SOUND DESIGN THROUGH LARGE AUDIENCE INTERACTION

Kjetil Falkenberg Hansen
KTH Royal Institute of Technology,
Stockholm, Sweden

Martin Ljungdahl-Ericsson
Edsbyn/University West,
Sweden

Ricardo Atienza
University College of Arts, Crafts
and Design, Stockholm, Sweden

ABSTRACT

In collaboration with Volvo Cars, we presented a novel design tool to a large public of approximately three million people at the three leading motor shows in 2017 in Geneva, Shanghai and New York. The purpose of the tool was to explore the relevance of interactive audio-visual strategies for supporting the development of sound environments in future silent cars, i.e., a customised sonic identity that would alter the sonic ambience for the driver and by-passers. This new tool should be able to efficiently collect non-experts' sonic preferences for different given contexts. The design process should allow for a high-level control of complex synthesised sounds. The audience interacted individually using a single-touch selection of colour from five palettes and applying it by pointing to areas in a colour-book painting showing a road scene. Each palette corresponded to a sound, and the colour nuance in the palette corresponded to certain tweaking of the sound. In effect, the user selected and altered each sound, added it to the composition, and finally would hear a mix of layered sounds based on the colouring of the scene. The installation involved large touch screens with high quality headphones. In the study presented here, we examine differences in sound preferences between two audiences and a control group, and evaluate the feasibility of the tool based on the sound designs that emerged.

1. INTRODUCTION

There is a growing interest in the field of sound design in the car industry; not just regarding branding and the traditional design of sounding objects (doors, motors, etc.) [1] and modern car-audio systems [2], but also the conception of outdoor/indoor sonic atmospheres linked to the emergence of silent electric cars [3]. Once the sonic trace of the combustion motor is removed, how to sonically signal the presence of the car for neighbouring pedestrians or bikers and how to conceive the new indoor sonic ambiences of the vehicle as a main qualitative component of the travel experience is a matter of regulations [4, 5], safety [6], as well as aesthetics [7].

Regarding the external sonic presence of the car, it constitutes first a key safety issue, in particular in urban environments. This sonic print can equally give support to different forms of encounter between the traversed environments and the vehicles/passengers, including informational but also masking as well as aesthetic components. A customised sound identity could operate here as a relevant mediator between car and place, and its design should thus take these multiple components into account.

This "additive" approach to everyday sonic environments (i.e., not only through traditional subtractive acoustic methods such as insulation, absorption or noise cancelling) is an expanding field of research that has been explored in the last years in different research case studies and practice-based interventions, focusing in particular on public space in dense urban contexts (e.g. [8–10]). The authors have equally explored this question of additive sound design in the context of different mobility modes e.g. in the research project ISHT (*Interior Sound Design of High Speed Trains* [11, 12]) in collaboration with the train manufacturer Bombardier. The authors have also been involved with sound design of complex shared working spaces such as flex- and activity-based offices [13]. In both examples the main focus was on how to actively improve the experience of place and situation through subtle additive sonic interventions.

Volvo Cars are focusing today on electric cars and particularly interested in the new sonic needs and possibilities brought by this more silent mobility. The world's major motor shows—the Geneva International Motor Show, New York International Auto Show, and Auto Shanghai—were regarded as a relevant opportunity for a first exploration of the users' requirements, preferences, potential desires and customisation skills. To that aim, we were asked to design the interactive environment to be used in these three locations with a visitor count of almost three million people. A general frame was provided by Volvo: *to find an intuitive and efficient way for visitors to design, or better to sketch their own sound atmosphere*; manipulating sound can certainly be regarded as an abstract and complex activity for common visitors.

As a first step towards this ambitious aim, we wanted to study here the potential efficiency of an audio-visual interactive environment for the collection of users' preferences and exploration of designing skills. However, we do not aspire to propose a finished design tool. As such, the study presented deals more with socio- or ethno-cultural differences in listening and interaction than a sound design that can readily be applied by the industry.

A simple screen-based interactive visual/graphic environment was chosen to provide support to this fast and individual prototyping process. Simplicity and robustness are central attributes in successful large-scale installations, and large touchscreens have been found to engage and captivate people [14]. Despite this simple frame, this parallel visual/sonic interaction mode represents also in a more appropriate way the multi-sensory nature of a driving experience; a purely sonic exercise would be in this sense an excessively reduced model. There are more nuanced discussions about site-specific challenges for installations [15], recommendations for design such as ease-of-use [16], fun [17], and for intuitive interaction and matching people's expectations [18].

