Human Factors in Virtual Environments and Game Design, Vol. 137, 2024, 186–196
https://doi.org/10.54941/ahfe1005001

Measuring Detection and Habituation of Olfactory Stimuli in Virtual Reality for Improved Immersion

Scott Ledgerwood, Erika E Gallegos, and Marie Vans

Department of Systems Engineering Colorado State University, Fort Collins, CO 80523, USA

ABSTRACT

Odor displays for virtual reality (VR) have been used to enhance immersion, mood, and cognition, allowing for multi-sensory integration in VR applications from entertainment to education/training use cases. These odor displays can improve the overall human systems integration of VR, particularly for VR as a training tool with digital twins. However, foundational research is needed to understand thresholds for detection and overstimulation of odor in a VR environment. The objective of this paper is to evaluate 1) how various levels of scent intensity and duration effect detection of scent and 2) if habituation of intensity and pleasantness of scent occurs after prolonged exposure. To address this, a study was conducted on 34 participants using a wearable VR headset with an attached odor display. The study used a 5 (intensity: 105, 150, 225, 300, 600 ms; within-subject) x 3 (delay: 15, 30, 60 sec; within-subject) x 2 (scent: pleasant, unpleasant; within-subject) factorial design. Generalized linear mixed models were used to assess the ability to detect two different scent stimuli and ratings of perceived pleasantness and intensity levels over time. This article describes the odor display, scent selection, experimental design, and findings from the study. The results show significant effects on detecting the scent at a given intensity level and participant age, but none for gender or the selected delays. This framework and the associated conclusions can be used to guide multi-sensory integration in VR environments for improved human interaction.

Keywords: Extended reality, Human systems integration, Multi-sensory integration, Scent, User experience

INTRODUCTION

Virtual reality (VR) is a powerful tool that allows humans to interact with systems at scale, often with more feasibility and/or accessibility. This can be useful during requirements, design, evaluation, and training stages of a system's lifecycle. However, the usefulness of these digital counterparts can be limited by the immersiveness of the experience, especially when the human-system interaction is of focus (e.g., VR for training). Therefore, increasing the sense of presence felt by users can improve the utility of these VR tools.

VR simulation requires effective integration of senses beyond just the user's hearing and vision to increase their sense of presence (Slater et al., 1997). For example, odor displays (i.e., scent dispersion) have been integrated with VR applications to supplement training experiences (Richard et al., 2006) and have been successfully used to promote behavioral reactions (Flavian et al., 2021), reduce VR motion sickness (Ranasinghe et al., 2020) and for stress relief (Pizzoli et al., 2021; Serrano et al., 2016).

Integrating scent stimuli can improve a trainee's sense of presence and immersion, thus increasing the fidelity of the VR experience (Baus et al., 2019). For example, incorporating a smoke-like scent into a VR firefighter training program can create a realistic smoke-filled dwelling scenario. The use of scent coupled with the 3D displays and interactions of VR can better replicate the stressors of the actual situation in ways traditional 2D training cannot (Moore et al., 2015). Moreover, the inclusion of odor in VR has been associated with achieving intended training outcomes and improved knowledge/skill capture (Whitlock et al., 2022).

While literature suggests that the inclusion of scent can be beneficial in VR, gaps still exist in formalizing the integration and deployment of olfactory in VR modules. This is in large part because of the variability across delivery methods for odor displays, which can include airflow, vortex ring, natural diffusion, or direct injection (Yanagida, 2012). There are tradeoffs between these approaches, where ambient dispersion offers a less intrusive approach, while direct injection can provide more controlled and precise perception. An odor display connected to a VR head-mounted display (HMD) is likely more user friendly and easier to deploy, hence this design was selected for use in this study. Previous research has explored the use of odor displays in VR HMDs but has not considered the intensity associated with these displays (Covarrubias et al., 2015).

In addition to the delivery mechanism, temporal effects are also relevant considerations. Previous research has used direct injection olfactometers to assess adaptation and habituation for human subject responses (Wang et al., 2002; Croy et al., 2013). These have found that continuous olfactory stimulation creates a phenomenon where sensory activity decreases and prohibits cognitive perception of the scents (Wang et al., 2002). Although habituation across scent types can vary, where users have shown to habituate to an unpleasant scent (hydrogen sulfide) after repeated exposure, but not so for pleasant scents (phenethyl alcohol and peach) (Croy et al., 2013). Minimizing this adaptation by determining the smallest amounts of scent disbursement can mitigate scents lingering in the air and presenting conflicting information (Noguchi et al., 2011).

