# DESIGN CONSIDERATIONS FOR SHORT ALERTS AND NOTIFICATION SOUNDS IN A RETAIL ENVIRONMENT

**Gustav F. ARFVIDSSON** (garf@kth.se) [1], **Martin L. ERIKSSON** (martin@soundmark.se) (0000-0002-7951-2089) [2], **Håkan LIDBO** (hakan@hakanlidbo.com) [3], and **Kjetil FALKENBERG** (kjetil@kth.se) (0000-0003-4259-484X) [1]

[1] *School of Electrical Engineering and Computer Science*, **KTH Royal Institute of Technology**, Stockholm, Sweden
[2] *Media and Design*, **University West**, Trollhättan, Sweden
[3] **Rumtiden**, Stockholm, Sweden

## ABSTRACT

The design and noticeability of alert sounds have been widely researched and reported, and not least, notification sounds are ubiquitous in both software and hardware product development. In an ongoing research project concerning the retail industry, we aim at designing short alert sounds that only grab attention from one group of customers, while others do not register the alerts: this particular aspect has to our knowledge not yet been studied. To establish design guidelines for such alert sounds, we conducted an experiment where test subjects would experience ordinary shopping activity including background music and an ambient soundscape in a virtual reality clothing store, but with added alert sounds. We tested, specifically, six differently designed sound alerts belonging to two classes: contextual-specific congruent sounds, and incongruent sounds that did not fit the sonic context. The results disproved our assumptions that incongruent sounds would outperform the congruent and thus in the context more anticipated sounds. The findings suggest that alert sounds can be designed with subtlety and still be noticeable and that customers will not necessarily be annoyed. We present here a first approach towards design guidelines for short alert sounds in a shop environment.

## 1. INTRODUCTION

Notification sounds are omnipresent in our lives: we get exposed to alert sounds for communication, interaction, status, and safety situations, and we get these from our mobile devices and computers, vehicles, household machines, and public buildings. Furthermore, the alert sound designs span from clicks and beeps to human voices and music excerpts.

In an ongoing research project "Sonification of store goods" involving retail, theft, and shopping, we need to reconsider attentiveness towards notification sounds: we aim at designing short alert sounds for non-critical contexts that only grab attention from targeted actors in an environment. The main aim for our project is to discourage and prevent shoplifting by playing alerts as sonifications of interactions with goods in shops without discomforting regular customers or distracting employees. Ideally, the sonifications should not attract attention from others than store clerks and shoplifters, and the sounds should not reduce the overall shopping experience.

Using sonification for monitoring state can free up cognitive resources [1], cutting back costs on expensive video surveillance systems, and open for live monitoring where information on what goes on in the store can be conveyed to the staff in real time. It also solves the ethical question of storing customer information in the form of video material, with reduced impact on personal integrity. However, not much has been done in terms of using sonification in store environments.

We have identified knowledge gaps in several aspects of this particular challenge, for instance, how fast do listeners react to sounds (with head movements), how do sounds that have either context-specific congruence or incongruence to the store's sonic ambience differ in terms of grabbing attention, do sounds with early onsets perform better for notification and localization than slow onsets, and will repeated exposure increase or decrease attention.

For this present study, our goal is to investigate the effects that exposure to different sound types have on customers and clerks in a virtual store. The research question is to find if visitors to a (virtual) store will have their attention drawn towards sound alerts being played depending on the type of sound. Our assumption was that congruent sounds would draw less attention than incongruent sounds which diverge from the sonic store environment. Especially, we expected a recording of chirping birds and a metallic wind chimes to be overrepresented in terms of detection as these sounds were chosen intentionally to be detached from the context.

In the next section, we present the contextual framework of the project, namely loss prevention in retail, and also necessary theory on perception, sonification, sound design, and the experimental environment. The method section describes both the sound design and the practical test design in a semi-controlled experimental setting. The result section focuses to some extent on head movement data, while in the discussion we interpret the results from a practical sound design perspective. The paper concludes with a first approach towards alert sound design considerations and implications in a non-critical context.