In the last years' growing body of work on how to engage with sound and music through technology, most notably in sonic interaction design [19], sonification [20], and in new interfaces for musical expression [21], we could not find any closely related studies on designing sounds through a typical drawing task nor guidelines for such. Sketching is however a general, emerging research field within the CHI community (e.g., [22]). We are also well aware that the coupling between sound and colour or shapes have been widely investigated; in for example [23], children were drawing trajectories to fit a musical stimuli on paper, and they found support for cross-modal mapping of drawing and sound. Sonification of colours has also been found useful for visually impaired and in other applications [24], to name only one example. Our designed is also loosely inspired by a DJ system developed for creating music based on coloured discs [25]; a camera traced the spinning platter and controlled a synthesizer with the registered colours passing a tangent.

Adjacent research in synesthesia deals with cross-coupling of sensory inputs. For instance, the connection between brightness in colour and musical timbre has been shown [26], and it has been found that going from music to visual stimuli is most common [27]. Both synesthesia and related questions concerning deviation and cultural differences in colour perception [28] are somewhat relevant for this study but intentionally excluded because the method design does not really allow to take these into account.

A soundscape is according to [29] the auditory counterpart to a visual landscape. Soundscapes can exist as perceptual constructs, but also as physical phenomena [30]. When perceived in a shared context, the constituents of the acoustic and the visual interrelate [31].

By essentially using a number of shared composition materials, we want to explore how different set contexts influence the choices realised by participants. Is it possible to trace case-specific preferences? Along this process, the user could only manipulate sound by interacting with a colour book interface. All interaction with the system and the resulting compositions were logged in the form of text files and recordings collecting the series of actions executed by visitors on the tactile screen.

2. METHOD

The interactive system we built, called *Volvo Sound Studio*, was based on an audio-visual environment including touch screens, high-end headphones, and software built entirely in Pure data and the graphic library GEM (for details see Table 1). While the hardware was decided by or in collaboration with Volvo and Bower & Wilkins (Volvo's partners in the design of their car-audio systems and main providers of audio components) the programming environment and the interaction design were the responsibility of the authors. The system has previously been described in brief [32], but then without results from the interaction.

Hardware/ software	**Description**	**Link**
Windows	Surface Studio PixelSense™ 28" touch display, Windows 10	↗
Bowers & Wilkins	P9 Signature over-ear headphones	↗
Audio-Quest	DragonFly Red 32 bit ESS 9016 DAC and headphone amplifier	↗
Pure Data *0.47-1*	Real-time programming environment for audio	↗
GEM *0.93.3*	Graphics environment for multimedia	↗

Table 1. Overview of the technical set-up and system components.

2.1 Interface and Interaction Design

During several brainstorming and development sessions together with representatives from the partner companies, the concept of having a simplistic colour-book interface was decided. A strong argument was to create an experience that did not resemble sound design or mixing tasks, and something that would appear novel, which is of utmost importance for a large-scale public event like this. Another key characteristic of this concept was the accessible and intuitive nature of the interaction to happen, requiring no previous experience in editing or designing sound, no specific introduction to the user and very limited time before reaching a meaningful result. The main focus was thus placed on the driving contexts explored (graphically presented on screen and sonically evoked through a soundscape background track) and their specific impact on the sound choices operated by visitors.

The concept of the colour-book, with its deep abstraction of a represented environment and its simple interactive mode, was intended to allow the user not to focus exclusively on the visual information, but also on the sound textures under development, i.e., in search of a balance between both senses involved. A highly demanding visual experience could more easily become a mono-sensorial exercise from an attention point of view. In the same sense, a broader palette of graphic possibilities and realism would equally require a careful analysis of the correlations proposed between colour and sound; this complex aspect was avoided as this early study focused on designing potentials more than final sketching tools.

Participants were invited to "paint" three different land-

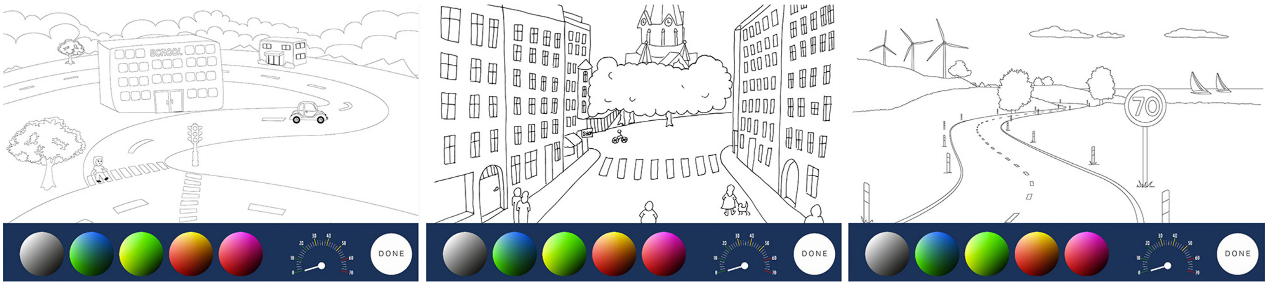

Figure 1. The three scenes presented to the audience in colour-book style: from left the school area, city centre and countryside road scenes. Colour is chosen in the palettes and then applied to an empty area. Each palette has a sound assigned, and the sounds from the painted areas are mixed. The speedometer changes a global filter on the mixed sound.