Previous research has laid a foundation for olfactory integration into VR for improved immersion. Our study aims to build on this by investigating the effectiveness of a wearable odor display for detecting scent stimuli in a VR simulation, considering variations in scent type, intensity, and delay. Our primary goal is to determine the optimal parameters under which a participant can accurately detect a scent stimulus using the wearable odor display in VR. In this paper, we seek to answer the following objectives: 1) how do various levels of scent intensity and duration effect detection of scent, and 2) does habituation towards a neutral scent occur after prolonged exposure to different scents.

METHODS

A VR study was conducted to evaluate factors that affect reliable detection and habituation of olfactory stimulation. Participants experienced a 20-minute virtual simulation, during which they were exposed to two different scents across various scent intensities and delays (i.e., interstimulus intensity, ISI). The study had Institutional Review Board (IRB) approval and informed consent was obtained from each participant prior to data collection.

Participants

There were 34 participants included in this study (12 females, 22 males). Based on previous literature of homogenous olfactory perception groups (Noguchi et al., 2011), participants were binned into two age groups: 35 and under (N = 20) and over 35 (N = 14).

Data collection occurred near the end of the COVID-19 pandemic (November – December of 2022), which has been associated with loss of smell. As seen in Table 1, a total of 22 participants shared that they had COVID-19 with recoveries ranging from 1 to greater than 6 months prior to study involvement. Moreover, five participants reported that they had lost their scent at one point due to COVID-19. However, these individuals performed similar to other individuals and thus we did not consider it as a concern for our study.

Table 1. Participant recovery time since last testing positive for COVID-19.

Time	Count	Percent
< 1 month	0	0%
1 – 3 months	2	5.9%
4 – 6 months	9	26.5%
> 6 months	11	32.3%
Never / Not Sure	12	35.3%

Participants were recruited through flyers and emails at Colorado State University. Exclusion criteria for the study excluded participants under 18, over 55 due to diminished ability to detect odor stimuli (Doty et al., 1984), and pregnant women due to potential heighted sense of smell. During recruitment, participants were also shown the odor display and given bursts of a scent, if they could not detect the scent, they were excused from the study.

Equipment

An Oculus Quest 2 was used for the VR HMD. A commercial odor display from OVR Technology was fit to the VR HMD, approximately 1-inch from the users' nose, see Figure 1. The display integrates into the VR experiment over Bluetooth, such that scent is dispersed based on the VR environment using temporal or spatial cues; in this study we dispersed scent at specific time intervals.

Figure 1: VR HMD with odor display attached.

The wearable odor display provided preconfigured scent kits, which contained relaxation or trauma-related odors. The scent cartridge, comprised of nine capsules, each contained a water-based scent solution with up to 5% fragrance, 1% of Benzisothiazolinone, and a surfactant. New cartridges were provided prior to the study as the lifetime of the scents is two months before it begins to degrade.

Two scents were used in this study: pleasant (smoke) and unpleasant (body odor). A pilot test with eight participants and six scents was conducted to determine the pleasantness ratings of various odors on a 7-point scale. The odor with the highest pleasantness rating and the odor with the lowest pleasantness rating were selected for inclusion in this study.

In addition, a water-based solution with no scent was dispersed every 5-seconds throughout the experiment. This neutral [water] scent was used to mitigate prior scents from lingering, as well as to minimize the likelihood of a participant initiating their detection each time they felt a water burst or heard a tiny click.

The physical room that the study was carried out in was an enclosed lab space, such that the participant sat at a table using the VR alone in the room with a closed door. The room was quiet, with a slight white noise generator, and good ventilation. The researcher sat right outside, next to a window into the room.

VR Environment

Within the VR space, participants were in an enclosed room with a large screen on one wall. The screen played a 20-minute video of the Great Barrier Reef, see Figure 2. This video was selected based on prior research that showed similar footage did not evoke an emotional state (Han et al., 2012). Audio was disabled for the video to remove a potential confounding factor. A questionnaire was used to collect consent, demographics, and detection of the scents; this was incorporated into the VR environment and users interacted with it using the VR controller.

Figure 2: Clip of the great barrier reef video in the VR environment.