## 2. BACKGROUND

There is a documented need for surveillance in stores. The total loss due to shoplifting has reached US$10 billion yearly in the United States [2]. To counter this, stores adopt different methods of surveillance such as increasing monitoring staff, security guards and cameras, and electronic alarm systems such as electronic article surveillance and RFID tags; the most effective countermeasures generally involve human factors [3].

In addition to anti-theft alarms in stores using RFID tags and alarm noises when someone walks out of the store, using loud sounds to prevent, disrupt, or rectify undesired situations are widely implemented in the society today. Car alarms, for example, typically appropriate the car horn; however, there are also suggestions for alternative designs to be found in the literature, such as musically informed car alarms [4]. Another common case is in hospital environments where doctors are constantly exposed to a great number of different alerts and warnings from apparatus, often similarly-sounding, with the result that alarms are missed, ignored, or even turned off [5].

### 2.1 Perception and Localization of Alert Sounds

The perception of sound is a heavily researched area and has been much influenced by Lord Rayleigh's "duplex theory" of sound localization as a combination of interaural differences from sound pressure level and phase [6]. Building on his observations, studies of localization of sound in the horizontal plane have found that, although stimuli could be located with reasonable accuracy, test subjects confuse stimuli presented in front and from back. Specifically, most errors occur around 3000 Hz and decline at higher or lower frequencies [6].

Directional hearing is the ability to locate the position of a sound source. This ability depends on comparisons between the acoustic inputs from the two ears, while pitch and intensity can be derived from only one ear [7]. The direction is determined binaurally from the time difference and the loudness difference of the sound waves reaching the two ears. The onset and beginning part of a sound is more important for our perception than later parts of a sound [8].

In a study on localization of sound in rooms, Rakerd and Hartmann found that impulsive tones with short onset and offset were more accurately located than those with slow onset due to the precedence effect; also, tones of longer duration that gave no precedence effect showed large individual differences [9]. There were no measurable effects on pulse durations ranging from 5–2000 milliseconds, instead Rakerd and Hartmann proposed the "plausibility hypothesis" where listeners ignore ongoing location cues after the onset when these are implausible.

### 2.2 Perceptual Attention and Urgency

Perceptual attention is defined as the ability to extract relevant information from complex surroundings that cueing, i.e., playing a sound or a warning about a sound before the actual sound is played, can improve our ability to detect and to locate a sound [10]. Therefore, the use of the same sound, or a sound the one is used to hear, would increase the ability to detect and locate it. However, with simultaneously played sounds there is a risk of a performance reduction in terms of reaction time [11], and with repeated sounds there is a risk of increased annoyance [12].

Alert sounds typically have been designed and implemented to communicate urgency and attract attention towards critical events that require immediate action [13]. It is reasonable to say that attentiveness to urgent notification sounds has been widely researched and reported; one example is the concept of "attensons" as put forward by Hellier and Edworthy, which are attention sounds designed from established perceptual and psychophysical principles such as signal-to-noise ratios [14].

The auditory system is built to process simultaneous and overlapping stimuli, although dependent on attention [15]. One study showed that people working in security operations centers with computer-network security were aided by sonification in ways that enabled peripheral monitoring in busy multitasking environments [16]. Sound alerts also has the benefit of utilizing the fastest of the human senses [17].

Our perceptual and cognitive knowledge of an environment is grounded in our ability to learn from previous experiences. We use this knowledge in relation to a context to predict which sounds that are likely to appear, but also to reject interpretations incongruous with a context [18]. When audiovisual sensory information is unrelated it leads to an uncertainty of interpretation, causing an attentional focus on identifying what is incongruent [19].

### 2.3 Sound Design and Contextual Sounds

Notifications can be realized with sonification: here we define sonification as systematically translating sensor data to non-vocal sounds. In particular, we use an event-based approach for monitoring state in a multimodal environment [1], by playing sound recordings to describe the interaction taking place.

