

TITLE PAGE

Final Progress Report:

A Formal Approach to Detecting and Correcting Simultaneous Masking in the IEC 60601-1-8 International Medical Alarm Standard

Principal Investigator:

Matthew L. Bolton
University at Buffalo

Team Members:

Senior Personnel

Judy Edworthy
University of Plymouth (UK)

Andrew Boyd
University of Illinois at Chicago

Scott Erdley
University at Buffalo

Graduate Research Assistants

Meng Li

Xi Zheng

Jiajun Wei

Eliot Biltekoff

Research Volunteer

Thomas Darget
École nationale supérieure d'électrotechnique, d'électronique, d'informatique,
d'hydraulique et des télécommunications (ENSEEIH; France)

Project Dates:

09/30/2018 – 09/29/2019

Federal Project Officer:

David Rodrick

The research reported in this paper was supported by the Agency for Healthcare Research and Quality under award number R18HS024679. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at the University at Buffalo. Informed consent was obtained from each participant.

Grant Award Number:

R18HS024679

STRUCTURED ABSTRACT

Purpose: In this work, we developed a computational method that allowed us to systematically evaluate the reserved alarm sounds of the IEC 60601-1-8 international medical alarm standard to determine when and how they can be totally and partially masked. **Scope:** IEC 60601-1-8 gives engineers instruction for creating human-perceivable auditory medical alarms. This includes reserved alarm sounds: common types of alarms in which each is a tonal melody. Even when this standard is honored, practitioners still fail to hear alarms. Simultaneous masking, a condition in which one or more alarms is imperceptible in the presence of other concurrently sounding alarms due to limitations of the human sensory system, is partially responsible for this. **Methods:** In this research, we use automated proof techniques to determine if masking can occur in a modeled configuration of medical alarms. This allows us to determine when and how reserved alarm sounds can mask other reserved alarms and to explore parameters to address discovered problems. **Results:** Significant masking problems were found for both the total and partial masking of high-, medium-, and low-priority reserved alarm sounds. We also showed that discovered problems can be mitigated by setting alarm volumes to standard values based on priority level and by randomizing the timing of alarm tones. We also created an easy-to-use tool so that medical alarm designers can apply our method and findings to the design of alarm sounds. **Key Words:** Medical devices and technologies, Medical Alarms, Audition, Patient safety, Psychophysical methods, Computational modeling

PURPOSE

This research used novel, formal, proof-based analyses to investigate and improve the audibility of concurrently sounding medical alarms from the IEC 60601-1-8 international medical alarm standard.

The failure of humans to respond to auditory medical alarms has resulted in numerous patient injuries and deaths and is thus a major safety concern.^{13,14,35,36} A relatively understudied source of response failures has to do with simultaneous masking, a condition in which concurrent sounds interact in ways that make one or more of them imperceptible due to physical limitations of human perception.²⁰ The widely used IEC 60601-1-8 international medical alarm standard was created to improve alarm discernibility and identification.²⁴ It gives designers a collection of standard reserved sounds for use in common alarm conditions. It also provides instructions for designing new alarm sounds. Unfortunately, the melodic tonal patterns of IEC 60601-1-8's alarms are particularly susceptible to simultaneous masking. This is very dangerous because, in medical environments, where seconds can mean the difference between the life and death of a patient, a human's ability to hear an alarm is necessary for him or her to appropriately react to it. Thus, there was a real and urgent need to address this limitation of the IEC 60601-1-8.

Experts have acknowledged that masking is a problem in medical alarms,^{15-17,25,29,31,32} and instances of it have been detected experimentally.^{30,37} However, masking may only occur with very specific combinations of multiple alarms and temporal overlaps between them. It is therefore virtually impossible to use observational or experimental methods to evaluate all the combinations of alarm interactions to find masking. To address this, we developed a novel method that uses formal modeling, psychophysical modeling, and proof-based verification analysis to automatically examine all the possible interactions within a modeled alarm configuration to determine if masking is possible.^{4,21-23} In this work, we updated this method and evaluated the masking potential of IEC 60601-1-8 medical alarms. We then used the method to analyze the standard and identify changes to it that would prevent masking from occurring. This was achieved by accomplishing three specific aims:

AIM 1: EXTEND THE MASKING DETECTION METHOD. To enable our method's use to evaluate the alarms of IEC 60601-1-8, we needed to extend and validate it. To accomplish this, we improved the method's scalability (to allow for the analysis of more concurrently sounding alarms); added the ability to search through the standard's open parameters to find conditions that both produce and avoid masking; enabled the method to account for additional harmonics in alarm sounds; and validated method predictions with a human subjects study.

AIM 2: DETECT AND CORRECT MASKING IN IEC 60601-1-8 ALARM SOUNDS. We used the extended method from Aim 1 to evaluate the reserved alarms of IEC 60601-1-8 to determine if each alarm is capable of being masked. We then used the method to identify values of the standard's open parameters that avoided masking. We used these findings to recommend standard modifications.

AIM 3: DEVELOP A PROCEDURE TO DESIGN AND TEST NEW ALARMS. To enable our method's use in the design of IEC 60601-1-8-compliant alarms, we will define a procedure. This procedure will allow engineers to both test whether an alarm design contains masking and use the method's new search capabilities to find design parameters that avoid it. Case studies will be used to show that the method supports its intended design goals.

SCOPE

BACKGROUND. Below, we cover the background on our method and the IEC 60601-1-8.

Our Method. Our method^{4,21-23} uses a unique combination of psychoacoustics and model checking to determine if simultaneous masking is possible in a modeled configuration of medical alarms. The psychoacoustics of simultaneous masking⁹ mathematically represent how the volume and tone/frequency of a sound cause masking. These are based on how the sensitivity of sensory cells on the inner ear's basilar membrane changes in the presence of other sounds. This threshold shift is represented as a masking curve. For tonal sounds, like those used in the IEC 60601-1-8, the masking curve for a given masking sound (the masker) is represented as

$$\text{curve}_{\text{masker}}(z_{\text{maskee}}) = \text{spread}_{\text{masker}}(z_{\text{maskee}} - z_{\text{masker}}) \cdot v_{\text{masker}} - 6.025 - 0.275 \cdot z_{\text{masker}}, \quad (1)$$

for which z_{maskee} and z_{masker} are the frequency of the potentially masked sound (the maskee) and the masker, respectively, on the Bark scale¹² and v_{masker} is the volume of the masker in dB. The spreading function $\text{spread}_{\text{masker}}$ represents how the masking effect changes as the frequency distance between the maskee and masker ($z_{\text{maskee}} - z_{\text{masker}}$ or δz) changes:

$$\text{spread}_{\text{masker}}(\delta z) = \begin{cases} -17 \cdot \delta z + 0.15 \cdot v_{\text{masker}} \cdot (\delta z - 1) & \text{for } \delta z \geq 1 \\ -17 \cdot \delta z & \text{for } 0 \leq \delta z < 1 \\ -(6 + 0.4 \cdot v_{\text{masker}}) \cdot |\delta z| & \text{for } -1 \leq \delta z < 0 \\ -6 \cdot |\delta z| - 0.4 \cdot v_{\text{masker}} - 11 & \text{for } \delta z < -1 \end{cases}. \quad (2)$$

Masking effects are additive, such that masking potential increases with the number of sounds in the environment.²⁷ The additive process produces a new absolute threshold of hearing (in dB) of a sound (the potential maskee) in the presence of N masker sounds with the formulation⁹:

$$\text{mthresh}_{\text{maskee}} = 10 \cdot \log \left(10^{(\text{abs}_{\text{maskee}})/10} + \left(\sum_{n=1}^N \left(10^{(\text{curve}_{\text{masker}_n}(z_{\text{maskee}}))/10} \right)^\alpha \right)^{1/\alpha} \right). \quad (3)$$

In this, α is a constant that can vary based on the type of sounds. In our work, we use $\alpha = 0.33$, because it is most appropriate for tonal sounds masking other tonal sounds. Furthermore, $\text{abs}_{\text{maskee}}$ is the original absolute threshold of hearing at the maskee's frequency. This is represented by

$$\text{abs}_{\text{maskee}} = 3.64 \cdot (f_{\text{maskee}}/1000)^{-0.8} - 6.5 \cdot e^{-0.6(f_{\text{maskee}}/1000-3.3)^2} + 10^{-3} \cdot (f_{\text{maskee}}/1000)^4, \quad (4)$$

in which f_{maskee} is the maskee's frequency in Hz.

