, Nichols AC. 1953. Axillary odor: experimental study of the role of bacteria, apocrine sweat, and deodorants. Arch Dermatol Syph. 68:430–446.

Singh D, Bronstad M. 2001. Female body odor is a potential cue to ovulation. Proc R Soc Lond B Biol Sci. 268:797–801.

Soussignan R, Schaal B, Marlier L, Jiangs T. 1997. Facial and autonomic responses to biological and artificial olfactory stimuli in human neonates: re-examining early hedonic discrimination of odors. Physiol Behav. 62: 745–758.

Wedekind C, Escher S, Van de Waal M, Frei E. 2007. The major histocompatibility complex and perfumers' descriptions of human body odors. Evol Psychol. 5:330–343.

Wedekind C, Füri S. 1997. Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? Proc R Soc Lond B Biol Sci. 264:1471–1479.

Wedekind C, Seebeck T, Bettens F, Paepke AJ. 1995. MHC-dependent mate preference in humans. Proc R Soc Lond B Biol Sci. 260:245–249.

Accepted October 15, 2008