There is a need for design of more aesthetically pleasing sonification designs and alert sounds [20]. Several authors have also stressed the importance of designing alert sounds with a high level of ecologic validity [21, 22] to match the function. In this study, sounds that would grab attention but not be disturbing were applied in a virtual store. This type of sound design where sounds that should be heard without being too cognitively demanding has been approached in a previous study [23].

Ecological validity in terms of good recordings or realistic sound simulations are not necessarily the most efficient design strategy to convey information. Instead, context-specific congruent low-level models that sacrifice realism for plainness have proven to be effective in communication, as shown in the research literature on sound objects and cartoonification [24]. Our assumption for the study was that incongruent sounds unrelated to the actions and the environment would inevitably draw more attention than the congruent sounds, and would also be perceived as more disturbing.

In the present work, we are not designing alerts or sonifications that communicate immediacy or urgency, but arguably with more coherence than the range of less urgent notifications that has resulted from the growing all-purpose use of smartphones and other technologies.

### 2.4 Virtual Reality Environment

Virtual reality (VR) is a computer generated interface that realistically simulates a physical environment. It is typically experienced through a head-mounted display (HMD) such as the Oculus Quest [1]. There are many advantages with using VR in research studies, such as increased experimental control, isolating test variables, and of approaching multiple variables in controlled conditions. Commercially available VR products like the Oculus Quest makes it possible to track body motion and head movements, which in many circumstances facilitate running complex experiments which would otherwise be difficult [25].

The kind of VR used in this study, HDM, is defined by providing 3D stereo vision, surround vision and user dynamic control of viewpoint [25]. In addition, sound was played uncompressed through Audio-Technica ATH-M50X stereo headphones which fit comfortably on the HDM. The spatialization mode in the software was without corrections for vertical head displacement, only horizontal movements. These features, when implemented together, provide for an immersive experience where the user is perceptually shielded from the surroundings, but where the experience matches a real world. Studies have showed that the sense of presence and immersion is generally high [26].

## 3. METHOD

We designed the experiment such that test subjects would experience a visit to a virtual clothing store including moving, autonomous customers, background music, recorded store ambience, and added alert sounds. The main data collected for analysis were head movements and interviews, while all audio events were variables under our control. The position and rotation of the test subjects' head movements and in-game movements were sampled and saved in the Oculus Quest HMD at 10 Hz.

The clothing store VR environment and sound programming were implemented using the game engine Unity 3D. The store measured approximately 700 $m^2$ and its merchandise consisted of shirts, pants, hats, backpacks and belts, among others, see Fig. 1. In addition to that, six avatars, two men and four women, would walk around in the environment and interact with the merchandise.

### 3.1 Experiment Design

The experiment included 16 test subjects (9 female, 7 male, age 24–53). Most of them had little to none previous experience with VR environments. The subjects were randomly assigned to one of two groups, *Knowing* and *Unknowing*, and were told they would play the part of store clerk in an informal game or VR experience where you cannot win or lose. Then, the groups were given different instructions: The knowing group was informed that alert sounds may occur in the store, which signalled that one of the avatars picked up and looked at some merchandise. They were given the instruction to experience the store and possibly pay attention to what the avatars were doing. The unknowing group was not informed about the alert sounds. They were given the instruction to experience the store and that we would conduct an interview to evaluate the "quality of the avatar's AI" without explaining what that meant.

Figure 1. The virtual store environment, which has one large space with shelves and clothing racks, one small adjacent room with more items, and one fitting room. The avatars on the platform walk around and look at items during the test.

The reason for having two conditions was to compare across the participants. In this study, the focus is on reactions to different sound types, congruent and incongruent, while attentiveness between knowing and unknowing participants are explored in more detail in a related paper [27].

To let all participants experience the store in a comparable way, they were not in control of their avatar's motion across the room, but only the head movement. In order to avoid having a strange VR sensation, the subjects were instructed to hold on to and position themselves between two chairs and follow the avatar's motion through walking and turning on spot, which through testing proved to be very helpful.