With these psychoacoustics, if the potential maskee’s volume is less than or equal to the new threshold ($m_{thresh_{maskee}}$ Eq. [3]), the maskee will be simultaneously masked. These psychoacoustics have been validated across many applications^{1,10} and are the basis of the MPEG and other “lossy” audio codecs.⁹ More specifics about the psychoacoustics can be found in other publications.^{5,23}

Model checking¹¹ is an automated approach to performing mathematical proofs (called formal verification), which comes from the larger discipline of formal methods. To perform model checking, an analyst must create a formal model that captures a target system’s behavior. This is usually represented as a collection of concurrently executing state machines: a set of variables and transitions between variable values. Specification properties are used to assert desirable system conditions using a combination of model variables, Boolean logic operators, and temporal operators.¹⁹ The model check approach to formal verification proves whether or not the model satisfies the specification by exhaustively searching the system model’s statespace. The specification is proven true if no violation is found. If one is discovered, the model checker produces a counterexample: a trace through the model that shows how the violation manifested.

When psychoacoustics and model checking were combined together in our method (Fig. 1),^{4,23} analysts used an Excel spreadsheet to model alarms. Software then automatically generates formal models of the represented alarms along with the specification properties for checking whether masking can occur. The method supported the ability to detect both the partial and total masking of each modeled alarm in the configuration. When one of these properties was checked, the model checker considered all the possible alarm sound interactions to see if it could discover masking.

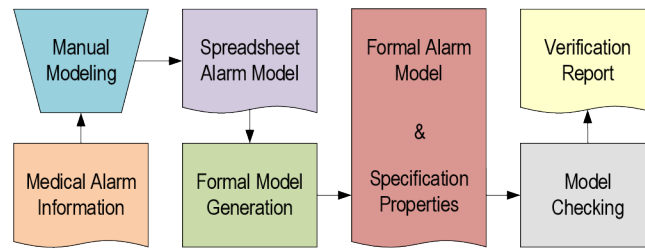


Fig. 1: The method for using model checking to discover masking between concurrently sounding medical alarms.^{5,6}

Older versions of our method^{4,21-23} were used to evaluate realistic applications. However, analyses took prohibitively long to complete. For example, partial masking analyses for one alarm of a telemetry monitor took 4.57 days. Thus, scalability improvements were required.

IEC 60601-1-8 and its Reserved Alarm Sounds. To help alarm designers and improve alarm recognition across the healthcare industry, the IEC 60601-1-8 international standard has a collection of reserved alarm sounds. These represent common types of alarms that can be used across devices. Each tone is described using a Helmholtz name (a letter representing a piano note). There are eight high-priority alarms, eight medium-priority alarms, and one low-priority alarm (Table 1). The high-priority alarm melodies have 10 sequential tones, with pauses between each tone. The medium-priority alarms have three sequential tones separated by pauses; these three tones of medium-priority alarms are the first three tones of their high-priority counterparts (though not necessarily with the same timings and volumes). The low-priority alarm has two sequential tones, with a pause in between.

The standard does not require specific volumes and timings of these sounds. Rather, there is a range of acceptable values that conform with the general alarm design guidelines of the standard. The relevant parameters are summarized in Table 3.

Table 1: Melodies used for IEC 60601-1-8 Reserved Alarm Sounds

Name	Alarm Priority												
	Medium					High							
General	c	c	e	-	c	c	—	c	c	e	-	c	c
Cardiac	c	e	g	-	g	C	—	c	e	g	-	g	C
Perfusion	c	f [#]	e	-	c	f [#]	—	c	f [#]	e	-	c	f [#]
Ventilation	c	a	f	-	a	f	—	c	a	f	-	a	f
Oxygen	C	b	a	-	g	f	—	C	b	a	-	g	f
Temperature	c	d	e	-	f	g	—	c	d	e	-	f	g
Drug Delivery	C	d	g	-	C	d	—	C	d	g	-	C	d
Failure	C	c	e	-	C	c	—	C	c	e	-	C	c
Low	e	c											

1 2 3 4 5 6 7 8 9 10

Note. Numbers under the table indicate the order in which a tone is played in an alarm’s melody. The first three tones of each high-priority alarm are the same as the corresponding medium-priority alarm. The last five tones of a high-priority alarm (tones 6–10) are a repetition of the melody from the first five tones of the alarm. Letters are musical pitches (Table 2). There is a pause between each alarm tone. A - indicates a slightly longer pause, and a — represents an even longer pause. The timing of tones and pauses is described in Table 3.

Table 2: Musical Pitches

Note	Frequency (Hz)
c	261.63
d	293.66
e	329.63
f	349.23
f [#]	369.99
g	392.00
a	440.00
b	493.88
C	523.25

Table 3: IEC 60601-1-8 Alarm Design Parameters

Characteristic	Alarm Priority		
	High	Medium	Low
Num. Tones	10	3	1 or 2
Tone Vol. (dB)	v_H	v_M	v_L
Max. Tone Vol. Diff. (dB)	10	10	10
Min. Num. Add. Harmonics	4	4	4
Add. Harmonics Freq. (Hz)	[300, 4000]	[300, 4000]	[300, 4000]
Add. Harmonics Vol. (dB)	$[v_H - 15, v_H + 15]$	$[v_M - 15, v_M + 15]$	$[v_L - 15, v_L + 15]$
Tone Duration (s)	t_H	t_M	t_L
Tone Spacing (s):			
b/w 1 st and 2 nd	x	y	y
b/w 2 nd and 3 rd	x	y	
b/w 3 rd and 4 th	$2x + t_H$		
b/w 4 th and 5 th	x		
b/w 5 th and 6 th	z		
b/w 6 th and 7 th	x		
b/w 7 th and 8 th	x		
b/w 8 th and 9 th	$2x + t_H$		
b/w 9 th and 10 th	x		
Time Between Repeat (s)	[2.5, 15]	[2.5, 30]	[15, ∞] or no repeat
where:	$v_H \geq v_M \geq v_L$	$t_H \in [0.075, 0.2]$	$t_M \in [0.125, 0.25]$
	$y \in [0.125, 0.25]$	$t_M + y \geq t_H + x$	$x \in [0.05, 0.125]$
	5% max. x and y variance in an alarm		$z \in [0.35, 1.3]$
			A blank indicates nonapplicability

