# A Study on The Quality of Experience of Crossmodal Mulsemedia

Alexandra Covaci
University of Kent
Canterbury, United Kingdom
a.covaci@kent.ac.uk

Gebremariam Mesfin
Brunel University
London, United Kingdom
gebremariam.assres@brunel.ac.uk

Nadia Hussain
Brunel University
London, United Kingdom
nadia.hussain@brunel.ac.uk

Elahe Kani-Zabihi
University of West London
London, UK
Elahe.Kani@uwl.ac.uk

Frederic Andres
National Institute of Informatics
Tokyo, Japan
andres@nii.ac.jp

Gheorghita Ghinea
Brunel University
London, United Kingdom
george.ghinea@brunel.ac.uk

## ABSTRACT

Media content is consumed more and more on digital and mobile devices. Internet of Things enables media providers to move beyond the traditional content and entertainment space and offer a plethora of new services that creates an engaging user experience around content. However, one of the questions that remain open is how to design content that enhances users' quality of experience by widening the classic audio-visual multimedia content spectrum to stimulate also other human senses such as olfaction or touch. In this paper we look at systematic associations made between different sensory modalities (crossmodal correspondences) in a digital setup and on their influence on the quality of experience. Our experiments target various senses: audio, visual, olfaction and tactile and show that creating untraditional content that matches the visual dominants of the video does not contribute to an improved experience of mulsemedia.

## CCS CONCEPTS

• **Human-centered computing → Interaction paradigms**; **Interaction devices**; **Interaction techniques**;

## KEYWORDS

mulsemedia; digital content; IoT services; quality of experience; crossmodal correspondences

## 1 RELATED WORK

As highlighted in [15], one of the key challenges for olfaction-based mulsemedia applications is how to integrate the various modalities as a step towards truly immersive experiences and enhanced user QoE.

### 1.1 Internet of Mulsemedia Things

As an increasing number of devices join IoT, the potential of IoT applications is rising significantly stimulating research on security implications and logistical implementations. But IoT is not only about efficiency, this shift towards sheer inter-connectivity dramatically changes our leisure and entertainment habits.

Inspired by the IoT revolution, multisensory digital content (mulsemedia) design can benefit from emerging new ways of engaging with audiences. Holographic displays based on finger gestures, creating the illusion of objects floating in air, have been proposed in [27]. Gaming devices are becoming more IoT friendly. One example are the Sifteo cubes, a tangible and graphical user interface platform, that make experiences more interactive and socially-focused [13]. They offer users the opportunity to directly engage with other objects and to create collaborative media. Google's Project Jacquard creates innovative textiles in collaboration with Levi's allowing users to be alerted through haptic sensations [1]. Fragrances provide interesting opportunities with good results on increasing focus or relaxation. Pium [2] is a new device that combines aromatherapy with IoT automating the scent release process based on a user's profile.

Considering all these opportunities, IoT enables moving beyond the traditional content and entertainment space by allowing new ways of interaction. Sensor technologies like GPS and RFID can be programmed in a personalized fashion. Live events or football games can be enriched through the interconnection of smart devices with devices that deliver additional sensory effects to the user. The QoE of IoT applications considering factors like the video coding, the data accuracy and the network delay have been investigated in [6]. However, the focus of this paper is different. We look on the creation of content based on cognitive principles and we analyze how they affect the perception of digital mulsemedia and the user's experience. Our research questions fit the spectrum of content creation with an application in IoT mulsemedia products.

[1]http://atap.google.com/jacquard/
[2]https://www.pium.co

## 1.2 Olfactory-visual crossmodal correspondence

Crossmodal correspondences modulate multisensory integration / perception. They refer to mappings between stimulus properties from different sensory modalities in a manner that is surprisingly consistent (e.g., pitch in audition and brightness in vision). While the focus in our paper is primarily on effects of the crossmodal correspondences between olfactory and visual stimuli, crossmodal associations were documented between many pairs of sensory modalities such as: vision and touch [18], audition and touch [28], flavors and sounds [2], flavors and vision [7]. All these were studied in a non-digital context.

Crossmodal correspondences were shown to influence people's performance under different experimental paradigms: direct crossmodal matching, faster classification tasks, faster simple detection tasks, implicit association tests, spatial localization tasks, and perceptual discrimination tasks [20].