Questionnaire

In addition to demographics, the questionnaire was used to capture scent detection and ratings of intensity and pleasantness. Participants were instructed to click a button on the VR controller when they detected a scent. This would create a pop-up window asking them to rate the intensity (7-point Likert scale, from weak to strong) and pleasantness (7-point Likert scale, from unpleasant to pleasant) of the scent. Participants could click "submit" after selecting their ratings or "cancel" if they accidentally clicked the button to bring up the questionnaire.

Procedure

The study was first piloted on four participants to assess study procedures and data logging; data from these participants were not used in analysis. Then, 34 participants successfully completed the study. Participants were first shown the wearable odor display and provided a sample scent not used in the study to understand how the device functioned. They were instructed on the VR controls and the task of detecting and rating a scent when detected. Then, users donned the headset, completed a questionnaire on demographics, and then the video of the Great Barrier Reef began. They would continue watching the video, waiting for the next scent to be detected, and repeat the process of rating the intensity and pleasantness. This continued until all scents were dispersed (15 pleasant and 15 unpleasant; 30 in total). The study took approximately 25-minutes per participant.

The study used a 5 (intensity: 105, 150, 225, 300, 600 ms; within-subject) x 3 (delay: 15, 30, 60 seconds; within-subject) x 2 (scent: pleasant, unpleasant; within-subject) factorial design. Hence, all participants experienced all levels of intensity and delay for each scent. The order of intensity x delay x scent was randomized across participants. These levels of intensity and delay were selected based on previous research (Wang et al., 2002), which indicated that stimulus strength of 35 ms to 200 ms at 5, 10, and 60 second delay levels, based on a direct injection odor display, had significant effects at $p < 0.01$. Larger delay and intensity values were selected for our study based on the open-air design of our unit compared to their direct injection olfactometer (Lorig et al., 1999).

Data Analysis

The R statistical software program (version 4.2.2) was used for data cleaning and analysis. Statistical significance was assessed at $\alpha < 0.05$.

Scent detection (hit/miss) was modeled using a binary logistic mixed model. Random effects on the participant level were controlled for in the model, to account for the repeated observations within subjects. Post-hoc analysis was used to analyze marginal effects.

Measures of scent habituation (self-reported intensity and pleasantness) were evaluated using repeated measures analysis of variance (ANOVA), where data was blocked on the participant level. Tukey honest significant difference post hoc tests were used to further evaluate significant factors.

Dependent Variables

Scent detection, coded as a hit or a miss, was based on the participant response to the scent stimuli. A hit was defined as the participant pressing the controller button to initiate the intensity and pleasantness questionnaire within 2 to 17 seconds after the scent was dispersed. A miss was defined as no button press, a button press earlier than 2 seconds, or a button press longer than 17 seconds after a scent was dispersed.

Scent habituation was based on the responses over time to the intensity and pleasantness questionnaire that appeared after the participant recorded a hit. After each detection, participants reported their rating of intensity and rating of pleasantness, both on a 7-point Likert scale.

Independent Variables

There were five fixed effects considered in the models:

- Scent: smoke or body odor.
- Intensity: duration that the scent was dispensed for, where a longer intensity corresponded to a stronger scent, with levels of 105 ms, 150 ms, 225 ms, 300 ms, and 600 ms.
- Delay: time between scent dispersion, with levels of 15 seconds, 30 seconds, or 60 seconds.
- Order: sequence of the combinations of scent x intensity x delay presented to the user over the 30 disbursements. Order was used to assess if there were any habituation effects to the subjects' perceived intensity or pleasantness of the two scents.
- Age: categorized as $\leq$ 35 years old or > 35 years old.
- Gender: categorized as male or female.

RESULTS

Scent Detection

A generalized linear mixed model was used to predict scent detection (hit/miss), see Table 2. There was a significant effect for intensity at 600 ms, age, and the interaction of smoke x intensity at 600 ms. Specifically, that detection for a 600 ms intensity was 4.41 (i.e., $e^{1.483}$) times more likely than for an intensity of 105 ms, but there was no differentiation in the probability of detection between the other intensities. Similarly, individuals 35 years old and under were 2.17 times more likely to detect a scent compared to individuals over 35 years old. There was also an additional decreased likelihood of detecting smoke at 600 ms, suggesting that body odor was even more detectable at this high intensity. Interestingly, there was no effect of delay or gender on detection.

Table 2. Main effects from model predicting scent detection.