METHODS

AIM 1: METHOD UPDATES. Our new method version uses the same process shown in Fig. 1. However, whereas the original method used a very different architecture, in which time was explicitly represented by a clock and simulation steps had to account for the number of time events that were considered in a given analysis, the new method took a very different approach, using the architecture that is summarized in Fig. 2 (details can be found elsewhere⁵). Specifically, the formal model contains variables to represent the state of each of the N alarms as they change over time. Each state when the alarm is producing sound is given a unique identifier (i.e., $Alarm_{i,2}$ is the state when alarm one is

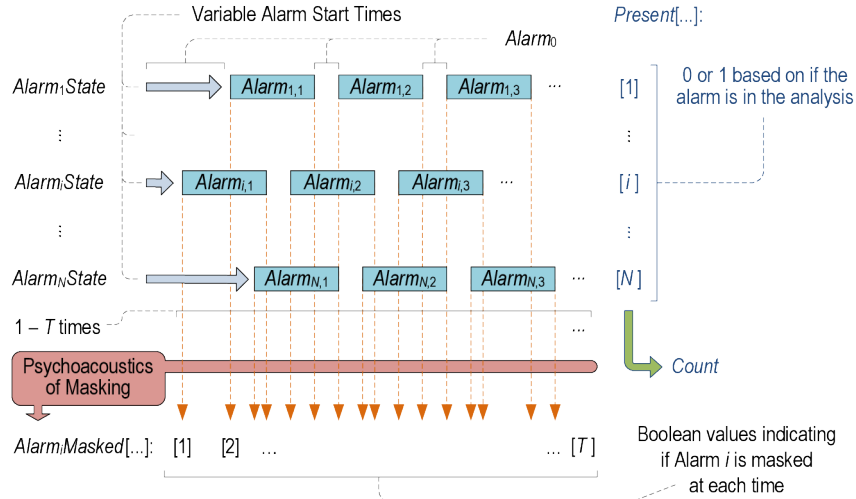


Fig. 2: Conceptual overview of how the formal alarm model from Fig. 1 works.

sounding its second tone). Whenever any alarm is not producing sound (it is not sounding or is in a pause), it is in state $Alarm_0$. The start time of each alarm is an open parameter that can assume any real-valued time that is ≥ 0 (thus allowing the model checker to consider any potential overlapping of alarms). Based on a given set of alarm start times, across all alarm states, the model computes a set of T discrete times based on the total number of times when change events happen in the alarms. Then, using the psychoacoustics of simultaneous masking and the physical properties (volumes and frequencies) of all the alarms' states at each time, the model computes an array of Boolean variables called $Alarm_iMasked$ (indexed from 1 to T corresponding to each discrete time) to indicate if a given alarm ($Alarm_i$) is masked at each discrete time. Specification properties can then be asserted over this array so that model checking can determine if a given alarm ($Alarm_i$) can ever be partially or totally masked.

In updating the method, we enabled it to account for the masking effects of additional harmonics. Finally, because we expected significant scalability improvements with this new approach, we implemented parameter exploration by allowing analysts to use scripts to generate multiple versions of a model with iterative differences in parameters between models. Thus, across multiple model-checking analyses, the analyst could explore parameters to determine how to address discovered masking.

This new architecture did not require checking across multiple model steps, and state information was completely derived from an array of alarm times. Thus, we hypothesized that the new method would be both more computationally efficient than the previous approach and less likely to miss masking conditions.

We evaluated our new method with two different types of tests. In the first, we evaluated cases originally reported in another reference²³ to determine if the new version could detect the same masking conditions as the old one and to compare computational efficiency (verification time) of the two approaches. In the second, we set out to fully characterize the scalability of the new method using a series of synthetic case studies based on the parameters of IEC 60601-1-8. Verifications were performed in parallel on a desktop workstation computer with a 12-core 3.60 GHz Intel® Xeon® E5-1650 CPU with 128 gigabytes of RAM running Linux Mint.

AIM 1: VALIDATION APPROACH. The psychoacoustics of simultaneous masking have been well validated,^{1,9,10} but they have never been shown to be valid for predicting whether someone will be able to hear IEC 60601-1-8 medical alarms. To address this, we conducted an IRB-approved study to determine if the predictions made with our method were accurate. In two standard, 200-trial, nonparametric signal detection experiments,²⁸ 28 nursing students judged whether alarm sounds were present in collections of concurrently sounding standard-compliant tones (see the Appendix for enrollment reports). The first experiment used alarm sounds with single-frequency (primary harmonic)

tones. The second experiment's sounds included the additional, standard-required frequencies (often called subharmonics). T tests compared miss, false-alarm, nonparametric sensitivity, and nonparametric bias measures between masking and nonmasking conditions and between the two experiments.

AIM 2: RESERVED SOUND ANALYSIS. For the reserved sound analyses, we used our updated version of the method to determine the minimum number of alarms required to mask each reserved alarm sound in all the following conditions: (1) each high-priority reserved alarm sound was analyzed in the presence of all the other high-priority alarm sounds; (2) each medium- and low-priority alarm sound was analyzed in the presence of all the other medium- and low-priority alarm sounds; and (3) each high-priority reserved alarm sound was analyzed in the presence of all the other medium- and low-priority alarm sounds.

We created three base configurations of alarms that were used for the analyses: one for each of the (1)–(3) conditions above. Each of these configurations contained a set of reserved alarm sounds from Table 1. The first, which was used to evaluate masking between high-priority reserved alarm sounds (Condition 1), contained all the high-priority reserved alarm sounds. The second, which was used to evaluate masking between medium- and low-priority reserved alarm sounds (Condition 2), contained all the low- and medium-priority reserved alarm sounds. The third, which was used to evaluate the masking of high-priority reserved alarm sounds by lower-priority sounds (Condition 3), contained all the low- and medium-priority reserved alarm sounds along with a blank entry used for representing a single high-priority alarm sound. In all three base models, frequencies were assigned the value corresponding to the presented Helmholtz names (Table 2). All sounds in each configuration within an alarm were given a standard volume of 80 dB ($v_H = 80$ and $v_M = 80$ for high- and medium-priority alarms, respectively; see Table 3). Furthermore, to maximize the potential of alarms masking other alarms (and to be consistent with standard alarm design practice), timings were kept consistent between configurations for every alarm, with tone times of $t_H = t_M = 0.2$ seconds. Pause times of $x = y = 0.125$ seconds and $z = 0.35$ seconds were used.

With these base configurations, we used the method's parameter exploration so that the volume of each analyzed (masked) alarm would vary while keeping the volume of the other (masking) alarms at the standard level of 80 dB. For the first two configurations, the analyses iteratively set the analyzed alarm to each of the alarms contained in the configuration. For the third configuration, the analyses would iteratively insert a different high-priority reserved alarm sound into a blank entry; this was always the analyzed high-priority alarm. Using this approach, we automatically generated formal models for each alarm, in which the volume of the analyzed alarm would range between 35 and 80 dB (in increments of 1 dB) between models. This volume range was used specifically because 80 dB represents an expected upper bound on the volume of a medical alarm. Furthermore, the considered range accounts for the 10-dB range in variance in alarm volumes within a given designed alarm system. The additional variance further accounts for potential differences that could arise between the volumes of alarms from independently engineered systems. The 35-dB minimum was chosen because it should be sufficiently low enough to ensure the maximum potential for alarm masking. Finally, although the actual volume level of maskers does impact the shape of the masking curve, in which higher volumes have more masking than lower volumes, this variation is very minor.⁹ Thus, results of analyses presented with this method should provide a good approximation (with a minor bias toward detection) of how the relative volumes of maskers at any given relative volume will manifest.