The rotation of the test subjects head movements during the test were compared to the location and time of the alert sounds, and we could see if the alerts triggered any reaction with the test subject. Data was analyzed using t-tests and Chi-square with 5% significance level. The movement data is available online. [2] The VR session lasted for 10 minutes. After the experiment, the participants were interviewed about their experience.

### 3.2 Sound Design

The sounds in the environment, apart from the actual alert sounds, consisted of generic background music and clothing store background noise/ambience. These sounds are also used in several associated experiments not reported here. Store ambience sound and the background music

[1] https://www.oculus.com/quest/

[2] https://annexes.smcresearch.se/2021-SMC-AELF

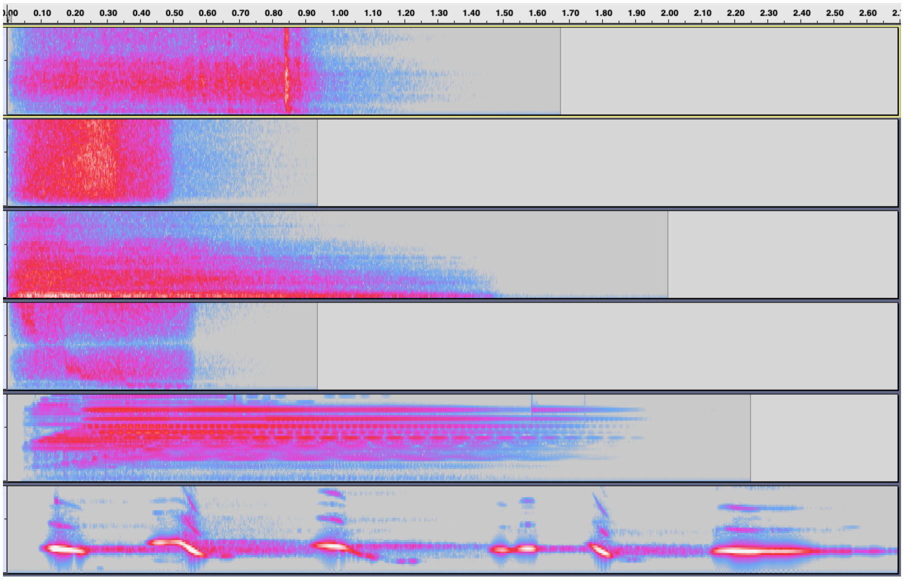


Figure 2. Spectrograms of the six alert sounds. From the top: The three congruent sounds, then the incongruent sounds sweep, chimes and bird. The sounds are also available online (see Footnote 2 ).

were omnipresent and played from virtual speakers placed all over the ceiling. The alert sounds were spatially separated in the VR environment, where each sound event was played from a virtual speaker close to the place of interaction. However, the alert sounds were not acoustically affected by walls and other objects. Therefore, the acoustic environment can be considered as an open space; although within the VR context, the experience is simply that of standing in a room without audible reverberation.

Six alert sounds were designed by two professional sound designers through iterations based on sound qualities such as attack, length and intensity, and did not have harmonic or tonal qualities that would conflict with the background music. The sounds belonged to one of two groups: three congruent sounds corresponded contextually with the clothing store environment and three incongruent sounds were contextually detached.

The three incongruent sounds—bird song, a time stretched sweep sound, and wind chimes—were selected on grounds of their disassociation from a store environment, and these three did not bear any internal resemblance. In particular, the wind chimes and bird song were intentionally distinctly detached from the context. See Fig. 3.2 for spectrogram representations; all sounds are also available for listening online (see Footnote 2 ).

The sounds considered as congruent were two recordings of a clothing hanger and a sweep-like sound, based on them mimicking the action of removing a piece of clothing from a hanger. The sweep sound was clearly resembling the hanger sounds in terms of structure and timbre, but in a cartoonified manner.