scapes represented by three stylised scenes (see Fig. 1), namely a ***school area scene*** with a school building surrounded by curvy small roads and a low density area, a ***city centre scene*** with a dense urban environment, and a ***countryside scene*** depicting a curvy road in an open semi-natural landscape. None of the scenes included presence of other cars, neither traffic or congestion, as we wanted users' attention to be only on their "own" car. In Fig. 1, the car in the school scene can however indeed be interpreted to be another sounding car.

The graphic language employed is intentionally simple, based on the model of the colour-book, as well as the colouring mode in action: each tone is applied at once on entire predefined areas of each drawing, clearly delimited by black lines on a white canvas. In the version shown to the public, the user could choose from five colour palettes (gradations of grey, green–blue, green–yellow, red–yellow and red–magenta) and apply the chosen tone to seven different areas in the drawing. The selection of the colour tone was providing via headphones an immediate sonic feedback varying as the user was exploring the different possible colour nuances; different filters and sound effects followed, in sonic terms, the movement of the arrow on the colour selection area (see Fig. 2).

Each colour palette was linked to a different sound material, the same per palette for the three scenes explored (see Table 2). While selecting the colour, i.e., working on a colour palette, only the corresponding sound can be heard, and when finally applied to the main canvas, the user gets the entire sonic mix or composition, including a background atmosphere specifically designed for each individual scene configured by a number of typical sonic components of the depicted landscape (activities, textures, soundmarks, etc.). This background atmosphere or soundscape aims at providing a coherent and easily understandable sound supporting layer for a more flexible and free exploration of the new materials to insert which essentially present no specific relation to the different contexts. The exception is three specific contextual sounds presenting, *a priori*, an intentional relation to each context; these three sounds will be the object of particular attention along the analysis. At any time, the visitor can interact with a speedometer, described below.

When no interaction was provided, the screen automatically displayed an information video silently demonstrating how to use it,[1] and inviting passers-by to test the environment. For sound samples of both countryside scenes represented in Fig. 2, including the specific background soundscape (final compositions, produced by participants), see Table 3.

2.2 Sound Design and Synthesis

A total of five sounds were designed. Each was five seconds long and could be seamlessly looped at any given point. The composing of the sound material had to cater two design specifications. Each sound should easily and harmoniously have the ability to be intertwined with any other of the sounds. The sounds also needed to possess the right balance between being interesting without generating irritation during prolonged listening. Three of the sounds would also be contextually linked to the scene through similarity to the matching soundscape, one for each scene, and two would be contextually deviant. The divergent sounds were composed with the intention to be associated with a real and a fictional vehicle: one sound was an actual Volvo engine and the other a paraphrasing of how UFOs often are depicted in sci-fi movies. This was done to find out if the participants would connect the contextually connected sounds to the scene and if an actual car engine were preferred.

The sounds, except the engine sound, were created with an energy concentration around 1000 Hz, a frequency range not too crowded in the exhibition areas. The sounds were also designed through rhythm and pitch changes, making them fluctuate in order to be distinguished from other constituents emitting in the same frequency range of the soundscape. The intention of the dynamic character of the sounds was to be more cohesive with the dynamics of the ever-changing soundscape. The sounds would fit the surrounding rhythmic events and at the same time convey movement and grab attention, and avoid being static sounds that add to the denseness of the urban sonic environment.

The sound design had the aim of letting the user design its preferred car sound based on the location generating the driving situation, in this case the sonic and visual stimuli presented at the scene. E.g., the harmonic sound was designed to be connected to the city center scene, having a dynamic movement and a melodic nature to be discerned

