EFFECTS OF AMPLITUDE ENVELOPE ON REDUCING ALARM ANNOYANCE

EVALUATING THE EFFECTS OF AMPLITUDE ENVELOPE MANIPULATION ON REDUCING AUDITORY ALARM ANNOYANCE

By SHARMILA SREETHARAN, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2019) Hamilton, Ontario (Psychology, Neuroscience & Behaviour)

TITLE: Evaluating the effects of amplitude envelope manipulation on reducing auditory alarm annoyance

AUTHOR: Sharmila Sreetharan, B.Sc. (McMaster University)

SUPERVISOR: Professor Michael Schutz

NUMBER OF PAGES: vii, 27

Lay Abstract

Alarms in intensive care units are perceived as annoying, ultimately negatively affecting both clinicians and patients. These alarms are mandated by the International Electrotechnical Commission to have sustained or flat amplitude envelopes (i.e., referring to the change in loudness over time), which does not reflect naturally occurring stimuli that typically have decaying or percussive amplitude envelopes. The current experiments assessed the effect of percussive envelopes on alarm learnability and annoyance. We showed in a series of experiments that there is no difference in learning alarms with flat or percussive envelopes. However, we showed that alarms with percussive envelopes are perceived to be less annoying than alarms with flat envelopes. These results offer one potential solution to reduce alarm annoyance in intensive care units without harming the learnability of these alarms.

Abstract

Auditory alarm annoyance plagues clinicians, which results in alarms desensitization and ultimately affects patient care. Contributing to this problem are the International Electrotechnical Commission (IEC) 60601-1-8 alarms, a standardized set of melodic alarms used to convey information to clinicians in intensive care units. By design, IEC alarms employ flat amplitude (i.e., amplitude invariant) envelopes and are not reflective of naturally occurring sounds with percussive amplitude (i.e., decaying) envelopes. We present a series of three experiments evaluating the effect of amplitude envelope manipulation (i.e., incorporating percussive envelope) on memory and annoyance in IEC alarms synthesized using pure tones (experiment 1), complex tones (experiment 2) and assessing annoyance pre and post memory assessment (experiment 3). For the memory assessment, participants were assigned to learn either the flat alarms or percussive alarms. During the memory assessment, participants were informed of the alarm–referent pairings (study phase), practised identifying alarms (training phase), had a short break, and tested on their ability to identify alarms (evaluation phase). The annoyance assessment was a two alternative forced choice task where participants identified which alarm they perceived to be more annoying from a pair of alarms differing in either envelope-type or alarm-type. Across all experiments there was no difference in alarm learnability between those learning either flat or percussive alarms during the memory assessment. Annoyance assessments revealed that all participants chose the flat alarms to be more annoying than the percussive alarms, independent of the memory assessment condition. These results showcase the potential of using percussive alarms to reduce alarm annoyance without harming learnability, a cost-efficient manipulation.

Acknowledgements

I would like to thank Dr. Michael Schutz for giving me the opportunity to be a part of the MAPLE lab and opening me to the wonderful world of music cognition research. Many thanks to my committee members Drs. Laurel Trainor and David Shore for providing constructive feedback and advice over the past two years. Special thanks to Dr. Joseph Schlesinger for providing constant feedback, support and potential research avenues throughout my graduate studies. I would like to also thank PNB office ladies and Christine Speare-Van Vugt for their administrative help throughout my graduate career and overall positive demeanour in our interactions. My experience in graduate school would not have been the same without the countless friends I've made in PNB (including Lauren Smail, Portia Kalun, David Prete, Hector Orozco, to list a few) who have made all the wonderful late nights at Friday socials worthwhile.

I would like all the members of the MAPLE lab who have put up with my sassy comments, terrible jokes and contributing to my owl obsession. It goes without saying that none of my data collection would have been possible without the QQs I've had over the past two years. Thank you to both Annilee Baron and Marsha Natadiria for all their efforts and patience as I constantly bothered them with random memes, videos, questions while their time as lab manager. Lastly, I'd like to thank my partner in crime, adoptive mother and only other graduate lab mate I've known, Aimee Battcock. I could write whole other thesis on the things I've learned through watching her mentorship students, her presentation skills and all the constructive feedback I've received despite my insecure attachment and constant sassiness.

Table of Contents

Lists of Figures

Declaration of Academic Achievement

Reducing Alarm Annoyance through Amplitude Envelope Manipulation

Authors: Sharmila Sreetharan, Joseph Schlesinger, Michael Schutz

Publication: In preparation

The design of this study was conceived primarily by SS, under the supervision of MS.

Experiments were conducted by SS with feedback from JS and MS. Manuscript written by SS.

1. Introduction

1.1 Alarm Annoyance

Auditory alarms in the hospital setting employ non-speech, short melodic sequences to convey messages to clinicians. Compared to speech-based alerts, auditory alarms have better temporal resolution, offer greater privacy for patients, and are universally recognizable (Sodnik & Tomazic, 2015). For example, the International Electrotechnical Commission (IEC) 60601-1-8 (henceforth 'IEC alarms'), which is the global medical standard, employs short melodic sequences to convey messages about patient and machine status to clinicians (Comission, 2006). Yet, despite their utility these alarms are very annoying, ultimately negatively impacting patient safety (Edworthy & Hellier, 2005).

Sound annoyance is defined as "an evaluative response towards [a] sound and its source, including both emotional ('nuisance, 'unpleasantness') and cognitive aspects ('disturbance', 'interference)" (Guski, Felscher-Suhr, & Schuemer, 1999).When creating alerting alarms, designers typically adopt a "better-safe-than-sorry" approach (i.e., through increasing the volume and frequency of alarms), even when action is not needed. Human perceptual systems have evolved to "tune in" to meaningful stimuli and "tune out" (or habituate) to whatever doesn't require action (Wilcox, 2011). Consequently, efforts to make alarms more annoying ironically also may accidentally make them less alerting. Thus, annoyance is an important factor of consideration in designing warning systems, including auditory alarms (Fagerlonn, 2011).

Alarm fatigue is defined as the desensitization to alarms as a result of sensory overload when exposed to an excessive number of alarms (Sendelbach & Funk, 2013). As a result, clinicians are more likely to ignore or even miss critical alarms. One report stated that on average two critical alarms are missed per hospital, per day (Donchin et al., 2003). Moreover, users might even turn off alarms due to their unpleasantness (Block, Nuutinen, & Ballast, 1999) as they believe they can work more effectively without the cacophony of the alarms. However, this can have dire consequences for patients. For example, in an unfortunate incident, while a patient was undergoing abdominal surgery, the clinicians turned off all alarms on equipment in the operating room because they found them to be annoying. During the procedure the respirator was temporarily turned off but never turned back on since the alarms did not sound when the patient was under duress. The patient suffocated and entered a vegetative state, eventually dying 11 days later ("Alert fatigue leads to OR fatalities", 2010). By creating less annoying and more ergonomic alarms, we can reduce the likelihood of alarms being ignored, missed, or turned off. In this paper, we investigate the potential of designing less annoying but equally effective alarms by varying the amplitude envelopes (i.e., the shape of sound over time) used in individual tones.