Variable	Coeff	SE	z-value	p-Value
Intercept	−1.488	0.346	−4.31	<0.001
Smoke *(ref. Body Odor)*	−0.627	0.396	−1.59	0.113 (ns)
Intensity 150 *(ref. 105 ms)*	0.254	0.337	0.75	0.452 (ns)
Intensity 225 *(ref. 105 ms)*	0.033	0.314	0.11	0.916 (ns)
Intensity 300 *(ref. 105 ms)*	0.494	0.296	1.67	0.095 (ns)
Intensity 600 *(ref. 105 ms)*	1.483	0.298	4.98	<0.001
Delay 30 *(ref. 15 sec)*	−0.095	0.249	−0.38	0.701 (ns)
Delay 60 *(ref. 15 sec)*	0.177	0.243	0.73	0.466 (ns)
Age $\leq$ 35 *(ref. > 35 y.o)*	0.777	0.251	3.09	0.002
Male *(ref. female)*	−0.023	0.254	−0.09	0.927 (ns)
Smoke x Intensity 150 ms	−0.262	0.498	−0.53	0.599 (ns)
Smoke x Intensity 225 ms	−0.003	0.481	−0.01	0.996 (ns)
Smoke x Intensity 300 ms	−0.357	0.466	−0.77	0.444 (ns)
Smoke x Intensity 600 ms	−1.201	0.463	−2.59	0.010
Smoke x Delay 30 sec	0.368	0.381	0.97	0.335 (ns)
Smoke x Delay 60 sec	0.130	0.380	0.34	0.733 (ns)
Model Fit	*AIC*	*LL*	*LRatio*	*p-value*
Model	1130.2	−548.1	75.8	<0.001
Null	1176.0	−586.0		

To further understand the differences of scent and intensity levels, an interaction plot of the estimated marginal means was created, see Fig. 3. The plot on the left shows the differences between intensities for the body odor scent and for smoke on the right. Specifically, the left plot shows that for the body odor scent, we can see a significant effect for the difference between an intensity of 600 ms and all other intensities, as denoted by the spike in predicted probability at 600 ms. The differences amongst the lower intensities had no significant effects. The plot for smoke shows similar trends, however not as pronounced.

Since older age has been associated with decreased scent detection, we further investigated the effects of age in our study, as shown in Fig. 4. Participants aged 35 and under had a probability of 43% (CI: 36% to 52%) for detecting the body odor scent and 26% (CI: 21% to 34%) for detecting the smoke scent across all conditions. In contrast, the over 35 years old group had a lower likelihood of detecting both scents, with a probability of 28% (CI: 21% to 37%) for body odor and 15% (CI:11% to 22%) for smoke.

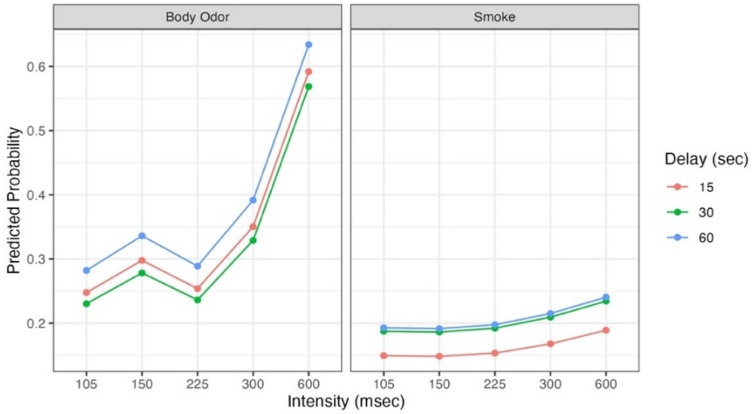

Figure 3: Estimated marginal means of each scent across levels of intensity and delay.

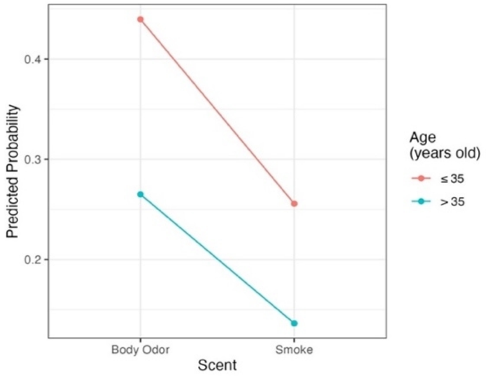

Figure 4: Estimated marginal means of each scent across age groups.