As demonstrated by researchers in different experimental approaches, nonarbitrary crossmodal correspondences exist between a variety of olfactory and visual stimuli.

Presentation of visual stimuli can influence olfactory information processing and vice versa. Visual information affects the olfactory perception in many aspects such as intensity [10], pleasantness [17] and quality evaluations [22]. Although different visual stimuli have been analyzed in previous research, color was the predominantly used feature ( [8–10]). Visual features are also matched with audio pitch [12].

Because this area has a practical relevance as well, it started to attract the interest of food sectors, marketers and advertisers who became interested in how they can convey information about the fragrance/flavor/taste of their products by stimulating matched senses like visual and auditory or touch. Thus, there are several studies trying to understand how a consumer can be influenced when making use of the five senses [21].

The work on crossmodal correspondences outside the field of cognitive sciences is limited. Understanding cognitive processes across all sensory modalities could play an important role in developing efficient multimodal interfaces that consider all the subtleties involved in the human perception [25].

Crossmodal correspondences support comprehension and retention of information through the accommodation of users' sensorimotor skills. Thus, their application in different contexts, such as interaction design, computer graphics, information retention, QoE could bring interesting insights.

In computer graphics, displaying crossmodaly-linked content has been shown to distract viewers from correctly identifying animation quality [1].

Other promising applications were illustrated in [5, 24], where the authors explore the efficiency of olfaction in introducing a new semantic layer into interaction design and HCI when engaged in driving or mindfulness training.

Having as starting point these promising results and the fact that in the same time, olfaction is closely linked to memory and emotion [3], in this paper we investigate whether delivering a crossmodally matching olfactory content in a mulsemedia experience has an effect of the user's QoE. Our hypothesis is that when the mulsemedia content is built following crossmodal principles, the experience of the users is enhanced.

## 2 USER STUDY - QUALITY OF EXPERIENCE IN CROSSMODAL MULSEMEDIA

The experiments we designed were aimed to investigate the potential influence of using crossmodal correspondences concepts in designing mulsemedia on the QoE experienced by the users. More specifically, this paper focuses on the analysis of olfaction - visual crossmodal correspondences in a digital mulsemedia setup. Our setup involves several senses like: visual, audio, olfactory and haptic. The visual content consists of six videos characterized by certain dominant visual features: color (blue, yellow), brightness (low, high), and shape (round, angular). These videos are enhanced with crossmodally matching scents and sounds. The sounds serve also as input to a haptic vest with vibration motors, worn by participants during experiments. We chose to use this vibrotactile display because of the users' increased emotional response to haptic-enhanced media, that was demonstrated in the literature [26].

### 2.1 Participants

12 participants (6 females, 6 males) took part in this experiment. They were aged from 18 to 41 years old and hailed from diverse nationalities and educational backgrounds (undergraduate and postgraduate students as well as academic staff). All participants spoke English and self-reported as being computer literate. They were randomly assigned to a control or an experimental group.

### 2.2 Experimental apparatus

The experiments took place in a laboratory environment with good ventilation, thus avoiding the problem of smell mixing. In Figure 1, we illustrate the display system we used for our experiments.

The olfactory content was provided by the Exhalia SBi4, that was considered by previous research more reliable and more robust than existing devices [16]. This was placed at 0.5m from the assessor, allowing her/him to detect the smell in 2.7-3.2s, as shown in [15] The SBi4 can store up to four interchangeable scent cartridges at a time, but we used a single slot in our experiments to prevent the mixing of scents. These cartridges contain scented polymer through which air is blown (through four built-in-fans). The synchronized presentation of the olfactory data was controlled through a program built using Exhalia's Java-based SDK.

The laptop used in our setup has an Intel CoreTM 2 Duo CPU @ 1.66GHz, 2GB RAM, and runs the Windows 10 professional operating system. The videos were displayed on the computer monitor with a resolution of 1366x768 pixels, with a viewing area of 1000x700 pixels in the center of the screen.