1.2 Amplitude Envelope

Amplitude envelope (hereafter referred to as 'envelope') is an acoustical property of sound referring to changes in a sound's amplitude (i.e., intensity or loudness) over time, specifically in reference to a sound's offset. We will focus on two types of envelopes: percussive and flat. Percussive sounds (i.e., sounds with percussive envelopes) have an abrupt onset reaching maximum amplitude quickly, followed by an exponential decay (Figure 1A). Percussive sounds are characteristic of naturally occurring impact sounds, such as the sound produced by two glasses clinked together. Flat sounds (i.e., sounds with flat envelopes) have both abrupt onsets and offsets that flank an amplitude invariant, with a period of sustain at the maximum amplitude level (Figure 1B). Most artificially synthesized sounds are flat, such as the sound emitted by a truck backing up or the emergency broadcast tone. Differences in envelope are salient, as they affect our ability to identify everyday objects (Giordano, Rocchesso, &

McAdams, 2010; Lutfi, 2001), make duration judgements (Vallet, Shore, & Schutz, 2014), and bind together multimodal stimuli (Chuen & Schutz, 2016; Grassi & Casco, 2009; Schutz & Bamikole, 2015; Schutz & Kubovy, 2009). Additionally, envelope can play a role in learning and memory of sound-object associations. A series of experiments showed those who heard percussive sounds learned sound-object associations faster and correctly recalled more associations than those who heard flat sounds (Schutz et al., 2017).

Previous work conducted by our group expanded on these associative memory findings investigating applications of envelope manipulation to IEC alarms. By design, the IEC alarms have a flat envelope shapes (Wee & Sanderson, 2008); in particular, each note in the tone sequence has a flat envelope shape. Previously, the effect of envelope manipulation on IEC alarm learnability was evaluated by changing the envelope shape of each note in the alarm to non-overlapping, percussive envelopes. Participants learned either flat or percussive versions of the IEC alarms and completed a memory assessment task to measure learning and recall. Envelope manipulation did not affect learning or recall of IEC alarms (Gillard & Schutz, 2016).

Here we evaluated the effect of differing envelopes on annoyance, as well as learning and short-term memory, but used overlapping percussive envelopes with varying lengths (see section 2.1.2 for details). Participants learned either flat or percussive alarms during the memory assessment task (consisting of a study phase, training phase, distractor phase and evaluation phase). Additionally, participants assessed annoyance for both flat and percussive alarms during a two-alternative forced choice (2AFC) task to evaluate the effect of envelope on alarm annoyance (see Figure 3 for overview of methods). In this paper, we present a series of three experiments looking at the effect of envelope manipulation on memory and annoyance in IEC alarms synthesized using pure tones (experiment 1), complex tones (experiment 2) and assessing annoyance pre and post exposure (experiment 3). We hypothesized that envelope would not play a role in learning and memory but would affect annoyance. Specifically, we predicted that there would be no difference in performance during the training and evaluation phases for those learning either flat or percussive alarms. We also predicted that flat alarms will be perceived to be more annoying, regardless of whether participants learned flat or percussive alarms.

2. Methods and Results

2.1 Experiment 1

2.1.1 Methods

We recruited 40 participants (35 female, 5 male) ranging in age from 18 – 23 years *(M* = 19.08, *SD* = 1.40) from McMaster University through the Psychology and Linguistics SONA online research participation systems. All participants had corrected-to-normal vision and received course credit as compensation.

We generated both flat and percussive versions based on the pitch, timing, and envelope parameters of the eight, medium-priority IEC 60601-1-8 alarms using the MAPLE Lab Auditory Exploration Suite for Teaching, Research, and Observation or MAESTRO (Ng & Schutz, 2017). The flat versions of the IEC alarms (hereafter referred to as “flat alarms”) consisted of pure tones (i.e., sine waves). We synthesized the flat alarms according to the envelope specifications outlined by the current IEC alarm guide (Block, Rouse, Hakala, & Thompson, 2000); each tone had a 20 millisecond (ms) onset/offset and 200 ms period of sustain. Successive tones had inter-stimulus onset intervals of 400 ms (See Figure 2 for visual depiction of stimulus timing). All notes in the flat alarm had an amplitude of “1.0”. The percussive versions of the IEC alarms (hereafter referred to as “percussive alarms”) consisted of pure tones with percussive envelopes. All tones within the percussive alarms had 20 ms onsets, no periods of sustain, and varying offsets. The offsets for the first, second, and third notes in each tone sequence were 1020 ms, 640 ms, and 380 ms, respectively. Consistent the flat alarms, successive tones had inter-stimulus onset intervals of 400 ms. Since percussive tones are perceived to shorter than equivalent flat tones (Vallet et al., 2014), we extended the length of the final note in the percussive sequence in order to roughly equate for similar perceived durations. We equated for perceived duration through an informal assessment.

Upon arrival, participants completed the consent form as well as a short survey inquiring about demographic information (e.g., age, gender) and musical training. A research assistant (RA) conducted the experiment in a sound-attenuating booth while physically separated from the participant by a partition to reduce experimenter bias. We randomly assigned participants to hear either the flat IEC alarms or the percussive IEC alarms during the memory assessment. The memory assessment consisted of four phases: a study phase, a training phase, a break phase, and an evaluation phase.

Study phase. The RA explained that, *"The alarms you will be asked to learn today are used in the ICU to inform doctors about a patient's condition or a machine's status. You will be asked to identify eight different alarms represented by melodic tone sequences."* Afterwards, the RA gave participants a short description of all eights alarms and gave the participant cue cards with the alarm names that they were allowed to manipulate in any way to help them remember the alarm–referent pairings. The RA played each alarm twice, mentioning the correct referent (i.e., the associated alarm name) after each presentation. White noise presented between different alarm presentations prevented practise. To avoid order effects, we randomized alarm presentation order.

Training phase. After being informed of the alarm–referent pairings, participants practised identifying the alarms. After each alarm was played, participants guessed the referent. At the end of each trial, participants were told whether they were correct/incorrect and were reminded of the correct alarm–referent pairing regardless of whether they were correct/incorrect. We played white noise between each alarm presentation to mask echoic memory to minimize rehearsal. This continued for each training block (consisting of all eight alarms) until participants reached our learning criterion of correctly identified 7/8 alarms correctly on two consecutive blocks or until a maximum of 10 blocks were reached. keep up morale, participants received positive reinforcement every other block (e.g., "you are doing very well") regardless of performance.