[1] Video: https://kth.box.com/v/smc2019-HLA-video

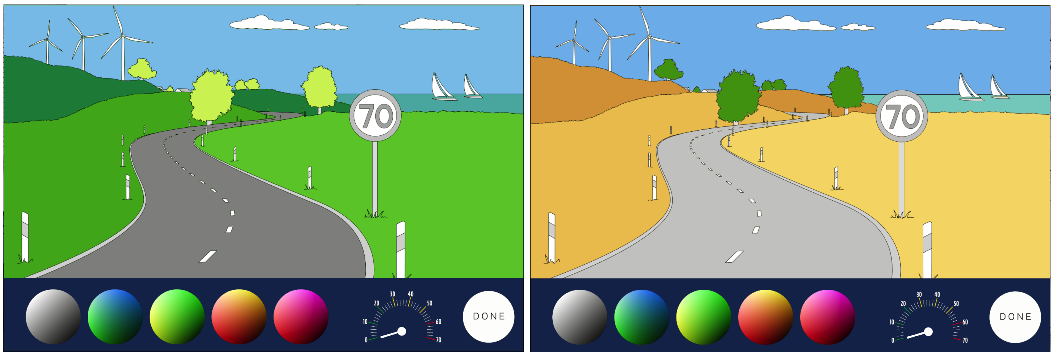

Figure 2. Examples of painted countryside scenes with different interpretation of the vegetation. The left image would be dominated by sounds from the green palette, namely the "rolling" sound, while the right image would be dominated by sounds from the orange palette, namely the "motor" sound.

in that context. The texture of a church bell used was softened to get a floating and non intrusive sound, that for the listener would alter between staying in the background and foreground.

The sounds described above were not played as is, but instead treated in a synthesis based on looping. Each instance of choosing colour tone in a palette and applying to an empty space would trigger the synthesizer. In order for the result to appear dynamic (like a motor sound), the sound was replayed with three very short loops (fractions of a second) of random start and length within the sound file from its start, middle and end. Onto each looped part, we added an audio effect controlled by the colour nuance (see Table 2). In total there could thus be 21 loops going simultaneously, creating a very dense but still dynamic composition.

In order to provide a coherent visual/aural experience, once a colour is applied onto a scene, it is placed in the stereophonic space according to its relative position in the drawing. Each area in a scene will thus correspond to a different left/right panning value. This effect will be particularly audible in the binaural space supported by headphones.

The speedometer aims at evoking the familiar glissando sound effect linked to variations of speed as the motor accelerates or decelerates. To that extent, the speedometer controls a non-linear pitch-shifting function and loop-speed effect where pitch and loop speed do not covary. Each context will present a speed limit, according to normal traffic regulations (urban, country road, etc.).

2.3 Data Collection

Data were collected from two of the three expositions we exhibited the Sound Studio in; Geneva and Shanghai, as well as from a supplementary laboratory experiment. [2] Due to reasons beyond our control, New York data became inaccessible for this study. We chose two dates for inclusion: both Sundays after the first open weekend (March 3 and April 4 2017, respectively). These days were less busy and we expected the installation to be stable (it turned out they ran without trouble for the entire duration).

[2] Log data: `https://kth.box.com/v/smc2019-HLA-data`

There are many parameters from such data collection that we have little or no control over, and several are important. The following list is inconclusive, but can serve as a careful suggestion for reading the results with some attention:

The user: We are confronted in this study to entirely anonymous participants: no personal information regarding the user and its profile was registered. That includes nationality, gender, age, physical characteristics, background, etc.

The intention: The users were not debriefed in order to understand their interaction. This means we know nothing about if there was preference for using colour instead of sound, the users seriousness, if they were happy with the results, or if they changed all parameters from their desired soundscape in the last moment before submitting. [3]

The interaction: Nothing surrounding the session apart from the time of day was registered. That includes if it was a returning user, if headphones were used correctly or even used at all, if only one user interacted, if there were disturbances elsewhere, etc.

However, there were always exhibition hosts attending the two screens and giving support to the guests, so we can reasonably argue in favour the exploitable nature of the data collected. In order to verify this hypothesis, we conducted—in addition to the exhibition setting—a laboratory experiment with twelve first year media technology students. They interacted without disturbance for as long as they wished. The group was balanced with regards to gender (6f/6m), but unbalanced with respect to age (between 19–25 years), nationality (Swedish), and all being technology students. The purpose of having a control group was to understand whether stress and contextual factors (time, sound environment, multiple surrounding presences and actions) would critically influence the exhibition audiences.

3. RESULTS

For the two chosen dates from the data collection period, we have 312 and 431 sessions from Geneva and Shanghai, respectively, or one session every two minutes. A session was defined as one completed soundscape design from selecting a scenery to clicking the 'Done' button. The av-

[3] Such behaviour would however be identified in the log data.