For each of the generated models, we performed multiple verifications to determine the minimum number of other alarms required to both partially and totally mask the analyzed alarm. As before, verifications were performed in parallel on a desktop workstation computer with a 12-core 3.60 GHz Intel® Xeon® E5-1650 CPU with 128 gigabytes of RAM running Linux Mint.

AIM 3: PROCEDURE AND DESIGN APPROACH. Although useful for analyzing the masking capabilities of medical alarm designs, the method required expertise in programming and formal verification. Thus, to make the technology approachable to the medical device designers and testers, we

have developed an easy-to-use desktop application called MAASC (Medical Alarm Audibility System Checker). This was implemented as a platform-independent Java program.

RESULTS

AIM 1: METHOD TESTING RESULTS. Our results focusing on testing the detection and scalability of our new version of the method showed that the new version replicated and improved detection capabilities compared with the legacy method and did so with significant reductions in verification times.

The legacy analysis results (using application from another publication²³) demonstrated that the new version of the method was capable of achieving the same level of masking detection seen with the previous version of the method. This was shown by the fact that the new method was able to find the positive masking results for all the considered cases. Due to the scalability limitations of the original method, the original verification analyses did not use a search depth sufficient enough to consider all possible model interactions. Thus, the analyses with the new method actually found masking conditions not previously discovered. This demonstrated the improved detection capabilities afforded by the completeness of the new method. This is a significant result, because it shows that the new method does not miss critical alarm interactions.

Not only was the new method complete but also it significantly improved scalability. In particular, total masking analyses across the legacy cases saw across-the-board reductions in verification times ranging from 66.67% to 99.97% for comparable analyses.

The method improvements are responsible for the performance observed in synthetic case studies used for assessing scalability. In particular, the analysis that took the longest time to complete (total masking of a 10-tone alarm in the presence of eight other similarly complex alarms) was completed in fewer than 3 days (249,676.89 s). These results were important for the project, because they showed that the method can be realistically used by designers to evaluate the masking potential of alarms from modern medical devices. Specifically, for a complex configuration with up to eight alarms, each with up to 10 tones (and 10 pauses), it will take fewer than 3 days of computational time to run an evaluation on each alarm. Furthermore, analyses of different alarms can be run in parallel. Thus, a full design could be evaluated in 3 days with enough conventional computational resources. It is worth noting that eight alarms (with 10 tones each) is a fair number of alarms to consider in a given analyses when evaluating a design. For example, even the IEC 60601-1-8 international medical alarm standard specifies alarms with a maximum of 10 tones, and the standard only contains eight reserved alarm sounds, which themselves can have up to 10 tones. Thus, our method is fully capable of evaluating the alarms consistent with IEC 60601-1-8 as well as its reserved sounds. As such, the work presented here has the potential to allow designers to reduce the likelihood that alarms in their devices will be masked. This should improve the probability that medical practitioners will hear the alarms, respond to them appropriately, and thus avoid adverse health outcomes. Furthermore, the results showed that we would be able to evaluate the masking potential between IEC 60601-1-8 reserved alarm sounds, the results of which are discussed subsequently.

More information on method testing results can be found in Bolton et al.⁵

AIM 1: EXPERIMENTAL VALIDATION RESULTS. Our results for the experimental validation of the ability of our method to predict whether or not humans will hear medical alarms was spread across two signal detection experiments, in which the second was a replication of the first with the inclusion of IEC 60601-1-8's required additional harmonics.

Experiment 1. The results of the comparisons of miss and false-alarm rates (M and F , respectively) are reported in Fig. 3. These analyses showed that miss rate (M) was significantly higher for masking trials than for trials without masking. There was no significant difference in false-alarm rates (F) between masking and nonmasking trials.

Sensitivity (A') and bias (B''_b)³⁴ results are reported in Fig. 3. These showed that sensitivity was significantly lower for masking trials than nonmasking ones. This means that people had a more difficult time distinguishing between signal and noise when masking was predicted than when it was

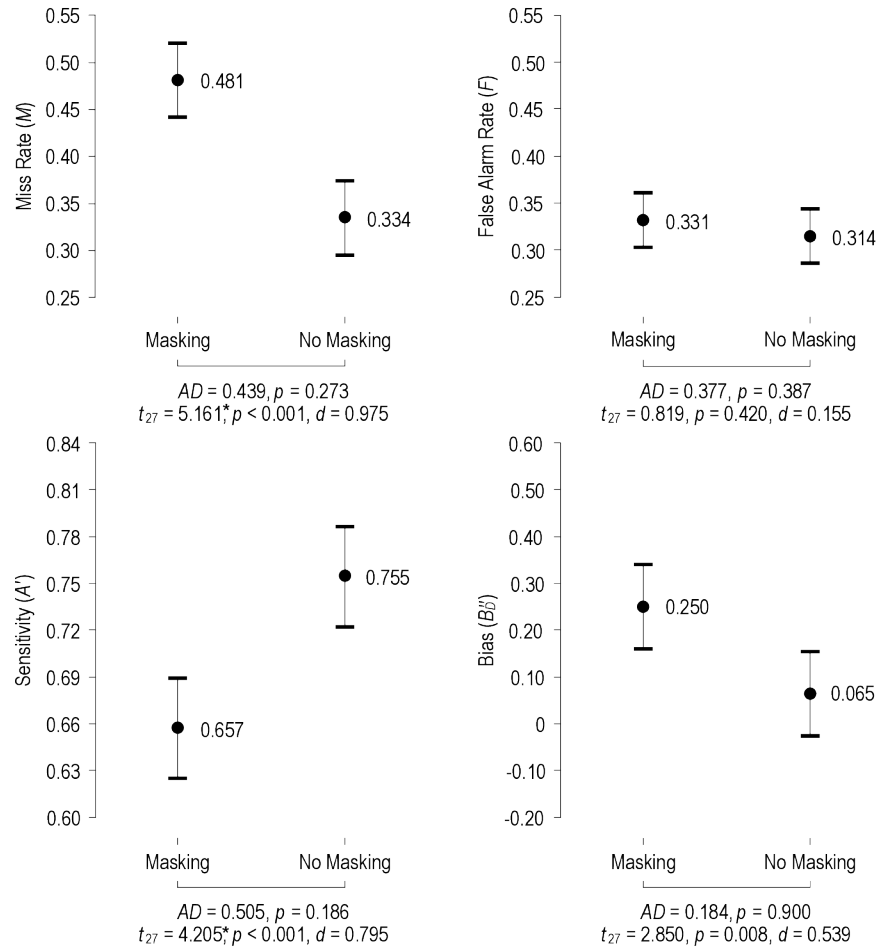


Fig. 3: Means (labeled circles) and 95% within-subject confidence intervals (horizontal bars) for miss rates (M), false-alarm rates (F), sensitivity (A'), and bias (B''_d) for both the masking and nonmasking conditions observed during Experiment 1. Rates are presented with Anderson Darling statistics that indicate that differences between the paired rates of participants followed a normal distribution. Rates are also presented with paired T-test results and their corresponding Cohen's d effect size. Statistical significance is indicated with an *.

not. On average, bias measures were positive. This indicates that participants tended to say “Yes” more often than they said “No.” People tended to say “Yes” more often for masking trials than for nonmasking trials. This difference would have met a 0.05 significance level ($p = 0.008$); however, this failed to meet the Bonferroni-adjusted level of statistical significance of $0.05/10 = 0.005$ based on the 10 T tests used in analyses.