Distracter phase. Participants had a five-minute break. During this time, they played a silent game of online mini-putt. We included this break to prevent practise and to measure the effects of short-term memory in the subsequent phase.

Evaluation phase. Participants were tested on their ability to remember the association after a single presentation of each tone sequence (i.e., one block of trials). Alarm presentation was randomized for each participant. After identification, participants reported the confidence of their response on a 6-point Likert scale (from *1 = "Not confident at all"* to *6 = "Very confident"*). Participants received no feedback after their response. At the end, we asked participants for any techniques they used to help them remember the alarm–referent pairings.

Annoyance task. After the memory assessment, participants completed an annoyance task. This involved making a 2AFC decision about annoyance for pairs of alarms. In half the trials participants heard alarms with matched pitch sequences but different envelopes (i.e., the percussive and flat versions of one pitch sequence). The other half of the trials used different pitch sequences with matched envelopes (e.g., two flat alarms). We fully randomized the trials to reduce the transparency of the task and subject bias. Participants identified whether they perceived the first or the second tone sequence in the pair as more annoying. Participants training on flat alarms had not previously heard the percussive alarms, and vice-versa. In this paper, we report the results of trials from when we forced assessment based on envelope (e.g., flat general alarm vs. percussive general alarm).

2.1.2 Results

We conducted Mann-Whitney tests evaluating the effect of envelope condition on alarm learnability as measured by the number of blocks required to meet our criterion (i.e., correctly identifying 7/8 alarms correctly on two consecutive blocks to a maximum on 10 blocks). We found no difference for those trained on flat alarms (*Mdn* = 10) and percussive alarms (*Mdn* = 10) on alarm learnability, $U = 187.5$, $p = .72$. We conducted Mann-Whitney tests evaluating the effect of envelope condition on alarm identification after a short break. We found no difference in alarm identification for those evaluated on the flat alarms (*Mdn* = 7.0) or percussive alarms (*Mdn* = 5.5), U=263, $p = .08$.

We performed chi-square tests of goodness-of-fit tests to examine whether flat alarms and percussive alarms are perceived to be equally annoying for those that learned the percussive or flat alarms during the memory assessment. Participants that learned the percussive alarms during the memory assessment chose the flat alarms to be significantly more annoying than the percussive alarms X^2(1, N=160) = 115.60, $p < .001$. Similarly, participants that learned the flat alarms during the memory assessment chose the flat alarms to be significantly more annoying than the percussive alarms X^2(1, N=160) = 11.03, $p < .001$.

2.2 Experiment 2

We conducted experiment 1 with pure tones to begin understanding the effects of envelope on alarm annoyance and learnability. For experiment 2, we replicated experiment 1 using complex tones instead of pure tones as complex tones allow for great localizability and provide greater resistance to masking (Edworthy et al., 2017; Patterson, 1982).

2.2.1 Methods

We recruited 40 participants (35 female, 5 male) ranging in age from 18 – 27 years *(M* = 18.75, *SD* = 1.56) from McMaster University through the Psychology and Linguistics SONA online research participation systems. All participants had corrected-to-normal vision. Upon completion, participants received course research credit as compensation. As in experiment 1, we generated both flat and percussive versions of the eight, medium-priority IEC 60601-1-8 alarms using MAESTRO. Uniquely, we synthesized "complex tones" with notes in the tone sequences containing the first three harmonics above the frequency and each harmonic exhibiting half the amplitude of the preceding harmonic. In other words the first harmonic had an amplitude of "0.5", the second harmonic had an amplitude of "0.25", and the third harmonic had an amplitude of "0.125". Amplitudes were consistent for both flat and percussive versions of the IEC alarms. We used the same procedure as in experiment 1.

2.2.2 Results

We conducted Mann-Whitney tests evaluating the effect of envelope condition on alarm learnability as measured by the number of blocks required to meet our criterion (i.e., correctly identifying 7/8 alarms correctly on two consecutive blocks to a maximum of 10 blocks). We found no significant difference between those learning flat alarms (*Mdn* = 6.0) and percussive alarms (*Mdn* = 9.5) on alarm learnability, $U = 130.5$, $p = .057$. We also conducted an independent samples t-test evaluating the effect of envelope condition on alarm identification after a short break. We found no significant difference in alarm identification for those evaluated on the flat alarms (M= 6.40, SD =1.54) and percussive alarms (M= 5.30, SD =1.92), $t(38)=2.00$, $p = .053$.

We performed chi-square tests of goodness-of-fit to examine whether flat alarms and percussive alarms are perceived to be equally annoying for those that learned the percussive or flat alarms during the memory assessment. Participants rated flat alarms significantly more annoying regardless of whether they heard percussive (X^2(1, N=160) = 108.90, $p < .001$) or flat (X^2(1, N=160) = 4.23, p =.04.) sequences during the memory assessment.

2.3 Experiment 3

To further explore the effect of exposure on alarm annoyance, we took two steps to control for potential confounds related to training prior the annoyance task. First, in experiments 1 and 2, participant's performance on the training phase determined the overall amount of exposure (i.e., participants performing worse took longer to reach our criterion and had more alarm exposure). In principle this differential amount of exposure (which varied by envelope condition) could affect annoyance ratings. In experiment 3, we assessed annoyance after a fixed amount of training to match exposure. We used seven training blocks—the average required to reach the criterion in experiment 2. Secondly, our previous experiments could not account for pre-existing differences in annoyance as we assessed it only after exposure. To account for this, we assessed alarm annoyance prior to the influence of exposure to either the flat or percussive alarms (i.e., memory assessment). However for consistency with previous experiments we also administered the annoyance task after the memory assessment. This allowed for exploration of changes due to exposure to the memory assessment.

2.3.1 Methods

We recruited 72 participants (48 female, 24 male) ranging in age from 18 – 23 years (*M* = 18.94, *SD* = 1.14) from McMaster University through the Psychology SONA online research participation system and the Introduction to Music Cognition course. All participants had corrected-to-normal vision. Upon completion, participants received course research credit as compensation. We used the same alarms synthesized with complex tones as in experiment 2.

