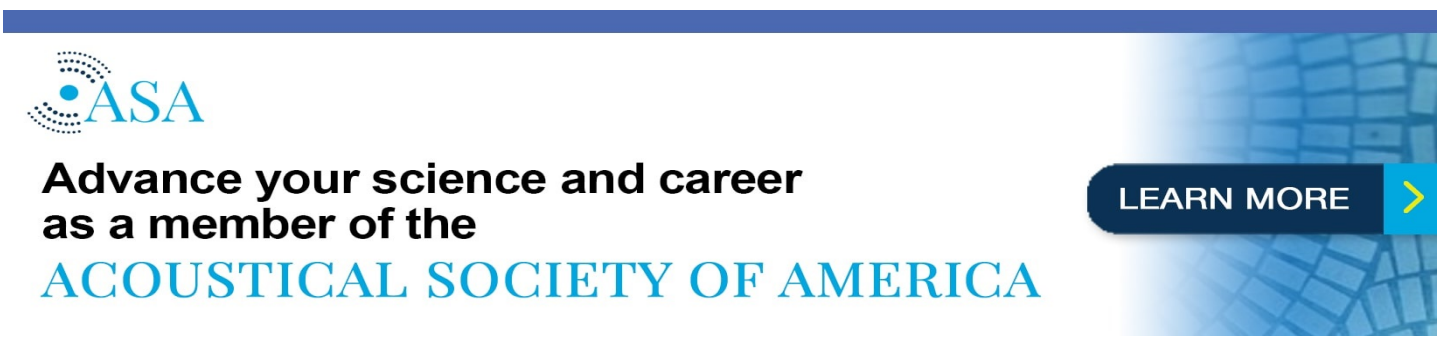
More detectable, less annoying: Temporal variation in amplitude envelope and spectral content improves auditory interface efficacy

Liam Foley, Joseph Schlesinger and Michael Schutz

Citation: The Journal of the Acoustical Society of America **151**, 3189 (2022); doi: 10.1121/10.0010447
View online: https://doi.org/10.1121/10.0010447
View Table of Contents: https://asa.scitation.org/toc/jas/151/5
Published by the Acoustical Society of America



JASA ARTICLE

More detectable, less annoying: Temporal variation in amplitude envelope and spectral content improves auditory interface efficacy

Liam Foley,[1,a] Joseph Schlesinger,[2,b] and Michael Schutz[3,c]
[1]Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Canada
[2]Anesthesiology Critical Care Medicine, Biomedical Engineering, Vanderbilt University Medical Center, Nashville, Tennessee 37212, USA
[3]School of the Arts, McMaster University, Hamilton, Canada

ABSTRACT:
Auditory interfaces, such as auditory alarms, are useful tools for human computer interaction. Unfortunately, poor detectability and annoyance inhibit the efficacy of many interface sounds. Here, it is shown in two ways how moving beyond the traditional simplistic temporal structures of normative interface sounds can significantly improve auditory interface efficacy. First, participants rated tones with percussive amplitude envelopes as significantly less annoying than tones with flat amplitude envelopes. Crucially, this annoyance reduction did not come with a detection cost as percussive tones were detected more often than flat tones—particularly, at relatively low listening levels. Second, it was found that reductions in the duration of a tone's harmonics significantly lowered its annoyance without a commensurate reduction in detection. Together, these findings help inform our theoretical understanding of detection and annoyance of sound. In addition, they offer promising original design considerations for auditory interfaces. © *2022 Acoustical Society of America*. https://doi.org/10.1121/10.0010447

(Received 2 November 2021; revised 21 April 2022; accepted 22 April 2022; published online 12 May 2022)
[Editor: James F. Lynch]] Pages: 3189–3196

I. INTRODUCTION

Understanding the acoustic factors used in parsing competing acoustic signals provides insight to the auditory system, as well as useful guidance in designing effective auditory interfaces—systems for communicating with electronic devices using sound. Auditory interfaces can provide confirmatory auditory feedback (Bodendörfer *et al.*, 2015), express data through sonification (Taenzer *et al.*, 2010), and auditory notifications through alerts and alarms (Phansalkar *et al.*, 2010). Medical device alarms are a particularly high-value auditory interface whose problems have been extensively documented (Edworthy, 2013). They are commonly designed to sound often and loudly in an attempt to ensure detection (Edworthy and Hellier, 2005; Schlesinger *et al.*, 2018); however, their numerosity leads to serious issues of annoyance undermining their efficacy (Rayo *et al.*, 2019).

Recent investigations into the use of temporal variation in amplitude envelope, the way a sound's amplitude changes over time, has shown promise in improving auditory alarm efficacy, reducing annoyance without harming learning or memory of alarms (Sreetharan *et al.*, 2021). Here, we build on that work by exploring how temporal variation in amplitude envelope (flat and percussive envelopes) and spectral content (differing durations of harmonic energy) affect not only annoyance but also a core issue of auditory perception—tone detection.

A. Auditory alarms: Detectable yet annoying

Auditory alarms are invaluable in many safety critical industries, including (but not limited to) industrial processing (Laberge *et al.*, 2014), aviation (Bliss, 2003), automobiles (Marshall *et al.*, 2007), railways (Edworthy *et al.*, 2011), and healthcare (Edworthy, 2013). These alarms are typically easily detectable in isolation; however, common design philosophy dictates that loud alarms must sound often to be detected (Schlesinger *et al.*, 2018), leading to an unmanageable sonic cacophony in their real world use (Sendelbach and Funk, 2013).