Experiment 2. Results for the miss rate (M) and false-alarm rate (F) analyses are shown in Fig. 4. These showed that the miss rate was significantly higher for the masking condition than for the nonmasking one and that there were no statistically significant differences in false-alarm rates.

Results Discussion. The results of the sensitivity (A') and bias (B''_d) analyses are also shown in Fig. 4. These showed that there were significant differences between sensitivity and bias. On average, participants were significantly less sensitive in the masking condition than in the nonmasking condition. Conversely, participants had a significantly higher bias (and thus tended to say “Yes”) more often in the masking condition.

The comparison of these SDT statistics to the comparable ones from Experiment 1 (see Fig. 5) revealed that there were no significant differences between any of them.

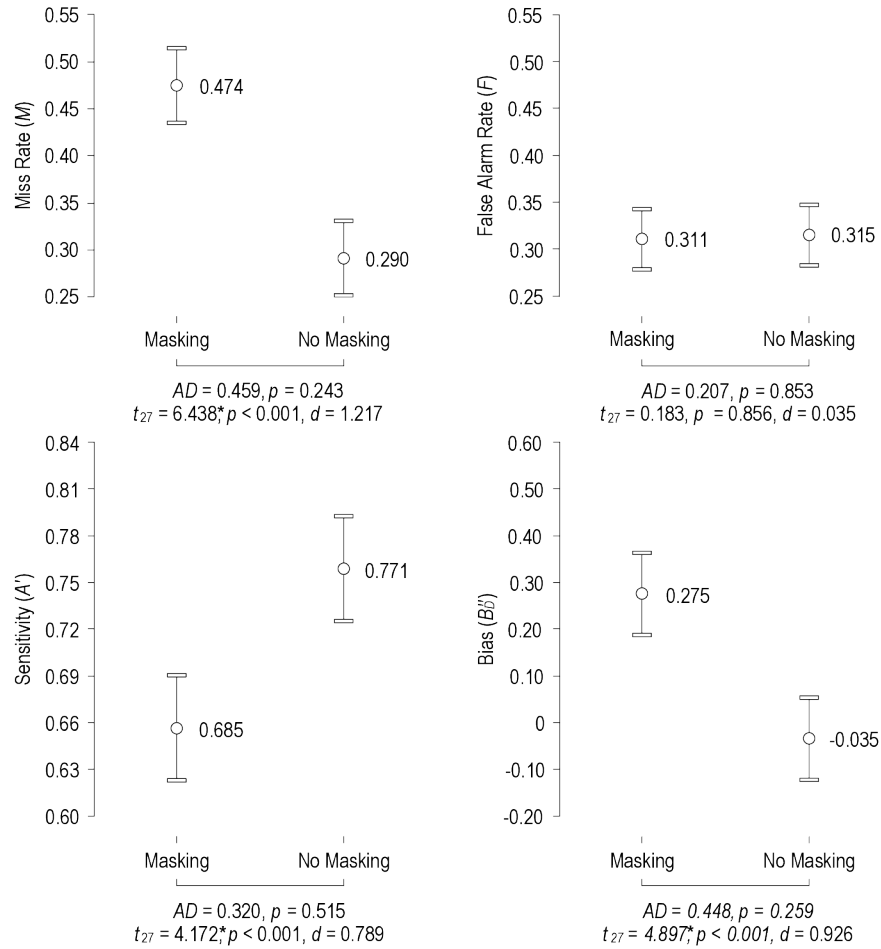


Fig. 4: Means (labeled circles) and 95% within-subjects confidence intervals (horizontal bars) for miss rates (M), false-alarm rates (F), sensitivity (A'), and bias ($B'd$) for both the masking and nonmasking conditions observed during Experiment 2. Rates are presented with Anderson Darling statistics that indicate that differences between the paired rates of participants followed a normal distribution. Rates are also presented with paired T-test results and their corresponding Cohen's d effect size. Statistical significance is indicated with an *.

These results were consistent with what we hypothesized. Specifically, participants made more misses when sounds were masked than when they were not. In the nonmasking condition, participants had misses only roughly one third of the time, but in the masking condition, participants had misses almost 50% of the time (which would be expected by random guessing). Conversely, there was no significant difference in false-alarm rates between the two conditions, which happened roughly 30% of the time. Furthermore, participants had reduced sensitivity for masking trials than they did for nonmasking ones. Collectively, these results show that participants had more trouble distinguishing between signal and noise in the masking condition than in the nonmasking one and that this was predominantly because masking makes it more likely that humans will miss alarms. This is an important result, because it validates that the psychoacoustics of simultaneous masking used in our method are able to accurately predict whether or not masking will contribute to alarm audibility. Furthermore, the fact that there were no significant differences between the results of the two experiments means that the additional harmonics required by the standard do not have any impact on masking.

These results are significant because, to the best of our knowledge, this is the first research to empirically show that masking is a problem for IEC 60601-1-8 alarms. Furthermore, our research showed that the masking effect is strong enough to make IEC 60601-1-8-compliant alarms inaudible,

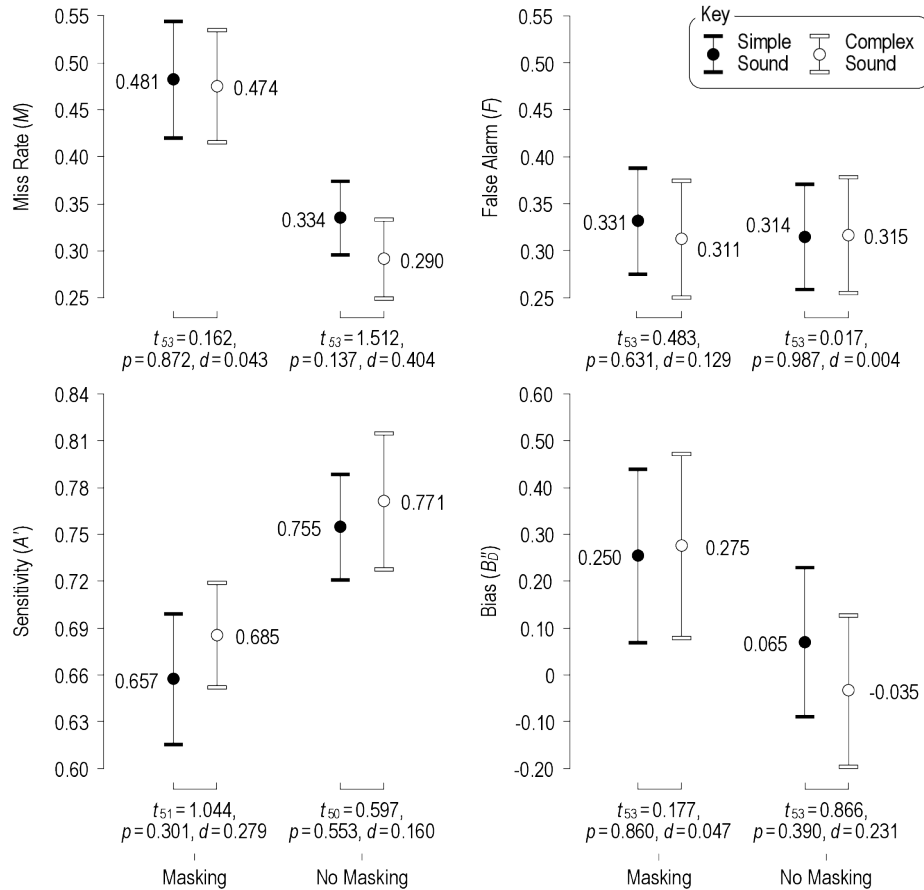


Fig. 5: Comparisons of miss rate (M), false-alarm rate (F), sensitivity (A'), and bias (B''_d) values measured in Experiments 1 and 2 (reported previously in Figs. 3 and 4). T-test statistics (reported with their corresponding Cohen's d effect size) show that there were no statistically significant differences observed between comparable rates of the two experiments. Note that, due to the nature of the comparisons being done, these plots are presented with between-subject confidence intervals, which differ from the within-subject confidence intervals presented in Figs. 3 and 4.