The memory assessment and the annoyance tasks were similar to the previous experiments with the following exceptions. First, participants completed the annoyance task, identical to the annoyance task in the previous experiment, hereafter referred to as time 1. After, participants completed the memory assessment comprised of the same four phases: study phase, training phase, distracter phase, and evaluation phase (see section 2.1.1 for a detailed breakdown). All participants completed a total of seven blocks, regardless of performance (i.e., even if they passed the criterion). At the end of the memory assessment, participants completed the annoyance task once again, hereafter referred to as time 2 (See Figure 3 for overview of paradigm for all experiments).

2.3.2 Results

We conducted Mann-Whitney tests evaluating the effect of envelope condition on alarm learnability as measured by the number of blocks required to meet our new criterion (i.e., correctly identifying 7/8 alarms correctly on two consecutive blocks to a maximum of *seven* blocks). We found no difference for those trained on flat alarms (*Mdn* = 5.5) and percussive alarms (*Mdn* = 5.0) on alarm learnability, $U = 691.5$, $p = .61$. We conducted Mann-Whitney tests evaluating the effect of envelope condition on alarm identification after a short break. We found no difference in alarm identification for those evaluated on the flat alarms (*Mdn* = 8.0) or percussive alarms (*Mdn* = 8.0), *U* = 647, *p* =1.00.

Time 1 (Pre Memory) Annoyance Assessment. We performed chi-square tests of goodness-of-fit to examine whether flat alarms and percussive alarms are perceived to be equally annoying for all participants for the time 1 annoyance task. Participants rated flat alarms significantly more annoying regardless of whether they were assigned to the percussive (X^2(1, N=288) = 141.68, $p < .001$.) or flat (X^2(1, N=288) = 159.01, $p < .001$) memory assessment conditions. Additionally, we performed a chi-square test to examine differences in annoyance ratings between participants before random assignment to the flat or percussive condition for the memory assessment. We found no difference in annoyance ratings between assigned groups (percussive, flat) before exposure, X^2(1, N=288) = 0.95, $p = .33$, indicating no bias in our random assignment.

Time 2 (Post Memory) Annoyance Assessment. We performed chi-square tests of goodness-of-fit to examine whether flat alarms and percussive alarms are perceived to be equally annoying for those that learned percussive or flat alarms during the memory assessment for the time 2 annoyance task. Participants that learned the percussive alarms during the memory assessment chose the flat alarms to be significantly more annoying than the percussive alarms X^2(1, N=288) = 144.5, $p < .001$. Similarly, participants that learned the flat alarms during the memory assessment chose the flat alarms to be significantly more annoying than the percussive alarms X^2(1, N=288) = 148.27, $p < .001$. We performed McNemar's tests to evaluate changes in annoyance ratings from time 1 to time 2 by memory assessment condition. We found no change in annoyance ratings for those than learned flat alarms, X^2(1, N=288) = 0.17, $p >.05$, or percussive alarms X^2(1, N=288) = 0, $p >.05$ from time 1 to time 2.

3. General Discussion

3.1. Alarm Annoyance

In experiments 1 and 2, we found that flat alarms are perceived to be more annoying than percussive alarms for those who learned either percussive or flat alarms (Figure 4A, 4B). According to the mere exposure effect, we would expect that participants in the flat condition, who had greater exposure to the flat alarms than the percussive alarms, would perceive the percussive alarms to be more annoying. However, the mere-exposure effect does not fully explain our results; both participants trained on percussive and flat sequences found flat more annoying in pure tones (experiment 1) and complex tones (experiments 2 and 3), implicating the role of envelope in annoyance judgements. In our experiments we did not directly compare the role of overtones in annoyance ratings since harmonics play an important role in addressing psychoacoustic concerns of localizability and providing resistance to masking (Edworthy et al., 2017; Patterson, 1982).

In experiment 3, when participants rated annoyance at time 1 (pre memory assessment), flat alarms were rated to be more annoying than the percussive alarms. This revealed that in the absence of alarm exposure, the flat alarms were perceived to be more annoying. At time 2 (post memory assessment), participants in both experimental groups rated the flat alarms to be more annoying than the percussive alarms. Furthermore, comparing annoyance ratings across times 1 and 2, participants consistently chose the flat alarms to be more annoying than the percussive alarms (Figure 4C), independent of memory assessment condition. This may suggest that after exposure to both flat and percussive alarms, participants have an intractable preference for percussive alarms unaffected by further exposure to one type of alarm; however, we speculate that we are observing carryover effects due to repeated annoyance measurements presented within a short time period. The inclusion of two identical annoyance tasks likely resulted in practise or carryover effects (Cleophas, 1999) as participants might have wanted to respond consistently during both annoyance assessments, effectively resisting any effects of memory assessment condition. To further understand the effect of envelope exposure in annoyance assessment, we contend that longer exposure periods should be examined to understand if envelope preferences are malleable.

Despite the potentially confounding carryover effects, experiment 3 demonstrates a preference for percussive alarms in the absence of exposure. Consistent with the results from experiments 1 and 2, it suggests that percussive alarms are less annoying than equivalent flat alarms—even when controlling for exposure time. Additionally, it illustrates that flat tones are more annoying than percussive even prior to exposure during the assessment condition. These results are consistent with previous work showing that stimulus intensity (i.e., loudness) accounted for more variability in preference ratings for pure tones than musical dissonance (i.e., harshness or lack or harmony). Specifically, loudness was the most important factor in participants' ratings of preference (Martindale & Moore, 1990). Although envelope is not synonymous with intensity, sounds with percussive envelopes exhibit exponential decaying levels in intensity and overall lower intensity levels compared to equivalent sounds with flat envelopes. We roughly equated perceived loudness between the sequences based on informal listening. Although future research could use a more formalized approaches (i.e. through pilot studies addressing perceived loudness explicitly), we do not think the strong differences in annoyance observed can be explained by subtle potential differences in perceived loudness. Since we held all other factors (e.g., pitch, timing, rhythm) constant and only varied envelope for the trials assessing annoyance, we suggest that envelope differences are an important factor in annoyance assessments.

Although further research is needed to clarify the basis for these differences in annoyance, we propose the preference for percussive alarms can be attributed in part to their prevalence in our natural environment. Percussive tones are prevalent in the natural world; many small animals use alarm vocalizations with varying percussive envelopes to convey signals to conspecifics without revealing their location (Bradbury & Vehrencamp, 1998). In our daily interactions with our environment, percussive sounds are heard as we produce impact-type events (e.g., placing an object on a surface) and are characteristic of many instruments (e.g., xylophone, piano) that have pleasant timbres (Gabrielsson & Juslin, 1996). We have become both habituated and reliant upon percussive tones to inform us about our surroundings and guide our behaviours whereas flat tones are novel sounds we have not yet habituated to providing little meaningful information.