Sound	Description and effect	Context	Palette	Link
“Harmonic”	Resonating and reverberating church bell texture fragments and a minor chord. Modified with pitch-shifted delay.	City centre	Grey–black	↗
“Rolling”	Filtering and delaying a minor 7^{th} chord using a recording of waves to trigger the delay channel for an organic rhythm. Modified with flanger.	Countryside	Green–blue	↗
“School bell”	A recording of wind chimes modulated to tremolate. Modified with a time-varied delay line [33].	School area	Green–yellow	↗
“Motor”	A recording of a Volvo V60 internal combustion engine high-pass filtered and moderately distorted to accentuate the engine sound. Modified with a comb-filter octaver.		Red–yellow	↗
“Sci-fi”	Oscillating a single tone which then is reverberated, and further flanged and resonated. Modified with voltage-controlled bandpass filter sweeps.		Red–magenta	↗
School soundscape	A mix of field recordings from a playground, a quiet urban environment, and a parking lot outside a shopping mall.			↗
City soundscape	A mix of field recordings of two cities recorded at different locations early in the morning, a recording from a busy pedestrian street, and a church bell.			↗
Countryside soundscape	A mix of field recordings of breaking waves, seagulls, cars driving at a distance, and soft wind on a field.			↗

Table 2. Sounds used in the interface with arbitrary names. The effect was added to the sound corresponding to the colour nuance. The links go to sound examples in the Soundcloud repository.

erage time for the interacting visitor was 81 seconds, so with 2–4 computers in the exhibition space, they were almost constantly in use. The longest session lasted 430 seconds, and all sessions shorter than 15 seconds were removed from analysis before the estimation made above. The control group generated 35 sessions of 36 planned.

There were significant differences ($p \ll 0.001$, two-tail t-test) in average time spent on interaction, where the Shanghai audience spend 19% more time, or around 14 seconds longer per completed session. Also, this audience painted in average two more areas in a session than the Geneva audience (the areas were painted 17.3 and 14.9 times, respectively, $p = 0.017$). In average, the Shanghai and Geneva groups adjusted and applied one new sound every five seconds. The control group had longer interaction time (44% longer, $p = 0.006$), but did not paint more areas; in effect, they listened more than two seconds longer to each sound before applying it.

The most prominent difference in interaction behaviour is connected to exploring the sound palette. While we cannot directly measure how the users listen for changes in sound, the logging function produces one line for each incremental value from dragging the finger within the palette. The exhibition audiences have similar values—Geneva almost twice as many as Shanghai—while the line count of the control group is more than six times higher than that.

In each scene, three of the five sounds (corresponding to the five colour palettes described above and shown in Table 2) were used similarly by the Geneva and Shanghai audience; the harmonic, the rolling, and the sci-fi sounds (see Fig. 3). The school bell sound was favoured only by the Geneva audience, and only in the School scene (by 89%, $p \ll 0.001$). We found no difference between Shanghai and the control group. The motor sound was favoured by the Shanghai audience, but only in the City scene (by 59%, $p \ll 0.001$). This preference was even stronger in the case of the control group (82% more, $p \ll 0.001$).

The three contextual sounds (the school bell, the harmonic, and the rolling sound) each had an intended connection to the scenes (the school, the countryside, and the city, respectively). For the school bell sound/school scene, the Geneva audience were, as shown above, significantly more likely to choose the contextual sound (almost twice as much). For the other scenes, there were no differences.

The audience in Geneva seemed to consistently choose a higher speed on the speedometer than those in Shanghai (in average 13.4% and 8.9% faster than “neutral” speed, respectively). However, the significance test did not confirm this observed tendency.

By shifting the colour nuance on the palette, each sound could be ‘tweaked’ in the sense that the sound was manipulated with an increasingly noticeable sound effect. Colour nuance was mapped to position in the circle, which showed a colour grading, while distance from the centre of the palette corresponded to sound effect strength regardless of absolute position in the circle. There was no significant difference between the audiences nor the control group in sound effect strength. In average, they tweaked the sounds to 0.56 distance from the centre, or just above halfway out towards the edge.

We did not find any difference in how varied the sound designs were in terms of how many palettes that were used

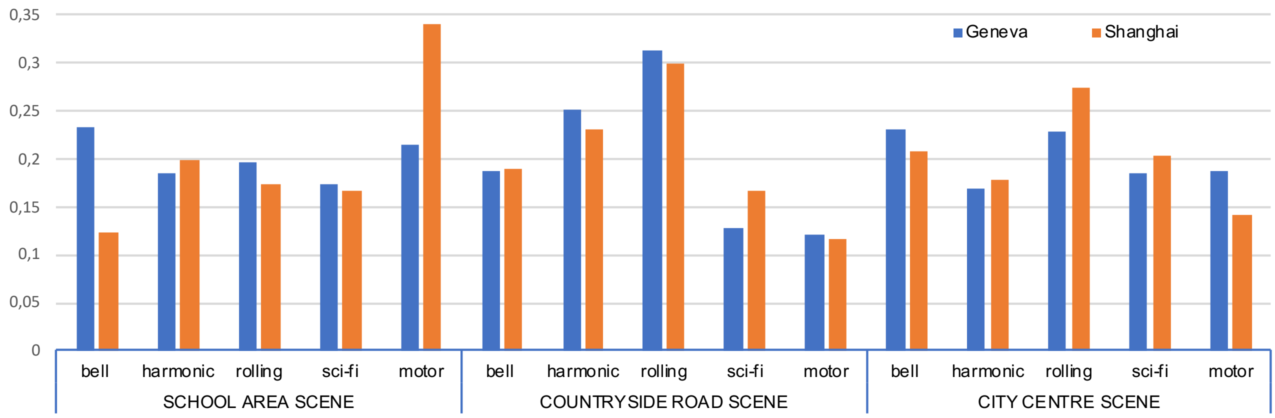

Figure 3. Proportion of sounds used by Geneva and Shanghai visitors for each of the three scenes.