The sounds were played, in total, 25 times per test at the exact same time, position, sound level, and in the same order. Each sound had its own virtual speaker in the store, connected to a physical object such as a shelf, a table, or a clothing rack, see Fig. 3. The volume was set, through testing, to be just a bit louder (a few dB) than the masking sound from the background music and ambience. The perceived sound volume of the alert would however depend on the position of the virtual speaker relative to the head rotation of the avatar, and to some extent on the background music at that very moment.

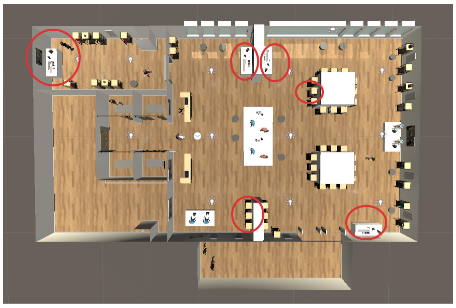

Figure 3. The layout of the virtual store environment with playing zones (virtual loudspeakers) for the sound alerts marked with red circles. Background music is played everywhere from virtual loudspeakers in the ceiling.

Notification sounds appeared from testing to have the same level when positioned both in the middle of the room and when following the avatar along its path. While the level decreased with distance to the speaker, the distances to each of the virtual speakers were identical for all subjects as they followed a set path through the store with the set actions and movements from the avatars. Only head rotation varied between subjects. The overall sound level, based on the background music, was set to a comfortable listening level by the subject.

## 4. RESULTS

First, we could confirm that head movements by the unknowing and knowing groups differ, see Fig. 4. During sound alerts, the average motion measured in angular speed and extent of rotation was almost twice as high for the knowing group, while in-between sound events being played, the amount of motion was almost the same.

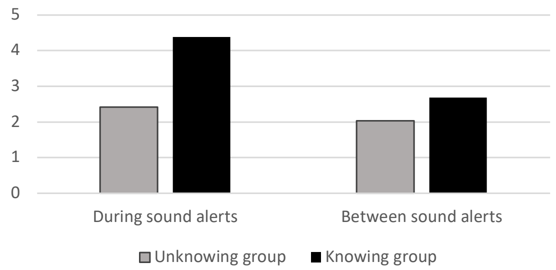


Figure 4. Head movements for the unknowing and knowing groups during sound alerts being played and in-between sound alerts, measured in speed and extent of the rotations.

The next step in the analysis of the head movement data was to determine "hits", or reactions to sounds where the head rotation pinpoints the sound source. An event would be considered a hit if the head rotation of the test subject existed in the range of 30 degrees from the direct line from the test subject to the event position. Experimenting with the angle and trying different ranges of hit area led to the conclusion that the range did not affect the number of hits considerably as the range increased, and therefore we set 30 degrees as default range.

We found that the reaction to a sound event typically came after two seconds from the start of the sound. The amount of hits would have a ceiling effect from three seconds and longer. Thus, we include hits identified between the start of the sound plus three seconds.

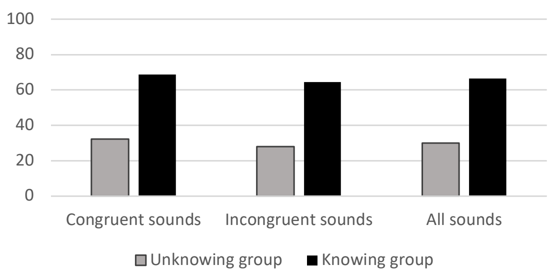


Figure 5. Reactions ("hits") for congruent, incongruent, and all alert sounds in percent for the unknowing and knowing groups.

Looking at the total amount of hits during the tests we can see that the group which were informed to look or listen for sounds outperformed the other group significantly ($\chi^2$, $p = 0.000$), see Fig. 5. On average, the unknowing group reacted to 30% of all sounds, while the knowing reacted to 67%.