Scent Habituation

Two ANOVAs were performed to assess if habituation towards the scent intensity and pleasantness within subjects occurred as the study progressed. That is, if self-reported ratings of intensity and/or pleasantness trended neutral by the end of the study. In both ANOVAs, the independent variables were order, scent, intensity, and delay, with random effects at the participant level.

The ANOVA on self-reported intensity indicated no effect of order or delay, but there was a within-subject effect of scent ($F(1, 253) = 29.49$, $p < 0.001$) and intensity ($F(4, 253) = 4.11$, $p = 0.003$). Tukey post hoc tests showed that body odor was rated as more intense (mean = 3.8, SD = 1.7) compared to smoke (mean = 3.1, SD = 1.4). Similarly, higher intensities were reported as more intense for both scents.

The ANOVA on self-reported pleasantness also showed no significant effect of order or delay, and in addition, intensity. However, there was a within-subject significant effect of scent ($F(1, 253) = 29.3$, $p < 0.001$); where smoke was rated as more pleasant (mean = 4.1) compared to body odor (mean = 3.4).

DISCUSSION

In this study, we examined the impact of odor intensity and delay on 34 participants using a water-based, wearable odor display attached to a VR HMD. The study employed a 5 × 3 x 2 within-subject factorial design with repeated measures. We saw no difference in participants ability to detect scents based on prior COVID-19 infection and gender.

Our results showed significant main effects on detection for intensity, scent, and age. Specifically, the unpleasant odor was detected at higher rates, which aligns with previous research that threats are more readily detected by the human olfactory system (Wang et al., 2002). There was also an effect on age, with subjects 35 years and under having a higher probability of a successful detection. This is consistent with research showing anosmia occurring after peak age of odor detection (30s and 40s) (Doty et al., 1984). Further, research has suggested interactions between age and gender with better performance for women aged 16–55 (Hummel et al., 2007), however, we did not see this effect in our results.

Our results suggests that the effectiveness of a scent may vary depending on the intensity, specific scent, and age of the user. It is possible that certain scents are more potent at lower intensities, while others may require higher intensities to produce a significant effect. The lower odds of smoke being detected at the fullest strength presented in this study indicate a possible high noise floor for the odor display. The smoke scent seems to be less distinguishable.

The results suggest that the ability to detect a scent across users does not appear to be affected by delays ranging from 15 to 60 seconds. It may be that the sense of smell is relatively quick to perceive and retain the information about the scent. The delay may not have been short enough to have a just noticeable difference (JND) for the wearable odor display. Prior research indicated 2.5 seconds as a JND, where users were not able to detect scents due to habituation but were successful at and above 10 second delay. The lack of effect for the 15, 30 and 60 second delays for the wearable odor display warrants additional research, which would be recommended in the 2.5 to 15 second delay range.

Croy et al. (2013) had found a trend for unpleasant scents to become more neutral after repeated exposure. Our study did not find that result based on the scents administered. It could be that the delays and intensities were different or that the unpleasant nature is scent dependent. Future research could seek to more closely mirror the levels and scents used in Croy et al. (2013) with a wearable odor display.

Limitations

The use of a water-based capsule and open cartridge design resulted in background noise when trying to detect a particular scent from the cartridge full of nine capsules. This can make it difficult to accurately measure or perceive a given scent. To reduce the noise floor and improve the accuracy of the odor display, steps should be taken to address the source of the noise and ensure that the conditions for displaying the scent are as controlled and consistent as possible.

While a general formula is provided for each scent in the cartridge, the exact formulation is not publicly available. Users of odor displays that reference and provide scents by name are advised to trial different intensity levels to ensure they are obtaining the desired response. A similar framework as presented in this paper could be used by these future studies.

The nature of this study focused on a psychometric signal detection task which relied on sequential bursts rather than spatial interactions. Humans traditionally engage with scents based on physically proximity to the source, i.e., where the intensity maximizes at the origin and dissipates as a user moves away. The results of this study provide a reference point for testing scent intensity levels and reveals that multiple scents can be disbursed at delays of 15 seconds with minimal effect.

Future Work

The use of multiple sensory integration into virtual training simulations could have a positive effect on knowledge outcomes and skill acquisition. This study highlighted considerations for deploying scents when using a wearable odor display with VR. Future research should apply these findings to odor integration in immersive simulations for evaluating training effectiveness compared to audio and visual only training. Additional findings may emerge on the user's sense of presence or willingness to adopt a technology with integrated scent (e.g., Net Promoter Score) (Witmer and Singer, 1998).