for each session. In average, 2.9 of the five palettes were used for one painting. However, the control group showed a larger variation with 3.5 palettes used in average ($p = 0.005$). Even for this measurement there is ambiguity concerning whether the user focused on sound or colour.

4. DISCUSSION

From the mentioned uncertainties concerning data collection, we cannot *know* if the users of our system among the audiences at either Geneva or Shanghai are representative for those geographically distant places. However, judging from the partner companies present, exhibition hosts, and popular media reports, we find it is reasonable to assume that visitors are representative for the general population of those places. In the following, we generalise the results accordingly. Furthermore, the statistical analysis must be read with caution as the experiment lacked in control. Thus, only a few distinct findings were reported above and included here to serve as a basis of discussion and future work.

Among the measurements extracted from the log files, there are far more similarities than differences between Geneva and Shanghai. They adjust the speed similarly, they tweak the sound to the same extent, and they use a comparable varied selection of palettes for a given scene. They do however differ in the amount of effort and time that is put into the interaction.

Five designed “base sounds” were assigned to the palettes, of which three were specifically suited for each of the three scenes. For some reason, the school bell sound was strongly associated with the school scene by the Geneva audience, but not the Shanghai. The other sound that the audiences applied differently was the motor sound. Generally, the Shanghai audience chose this for the school scene when the Geneva chose the bell sound.

One possible explanation is that the sound design was done by a European who has more experience with school bell sounds in Europe than in Asia. Another explanation is that a bell sound is not common in traffic situations in Shanghai, while low-frequency motor sounds are. Furthermore, although unlikely, we cannot overlook the possibility that users only chose a colour to match the scene which is different between Geneva and Shanghai.

The control group was in many respects more similar to the Shanghai audience, which contradicts the above-mentioned regional differences in how one experiences the school bell and motor sounds. Also, the control group favoured the rolling sound. An alternative explanation to these found differences could be that in the Geneva exhibition there were surrounding sounds which interfered with the low-frequency motor and rolling sounds.

A further perspective on choice of sounds is the sound feedback from the act of choosing a colour. For instance, to paint the sky blue, the user needs to find the colour in the second palette, or the “rolling” sound. Furthermore, to get to the blue, the audio effect will be strong (for sound examples demonstrating this, see Table 3). As a result, the user needs to compromise between image naturalness or intention and how it sounds. Similarly, the easiest way to paint a grey or black road is to use the first palette with the “harmonic” sound. However, none of these speculations can really explain why the Geneva public favoured the school bell sound (greens) for the school area.

Overall, the Shanghai audience spent more time and repainted more areas, despite that the number of completed sessions was higher. Without having time constraints or other distractions, the control group devoted considerably longer for finishing a session than the exhibition audiences, but without painting more areas. Also, the control group listened or searched far more attentively for the right nuance in the sound/colour palette. Regardless of this, it seems that most users could finish a sound design within less than 90 seconds. Considering the substantial variation in sound that was possible with the relatively simple tool, we are intrigued by this finding.

5. CONCLUSIONS

The installation *Volvo Sound Studio* globally fulfilled the expectations in terms of reliability as well as efficiency of planned interaction. Visitors to the major car exhibitions in Geneva and Shanghai (as well as New York, not covered in this paper) could complete rather complex sound designs within a period of time as brief as a minute. It was efficient also in regard to the proved intuitive nature of the interaction; granting universal access to visitors regardless of their skills or age. The graphic codes chosen and disposition of elements on the screen naturally guided the users in their explorations according to the observations realised. The stability of the system proved to be entirely reliable, which is not obvious when it comes to interactive systems running non-stop along several days and with no particularly qualified local maintenance.

We found that audiences at the Geneva and Shanghai motor shows had different preferences for the car's soundscape outside a school and in a city centre. The Geneva users favoured a school-bell sound which was specifically designed to match that scene, and for the city scene, the Shanghai users favoured a sound based on a combustion engine recording.