Furthermore, we found no significant effects ($\chi^2$-tests) of the distance from the sound source to the test subject, of the physical place of the sound, of the test subject's gender, or of the sound-causing avatar's gender. The knowing group were generally unaffected by the total length of the sounds; longer sounds gave slightly more hits while medium long sounds gave fewer, but the differences were small and not significant ($\chi^2$, $p = 0.24$). On the other hand, there was an effect of duration for the unknowing group ($\chi^2$, $p = 0.025$) where longer sounds resulted in more hits. Sounds longer than two seconds resulted in twice as many hits on average than sounds below one second. There were no significant differences between having an early or late (100–800 ms) sound amplitude peak ($\chi^2$, $p > 0.31$). There was no effect when the same stimulus was repeated for the unknowing group, but the knowing had a small increase of hits for a repeated sound ($\chi^2$, $p = 0.005$). We notice a slight decline in attentiveness among the test subjects, but there are no significant effects of exposure over time.

The Bird type sound generated most hits, followed by the three congruent sounds, see Fig. 6. These differences are significant ($\chi^2$, $p = 0.025$). The Bird sounds got most attention from the unknowing group, while the largest difference between the two groups was found for the Sweep sound. Chimes was the sound with overall least number of hits. However, these observations have not been evaluated statistically.

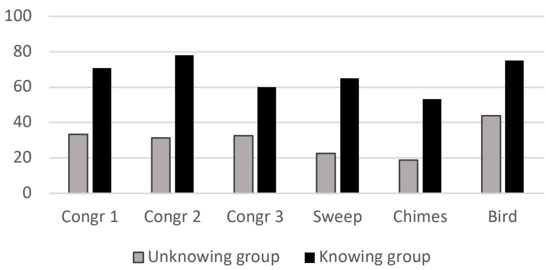


Figure 6. Reactions ("hits") to the different alert sounds in percent for the unknowing and knowing groups.

Finally, the results partly disproved assumptions that incongruent sounds would outperform congruent and more subtle sounds, see Figures 5 and 6. Instead, we notice certain patterns that will be discussed in the following.

## 5. DISCUSSION

Interpreting the results can give us a clue into what implications for the sound design it may result in. First off, it is possible to locate spatially separated sounds in a busy clothing store environment and directional sounds can be used. And more importantly, the sounds could be played at a lower volume overall. It was proven that the ability to detect and locate sounds in such an environment was not an overwhelmingly difficult challenge. And furthermore, the unknowing group tended to react less to the sounds than was expected. This leads us to believe that a lower volume could be used, where those who do not listen will be even less disturbed, but those who do listen will still be able to distinguish the majority of alerts.

Another finding we did not anticipate was that shorter sounds proved to be more suited compared to longer sounds. One might think that longer sounds would have a greater impact on our perception purely because of their length and longer exposure to our ears but when it comes to raising our attention, at least in this environment, shorter sounds are more effective. Furthermore, it was shown that the onset or attack in the shorter sounds had no effect on detectability. This was also in contradiction to the hypothesis, which could be due to the short durations overall. The onset differed with a maximum of 600 milliseconds and humans typically have a reaction time to audio stimulus of 140–160 milliseconds [28].

The results show that congruent sounds generally are noticed to a greater extent than the incongruent sounds. There are differences between groups in attention towards the sound types. This can be seen particularly for chimes and bird sounds compared to the three congruent sounds. Both bird and chimes stand out from the background music and ambient sounds because of their strong and distinct harmonics, but the chimes, which we expected to be the most noticed sound of all, scored the least. This can surely be partly explained by the length of the sound, but then the sweep sound should not fall in-between.

This leads us to two somewhat contradictory conclusions. Our hypothesis that incongruent sounds stand out from the context holds in part, with different reactions from the knowing and unknowing groups. However, the congruent sounds, although more subtle, are easily and consistently noticed by the knowing group, but not by the unknowing. This might have support in the previously mentioned “plausibility hyphotesis” for localizing sounds [9].

Alert sounds for store environments *could* thus be incongruent since customers, which are represented in this test by the unknowing group, will not notice the sounds as much as the knowing group, the store personnel. However, these incongruent sounds need to be carefully designed, while congruent sounds can be implemented with less care. This encourages discreet notification designs, but also opens up the design space for shops and allows for instance freedom to develop sonic branding as part of the store’s monitoring system.