### 2.3 Mulsemedia content

As illustrated in Table 2, the six videos used in our experiments were selected based on their dominant visual features such as color, brightness and angularity of objects. The accompanying sound was modified to respect the principles of auditory-visual crossmodal correspondences, following guidelines from the literature. The video with dominant yellow images from the Ethiopian sulphur springs (V1) was watched accompanied by high pitch sounds, while the
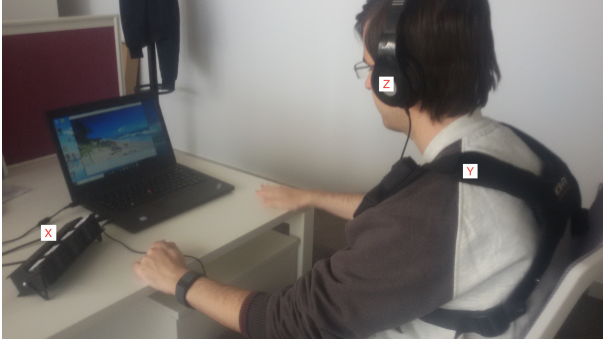
**Figure 1: Experimental setup for the visual, audio, olfactory, haptic media display system (X - Exhalia SBi4-radio scent emitter, Y - KOR-FX Haptic Vest, Z - Headphones)**

| Type of QoE | Questions |
|---|---|
| Olfactory | **Q1:** The smell was relevant to the video clip I was watching;<br>**Q2:** The smell came across strong;<br>**Q3:** The smell was distracting;<br>**Q4:** The smell was consistent with the video clip when released;<br>**Q5:** The smell was annoying;<br>**Q6:** The smell faded away slowly after watching the video clip;<br>**Q7:** The smell enhanced my viewing experience; |
| Auditory | **Q8:** The sound was related to the video clip I was watching;<br>**Q9:** The sound came across loud;<br>**Q10:** The sound was distracting;<br>**Q11:** The sound was annoying;<br>**Q12:** The sound enhanced the sense of reality whilst watching the video clip;<br>**Q13:** The sound enhanced my viewing experience; |
| Haptic | **Q14:** I enjoyed watching the video clip whilst wearing a Haptic Vest;<br>**Q15:** The Haptic Vest effects were relevant to the video clip I was watching;<br>**Q16:** The vibration was distracting;<br>**Q17:** The vibration was annoying;<br>**Q18:** The Haptic Vest effects enhanced the sense of reality<br>whilst watching the video clip;<br>**Q19:** The Haptic Vest effects enhanced my viewing experience; |
| Overall | **Q20:** Overall, I enjoyed the multisensorial experience. |

**Table 1: QoE questionnaire and the covered senses**

one dominantly blue, describing the sea, (V2) was delivered with low pitch sounds [19, 23]. The video with low level of brightness illustrating a solar eclipse (V3) was displayed accompanied by low

pitch sound, while V4, where sun was shining over the Arctic, was matched with high pitch sound. Angular skyscrapers were presented in V5 with high pitch sound, while the bouncing round shapes from V6 were accompanied by low pitch.
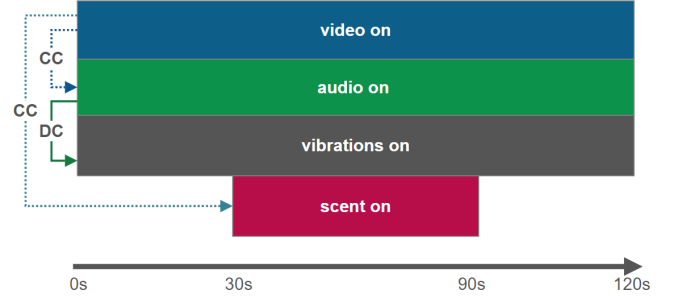


**Figure 2: Mulsemedia content delivery: CC - Crossmodal Correspondent content; DC - Direct Conversion content**

The olfactory content consisted of seven scents: rosemary, bergamot, lilial, clear lavender (low intensity), lavender (high intensity), lemon and raspberry.

Participants sat in a chair without armrests facing the screen. All participants wore headphones and a vibrotactile KOR-FX [3] gaming vest. To facilitate the vibrotactile experience, we chose the KOR-FX gaming vest that utilizes 4DFX based auditory-haptic signals to enable haptic feedback to the upper chest and shoulder regions. The vest is wirelessly connected to a control box meant to accept the standard sound output of the sound card of a computer. Users of this type of devices obtain additional information about environmental factors while becoming more immersed/involved in their experience [14].