In future studies we would like to explore a number of different choices regarding aspects such as alternative programming environments (GEM could readily be replaced by better performing graphic options such as Processing), data collected on participants (details on gender, age, would have been relevant criteria to explore, even perhaps the possibility of short surveys as initially suggested to our industrial partner) or even complementary methods focusing only the aural dimension to be compared with the results obtained here (e.g. employing the same sound materials during the design process).

With regards to carrying out research within the confines of a prominent publicity effort made at the most important venue for one of the leading car producers, there are naturally some compromises that had to be made. As a first general reflection, the design of the interactive modes and data collection protocols would have certainly been conceived differently within a purely research frame; efficiency was here a key concept, with a suggested time limit of one minute per user. Without doubt, such compromises are typically in conflict with research needs and requirements.

On a similar key, the installation we ended up building is probably not the ultimate for finding answers to how people design sounds: the colour-book concept could have been exchanged for other concepts. The design process was limited to the effort of the authors and with influence from a few involved persons at Volvo. Also, it would be impractical or even impossible to gather personal data from the visitors. On the other hand, there are very few opportunities of reaching out to such a number of people in another context.

Links to Sound Examples

All sound examples have been uploaded to the Soundcloud repository: https://soundcloud.com/kjetil-falkenberg-hansen/sets/sound-examples-smc2019; see Table 3.

Acknowledgments

The authors would like to thank Annika Hedin and Fredrik Hagman at Volvo Cars and Marie Hernmarck, Susanne Park and Robert Tenggren at SCP Grey for the productive conversations and helpful input. The project was sponsored by Volvo Cars.

Description	Linkname
The harmonic sound	harmonic-sound
The rolling sound	rolling-sound
The school bell sound	schoolbell-sound
The motor sound	motor-sound
The sci-fi sound	scifi-cound
School area soundscape	soundscape-school
City centre soundscape	soundscape-city
Countryside soundscape	soundscape-country
Finished scene Figure 2a	scene2a
Finished scene Figure 2b	scene2b
Speedometer demonstration from slow to fast to slow	speedometer
Audio effect demonstration on five sounds from minimum to maximum	audio-effects

Table 3. List of sound examples uploaded to the Soundcloud repository. All links have the format https://soundcloud.com/kjetil-falkenberg-hansen/*linkname*

6. REFERENCES

[1] E. Cleophas and K. Bijsterveld, *Selling Sound: Testing, Designing, and Marketing Sound in the European Car Industry*. Oxford University Press, dec 2011.

[2] R. Coppola and M. Morisio, "Connected car," *ACM Computing Surveys*, vol. 49, no. 3, pp. 1–36, oct 2016.

[3] D. S. Kim, R. W. Emerson, K. Naghshineh, and K. Myers, "Influence of ambient sound fluctuations on the crossing decisions of pedestrians who are visually impaired: implications for setting a minimum sound level for quiet vehicles," *Journal of Visual Impairment & Blindness (Online)*, vol. 108, no. 5, p. 368, 2014.

[4] National Highway Traffic Safety Administration (NHTSA), *Federal Motor Vehicle Safety Standards; Minimum Sound Requirements for Hybrid and Electric Vehicles*. US Government Federal Register, 2016, document Citation: 81 FR 90416. [Online]. Available: https://www.federalregister.gov/d/2016-28804

[5] Inland Transport Committee, *Proposal for a new Regulation concerning the approval of quiet road transport vehicles (QRTV)*. United Nations, 2016, eCE/TRANS/WP.29/2016/26. [Online]. Available: https://www.unece.org/fileadmin/DAM/trans/doc/2016/wp29/ECE-TRANS-WP29-2016-026e.pdf

[6] A. Zeitler, "Psychoacoustic requirements for warning sounds of quiet cars," *SAE International Journal of Passenger Cars - Electronic and Electrical Systems*, vol. 5, no. 2, pp. 572–578, jun 2012.

[7] N. Misdariis, A. Cera, E. Levallois, and C. Locqueteau, "Do electric cars have to make noise? An emblematic opportunity for designing sounds and soundscapes," in *Acoustics 2012*, S. F. d'Acoustique, Ed., Nantes, France, Apr. 2012. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00810920

[8] B. Hellström, "Noise design : Architectural modelling and the aesthetics of urban acoustic spase," Ph.D. dissertation, KTH, School of Architecture, 2003, nR 20140805.

[9] M. E. Nilsson, J. Alvarsson, M. Rådsten-Ekman, and K. Bolin, "Auditory masking of wanted and unwanted sounds in a city park," *Noise Control Engineering Journal*, vol. 58, no. 5, p. 524, 2010.

[10] B. Hellström, M. E. Nilsson, Ö. Axelsson, and P. Lundén, "Acoustic design artifacts and methods for urban soundscapes : A case study on the qualitative dimensions of sounds," *Journal of Architectural and Planning Research*, vol. 31, no. 1, pp. 57–71, 2014, qC 20140625.