even with the inclusion of the requisite additional harmonics. These are powerful results, because they mean that the psychoacoustics of simultaneous masking can be used to make predictions about whether people will be able to hear alarms from the IEC 60601-1-8 international standard and that this can be done with only the primary harmonics of the alarms.

Our results are of import to our method,^{4,6,21-23} which in turn has important implications for alarm design and masking in healthcare environments. First, by validating the predictive capabilities of the psychoacoustics that our method uses, we enable both the predictive power of our method to be used effectively to design and evaluate medical alarms and its use in our effort to evaluate the international medical alarm standard. Second, although our method can account for additional harmonics, doing so requires more computational time and resources. Thus, by showing that we only need to account for the primary harmonics in analyses, our results expand the potential usefulness and approachability of our method. Third, our results validate the previous analyses done with the method and the results that will be presented in the next section.

More information on experimental validation of the method can be found in a prior publication.³

AIM 2: RESERVED SOUND ANALYSIS RESULTS. For the analysis of the masking potential between the reserved alarm sounds of IEC 60601-1-8, verifications for a given analysis took between

3.52 seconds and 335,168.36 seconds (≈ 4 days) each, with a median time of 455.63 seconds (≈ 7.6 minutes). Below, we present results across all three of the considered analysis conditions.

Condition (1): High-priority sounds masking high-priority sounds. Results showing the minimum number of alarms required to both totally and partially mask each of the high-priority IEC 60601-1-8 reserved alarm sounds by other high-priority alarms at each of the considered volumes is shown in Fig. 6(a).

In the total masking results, if all the alarms are at the same volume (80 dB), only the Temperature alarm could be totally masked by fewer than five other alarms. With the 10-dB range allowed by the standard (Table 3), three of the alarms (General, Drug Delivery, and Failure) can be totally masked by two other alarms. All the other alarm sounds can be totally masked by three other alarm sounds. All the high-priority alarm sounds could be masked by a minimum of two other alarms for volumes of 63 dB or lower (constituting a volume difference of 17 dB). At volumes of 57 dB or lower, all but the General alarm could be totally masked by one other high-priority sound. There were differences in the performance of the alarms. For example, at higher volumes, the Oxygen alarm was the least susceptible to masking.

In the partial masking results, all high-priority alarm sounds could be partially masked by as few as two other high-priority sounds for volumes between 73 and 80 dB. All alarms could be partially masked by at least one other alarm for volumes of 72 dB or lower. Note that this occurred within the 10-dB range in alarm volumes allowed by the standard.

Condition (2): Medium- and low-priority sounds masking medium- and low-priority sounds. Results showing the minimum number of alarms required to both totally and partially mask each of the medium- and low-priority IEC 60601-1-8 reserved alarm sounds by other medium- and low-priority alarm sounds at each of the considered volumes is shown in Fig. 6(b).

For the total masking analyses, when all the alarms were at the same volume (80 dB), all but the Cardiac and Low alarms could only be totally masked by four or more other alarms. Within the allowable 10-dB range of variance in alarm volume (Table 3), all but one of the alarms (Oxygen) could be totally masked by at least two other alarms. All analyzed alarms below 65 dB could be totally masked by just one other alarm. Finally, there were clear differences in alarm masking susceptibility at different volume levels. At higher volumes, the Oxygen and Failure alarms were the least susceptible, and the Cardiac and Low alarms are the most. At lower volumes, the Oxygen, Failure, and Low alarms were the least susceptible, whereas the Cardiac and Temperature alarms were the most susceptible. The Low alarm could never be totally masked by fewer than two other alarms.

The partial masking results match those of the high-priority alarm sounds, in that every alarm could be partially masked by at least two other alarms for volumes ranging from 74 to 80 dB and one other alarm for volumes of 73 dB or lower.

High-priority alarms were either just as susceptible or less susceptible to masking than their medium-priority counterparts. This is shown in Fig. 6(c), in which the comparable number from Fig. 6(b) is subtracted from its counterpart in Fig. 6(a), showing that all the numbers are greater than or equal to 0. Thus, the minimum number of alarms required to mask high-priority alarms is always the same or greater than the minimum number required to mask the respective medium-priority sounds.

Condition (3): Medium- and low-priority sounds masking high-priority sounds. The analyses of the ability for the medium- and low-priority sounds to mask the high-priority alarm sounds revealed that the medium- and low-priority alarms could never totally mask the high-priority sounds (Fig. 6[d]). This remained true even when the high-priority sounds' volumes were as low as 35 dB. The partial masking results (Fig. 6[d]) replicated those observed for the high-priority alarms: every alarm could be partially masked by one other alarm for volumes of 73 dB or lower or two alarms for higher volumes.

Results Discussion. The results show that masking is a concern for the analyzed alarm sounds. Below, we discuss our results, their generalizability, and the implications for IEC 60601-1-8.

Total Masking. The literature does not indicate how many overlapping alarms people are able to differentiate, even if none of them are masked. However, it is clearly better for alarms to have a higher number of simultaneous alarms required for masking to occur. In any case, it seems reasonable that a

human would fail to hear one alarm when five or more are sounding. In this respect, if the alarms are kept at the same level (80 dB), the results of the total masking analyses (Fig. 6) are encouraging for three reasons. First, for the high-priority alarms, it takes between four and six other alarms (five and seven total alarms) to completely mask any given alarm (Fig. 6[a]). For the medium- and low-priority alarms, it takes a minimum of between three and six other alarms (four and seven total alarms) to completely mask any given alarm. Of these, only two (Cardiac and Low) could be masked by a minimum of three others (Fig. 6[b]). The low number observed for the Low alarm is not particularly concerning, because it is less important than any of the other alarms that may be masking it. However, an examination of the counterexample produced in the Cardiac alarm analysis shows that it is of equal importance to its three other masking alarms. Thus, masking of the medium-priority Cardiac alarm is a higher risk at this volume. Second, the high-priority alarms were consistently as or less susceptible to simultaneous masking as the lower-priority alarms were (as shown in Fig. 6[c]). Third, none of the high-priority alarms could be total masked in the presence of the medium- and low-priority alarms (Fig. 6[d]).