The result of finding no growing sensitivity to the alerts adds to the argument that sounds could be designed in a way that is directed towards those who listen for it without them being disturbing for those who do not. These findings also suggest that alert sounds can be designed with subtlety and still be noticeable, and even that customers will not be increasingly annoyed. However, more research on how alert sounds are perceived in a real store environment in relation to pleasantness, fatigue and function over longer time is needed.

The study included only a small number of participants, which jeopardizes using statistical methods and making conclusions from these. Also, the sound stimuli design was not formally evaluated before the experiment, nor how and how often these were presented. There was no randomization of stimuli presentation. The participants did not get any training in visiting a VR environment, and the store layout was not evaluated for realism. As such, there are many uncertainties present in the study, and the results should therefore be considered as preliminary.

## 6. CONCLUSIONS

Based on the findings presented here, a first approach towards design guidelines for short alert sounds in a retail environment are stated as follows. Alerts can:

– be congruent with and contextually fit the environment of where they are played,
– be played at a lower volume than the background music,
– be short in length, around one second,
– be designed without much attention of attack sharpness,
– be used without concerns of growing sensitivity over time,
– be incongruent, if designed with care.

This first approach towards design guidelines will be evaluated and developed in forthcoming experiments. As such, the particular sound designs that were tested in this experiment will serve as inspiration, but should not be considered to be general design recommendations.

We believe that there are promising opportunities for sound design for marketing purposes as well as for increasing customer shopping experiences and working conditions for employees. VR was used successfully in this exploratory study where a real store would introduce a number of uncontrollable variables.

**Acknowledgments**

The research was funded by the Hakon Swenson Foundation and The Swedish Retail and Wholesale Council (Handelsrådet).

## 7. REFERENCES

[1] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Verlag Berlin, 2011.

[2] Y. Yamato, Y. Fukumoto, and H. Kumazaki, “Proposal of shoplifting prevention service using image analysis and ERP check,” *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 12, no. S1, Jun. 2017.

[3] A. Lindblom and S. Kajalo, “The use and effectiveness of formal and informal surveillance in reducing shoplifting: A survey in Sweden, Norway and Finland,” *The International Review of Retail, Distribution and Consumer Research*, vol. 21, no. 2, pp. 111–128, May 2011.

[4] A. Sigman and N. Misdariis, "Alarm/will/sound: perception, characterization, acoustic modeling, and design of modified car alarms," in *ICMC - International Computer Music Conference*, Athenes, Greece, Sep. 2014. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01546060

[5] J. Frank E. Block, L. Nuutinen, and B. Ballast, "Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance," *Journal of Clinical Monitoring and Computing*, vol. 15, no. 2, pp. 75–83, 1999.

[6] J. C. Middlebrooks and D. M. Green, "Sound localization by human listeners," *Annual Review of Psychology*, vol. 42, no. 1, pp. 135–159, Jan. 1991.

[7] S. Kuwada and T. C. T. Yin, "Physiological studies of directional hearing," in *Proceedings in Life Sciences*. Springer US, 1987, pp. 146–176.

[8] D. Oberfeld, J. Hots, and J. L. Verhey, "Temporal weights in the perception of sound intensity: Effects of sound duration and number of temporal segments," *The Journal of the Acoustical Society of America*, vol. 143, no. 2, pp. 943–953, Feb. 2018.

[9] B. Rakerd and W. M. Hartmann, "Localization of sound in rooms, III: Onset and duration effects," *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1695–1706, Dec. 1986.

[10] E. R. Hafter, A. Sarampalis, and P. Loui, "Auditory attention and filters," in *Auditory Perception of Sound Sources*, W. A. Yost, A. N. Popper, and R. R. Fay, Eds. Springer US, 2008, pp. 115–142.

[11] R. Zajdel, J. Zajdel, A. Zwolińska, J. Śmigielski, P. Beling, T. Cegliński, and D. Nowak, "The sound of a mobile phone ringing affects the complex reaction time of its owner," *Archives of Medical Science*, vol. 5, pp. 892–898, 2012.