The sheer numerosity of alarms prevalent in healthcare (Varpio *et al.*, 2012) and industrial settings (Laberge *et al.*, 2014) exacerbates problems with loud volume levels with reports of over 350 alarms per-patient per-day (Sendelbach and Funk, 2013). Other reports have shown that only 0.5% indicate life threatening events (O'Carroll, 1986). Numerous alarms combined with low positive predictive value create an unfortunate vicious cycle. Making alarms annoying helps reduce misses, but large numbers of attention-grabbing alarms incentivize users to tune them out or even turn them off—occasionally resulting in disastrous consequences (Bliss *et al.*, 1995; Block *et al.*, 1999), such as one patient who died despite a warning alarm that was sounding for over an hour (Cvach, 2012; Sendelbach and Funk, 2013).

B. Role of temporal structure in recognition and perception

Temporal variation is a prominent feature in environmental (Foley and Schutz, 2021; Gygi *et al.*, 2004) and

[a]Electronic mail: ltimfoley@protonmail.com
[b]Also at: Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN 37212, USA.
[c]Also at: Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Canada.

musical (Schutz, 2019) sounds. These temporal fluctuations play an important role in stream segregation, our ability to perceive multiple distinct auditory units (Bendixen *et al.*, 2010; Iverson, 1995; Moore and Gockel, 2002), and masking, which occurs when one sound renders another inaudible (Buus, 1985). It also is important for recognition of speech (Shannon *et al.*, 1995) and environmental sounds (Gygi *et al.*, 2004), even in the absence of detailed spectral information.

As many natural sounds exhibit great temporal variation, our auditory system often uses this information to aid in understanding our environments. For example, temporal properties are important in modulating acoustic startle responses (ASR) in humans, a low-level reflexive response to acoustic stimuli. As a sound's rise time, or onset, increases, the amplitude and probability of ASR decreases (Blumenthal, 1988). Biases toward specific envelopes may reflect evolutionary pressures as looming sounds (i.e., those that are increasing intensity) are perceived as closer in space to the listener than receding sounds, which are decreasing intensity (Neuhoff, 2001). This could reflect an adaptive bias toward approaching (e.g., predator, falling tree) vs receding objects.

Amplitude envelope appears to play an integral role in many tasks, such as audio-visual integration (Chuen and Schutz, 2016; Grassi and Casco, 2009; Grassi and Pavan, 2012; Schutz and Kubovy, 2009), duration assessment (Grassi and Darwin, 2006; Vallet *et al.*, 2014), determining materiality in impact events (Giordano and McAdams, 2006), product preferences (Schutz and Stefanucci, 2019), and associative memory (Schutz *et al.*, 2017). Exploring its role in both detection in noise and annoyance will add to this growing theoretical understanding of envelope, in addition to applied uses, such as auditory interfaces.

C. Role of harmonic structure in detection and perception

Harmonically "complex" tones are known to be more detectable than pure tones (Buus *et al.*, 1997; McPherson *et al.*, 2020). Additionally, harmonic sounds (such as those consisting of the fundamental plus four harmonic components) are more detectable than otherwise equivalent inharmonic tones (McPherson *et al.*, 2020), suggesting that advantages in detection may not simply be related to more acoustic energy but rather the presence and relation of spectral components.

Natural sounds exhibit innumerable complexities in spectral content. Musical sounds, for example, often exhibit temporal shifts in harmonic content, where the spectral energy is shifting over the course of a tone's duration (Schutz, 2019). These differences in energy concentration are associated with a different perceptual center (*p*-center; Howell, 1988), which is known to play an important role in the timing of musical (Danielsen *et al.*, 2019) and speech (Marcus, 1981; Morton *et al.*, 1976) sounds. However, its role in tone detection has not been examined.

D. Present experiment

Here, we explore how detectability and perceived annoyance of sounds are affected by changes in their amplitude envelope and harmonic content. This manipulation is essentially a variation on "delayed harmonics," where a tone's fundamental sounds before its harmonics (Edworthy *et al.*, 1991). Additionally, we explore whether changes to a sound's envelope, reducing annoyance, affect their detectability. Based on our team's previous work (Anderson, 2019; Sreetharan *et al.*, 2021), we hypothesize that percussive sounds will be perceived as less annoying than traditional flat tones.

II. METHODS

A. Experiment 1

1. Participants

Introductory psychology students (83) at a large university in Ontario, Canada received course credit in return for their participation. All participants self-reported normal hearing and gave informed consent prior to the study in accordance with the university ethics board protocols.

2. Apparatus

We created the task in *Psychopy* version 2020.1.3 (Peirce *et al.*, 2019) on *Pavlovia*, an online study hosting service developed by *Psychopy* (Peirce *et al.*, 2019). Participants completed the experiment using their own computers and headphones.

3. Stimuli

Our experiment used two types of stimuli: (a) target tones and (b) speech noise. We created target tone stimuli in MAESTRO (Ng and Schutz, 2017), a graphical user interface (GUI)-based sound creation program built on top of the sound synthesis programming language Supercollider (McCartney, 1996). The tones consisted of a 400 Hz fundamental with five harmonics (800, 1200, 1600, 2000, and 2400 Hz). This frequency was chosen in relation to applications such as medical alarms standards, which have fundamental frequencies from 262 to 523 Hz (IEC, 2006). Percussive and flat tones exhibited a 5 ms onset; however, percussive tones exponentially decay (for all harmonics), whereas flat tones sustained until 5 ms before offset at which point they decayed. Based on previous findings that flat tones sound longer in duration than equivalent duration percussive tones (Vallet *et al.*, 2014), we created 600 ms percussive tones and 360 ms flat tones to roughly equate this. To control for differences in perceived loudness [see Bruder *et al.*, 2022 and Wong *et al.*, 2007 for similar uses of root-mean-squared (RMS) equating] between percussive and flat tones, we synthesized each tone with a total RMS amplitude of −27.3 dB (unweighted) relative to 0 dB (20 μPa).