All videos in our experiment were 120 seconds long. For the experimental group, the audio was adjusted to a frequency of 328Hz (high pitch condition) and 41Hz (low pitch condition). The vibrotactile feedback was provided for the whole length of the video and was derived automatically from the audio feedback. The olfactory content was delivered over a 60s interval, in the middle of the video, as shown in Figure 2. For this, a software framework has been developed to control the presentation of olfactory data and video. This allowed us to prevent the existence of any lingering scents. Additionally, in between the videos, while participants were filling in a questionnaire, the windows were open allowing new tests to take place in neutral conditions.

## 2.4 Procedure

**Pre-experiment study - audio pitch.** Before the experiments, we carried out a small pilot study with two participants, to get feedback on their thoughts and experience while trying our system. This was aimed to inform us about any disturbing or distracting factor. Since they reported that the high pitch audio was distracting, we lowered its volume to increase their comfort during the experiment.
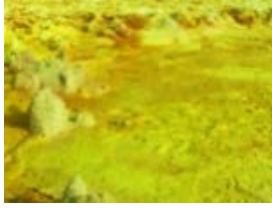
[3]http://korfx.com/products

| Video Snapshot | Description | Video Snapshot | Description |
|---|---|---|---|
| | **ID: V1** <br> **Theme:** Yellow sulphur springs <br> **Visual dominant:** Yellow color <br> **Audio:** High pitch <br> **CG scent:** Rosemary <br> **EG scent:** Bergamot | | **ID: V2** <br> **Theme:** Blue waves <br> **Visual dominant:** Blue color <br> **Audio:** Low pitch <br> **CG scent:** Rosemary <br> **EG scent:** Lilial |
| | **ID: V3** <br> **Theme:** Solar eclipse - sky turns dark <br> **Visual dominant:** Low brightness <br> **Audio:** Low pitch <br> **CG scent:** Rosemary <br> **EG scent:** Clear lavender | | **ID: V4** <br> **Theme:** Bright sun shining upon the Arctic <br> **Visual dominant:** High brightness <br> **Audio:** High pitch <br> **CG scent:** Rosemary <br> **EG scent:** Lavender |
| | **ID: V5** <br> **Theme:** Fly over angular skyscrapers <br> **Visual dominant:** Angular shape <br> **Audio:** High pitch <br> **CG scent:** Rosemary <br> **EG scent:** Lemon | | **ID: V6** <br> **Theme:** Bouncing balls <br> **Visual dominant:** Round shape <br> **Audio:** Low pitch <br> **CG scent:** Rosemary <br> **EG scent:** Raspberry |

**Table 2: Snapshots from the videos used during the experiment with details about their visual, audio and olfactory content**

| $Q_{ID}$ | Group $df$ | Group $F$ | Group $Sig.$ | Video $df$ | Video $F$ | Video $Sig.$ | Video*Group $df$ | Video*Group $F$ | Video*Group $Sig.$ |
|---|---|---|---|---|---|---|---|---|---|
| Q1 | 1 | 0.561 | 0.457 | 5 | 0.506 | 0.771 | 5 | 0.451 | 0.811 |
| Q2 | 1 | 0.560 | 0.457 | 5 | 0.233 | 0.947 | 5 | 0.454 | 0.809 |
| Q3 | 1 | 0.571 | 0.453 | 5 | 0.181 | 0.969 | 5 | 0.305 | 0.908 |
| Q4 | 1 | 2.589 | 0.113 | 5 | 0.390 | 0.853 | 5 | 0.707 | 0.621 |
| Q5 | 1 | 0.459 | 0.501 | 5 | 0.456 | 0.808 | 5 | 0.111 | 0.989 |
| Q6 | 1 | 0.049 | 0.826 | 5 | 1.424 | 0.229 | 5 | 0.839 | 0.527 |
| Q7 | 1 | 3.103 | 0.083 | 5 | 1.048 | 0.399 | 5 | 0.825 | 0.537 |
| Q8 | 1 | 0.287 | 0.594 | 5 | 1.701 | 0.150 | 5 | 2.324 | 0.055 |
| Q9 | 1 | 0.614 | 0.434 | 5 | 1.186 | 0.328 | 5 | 0.925 | 0.472 |
| Q10 | 1 | 12.834 | **0.001** | 5 | 0.767 | 0.577 | 5 | 1.441 | 0.224 |
| Q11 | 1 | 9.400 | **0.003** | 5 | 1.675 | 0.155 | 5 | 2.217 | 0.065 |
| Q12 | 1 | 2.076 | 0.155 | 5 | 1.939 | 0.101 | 5 | 2.430 | **0.045** |
| Q13 | 1 | 5.051 | **0.028** | 5 | 1.626 | 0.167 | 5 | 0.232 | 0.947 |
| Q14 | 1 | 1.739 | 0.192 | 5 | 1.080 | 0.381 | 5 | 0.580 | 0.715 |
| Q15 | 1 | 2.507 | 0.119 | 5 | 1.985 | 0.094 | 5 | 1.510 | 0.200 |
| Q16 | 1 | 1.981 | 0.165 | 5 | 0.807 | 0.549 | 5 | 0.554 | 0.734 |
| Q17 | 1 | 2.038 | 0.159 | 5 | 0.849 | 0.521 | 5 | 0.420 | 0.833 |
| Q18 | 1 | 0.601 | 0.441 | 5 | 1.975 | 0.096 | 5 | 0.752 | 0.588 |
| Q19 | 1 | 2.891 | 0.095 | 5 | 1.062 | 0.391 | 5 | 0.746 | 0.592 |
| Q20 | 1 | 0.000 | 1.000 | 5 | 0.722 | 0.574 | 5 | 0.281 | 0.922 |