[11] R. Atienza and N. Billström, "Fighting "noise"=adding "noise"? – Active improvement of high-speed train sonic ambiances," in *Ambiances in action/Ambiances en acte (s)-International Congress on Ambiances, Montreal 2012*. International Ambiances Network, 2012, pp. 235–240.

[12] K. F. Hansen and R. Bresin, "Use of soundscapes for providing information about distance left in train journeys," in *Proceedings of the 9th Sound and Music Computing Conference*, Copenhagen, jul 2012, pp. 79–84. [Online]. Available: http://www.speech.kth.se/prod/publications/files/3764.pdf

[13] M. L. Eriksson, R. Atienza, and L. Pareto, "The Sound Bubble: A context-sensitive space in the space," *Organised Sound*, vol. 22, no. 1, pp. 130–139, mar 2017.

[14] C. Ardito, P. Buono, M. F. Costabile, and G. Desolda, "Interaction with large displays," *ACM Computing Surveys*, vol. 47, no. 3, pp. 1–38, feb 2015.

[15] K. Beilharz and A. Martin, "The 'interface' in site-specific sound installation," G. Essl, B. Gillespie, M. Gurevich, and S. O'Modhrain, Eds. Ann Arbor, Michigan: University of Michigan, May 21-23 2012.

[16] J. Hochenbaum, O. Vallis, D. Diakopoulos, J. Murphy, and A. Kapur, "Designing expressive musical interfaces for tabletop surfaces," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Sydney, Australia, 2010, pp. 315–318. [Online]. Available: http://www.nime.org/proceedings/2010/nime2010_315.pdf

[17] B. Shneiderman, "Designing for fun," *interactions*, vol. 11, no. 5, p. 48, sep 2004.

[18] L. Hespanhol and M. Tomitsch, "Strategies for intuitive interaction in public urban spaces," *Interacting with Computers*, vol. 27, no. 3, pp. 311–326, jan 2015.

[19] K. Franinović and S. Serafin, Eds., *Sonic Interaction Design*. MIT Press Ltd, 2013.

[20] S. Pauletto, H. Cambridge, and P. Susini, "Data sonification and sound design in interactive systems," *International Journal of Human-Computer Studies*, vol. 85, pp. 1–3, jan 2016.

[21] R. Medeiros, F. Calegario, G. Cabral, and G. Ramalho, "Challenges in designing new interfaces for musical expression," in *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*. Springer International Publishing, 2014, pp. 643–652.

[22] M. Sturdee, M. Lewis, and N. Marquardt, "Feeling SketCHI?" *Interactions*, vol. 25, no. 6, pp. 64–69, oct 2018.

[23] E. Frid, R. Bresin, P. Alborno, and L. Elblaus, "Interactive sonification of spontaneous movement of children – cross-modal mapping and the perception of body movement qualities through sound," *Frontiers in Neuroscience*, vol. 10, nov 2016.

[24] S. Cavaco, J. T. Henriques, M. Mengucci, N. Correia, and F. Medeiros, "Color sonification for the visually impaired," *Procedia Technology*, vol. 9, pp. 1048–1057, 2013.

[25] N. Villar, H. Gellersen, M. Jervis, and A. Lang, "The ColorDex DJ system: a new interface for live music mixing," in *Proc. of the Conference on New interfaces for Musical Expression*. New York: ACM, 2007, pp. 264–269.

[26] L. E. Marks, "Synesthesia and the arts," in *Cognitive Processes in the Perception of Art*. Elsevier, 1984, pp. 427–447.

[27] G. Berman, "Synesthesia and the arts," *Leonardo*, vol. 32, no. 1, pp. 15–22, feb 1999.

[28] R. Pettersson, "Cultural differences in the perception of image and color in pictures," *ECTJ*, vol. 30, no. 1, pp. 43–53, 1982.

[29] M. Raimbault and D. Dubois, "Urban soundscapes: Experiences and knowledge," *Cities*, vol. 22, no. 5, pp. 339–350, oct 2005.

[30] J. Appleton, *The experience of landscape*. New York: J. Wiley & Sons, 1996.

[31] J. L. Carles, I. L. Barrio, and J. V. de Lucio, "Sound influence on landscape values," *Landscape and Urban Planning*, vol. 43, no. 4, pp. 191–200, jan 1999.

[32] K. F. Hansen, M. L. Eriksson, and R. Atienza, "Large-scale interaction with a sound installation as a design tool," in *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences - AM '17*. ACM Press, 2017.

[33] P. Massat, "GuitarExtended: DownTheDrain," Pure data effects patch, 2012.