CONCLUSION

The use of VR as a tool for digital twins has become increasingly popular, as it offers a safe and feasible modality for training; where more immersive VR experiences can improve training effectiveness and the overall human-system interaction. One such solution to increasing fidelity is incorporating multisensory integration in the VR experience. This study explored the effects of incorporating olfactory into a VR simulation. We identified thresholds of scent intensities and temporal delays that can be reliably detected, as well as evaluated demographic factors, such as age and gender. The framework and associated conclusions presented in this paper can guide olfactory integration into VR environments, for improved human interaction with VR digital twins.

REFERENCES

Baus O., Bouchard S., Nolet K. (2019). Exposure to a pleasant odour may increase the sense of reality, but not the sense of presence or realism. *Behaviour & Information Technology,* 38(12): 1369–1378.

Covarrubias M., Bordegoni M., Rosini M., Guanziroli E., Cugini U., Molteni F. (2015). VR system for rehabilitation based on hand gestural and olfactory interaction. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST)*, 117–120.

Croy I., Maboshe W., Hummel T. (2013). Habituation effects of pleasant and unpleasant odors. *International Journal of Psychophysiology,* 88(1): 104–108.

Doty R. L., Shaman P., Applebaum S. L., Giberson R., Siksorski L., Rosenberg L. (1984). Smell identification ability: Changes with age. *Science*, 226(4681): 1441–1443.

Flavian C., Ibáñez-Sánchez S., Orús C. (2021). The influence of scent on virtual reality experiences: The role of aroma-content congruence. *Journal of Business Research, 123:* 289–301.

Han S., Lerner J. S., Zeckhauser, R. (2012). The disgust-promotes-disposal effect. *Journal of Risk and Uncertainty,* 44: 101–113.

Hummel T., Kobal G., Gudziol H., Mackay-Sim A. (2007). Normative data for the "Sniffin' Sticks" including tests of odor identification, odor discrimination, and olfactory thresholds: an upgrade based on a group of more than 3, 000 subjects. *European Archives of Oto-Rhino-Laryngology,* 264: 237–243.

Lorig T. S., Elmes D. G., Zald D. H., Pardo J. V. (1999). A computer-controlled olfactometer for fMRI and electrophysiological studies of olfaction. *Behavior Research Methods, Instruments, Computers,* 31(2): 370–375.

Moore A., Herrera N., Hurst T., McMahan R., Poeschl S. (2015). The effects of olfaction on training transfer for an assembly task. *In Proceedings of the 2015 IEEE Virtual Reality (VR)*, 237–238.

Noguchi D., Sugimoto S., Bannai Y., Okada K. (2011). Time characteristics of olfaction in a single breath. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*, 83–92.

Pizzoli S. F. M., Monzani D., Mazzocco K., Maggioni E., Pravettoni G. (2021). The power of odor persuasion: The incorporation of olfactory cues in virtual environments for personalized relaxation. *Perspectives on Psychological Science*, 17(3): 652–661.

Ranasinghe N., Jain P., Tolley D., Tailan S. K., Yen C. C., Do E. Y.-L. (2020). Exploring the use of olfactory stimuli towards reducing visually induced motion sickness in virtual reality. In: *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*, 12: 1–9.

Richard E., Tijou A., Richard P., Ferrier J.-L. (2006). Multi-modal virtual environments for education with haptic and olfactory feedback. *Virtual Reality*, 10(3–4): 207–225.

Serrano B., Baños R., Botella C. (2016). Virtual reality and stimulation of touch and smell for inducing relaxation: A randomized controlled trial. *Computers in Human Behavior, 55*(Part A): 1–8.

Slater M., Wilbur S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6(6): 603–616.

Wang L., Walker V. E., Sardi H., Fraser C., Jacob T. J. C. (2002). The correlation between physiological and psychological responses to odour stimulation in human subjects. *Clinical Neurophysiology*, 113(4): 542–551.

Whitlock M., Chelen J., Ledgerwood S. (2022). Toward VR-based training modules for public safety scenarios. In *ACM CHI'22 – VR [we are] Training Workshop* 2022.

Witmer B. G., Singer M. J. Measuring presence in virtual environments: A presence questionnaire. (1998). *Presence: Teleoperators and Virtual Environments*, 7(3): 225–240.

Yanagida Y. (2012). A survey of olfactory displays: Making and delivering scents. *In Proceedings of IEEE SENSORS,* 1–4.