However, results at the 70-dB level and below are less encouraging. When an alarm is 10 dB below baseline (70 dB), three of the high-priority alarms (General, Drug Delivery, and Failure), all but one medium-priority alarm (Oxygen), and Low, can be totally masked in the presence of only two other alarms at the same priority level. This is concerning, because the international alarm standard allows for a 10-dB variation in alarm volumes within a given designed configuration (Table 3). Thus, even within the standard alarm sounds of a given device, there are ways for some high-priority alarms and nearly every lower-priority alarm to be masked when three alarms sound concurrently. Moreover, alarm resilience to masking decreases with the volume of the masked alarm. For example, it only takes two alarms at the same priority level to mask all the medium- and low-priority alarms at 69 dB (11 dB below the others), the high-priority Perfusion and Delivery alarms at 68 dB (12 dB below the others), and the Cardiac and Ventilation alarms at 66 dB (14 dB below the others). Given that the standard does not provide a reference volume level for alarms, it is completely reasonable to expect such volume differences to manifest between alarms from different devices in a medical environment.

At even lower volumes, many alarms can be totally masked by a minimum of one other alarm at volumes that could be possible between devices. The medium-priority Cardiac alarm (at 64 dB) can be completely masked by a medium-priority Temperature alarm. The medium-priority Temperature (at 63 dB) alarm can be totally masked by medium-priority General alarm. Medium-priority Ventilation (at 60 dB) can be totally masked by medium-priority Perfusion. At lower volumes, all but the low-priority alarm (Low) and the General high-priority alarm can be totally masked by one other alarm at the same level.

Our results also show variation in the ability of alarms to be totally masked. For high-priority alarms, General, Oxygen, and Ventilation were the least susceptible to total masking at higher volumes, whereas Temperature was the most susceptible. However, performance varies at lower volumes. For the medium- and low-priority alarms, Oxygen is the alarm that is the most robust to simultaneous masking, with Failure and Ventilation also being fairly robust at higher volumes. Cardiac is clearly the most susceptible medium-priority alarm. In fact, the Cardiac medium-priority alarm was more susceptible to masking than Low, indicating that a mismatch exists between alarm masking susceptibility and priority.

Partial Masking. The literature does not provide guidance about how much an alarm needs to be partially masked to be rendered inaudible. However, it is clear that any masking will impair people's ability to perceive an alarm, especially given that there have been noted problems with the alarms in the standard even without the presence of masking.^{26,33} In this regard, the results from partial masking are problematic, because they show that it only takes two concurrently sounding alarms (either at the same priority level or between priority levels) to partially mask each of the analyzed alarms. At 73 dB and lower (7 dB below the others, and well within the range allow by the standard), it only took one other alarm to partially mask all the analyzed alarms.

Generalizability. Although all our analyses used 80 dB (a realistic upper bound) as the baseline volume and although the volume of a masker does impact masking curve shape (more masking is afforded by

higher volumes), these variations are minor.⁹ Thus, we would not expect results to significantly change with decreases in the base level volume with relative decreases in the volume of the analyzed alarm.

Note that IEC 60601-1-8 specifies that every alarm sound (including the reserved ones) have at least four additional harmonics (or subfrequencies): tones with frequencies at whole-number multiples of the primary frequency that are at lower volumes. These are intended to improve alarm audibility. However, previous work showed that our method validly predicts simultaneous masking using the primary alarm frequencies irrespective of whether the additional harmonics were included. Thus, the results presented in this paper should be valid even though we only considered the primary frequencies of the reserved alarm sounds.

Given this information, our results collectively indicate that there is potential for masking being a serious problem for devices that are compliant with IEC 60601-1-8. This is because total masking can occur in the presence of three or fewer alarms, and partial masking can occur in the presence of one alarm, with variations in volume that are consistent for alarms within and between devices. If practitioners are unable to hear an alarm, they will not be able to respond to them. In a medical environment, where seconds can mean the difference between life and death, this could have profound implications for patient safety and health.

Implications for the Standard. Our results suggest means of interpreting the IEC 60601-1-8 reserved sounds that will minimize the effect of masking. First, for a designed configuration of alarms, alarms at a given priority should be kept as close to the same volume as possible. Second, for alarms between devices, clinical engineers should try to make similar-priority alarms as close to the same volumes as possible. They should not increase the volume of a particular alarm to improve its perceivability. These recommendations will reduce the chances that any particular alarm will be masked.

Note that this is not suggesting that all alarms should be kept at the 80-dB base level. This is because such volumes will likely contribute to problems associated with hospital noise.²⁵ Rather, the same or better performance (due to the nature of the psychoacoustics)⁹ will be achieved by keeping similarly prioritized alarms at a consistent, lower volume level. Third, IEC 60601-1-8 allows designers to use different timings of alarm tones and pauses between them—a feature that is not often employed (Table 3). Using different timings across alarms will make it harder for alarms to perfectly overlap and thus reduce the chance of total masking.

To assess this third point, we created new models of the high- as well as the medium- and low-priority alarms from analysis conditions (1) and (2), respectively. These models were identical to the originals except that the timing of each alarm's pauses and tones (each alarm's x and tH or y and tM ; Table 3) were assigned unique random numbers that were consistent with the requirements of the standard (Table 3). The generated values of these are shown in Table 4. The analyses for conditions (1) and (2) were then re-run. The results of these are shown in Fig. 7.

Although partial masking results were not affected, using these unique random timings dramatically improved the resilience of the alarms to total masking. The random timings completely eliminated the ability of high-priority alarms to totally mask other high-priority alarm regardless of the volume of the potentially masked alarm or the number of concurrently sounding alarms (Fig. 7[a]). Furthermore, the timings either completely eliminated the total masking of an alarm (for volumes of 55–80 dB for Oxygen and volumes of 75–80 dB for Drug Delivery) or increased the number of alarms required for total masking for the medium- and low-priority alarms (Fig. 7[c]). These improvements are specifically illustrated in Fig. 7(b) and Fig. 7(d), which show how the number of alarms required for masking from Fig. 7(a) and Fig. 7(c) improved from the analyses under conditions (1) and (2), respectively. Thus, assigning unique volumes to alarm timing parameters, even within the parameters allowed by the standard, has the potential to significantly improve alarm audibility.

It is well established that the alarms specified in IEC 60601-1-8 have a number of problems. This work demonstrates that simultaneous masking is also a concern. The uniformity of these sounds (their rhythm and harmonic structure) is the source of most of these problems. This paper demonstrates that even slight changes in temporal structure significantly reduce the probability of masking.

More information on the reserved alarm analysis results can be found in other papers.^{6,8}