**Table 3: Statistical analysis of self-reported measures for the questions in the QoE questionnaire**

**Conditions.** Participants were randomly divided in two groups of six each (experimental group: **EG**, and control group: **CG**) and watched the six videos in a random order in both conditions.

The two conditions tested by the EG and CG differed in the provided olfactory content. The audiovisual setup described in Subsection 2.3 was identical for both of these conditions. The haptic

content, automatically generated from the sound card of the computer, coincided as well. A detailed description of the experimental setup is presented in Table 2.

- For the **CG**, all the six videos were enhanced with rosemary scent. We chose this scent because of its demonstrated benefits on increasing alertness in tasks [4].
- For the **EG**, scents were selected based on olfactory-visual crossmodal principles investigated previously in the literature [11]. Accordingly, V1 was displayed in the presence of bergamot scent, and V2 was matched with lilial scent. As mentioned in Section 1, visual brightness is correlated with the intensity of odors [10]. Thus, V3 was displayed in the presence of weak lavender scent (clear lavender), while V4 was presented together with strong lavender scent. Because angular shapes are associated with lemon scent and round shapes with raspberry scent [9], these fragrances were used for V5 and V6, respectively.

**Collected data.**

Participants completed a subjective questionnaire consisting of 20 questions at the end of each video. This questionnaire is presented in Table 1 and was built so that it covers the experience in all the sensory dimensions employed in the experiment. Each question was answered on a 5-item Likert scale, anchored at one end with "Strongly Disagree" and with "Strongly Agree" at the other. The questions in the questionnaire are grouped to measure different types of experience (olfactory, audio, haptic and overall).

## 3   RESULTS

In this section of the paper we will discuss the findings with respect to subjective data captured during our crossmodal experiments.

To test whether the QoE evaluation was induced by visual characteristics or olfactory manipulations, a two-way, mixed-design analysis of variance (ANOVA) was performed having as factors: (i) **visual dominant features** - with six levels (yellow, blue, dark, bright, angular, and round) and (ii) **olfactory content** - with two levels (crossmodally matched olfactory content for EG and not-crossmodally matched olfactory content for CG). The QoE questions were also grouped by topic (olfaction, audio, haptic and overall experience). The tests were performed on the data with 95% confidence level, using the IBM statistical analysis software package SPSS. Results are presented in Table 3 with significant main and interaction effects in bold.

There was a significant main effect of the crossmodality of the olfactory content for the user experience reported in Q10, Q11, Q13: $F(1, 57) = 12.834$, $p = 0.001$; $F(1, 58) = 9.4$, $p = 0.003$; $F(1, 60) = 5.051$, $p = 0.028$. The interaction level was significant only for Q12 $F(5, 60) = 2.430$, $p = 0.045$. As shown in Figure 4, the crossmodally-matched olfactory content influences the QoE reported by users for certain visual content (high auditory QoE for blue and yellow content combined with crossmodal olfactory content). This indicates that the sense of reality was significantly enhanced for V1 and V2 in the crossmodal sound condition. Based on this, we could affirm that there is a stronger interaction between color and scent with effects on the auditory QoE.