3.2. Learning and Memory

Given that medical professionals must recall the meaning associated with each alarm, it is important to assess whether changes in their envelope structure harm the learning and recall of alarm messages. To assess this crucial issue, each experiment included assessment of learning and memory in addition to annoyance. Our results suggest that envelope does not harm learning of IEC alarms. Across all three experiments, we found no difference in the number of blocks required to reach our learning criterion (i.e., alarm learnability) between those that learned percussive or flat alarms. Additionally, we found no difference in performance after a short break between those that learned percussive or flat alarms (i.e., short-term memory) across all three experiments. However, we find a strong trend suggesting that both learning and short-term memory for flat alarms is better than percussive alarms. We suspect that the overlapping tones present in the percussive alarms may contribute to alarm confusion leading to poorer results. Although we did not test whether non-overlapping percussive alarms harm learnability, previously Gillard and Schutz (2016) showed that non-overlapping percussive alarms do not harm learning or short-term memory.

In contrast to our results, previous work demonstrated that percussive tone sequence–object associations facilitate faster learning and retention than flat tone sequence–object associations (Schutz et al., 2017). Whereas Schutz et al. (2017) paired tone sequences with common household objects (e.g., keys, books), we paired tone sequences with their alarm–referent (i.e., alarm name). An important distinction between theses associations is that household objects are considered to be concrete objects (i.e., having physical referents) whereas most of the IEC alarm–referents are either abstract concepts (power failure) or specific patient states (blood pressure is problematic). There is generally stronger representation for concrete concepts than abstract concepts; concrete concepts are supported by greater semantic features (Plaut & Shallice, 1991), more contextual information (Schwanenflugel & Shoben, 1983), and have structurally different representational frameworks (Crutch & Warrington, 2004). In the present work, it is possible that these object categorization differences led to a null result between short-term learning and memory of sound-object association.

Despite their utility, the IEC alarms are widely recognized to be difficult to learn (Edworthy, 2011, 2013; Edworthy et al., 2017; Sanderson, Wee, & Lacherez, 2006; Wee & Sanderson, 2008). It is likely the difficulty in learning IEC alarms can be attributed to the homogeneity in the alarm tone sequences. All medium-priority IEC alarms use three-note tone sequences, are synthesized within a narrow frequency range (262-523 Hz) and have similar melodic contours (i.e., rise/fall pattern of tones). In adherence to IEC 60601-1-8 standard, we maintained the same pitch, rhythm, and timing in our stimuli. Of the nine possible permutations of rise/fall patterns for three-note tone sequences (Block et al., 2000), the IEC alarms employ only six of these permutations. Due to the importance placed on melodic contour in learning these alarm–referent pairings (Dowling & Fujitani, 1971; Gillard & Schutz, 2013), the similarity in melodic contour is one possible reason that we did not find any effect of envelope.

IEC alarms are designed with the mnemonics to aid in learning (Block et al., 2000), but we did not mention these mnemonics to participants as previous studies have documented that these mnemonics are not effective at reinforcing alarm–referent pairings (Edworthy & Hellier, 2006; Wee & Sanderson, 2008; Wee & Sanderson, 2005). One study investigating the ability of 22 critical care nurses (i.e., clinicians responsible for responding to IEC alarms) to learn and discriminate IEC alarms over two training sessions with and without mnemonics. IEC alarm identification accuracy was poor; only one nurse identified the alarm accurately after two training sessions. Furthermore, they found no difference in overall identification accuracy for those that learned the alarms with mnemonics (56%) and without mnemonics (55%) (Wee & Sanderson, 2008). It is unlikely that a lack of learning differences in our experiments can be attributed to the absence of mnemonics.

3.3. Implications

Our results show that alternative envelopes to those mandated in IEC standards such as percussive envelopes lower alarm annoyance without harming the alarm learnability. Although this insight holds potential for applied use in alarm design, some may be concerned about volume reduction; percussive alarms have less energy than flat alarms (although this does not mean they are less salient—merely that their decay allows for less overall energy in the tone).

Although contrary to the “better-safe-than-sorry” approach taken by alarm designers by making alarms as loud as possible ensuring they are heard above background noises, softer alarms are equally effective. One study found that clinicians’ performance on a cognitively-demanding and clinically-relevant task (i.e., requiring responses to alarms) when the alarm volume was +4 dB above background noise levels was comparable to when the alarm volume -11 dB below or softer than background noise levels (Schlesinger et al., 2018). Moreover, at a larger scale, reducing noise levels has potential benefits for clinicians and patients alike. In an observational study, nurses provided saliva samples and subjective stress ratings every 30 minutes while continuous measurements of sound level and heart rate were recorded for a three-hour period. Higher sound levels were predictive of subjective stress and annoyance in addition to being correlated with faster heart rates (Morrison, Haas, Shaffner, Garrett, & Fackler, 2003). Excess noise also affects a patient’s healing processes and negatively impacts sleep quality (Konkani, Oakley, & Penprase, 2014).

Alarms are ubiquitous; an analysis of the alarms at the John Hopkins Hospital indicated that on average each bed alerts 350 alarms per day (Jones, 2014). Although amplitude envelope manipulation is highly unlikely to solve the alarm fatigue problem, any improvement, even incremental is beneficial. Between 2005 and 2008, there were 566 reports of alarm-related deaths (Cvach, 2012). Through creating less annoying and ergonomic alarms, we can help reduce the prevalence of alarm fatigue and create a safer environment for both patients and clinicians.

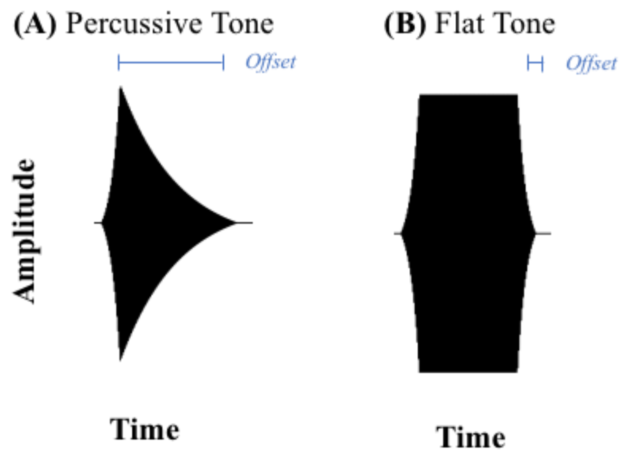

Figure 1. Sample amplitude envelope shapes. Each panel represents changes in amplitude (i.e., loudness) over time highlighting the difference in offset between (A) percussive tones and (B) flat tones.