Traditional approaches to synthesizing complex sounds (including our team's previous work) synthesize components

https://doi.org/10.1121/10.0010447

varying in lock-step with the fundamental, i.e., rising and falling at the same time and in the same proportion across the tone's duration. We varied the length of overtones relative to the fundamental. Specifically, we synthesized the harmonics with five lengths relative to the fundamental's duration: 100%, 50%, 25%, 10%, and 5%. We matched the amplitudes of all harmonics to the tone's fundamental at onset. Therefore, flat tones exhibited all components in equal amplitudes for either 5%, 10%, etc. of the tone at which point, only the fundamental sounded for the remainder of the tone's nominal duration. For the percussive tones, harmonic components decayed after 5%, 10%, etc. of the tone's duration, while the fundamentals amplitude had a half life of 120 ms, creating more dynamic variation in the relative mix of the components as the tone sounded. Figure 1 shows a three-dimensional (3D) spectrogram of the flat-100%, percussive (perc)-100%, flat-50%, and perc-50%.

We presented the tones at three signal-to-noise ratios (SNRs; 4 dB, −11 dB, and −27 dB). The 4 dB SNR is illustrative of the current status quo in many alarm implementations where alarms are designed to sound loudly, while −11 dB, and −27 dB were used to test audibility below the level of the background noise. Similar to past approaches (Bruder *et al.*, 2022; Schlesinger *et al.*, 2018), these latter two conditions assess signal detection at lower volumes, which has, in some cases, led to missed signals and adverse patient outcomes (Cvach, 2012).

We used background noise based on speaker-babble from track 24 of the QuickSIN test, a speech discriminability test designed to test speech comprehension (Killion *et al.*, 2004). The babble consisted of four speakers (one male and three females) speaking English, and was designed to simulate background speaking in a crowded space. We combined the target tone and speech noise stimuli in Amadeus Pro sound editor (HairerSoft, 2019).

B. Procedure

1. Experimental procedure

After giving informed consent, participants completed a two-part experiment assessing (i) each tone's detectability in noise and (ii) relative annoyance within pairs of tones (drawn from those used in the first task). For the detectability task, participants completed an un-speeded, two interval forced choice (2IFC) tone detection task, indicating whether the target tone appeared in interval "1" or "2" of the stimulus. Each interval was indicated by the appearance of a "1" or "2" on the computer screen. We generated 60 unique target tones by factorially combining envelope (percussive or flat), harmonic duration (5%, 10%, 25%, 50%, 100%), SNR (4 dB, −11 dB, −27 dB) and interval ("1" or "2"). We presented each target tone twice, resulting in 120 experimental trials. Participants first heard samples of the tone and noise used in the experiment and completed two sample trials before the experimental trials.

For the second task (annoyance), participants indicated which of the tones within a pair presented individually was most annoying. Participants rated annoyance in two blocks. In the "within" harmonic block, participants compared tones matched in harmonic durations (e.g., flat-100% harmonics vs perc-100% harmonics). Participants made ten comparisons, where each of the five harmonic matched-pairs was presented twice, with each tone presented in balanced order (e.g., perc-flat, then flat-perc). In the "full" comparison block, participants compared two randomly selected tones from all possible combinations of envelope and harmonic

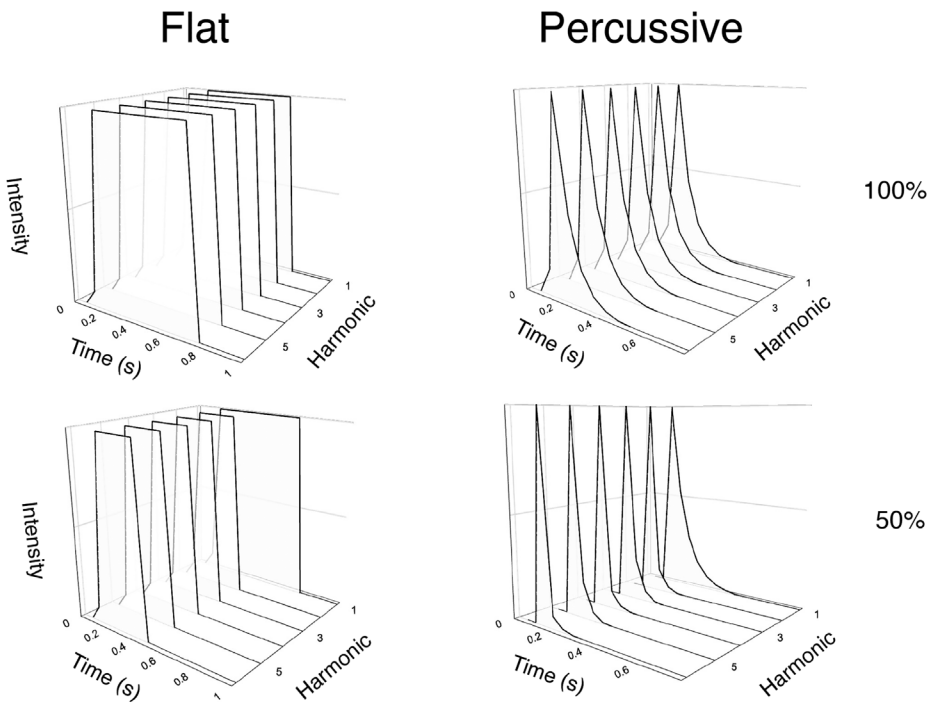


FIG. 1. The 3D spectrograms of four example stimuli (flat-100%, perc-100%, flat-50%, and perc-50%) are shown.

condition in ten comparison trials. We randomized the order of the between and random blocks across the participants.

C. Experiment 2

The within annoyance comparisons in the first experiment clearly showed that percussive tones are less annoying than their flat counterparts. However, that experiment offered less insight into relative annoyance across different harmonic durations as the full comparison block included only a randomized subset of all comparisons due to time constraints. To gain a clearer picture of relative annoyance, we conducted a fully balanced annoyance rating experiment pairing each of the ten tone types with one another. An additional 59 undergraduate students received course credit in return for their participation. All participants gave informed consent prior to the study in accordance with the university ethics board protocols. This experiment consisted solely of the fully balanced annoyance task. As in experiment 1, we factorially combined two types of envelope (percussive or flat) and five harmonic durations (5%, 10%, 25%, 50%, 100%), resulting in ten tones. We presented comparisons of each tone with the other nine tones twice, balancing presentation order. In contrast to experiment 1, which contained 120 signal detection trials and 20 annoyance trials, this experiment contained 90 annoyance trials and no signal detection trials.