Overall, as presented in Figure 3, there are no significant differences between the two experimental conditions and the reported

QoE. This seems to indicate that a crossmodally-matched olfactory content does not trigger an enhanced QoE with the exception of Q12, when the auditory experience was improved at the interaction between color and olfactory content.

## 4   DISCUSSION AND CONCLUSION

The results of the study reported in this paper show that designing a crossmodaly-matching mulsemedia content does not have a significant effect on improving the QoE. Participants in our experiments rated in a similar way their experience in both EG and CG. This indicates that the non-crossmodally-matched scent (rosemary) enhanced their QoE in a similar way with the crossmodally-matched scent.

In general, the audio that was also modified to respect visual-audio crossmodal principles was perceived as annoying, with scores for the auditory QoE lower than the other categories. Interestingly, it seems that the presence of crossmodally-matched scent had an effect on reducing this annoyance. Significant main effects related to the type of olfactory content appeared for Q10: The sound was distracting, Q11: The sound was annoying, and Q13: The sound enhanced my viewing experience. For all these, the QoE reported by the the EG was higher. Moreover, for Q12: The sound enhanced the sense of reality whilst watching the video clip, we observed an interaction effect between the video and olfactory factors showing that the videos with the color dominant had a significant higher QoE when combined with crossmodaly-matching scents.

This seems to indicate that although overall the experience of the users was similar, applying the crossmodal principles improved the reported auditory QoE.

In summary, our findings demonstrate that the crossmodal correspondence of scents has a positive effect on the hedonic ratings of sounds. Further study using various visual, olfactory and/or auditory stimuli is needed to generalize the current findings.

## REFERENCES

[1] Belma R. Brkic, Alan Chalmers, Kevin Boulanger, Sumanta Pattanaik, and James Covington. 2009. Cross-modal Affects of Smell on the Real-time Rendering of Grass. In *Proceedings of the 25th Spring Conference on Computer Graphics (SCCG '09)*. ACM, New York, NY, USA, 161–166. DOI:http://dx.doi.org/10.1145/1980462.1980494

[2] Anne-Sylvie Crisinel and Charles Spence. 2009. Implicit association between basic tastes and pitch. *Neuroscience letters* 464, 1 (2009), 39–42.

[3] Barry J Davies, Dion Kooijman, and Philippa Ward. 2003. The sweet smell of success: olfaction in retailing. *Journal of Marketing Management* 19, 5-6 (2003), 611–627.

[4] Miguel A. Diego, Nancy Aaron Jones, Tiffany Field, Maria Hernandez-reif, Saul Schanberg, Cynthia Kuhn, Mary Galamaga, Virginia McAdam, and Robert Galamaga. 1998. Aromatherapy Positively Affects Mood, Eeg Patterns of Alertness and Math Computations. *International Journal of Neuroscience* 96, 3-4 (1998), 217–224. DOI:http://dx.doi.org/10.3109/00207459808986469 PMID: 10069621.

[5] Dmitrijs Dmitrenko, Emanuela Maggioni, Chi Thanh Vi, and Marianna Obrist. 2017. What did I sniff? Mapping scents onto driving-related messages. In *AutomotiveUI'17 Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 154–163.

[6] A. Floris and L. Atzori. 2015. Quality of Experience in the Multimedia Internet of Things: Definition and practical use-cases. In *2015 IEEE International Conference on Communication Workshop (ICCW)*. 1747–1752. DOI:http://dx.doi.org/10.1109/ICCW.2015.7247433

[7] David Gal, S Christian Wheeler, and Baba Shiv. 2007. Cross-modal influences on gustatory perception. (2007).