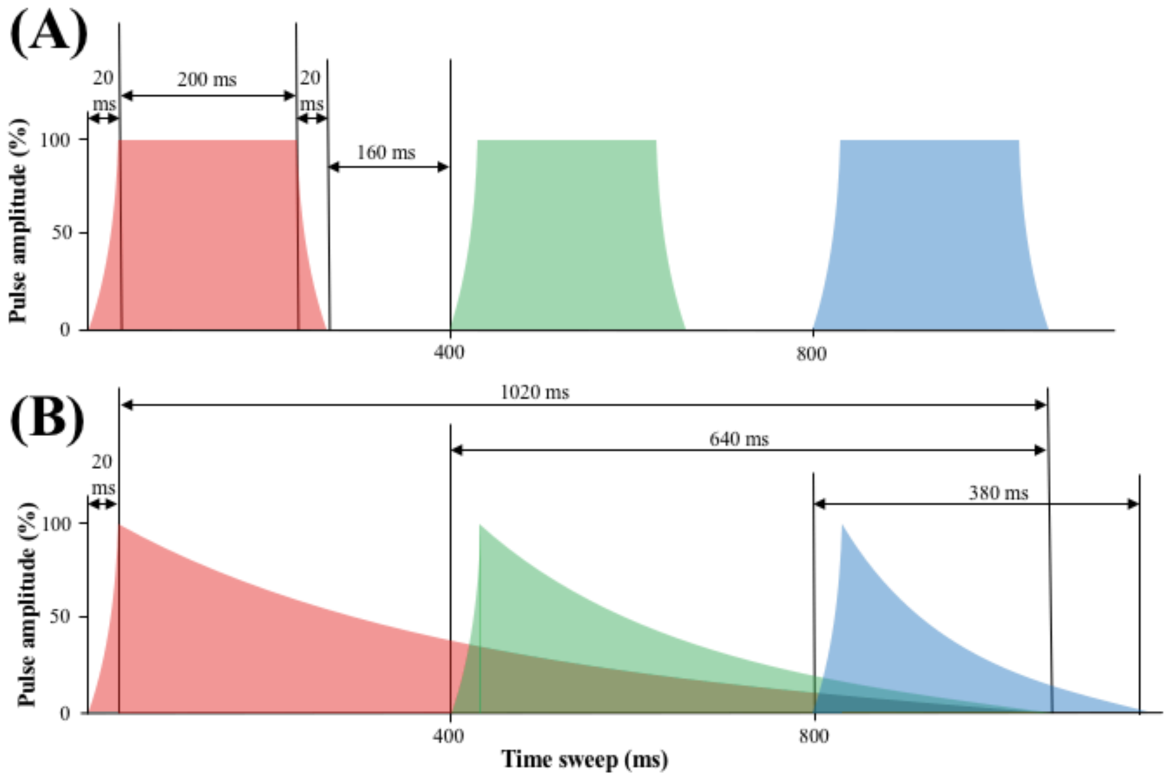

Figure 2. Timing information for (A) flat and (B) percussive alarms. Colours are used to distinguish the time and amplitude differences for each note during overlapping notes.

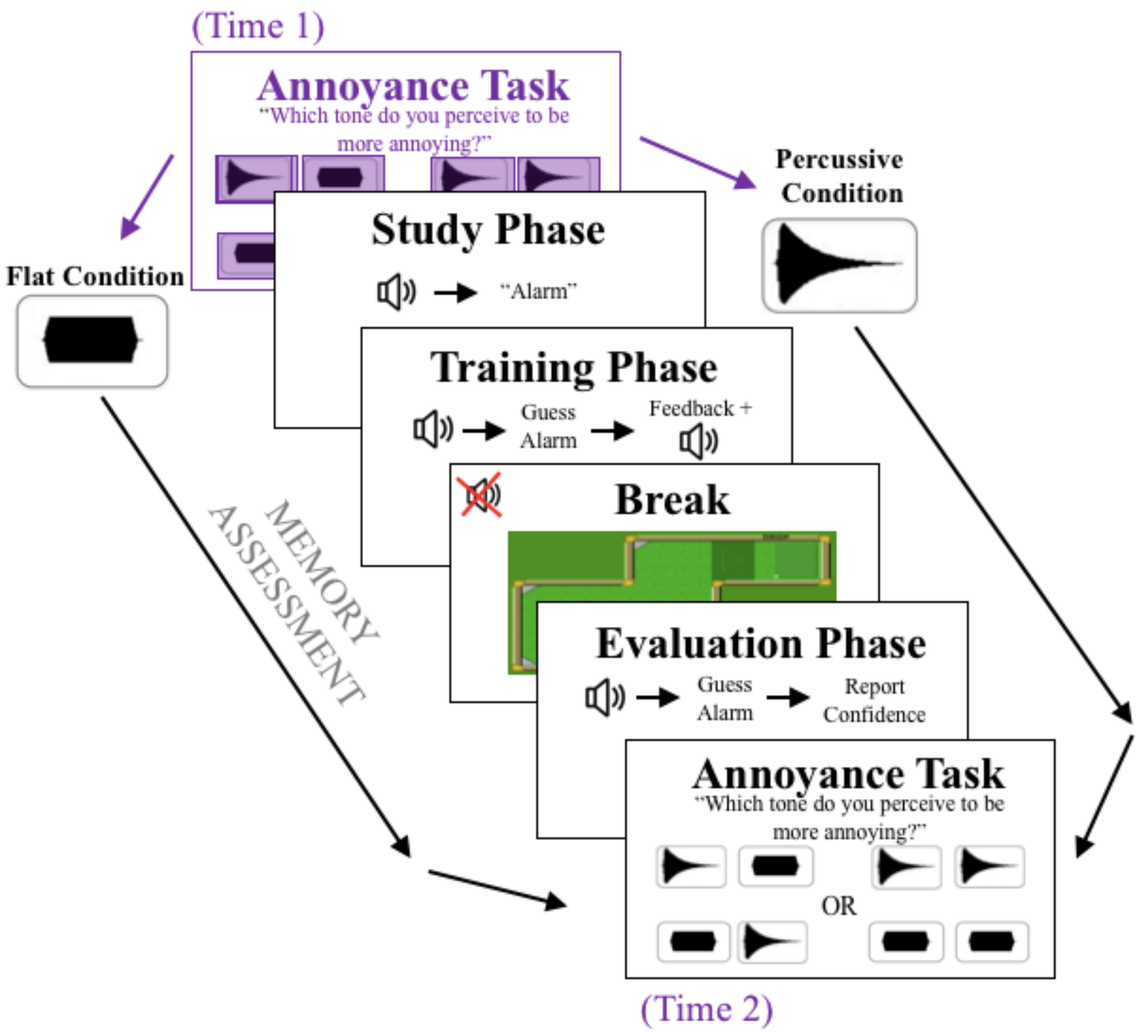

Figure 3. Overview of methods. Summary of the methods for all three experiments. Unique to experiment 3 (shown in purple) is the addition of the annoyance task before the memory assessment referred to as "Time 1" and the annoyance after memory assessment referred to as "Time 2".