III. RESULTS

A. Experiment 1

1. Signal detection

We used an analysis of variance (ANOVA) and generalized linear model (GLM) to independently assess the effects of different sound synthesis approaches on annoyance and detectability. These analyses give novel insight into the relative contributions of amplitude envelope and harmonic duration to these two dimensions of sound, which are important in improving auditory interface design and enhancing our understanding of basic auditory perception.

2. Detectability improves with percussive envelopes

We analyzed the effects of envelope and SNR on tone detection (Fig. 2) with a two way ANOVA, revealing a significant main effect of envelope, $F(1,9954) = 56.685$, $p < 0.001$, and SNR, $F(2,9954) = 676.360$, $p < 0.001$ on tone detection. Additionally, we found an interaction between these factors $F(2,9954) = 56.772$, $p < 0.001$. Given the interaction, we explored envelope's simple main effect at each SNR. Although we found no effect of envelope at 4 dB SNR and −11 dB SNR, we observed a significant effect at −27 dB SNR, $F(1,9954) = 172.8161$, where response accuracy for flat tones [$M = 0.56$, standard deviation (SD) = 0.49] was significantly lower than for percussive tones ($M = 0.72$, SD = 0.45).

Tukey-honest significant difference (HSD) *post hoc* comparisons on SNR revealed a significant difference between the −27 dB ($M = 0.64$, SD = 0.48) and −11 dB ($M = 0.91$, SD = 0.29; $p < 0.001$) conditions, as well as between the −27 dB ($M = 0.64$, SD = 0.48) and 4 dB ($M = 0.93$, SD = 0.26; $p < 0.001$) condition but not the 4 dB and −11 dB conditions ($p = 0.1$).

3. Detectability affected by harmonic duration

We analyzed the effects of envelope and harmonic condition on tone detection (Fig. 3) using a two way ANOVA, which found a significant effect of the main effects of envelope $F(1,9950) = 49.65 < 0.001$ and harmonic condition $F(4,9950) = 11.27$, $p < 0.001$. The envelope harmonic interaction was not significant $F(4,9950) = 1.99$, $p = 0.09$.

Tukey-HSD *post hoc* comparisons on the harmonic conditions collapsed across envelope, which revealed significantly worse detectability in the 5% harmonic condition ($M = 0.78$, SD = 0.41) compared to the 25% harmonic condition ($M = 0.85$, SD = 0.36; $p < 0.001$), 50% harmonic condition ($M = 0.85$, SD = 0.36; $p < 0.001$), and 100% harmonic

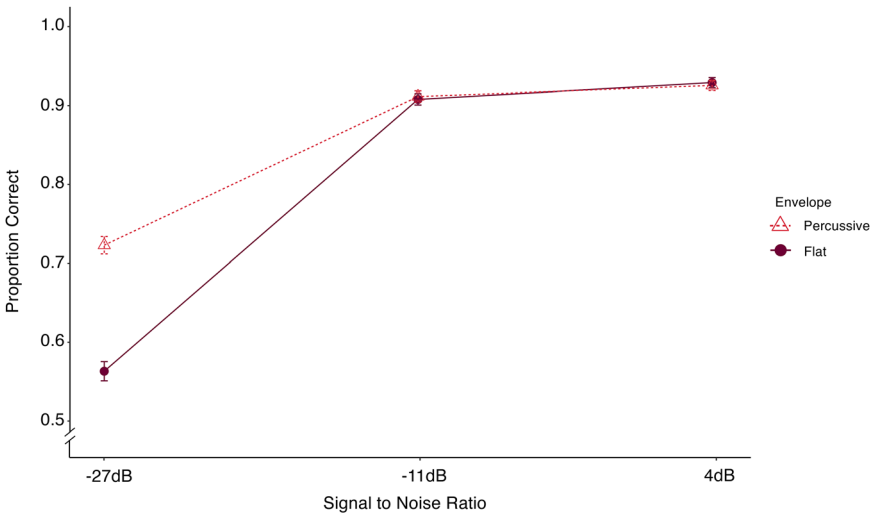


FIG. 2. (Color online) The proportion of correct responses in the signal detection task by envelope and SNR. The error bars denote the standard error.

https://doi.org/10.1121/10.0010447

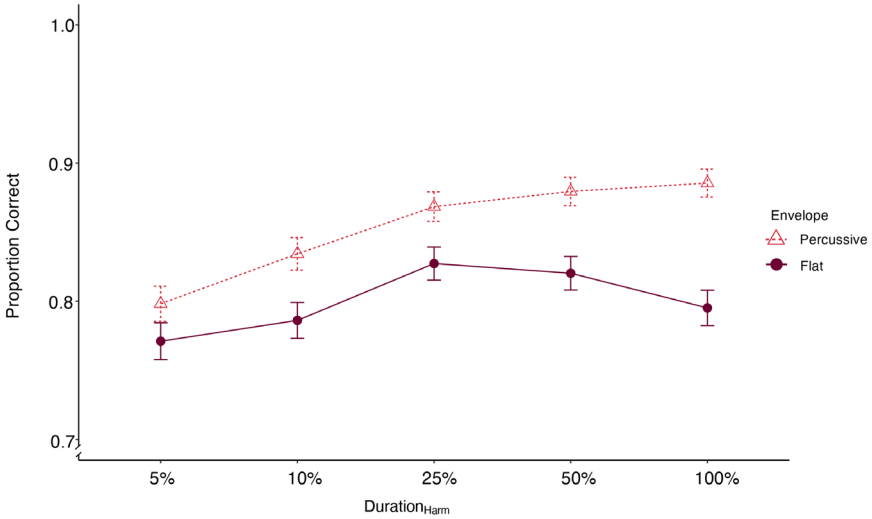


FIG. 3. (Color online) The proportion of correct responses in signal detection task by envelope and harmonic duration condition. The error bars denote the standard error.

condition ($M = 0.84$, SD $= 0.37$; $p < 0.001$), as well as between the 10% and 50% conditions ($p < 0.05$). We found no other significant differences in detection as a result of harmonic condition, indicating harmonic durations beyond 25% did not significantly improve detectability.