[12] Å. Skagerstrand, S. Köbler, and S. Stenfelt, "Loudness and annoyance of disturbing sounds – perception by normal hearing subjects," *International Journal of Audiology*, vol. 56, no. 10, pp. 775–783, May 2017.

[13] E. Hellier and J. Edworthy, "On using psychophysical techniques to achieve urgency mapping in auditory warnings," *Applied Ergonomics*, vol. 30, no. 2, pp. 167–171, 1999.

[14] ——, "The design and validation of attensons for a high workload environment," in *Human Factors in Auditory Warnings*. Routledge, 1999, pp. 283–304.

[15] R. P. Carlyon, "How the brain separates sounds," *Trends in Cognitive Sciences*, vol. 8, no. 10, pp. 465–471, Oct. 2004.

[16] L. M. Axon, B. Alahmadi, J. R. C. Nurse, M. Goldsmith, and S. Creese, "Sonification in security operations centres: What do security practitioners think?" *Workshop on Usable Security (USEC) at the Network and Distributed System Security (NDSS) Symposium 2018*, Jul. 2018.

[17] J. Shelton and G. P. Kumar, "Comparison between auditory and visual simple reaction times," *Neuroscience and Medicine*, vol. 01, no. 01, pp. 30–32, 2010.

[18] M. Bar, "The proactive brain: using analogies and associations to generate predictions," *Trends in Cognitive Sciences*, vol. 11, no. 7, pp. 280–289, Jul. 2007.

[19] P. Larsson, D. Västfjäll, P. Olsson, and M. Kleiner, "When what you hear is what you see: Presence and auditory-visual integration in virtual environments," in *Proceedings of the 10th annual international workshop on presence*. Citeseer, 2007, pp. 11–18.

[20] P. Vickers, "Ars informatica – ars electronica: Improving sonification aesthetics," in *Understanding and Designing for Aesthetic Experience: workshop at HCI 2005: The 19th British HCI Group Annual Conference*, 2005.

[21] W. W. Gaver, "What in the world do we hear?: An ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, mar 1993.

[22] P. Bergman, A. Sköld, D. Västfjäll, and N. Fransson, "Perceptual and emotional categorization of sound," *The Journal of the Acoustical Society of America*, vol. 126, no. 6, pp. 3156–3167, 2009-12.

[23] M. L. Eriksson, R. Atienza, and L. Pareto, "The Sound Bubble: A context-sensitive space in the space," *Organised Sound*, vol. 22, no. 01, pp. 130–139, mar 2017.

[24] F. Avanzini, M. Rath, D. Rocchesso, and L. Ottaviani, "Low-level models: resonators, interactions, surface textures," *The Sounding Object*, pp. 137–172, 2003.

[25] X. Pan and A. F. de C. Hamilton, "Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape," *British Journal of Psychology*, vol. 109, no. 3, pp. 395–417, Mar. 2018.

[26] J.-C. Servotte, M. Goosse, S. H. Campbell, N. Dardenne, B. Pilote, I. L. Simoneau, M. Guillaume, I. Bragard, and A. Ghuysen, "Virtual reality experience: Immersion, sense of presence, and cybersickness," *Clinical Simulation in Nursing*, vol. 38, pp. 35–43, Jan. 2020.

[27] K. Falkenberg, M. L. Eriksson, E. Frid, T. Otterbring, and S.-O. Daunfeldt, "Auditory notification of customer actions in a virtual retail environment: Sound design, awareness and attention," in *Proceedings of ICAD 2021*, 2021, in press.

[28] P. D. Thompson, J. G. Colebatch, P. Brown, J. C. Rothwell, B. L. Day, J. A. Obeso, and C. D. Marsden, "Voluntary stimulus-sensitive jerks and jumps mimicking myoclonus or pathological startle syndromes," *Movement Disorders*, vol. 7, no. 3, pp. 257–262, 1992.