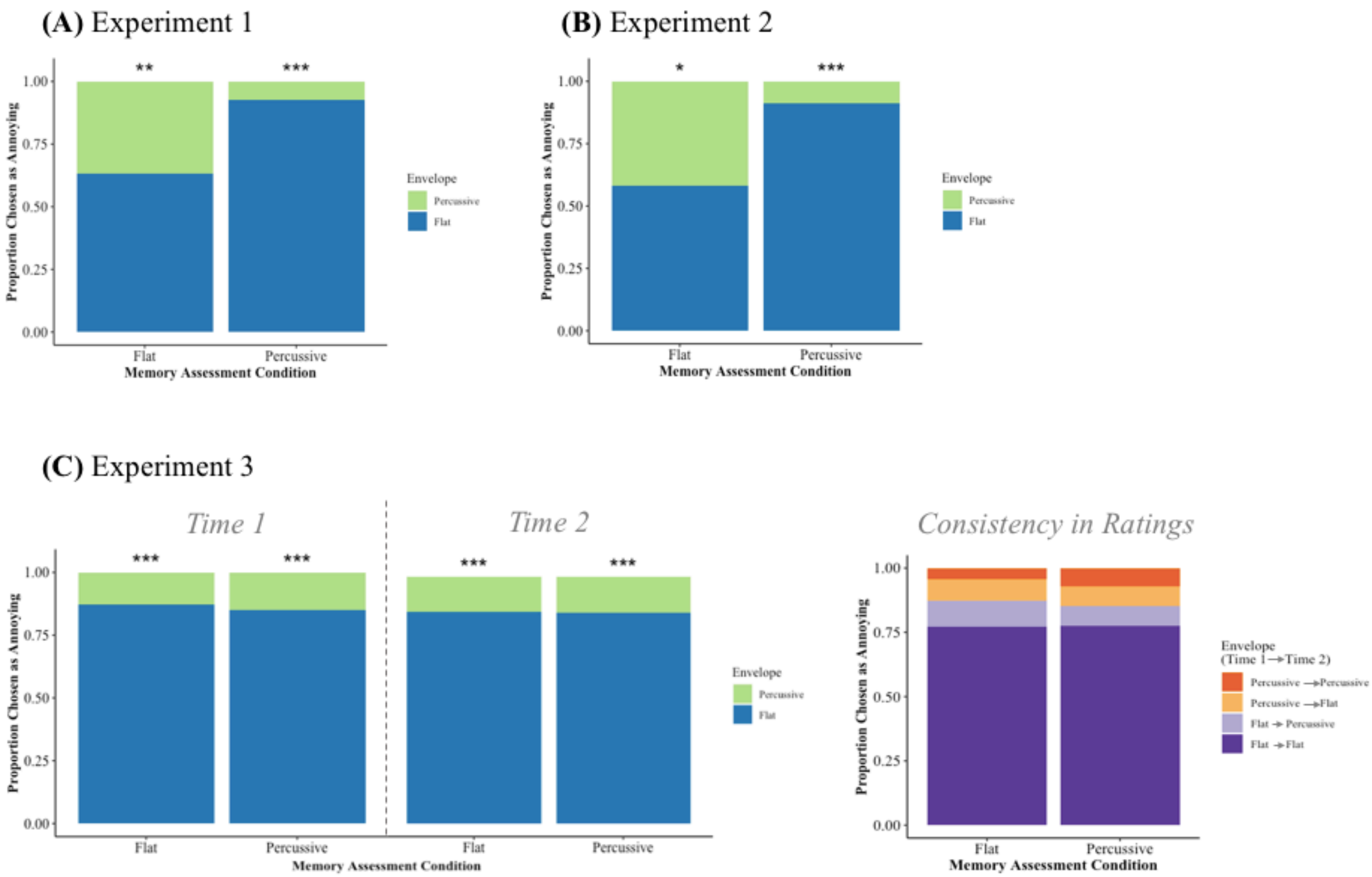

Figure 4. Annoyance task results for all experiments. In (A) experiment 1 and (B) experiment 2, participants in both flat and percussive memory assessment conditions chose the flat alarms to be more annoying than the percussive alarms. In (C) experiment 3, participants chose the flat alarms to be more annoying than the percussive alarms at both times 1 and 2. Participants were relatively consistent in their annoyance ratings, with no significant changes in envelope choices from time 1 to time 2 based on memory assessment condition. *$p < .05$, **$p < .01$, ***$p < .001$.

References

Alert fatigue leads to OR fatalities. (2010). *Same - Day Surgery*. Retrieved from http://libaccess.mcmaster.ca/login?url=https://search.proquest.com/docview/814644034?accountid=12347

Block, F. E., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing*, *15*(2), 75–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12578080

Block, F. E., Rouse, J. D., Hakala, M., & Thompson, C. L. (2000). A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. *Journal of Clinical Monitoring and Computing*, *16*(7), 541–546. https://doi.org/10.1023/A:1011471607530

Bradbury, J. W., & Vehrencamp, S. L. (1998). Principles of animal communication.

Chuen, L., & Schutz, M. (2016). The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: The role of spectral and amplitude cues. *Attention, Perception, & Psychophysics*, *78*(5), 1512–1528. https://doi.org/10.3758/s13414-016-1088-5

Cleophas, T. J. (1999). Carryover Effects in Clinical Research BT - Human Experimentation: Methodologic issues fundamental to clinical trials. In T. J. Cleophas (Ed.) (pp. 25–36). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-4663-0_4

Comission, I. E. (2006). International Standard IEC 60601: Medical electrical equipment. *Part 1-8: General Requirements for Safety. Collateral Standard: General Requirements, Tests and Guidance for Alarm Systems in Medical Electrical Equipment and Medical Electrical Systems*.

Crutch, S. J., & Warrington, E. K. (2004). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615–627. https://doi.org/10.1093/brain/awh349

Cvach, M. (2012). *Monitor Alarm Fatigue : An Integrative Review*. *Biomedical instrumentation & technology / Association for the Advancement of Medical Instrumentation* (Vol. 46). https://doi.org/10.2345/0899-8205-46.4.268

Donchin, Y., Gopher, D., Olin, M., Badihi, Y., Biesky, M., Sprung, C. L., … Cotev, S. (2003). A look into the nature and causes of human errors in the intensive care unit. *Quality and Safety in Health Care*, *12*(2), 143–148.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *The Journal of the Acoustical Society of America*, *49*(2B), 524–531. https://doi.org/10.1121/1.1912382

Edworthy, J. (2011). Designing effective alarm sounds. *Biomedical Instrumentation & Technology*, *45*(4), 290–294.