4. Annoyance affected by envelope and harmonic duration

To gain insight into the effect of envelope on annoyance (Fig. 4), we first evaluated its effect in each of the two conditions using separate chi square tests. In comparison to traditional flat tones, percussive tones received significantly lower annoyance ratings in the within harmonic $X^2(1,N = 83) = 34.005$, $p < 0.001$, and full comparison conditions $X^2(1,N = 83) = 17.349$, $p < 0.001$. Second, we evaluated the effect of harmonic duration on annoyance in the full condition using a GLM with a Poisson link function, employing the 5% harmonic as the reference condition for the harmonic variable and flat as the reference for the envelope variable. In relation to the baseline of the 100% harmonic condition and adjusting for envelope, the rate of annoyance significantly decreased for the 50% (19.3% decrease, $p = 0.03$), 25% (30.7% decrease, $p \leq 0.001$), 10% (39.0% decrease, $p \leq 0.001$), and 5% (46.9% decrease, $p \leq 0.001$) conditions. The percussive envelope had a significant decrease (25.2%, $p < 0.001$) of annoyance when compared to flat envelopes after adjusting for harmonics.

B. Experiment 2

The second experiment offered each participant the ability to evaluate every comparison as an independent experiment rather than comparisons done "within" envelope shapes or subsets of full comparisons as in the first experiment. We evaluated the effect of harmonic duration and envelope on annoyance (Fig. 5) with a GLM with a Poisson link function, using the 100% harmonic condition as the reference for the harmonic variable and percussive as the reference condition for the envelope variable. Consistent with experiment 1 in relation to the 100% harmonic condition and adjusting for envelope, the rate of annoyance

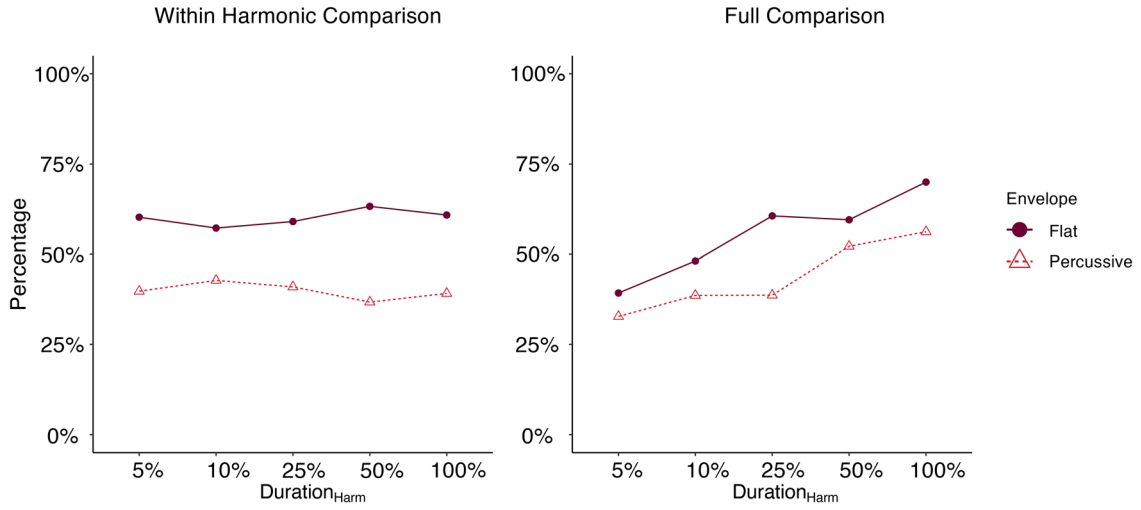


FIG. 4. (Color online) The annoyance ratings across "within harmonic comparisons" and "full comparisons" are shown.

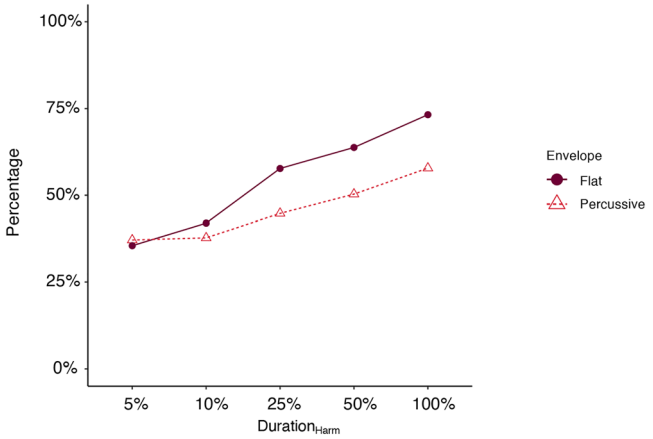


FIG. 5. (Color online) The annoyance ratings in the fully balanced condition are depicted.

significantly decreased for the 50% (12.9% decrease, $p < 0.001$), 25% (21.8% decrease, $p < 0.001$), 10% (39.2% decrease, $p < 0.001$), and 5% (44.6% decrease, $p < 0.001$) conditions. The percussive envelope had a significant decrease in the rate of annoyance ratings when compared to flat envelopes (16.3% decrease, $p < 0.001$) after adjusting for harmonics.