[8] Avery N Gilbert, Robyn Martin, and Sarah E Kemp. 1996. Cross-modal correspondence between vision and olfaction: the color of smells. *The American journal of psychology* (1996), 335–351.
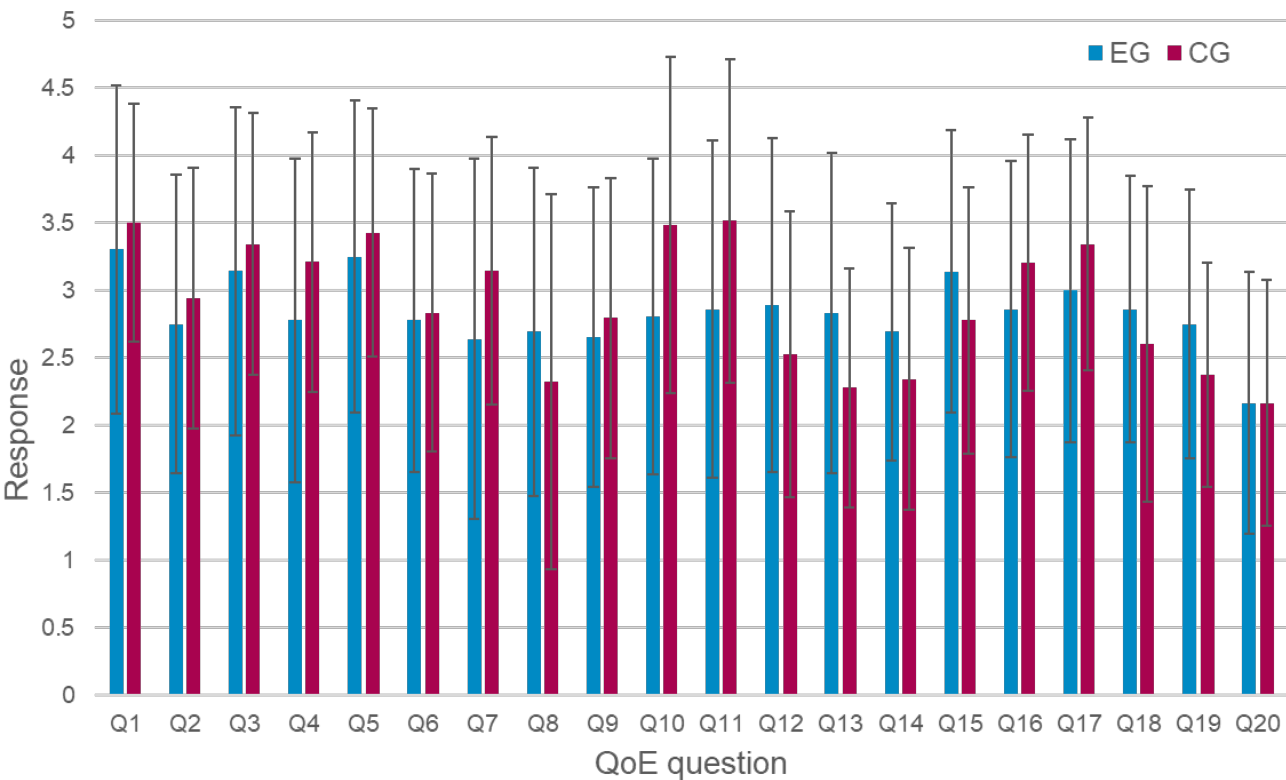
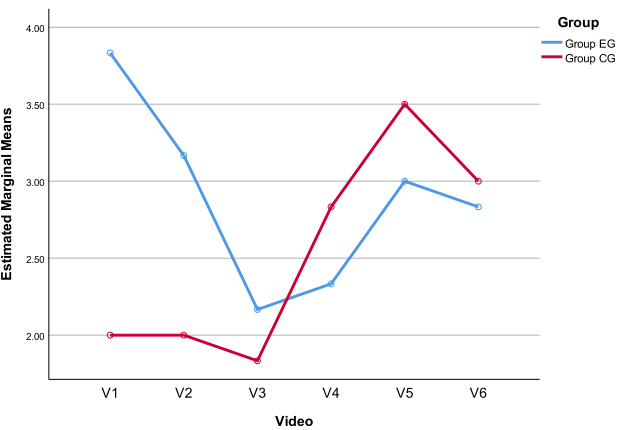**Figure 3: Mean values for the QoE questions**



**Figure 4: Interaction between the two factors (visual dominant and crossmodal olfactory content)**

[9] Grant Hanson-Vaux, Anne-Sylvie Crisinel, and Charles Spence. 2012. Smelling shapes: Crossmodal correspondences between odors and shapes. *Chemical senses* 38, 2 (2012), 161–166.

[10] Sarah E Kemp and Avery N Gilbert. 1997. Odor intensity and color lightness are correlated sensory dimensions. *The American journal of psychology* 110, 1 (1997), 35.