Edworthy, J. (2013). Medical audible alarms: A review. *Journal of the American Medical Informatics Association*, *20*(3), 584–589. https://doi.org/10.1136/amiajnl-2012-001061

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality and Safety in Health Care*, *14*(3), 212–215. https://doi.org/10.1136/qshc.2004.013052

Edworthy, J., & Hellier, E. (2006). Alarms and human behaviour: implications for medical alarms. *British Journal of Anaesthesia*, *97*(1), 12–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16698858

Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., … Pye, E. (2017). The

Recognizability and Localizability of Auditory Alarms: Setting Global Medical Device Standards. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *59*(7), 1108–1127. https://doi.org/10.1177/0018720817712004

Fagerlonn, J. (2011). Urgent alarms in trucks: Effects on annoyance and subsequent driving performance. *IET Intelligent Transport Systems*, *5*(4), 252–258. https://doi.org/10.1049/iet-its.2010.0165

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, *24*(1), 68–91. Retrieved from http://pom.sagepub.com/content/24/1/68.short

Gillard, J., & Schutz, M. (2013). Exploring melodic structure to increase heterogeneity of auditory alarm sets in medical devices. *Jounal of Medical Devices*, *7*(2), 12–13. https://doi.org/10.1115/1.4024309

Gillard, J., & Schutz, M. (2016). Composing alarms: considering the musical aspects of auditory alarm design. *Neurocase: The Neural Basis of Cognition*, *22*(6), 566–576. https://doi.org/10.1080/13554794.2016.1253751

Giordano, B. L., Rocchesso, D., & McAdams, S. (2010). Integration of acoustical information in the perception of impacted sound sources: The role of information accuracy and exploitability. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 462–476.

Grassi, M., & Casco, C. (2009). Audiovisual bounce-inducing effect: Attention alone does not explain why the discs are bouncing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 235–243. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19170485

Guski, R., Felscher-Suhr, U., & Schuemer, R. (1999). The Concept of Noise Annoyance: How International Experts See It. *Journal of Sound and Vibration*, *223*(4), 513–527. Retrieved from http://libaccess.mcmaster.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.31BAC0DC&site=eds-live&scope=site

Jones, K. (2014). Alarm fatigue a top patient safety hazard. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *186*(3), 178. https://doi.org/10.1503/cmaj.109-4696

Konkani, A., Oakley, B., & Penprase, B. (2014). Reducing Hospital ICU Noise: A Behavior-Based Approach. *Journal of Healthcare Engineering*, *5*(2), 229–246. https://doi.org/10.1260/2040-2295.5.2.229

Lutfi, R. A. (2001). Auditory detection of hollowness. *The Journal of the Acoustical Society of America*, *110*(2), 1010–1019.

Martindale, C., & Moore, K. (1990). Intensity, Dissonance, and Preference for Pure Tones. *Empirical Studies of the Arts*, *8*(2), 125–134. https://doi.org/10.2190/9X1D-0DDB-QRPH-AQQR

Morrison, W. E., Haas, E. C., Shaffner, D. H., Garrett, E. S., & Fackler, J. C. (2003). Noise, stress, and annoyance in a pediatric intensive care unit. *Critical Care Medicine*, *31*(1), 113–119.

Ng, M., & Schutz, M. (2017). Seeing sound: A new tool for teaching music perception principles. *Canadian Acoustics*, *45*(3). Retrieved from https://jcaa.caa-aca.ca/index.php/jcaa/article/view/3086

Patterson, R. D. (1982). Guidelines for Auditory Warning Systems on Civil Aircraft, (November

1982), 107. https://doi.org/10.1001/archderm.143.8.1080

Plaut, D. C., & Shallice, T. (1991). Effects of word abstractness in a connectionist model of deep dyslexia. In *Proceedings of the 13th annual meeting of the Cognitive Science Society* (pp. 73–78). Erlbaum Hillsdale, NJ.

Sanderson, P., Wee, A. N., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia*, *61*(2), 142–147.

Schlesinger, J. J., Baum Miller, S. H., Nash, K., Bruce, M., Ashmead, D., Shotwell, M. S., … Weinger, M. B. (2018). Acoustic features of auditory medical alarms—An experimental study of alarm volume. *The Journal of the Acoustical Society of America*, *143*(6), 3688–3697. https://doi.org/10.1121/1.5043396

Schutz, M., & Bamikole, P. (2015). Finding unity in impact events: Amplitude envelope and multi-modal integration. In S. Pecknold & M. Kiefte (Eds.), *Acoustics Week in Canada*. Halifax, Nova Scotia.

Schutz, M., & Kubovy, M. (2009). Deconstructing a musical illusion: Point-light representations capture salient properties of impact motions. *Canadian Acoustics*, *37*(1), 23–28.

Schutz, M., Stefanucci, J., Baum, S. H., & Roth, A. (2017). Name that percussive tune: Associative memory and amplitude envelope. *The Quarterly Journal of Experimental Psychology*, *70*(7), 1323–1343. https://doi.org/http://dx.doi.org/10.1080/17470218.2016.1182562

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 82.

Sendelbach, S., & Funk, M. (2013). Alarm fatigue: A patient safety concern. *AACN Advanced*

Critical Care, *24*(4), 378–386. https://doi.org/10.1097/NCI.0b013e3182a903f9

Sodnik, J., & Tomazic, S. (2015). *Spatial auditory human-computer interfaces*. https://doi.org/10.1007/978-3-319-22111-3

Vallet, G., Shore, D. I., & Schutz, M. (2014). Exploring the role of amplitude envelope in duration estimation. *Perception*, *43*(7), 616–630.

Wee, A. N., & Sanderson, P. (2008). Are melodic medical equipment alarms easily learned? *Anesthesia & Analgesia*, *106*(2), 501–508. https://doi.org/10.1213/01.ane.0000286148.58823.6c

Wee, A., & Sanderson, P. (2005). Testing New Alarms for Medical Electrical Equipment BT - Human-Computer Interaction - INTERACT 2005. In M. F. Costabile & F. Paternò (Eds.) (pp. 1146–1149). Berlin, Heidelberg: Springer Berlin Heidelberg.

Wilcox, S. B. (2011). A Human Factors Perspective: Auditory Alarm Signals. *Biomedical Instrumentation & Technology*, *45*(4), 284–289. https://doi.org/10.2345/0899-8205-45.4.284