IV. DISCUSSION

Our experiments offer new insight into how temporal variation (of a tone as well as its spectral content) can both improve auditory interface efficacy and offer insight into a basic auditory process—signal detection. For clarity, we will organize our discussion around the two main manipulations (amplitude envelope and harmonic content), discussing, in turn, implications for applied work and basic theoretical understanding.

A. Amplitude envelope

The quest to identify better candidate sounds for auditory alarms provided the primary motivation for this project, where tone annoyance remains a stubbornly common problem (Sreetharan *et al.*, 2021) despite decades of research (Block *et al.*, 1999; Gaver, 1997). Our data suggesting that moving beyond the simplistic temporal structures of current sounds can lower annoyance while *improving detection* in some circumstances holds significant potential. Although further research is needed, we suspect the greater temporal fluctuations in percussive tones offer increased contrast and segregation between target and noise.

In addition to how temporal variation in amplitude envelope can aid signal *detection*, we believe it could also improve *differentiation* between different alarms. Many alarms are temporally invariant, and we suspect this similarity in their amplitude envelopes plays a role in well-documented problems with alarm confusions (Edworthy *et al.*, 2017) and masking (Rayo *et al.*, 2019). As these problems are not limited to medical alarms, other auditory interfaces, such as input confirmation signals (Bodendörfer *et al.*, 2015) and data sonification, may also be improved through embracing this temporal variability.

In terms of theoretical implications, our results are consistent with work showing that complexities and idiosyncrasies in the temporal structure of musical sounds aid segregation (Iverson, 1995), where acoustic parameters such as attack time and the shape of attack are known to aid in differentiation. They complement previous findings showing envelope differences aid in differentiation between spectrally matched sounds (Cusack and Roberts, 2000). Additionally, the literature on ASR also provides useful context for the superior detection of percussive tones at low listening levels. Temporal parameters, such as attack time, can affect the probability, amplitude, and latency of ASR with effects evident at low stimulus intensities (Blumenthal, 1988).

Greater sensitivity to temporal changes has interesting connections to a body of work on differential sensitivity to so-called rising vs falling tones, which are perceived as differing in intensity changes (Neuhoff, 1998) and perceived distance (Neuhoff, 2001) despite being acoustic mirror images. Although our detection advantage may be indicative of similar perceptual biases, it is important to note that rising and falling tones mimic a sustained source either approaching or receding, whereas percussive tones are created by (and are generally perceived as) a stationary impact event (Schutz, 2016).

B. Harmonic content

Our second question explored the role of harmonic duration in detection as well as annoyance. Specifically, whether brief "splashes" of harmonics could achieve some of the detectability benefits of full-duration harmonics (i.e., normative approaches to sound synthesis). We found the 25% harmonic condition a potentially beneficial variation on traditional approaches with temporally invariant sounds. Specifically, reducing harmonic duration from 100% to 25% lowers annoyance (Fig. 5) without harming detection (Fig. 3). This suggests a new approach for addressing widespread concerns regarding alarm annoyance, without harming alarm detectability. Intriguingly, these benefits appear to supplement those of using percussive vs flat tones discussed in Sec. IV A.

To the best of our knowledge, previous investigations have not explored shortening the durations of harmonic content to end earlier than the fundamental. However, previous studies of delayed harmonic sounds note a reduction in their perceived urgency (Edworthy *et al.*, 1991). It should be noted, however, that annoyance and urgency do not perfectly correlate (Marshall *et al.*, 2007). Although further research using a range of temporally variable tones is needed, our findings have important implications for auditory interface design, where such changes, in tandem with more diversity in amplitude envelope shapes, could greatly reduce persistent issues with annoyance (Sreetharan *et al.*, 2021) and masking (Rayo *et al.*, 2019).

C. Limitations and future directions

The use of online data collection (necessitated by the COVID-19 pandemic) may have added more noise than traditional in-laboratory testing with high control. As an example, participants did not calibrate listening levels (see Cartwright *et al.*, 2016 for an online calibration methodology). Additionally, we could not restrict participants from changing their playback volume during the experiment. Although our randomized presentation order ensured they could not predict which tones they would hear next, these problems are indicative of the limitations of online testing. Nonetheless, this lack of control regarding headphones and other conditions could also be seen as useful, given the potential applied uses of this work. As alarm implementations inevitably differ, ensuring design principles hold true outside of acoustically pristine laboratories is important.

Our investigation does not clarify why the differences between flat and percussive envelopes are clearest at a lower SNR. It may be due to higher peak amplitude of percussive tones as changes in envelope have previously been recognized as affecting stream segregation (Cusack and Roberts, 2000; Iverson, 1995), and resistance to masking (Buus, 1985; Hall and Grose, 1988). Further experimenting with tones equated for different temporal envelopes (systematically matching tones for equivalence in peak amplitude vs total energy vs perceived loudness) will be important for gaining further insight.

V. CONCLUSION

Our data suggest interesting new considerations for constructing auditory interfaces that maximize detectability and minimize annoyance. This adds to a growing literature, showing the importance of moving beyond traditional flat tones for auditory perception research. As temporal variation is common in natural sounds, using temporally variable stimuli in our basic perception work will allow for a more complete understanding of our hearing abilities. Through this, we can, in turn, inform and improve auditory interfaces, which currently, paralleling much of basic auditory research, generally fixate on invariant flat tones (Schutz and Gillard, 2020).

From our results, it is clear that design decisions regarding the use of flat tones with full-duration harmonics are neither necessary nor desirable from a detection perspective. Therefore, our findings suggest two exciting ways in which temporal variation can improve alarm sounds to reduce annoyance without harming detection (a) using percussive amplitude envelopes and (b) exploring splashes of harmonic energy rather than full-duration harmonics. It appears that relative to proto-typical (100%-flat) tones, our 25%-percussive tone is considerably less annoying, yet, better detected. Although we see the greatest potential in the medical device industry, auditory interfaces are used in many safety critical areas, such as railways (Edworthy *et al.*, 2011), aviation (Bliss, 2003; Dehais *et al.*, 2016), the automotive industry (Marshall *et al.*, 2007), industrial processing (Laberge *et al.*, 2014), and healthcare (Edworthy, 2013). Therefore, even small improvements can lead to widespread implications, making many auditory interfaces more pleasant and effective.