[11] M Luisa Dematte, Daniel Sanabria, and Charles Spence. 2006. Cross-Modal Associations Between Odors and Colors. *Chemical Senses* 31, 6 (2006), 531–538. DOI : http://dx.doi.org/10.1093/chemse/bjj057

[12] Lawrence E Marks. 1987. On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance* 13, 3 (1987), 384.

[13] David Merrill, Emily Sun, and Jeevan Kalanithi. 2012. Sifteo Cubes. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM, New York, NY, USA, 1015–1018. DOI : http://dx.doi.org/10.1145/2212776.2212374

[14] Gene Munster, Travis Jakel, Doug Clinton, and Erinn Murphy. 2015. Next mega tech theme is virtual reality. *gene* 612 (2015), 303–6452.

[15] Niall Murray, Oluwakemi A Ademoye, Gheorghita Ghinea, and Gabriel-Miro Muntean. 2017. A Tutorial for Olfaction-based Multisensorial Media Application Design and Evaluation. *ACM Computing Surveys (CSUR)* 50, 5 (2017), 67.

[16] Niall Murray, Brian Lee, Yuansong Qiao, and Gabriel-Miro Muntean. 2014. Multiple-Scent Enhanced Multimedia Synchronization. *ACM Trans. Multimedia Comput. Commun. Appl.* 11, 1s, Article 12 (Oct. 2014), 28 pages. DOI : http://dx.doi.org/10.1145/2637293

[17] Nobuyuki Sakai, Sumio Imada, Sachiko Saito, Tatsu Kobayakawa, and Yuichi Deguchi. 2005. The effect of visual images on perception of odors. *Chemical Senses* 30, suppl_1 (2005), i244–i245.

[18] J Simner and V Ludwig. 2009. What colour does that feel? Cross-modal correspondences from touch to colour. In *Third International Conference of Synaesthesia and Art, Granada, Spain, April*.

[19] Ray H Simpson, Marian Quinn, and David P Ausubel. 1956. Synesthesia in children: Association of colors with pure tone frequencies. *The Journal of genetic psychology* 89, 1 (1956), 95–103.

[20] Charles Spence. 2011. Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics* 73, 4 (2011), 971–995.

[21] Charles Spence, Mary Kim Ngo, Bronwen Percival, and Barry Smith. 2013. Crossmodal correspondences: Assessing shape symbolism for cheese. *Food*

*Quality and Preference* 28, 1 (2013), 206 – 212. DOI:http://dx.doi.org/https://doi.org/10.1016/j.foodqual.2012.08.002

[22] Naomi L Streeter and Theresa L White. 2011. Incongruent contextual information intrudes on short-term olfactory memory. *Chemosensory Perception* 4, 1-2 (2011), 1–8.

[23] Xiuwen Sun, Xiaoling Li, Lingyu Ji, Feng Han, Huifen Wang, Yang Liu, Yao Chen, Zhiyuan Lou, and Zhuoyun Li. 2018. An extended research of crossmodal correspondence between color and sound in psychology and cognitive ergonomics. *PeerJ* 6 (March 2018), e4443. DOI:http://dx.doi.org/10.7717/peerj.4443

[24] Benjamin Tag, Takuya Goto, Kouta Minamizawa, Ryan Mannschreck, Haruna Fushimi, and Kai Kunze. 2017. atmoSphere: mindfulness over haptic-audio cross modal correspondence. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM, 289–292.

[25] Augoustinos Tsiros. 2017. The Parallels Between the Study of Cross-modal Correspondence and the Design of Cross-sensory Mappings. In *Proceedings of the Conference on Electronic Visualisation and the Arts (EVA '17)*. BCS Learning & Development Ltd., Swindon, UK, 175–182. DOI:http://dx.doi.org/10.14236/ewic/EVA2017.39

[26] Shafiq ur Réhman, Muhammad Sikandar Lal Khan, Liu Li, and Haibo Li. 2014. Vibrotactile TV for immersive experience. In *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*. IEEE, 1–4.

[27] Shota Yamada, Takashi Kakue, Tomoyoshi Shimobaba, and Tomoyoshi Ito. 2018. Interactive holographic display based on finger gestures. *Scientific reports* 8, 1 (2018), 2010.

[28] Jeffrey M Yau, Jonathon B Olenczak, John F Dammann, and Sliman J Bensmaia. 2009. Temporal frequency channels are linked across audition and touch. *Current Biology* 19, 7 (2009), 561–566.