ACKNOWLEDGMENTS

We would like to thank Ben Kelly, Julia Bissessar, Andres Elizondo Lopez, and Cameron Anderson for their help and support with this manuscript.

Anderson, C. J. (**2019**). "Examining melodic alarm recognition accuracy across contrasting timbres," Honour's thesis (McMaster University).
Bendixen, A., Denham, S. L., Gyimesi, K., and Winkler, I. (**2010**). "Regular patterns stabilize auditory streams," J. Acoust. Soc. Am. **128**(6), 3658–3666.
Bliss, J. P. (**2003**). "Investigation of alarm-related accidents and incidents in aviation," Int. J. Aviat. Psychol. **13**(3), 249–268.
Bliss, J. P., Gilson, R. D., and Deaton, J. E. (**1995**). "Human probability matching behaviour in response to alarms of varying reliability," Ergonomics **38**(11), 2300–2312.
Block, F. E., Nuutinen, L., and Ballast, B. (**1999**). "Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance," J. Clin. Monit. Comput. **15**(2), 75–83.
Blumenthal, T. D. (**1988**). "The startle response to acoustic stimuli near startle threshold: Effects of stimulus rise and fall time, duration, and intensity," Psychophysiology **25**(5), 607–611.
Bodendörfer, X., Kortekaas, R., Weingarten, M., and Schlittmeier, S. (**2015**). "The effects of spectral and temporal parameters on perceived confirmation of an auditory non-speech signal," J. Acoust. Soc. Am. **138**(2), EL127–EL132.
Bruder, A. L., Rothwell, C. D., Fuhr, L. I., Shotwell, M. S., Edworthy, J. R., and Schlesinger, J. J. (**2022**). "The influence of audible alarm loudness and type on clinical multitasking," J. Med. Syst. **46**, 5.
Buus, S. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**(6), 1958–1965.
Buus, S., Florentine, M., and Poulsen, T. (**1997**). "Temporal integration of loudness, loudness discrimination, and the form of the loudness function," J. Acoust. Soc. Am. **101**(2), 669–680.
Cartwright, M., Pardo, B., Mysore, G. J., and Hoffman, M. (**2016**). "Fast and easy crowdsourced perceptual audio evaluation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 619–623.
Chuen, L., and Schutz, M. (**2016**). "The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: The role of spectral and amplitude cues," Atten. Percept. Psychophys. **78**(5), 1512–1528.
Cusack, R., and Roberts, B. (**2000**). "Effects of differences in timbre on sequential grouping," Percept. Psychophys. **62**(5), 1112–1120.
Cvach, M. (**2012**). "Monitor alarm fatigue: An integrative review," Biomed. Instrum. Technol. **46**(4), 268–277.
Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M. R., and London, J. (**2019**). "Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds," J. Exp. Psychol. Hum. Percept. Perform. **45**(3), 402–418.
Dehais, F., Roy, R. N., Gateau, T., and Scannella, S. (**2016**). "Auditory alarm misperception in the cockpit: An EEG study of inattentional deafness," in *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience, Lecture Notes in Computer Science*, edited by D. Schmorrow and C. Fidopiastis (Springer, Cham), Vol. 9743, pp. 177–187.
Edworthy, J. (**2013**). "Medical audible alarms: A review," J. Am. Med. Inf. Assoc. **20**(3), 584–589.
Edworthy, J., and Hellier, E. (**2005**). "Fewer but better auditory alarms will improve patient safety," Qual. Saf. Healthcare **14**(3), 212–215.
Edworthy, J., Hellier, E., Titchener, K., Naweed, A., and Roels, R. (**2011**). "Heterogeneity in auditory alarm sets makes them easier to learn," Int. J. Ind. Ergonom. **41**(2), 136–146.

Edworthy, J., Loxley, S., and Dennis, I. (**1991**). "Improving auditory warning design: Relationship between warning sound parameters and perceived urgency," Hum. Factors **33**(2), 205–231.
Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., Khan, J., and Pye, E. (**2017**). "The recognizability and localizability of auditory alarms: Setting global medical device standards," Hum. Factors **59**(7), 1108–1127.
Foley, L., and Schutz, M. (**2021**). "High time for temporal variation: Improving sonic interaction with auditory interfaces," IEEE Signal Process. Mag. **24**(7), 4–9.
Gaver, W. W. (**1997**). "Auditory interfaces," in *Handbook of Human-Computer Interaction*, 2nd ed., edited by M. G. Helander, T. K. Landauer, and P. V. Prabhu (North Holland, Amsterdam), pp. 1003–1041.
Giordano, B. L., and McAdams, S. (**2006**). "Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates," J. Acoust. Soc. Am. **119**(2), 1171–1181.
Grassi, M., and Casco, C. (**2009**). "Audiovisual bounce-inducing effect: Attention alone does not explain why the discs are bouncing," J. Exp. Psychol. Hum. Percept. Perform. **35**(1), 235–243.
Grassi, M., and Darwin, C. J. (**2006**). "The subjective duration of ramped and damped sounds," Percept. Psychophys. **68**(6), 1382–1392.
Grassi, M., and Pavan, A. (**2012**). "The subjective duration of audiovisual looming and receding stimuli," Atten. Percept. Psychophys. **74**(6), 1321–1333.
Gygi, B., Kidd, G. R., and Watson, C. S. (**2004**). "Spectral-temporal factors in the identification of environmental sounds," J. Acoust. Soc. Am. **115**(3), 1252–1265.
HairerSoft (**2019**). Amadeus Pro, available at https://hairersoft.com/pro.html#1 (Last viewed August 2, 2021).
Hall, J. W., and Grose, J. H. (**1988**). "Comodulation masking release: Evidence for multiple cues," J. Acoust. Soc. Am. **84**(5), 1669–1675.
Howell, P. (**1988**). "Prediction of *P*-center location from the distribution of energy in the amplitude envelope: I," Percept. Psychophys. **43**(1), 90–93.
IEC (**2006**). 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" (International Organization for Standardization, Geneva, Switzerland).
Iverson, P. (**1995**). "Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes," J. Exp. Psychol. Hum. Percept. Perform. **21**(4), 751–763.
Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (**2004**). "Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **116**(4), 2395–2405.
Laberge, J. C., Bullemer, P., Tolsma, M., and Reising, D. V. C. (**2014**). "Addressing alarm flood situations in the process industries through alarm summary display design and alarm response strategy," Int. J. Ind. Ergonom. **44**(3), 395–406.
Marcus, S. M. (**1981**). "Acoustic determinants of perceptual center (*P*-center) location," Percept. Psychophys. **30**(3), 247–256.
Marshall, D. C., Lee, J. D., and Austria, P. A. (**2007**). "Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness," Hum. Factors **49**(1), 145–157.
McCartney, J. (**1996**). "SuperCollider: A new real time synthesis language," in *Proc. International Music Conference (ICMC'96)*, pp. 257–258.
McPherson, M. J., Grace, R. C., and McDermott, J. H. (**2020**). "Harmonicity aids hearing in noise," BioRxiv, 2020.09.30.321000 (Last viewed May 15, 2021).
Moore, B. C. J., and Gockel, H. E. (**2002**). "Factors influencing sequential stream segregation," Acta Acoust. Acoust. **88**, 320–332.
Morton, J., Marcus, S., and Frankish, C. (**1976**). "Perceptual centers (*P*-centers)," Psychol. Rev. **83**(5), 405–408.
Neuhoff, J. G. (**1998**). "Perceptual bias for rising tones," Nature **395**(6698), 123–124.
Neuhoff, J. G. (**2001**). "An adaptive bias in the perception of looming auditory motion," Ecol. Psychol. **13**(2), 87–110.
Ng, M., and Schutz, M. (**2017**). "Seeing sound: A new tool for teaching music perception principles," Can. Acoust. **45**(3), 104–105.
O'Carroll, T. M. (**1986**). "Survey of alarms in an intensive therapy unit," Anaesthesia **41**(7), 742–744.
Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. (**2019**). "PsychoPy2: Experiments in behavior made easy," Behav. Res. **51**(1), 195–203.
Phansalkar, S., Edworthy, J., Hellier, E., Seger, D. L., Schedlbauer, A., Avery, A. J., and Bates, D. W. (**2010**). "A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems," J. Am. Med. Inform. Assoc. **17**(5), 493–501.
Rayo, M. F., Patterson, E. S., Abdel-Rasoul, M., and Moffatt-Bruce, S. D. (**2019**). "Using timbre to improve performance of larger auditory alarm sets," Ergonomics **62**(12), 1617–1629.
Schlesinger, J. J., Baum Miller, S. H., Nash, K., Bruce, M., Ashmead, D. H., Shotwell, M. S., Edworthy, J. R., Wallace, M. T., and Weinger, M. B. (**2018**). "Acoustic features of auditory medical alarms—An experimental study of alarm volume," J. Acoust. Soc. Am. **143**(6), 3688–3697.
Schutz, M. (**2016**). "Clarifying amplitude envelope's crucial role in auditory perception," Can. Acoust. **44**(2), 42–43.
Schutz, M. (**2019**). "Acoustic structure and musical function: Musical notes informing auditory research," in *The Oxford Handbook on Music and the Brain*, edited by M. H. Thaut and D. A. Hodges (Oxford University Press, Oxford, UK), pp. 145–164.
Schutz, M., and Gillard, J. (**2020**). "On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli," Sci. Rep. **10**(1), 9520.
Schutz, M., and Kubovy, M. (**2009**). "Causality and cross-modal integration," J. Exp. Psychol. Hum. Percept. Perform. **35**(6), 1791–1810.
Schutz, M., and Stefanucci, J. (**2019**). "Exploring the effects of 'sound shape' on consumer preference," Ergonom. Des. **27**(1), 16–19.
Schutz, M., Stefanucci, J., Baum, S. H., and Roth, A. (**2017**). "Name that percussive tune: Associative memory and amplitude envelope," Q. J. Exp. Psychol. **70**(7), 1323–1343.
Sendelbach, S., and Funk, M. (**2013**). "Alarm fatigue: A patient safety concern," AACN Adv. Crit. Care **24**(4), 378–386.
Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**(5234), 303–304.
Sreetharan, S., Schlesinger, J. J., and Schutz, M. (**2021**). "Decaying amplitude envelopes reduce alarm annoyance: Exploring new approaches to improving auditory interfaces," Appl. Ergonom. **96**, 103432.
Taenzer, A. H., Pyke, J. B., McGrath, S. P., and Blike, G. T. (**2010**). "Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers: A before-and-after concurrence study," Anesthesiology **112**(2), 282–287.
Vallet, G., Shore, D. I., and Schutz, M. (**2014**). "Exploring the role of amplitude envelope in duration estimation," Perception **43**(7), 616–630.
Varpio, L., Kuziemsky, C., MacDonald, C., and King, W. J. (**2012**). "The helpful or hindering effects of in-hospital patient monitor alarms on nurses: A qualitative analysis," Comput. Inform. Nurs. **30**(4), 210–217.
Wong, L. L. N., Soli, S. D., Liu, S., Han, N., and Huang, M. W. (**2007**). "Development of the Mandarin Hearing in Noise Test (MHINT)," Ear Hear. **28**(2), 70S–74S.