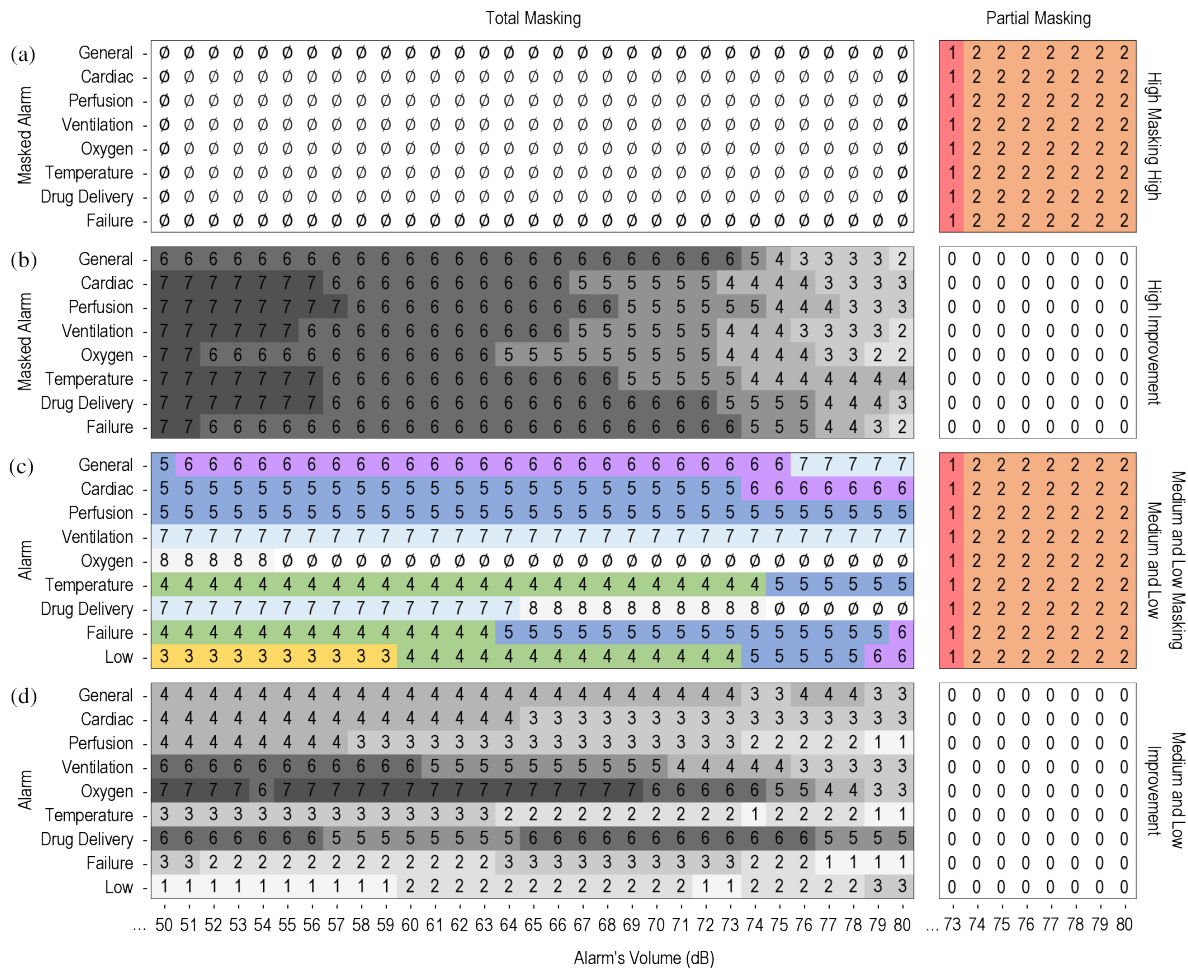


Fig. 7: Graphs in (a) and (c) show results from the replication of the analyses from conditions (1) and (2), respectively, but using alarm models with random times assigned to all alarm tone sounding lengths and pauses. A \emptyset indicates that no masking could occur. (b) and (d) show the increase/improvement in the minimum number of alarms to produce masking between these new results and the originals. (b) shows the improvement from the result in Fig. 6(a) to those in (a). A \emptyset was treated as an 8 in the production of this graph. (d) shows the improvement from the result in Fig. 6(b) to those in (c). A \emptyset was treated as a 9 in the production of this graph.

AIM 3: PROCEDURE AND DESIGN ANALYSIS RESULTS. MAASC (see Fig. 8) was implemented as an easy-to-use, platform-independent, Java-based, desktop application that supports all the features of our method. MAASC enables analysts to model a configuration of medical IEC 60601-1-8-compliant alarms with a simple point-and-click interface. Analysts can then use an integrated, industrial-grade model checker to determine if any of the alarms can be partially or totally masked by other modeled alarms. Built-in visualization allows analysts to identify exactly how the discovered masking conditions manifest (see Fig. 9). The tool further enables parameter exploration so that an engineer can find changes to an alarm's design that will correct discovered problems. The application also supports a feature in which random, standard-compliant timing parameters can be automatically generated for a loaded configuration. This is useful, because previously discussed results showed this to be an effective strategy for avoiding total masking. Beyond these capabilities, MAASC enables analysts to check the standard compliance of a modeled configuration of alarms and, in the case of a violation, identify the exact condition of nonconformance. The software also contains a number of examples and tutorials to help analysts understand and explore MAASC's features and help analysts check how their designed alarms will interact with the reserved alarm sounds of IEC 60601-1-8 (melodies that represent common types of alarms). As part of the tool's creation, a new modeling system

was developed, which enables partial masking detection to occur with computational times that are orders of magnitude smaller than in previously published methods,⁵ allowing analyses that previously could have taken days to occur in minutes. MAASC is freely available online (<http://fhsl.eng.buf-falo.edu/MAASC/>); runs on Windows, Mac, and Linux; and allows easy import and export capabilities with Microsoft Excel.

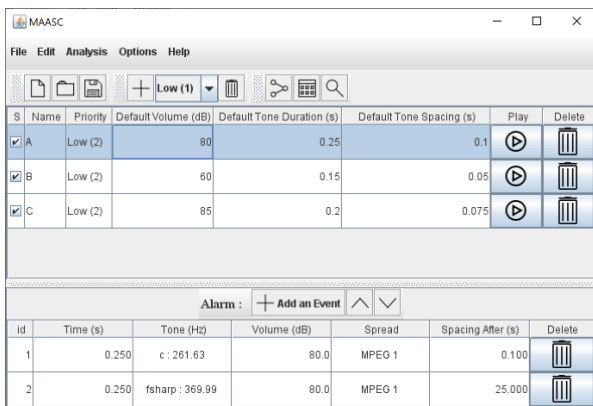


Fig. 8: MAASC main interface with three modeled alarms, each with two tones.

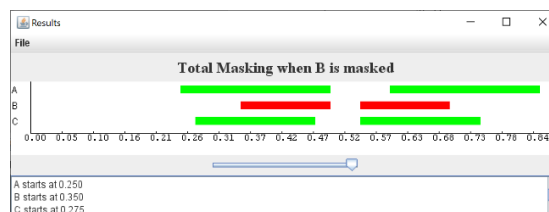


Fig. 9: Example visualization showing how alarm B can be totally masked by alarms A and C.

MAASC is a significant contribution, because it should enable medical device designers to easily exploit the power of our methods and the results documented in this report to create alarm designs that will avoid masking. In medical environments, where seconds can mean the difference between patient life and death, it is critical that medical practitioners be able to hear alarms. By giving engineers and hospitals a tool to design and deploy alarms that avoid audibility problems caused by simultaneous masking, MAASC has the potential to significantly improve patient safety.

More information on MAASC is available at <http://fhsl.eng.buffalo.edu/MAASC/> and in a separate article.²

LIST OF PUBLICATIONS AND PRODUCTS

JOURNAL ARTICLES. 4 total.

1. Bolton, Matthew L; Edworthy, Judy; Boyd, Andrew D. A computationally efficient formal method for discovering simultaneous masking in medical alarms. *Applied Acoustics*. 2018; **141**:403–415; Available from: <https://doi.org/10.1016/j.apacoust.2018.06.012>
2. Edworthy, JR; McNeer, RR; Bennett, CL; Dudaryk, R; McDougall, SJP; Schlesinger, JJ; Bolton, ML; Edworthy, JD; Vieira, EO; Boyd, AD; et al. Getting better alarm sounds into a global standard. *Ergonomics in Design*. 2018; **26**(4):4–13; Available from: <https://doi.org/10.1177%2F1064804618763268>
3. Bolton, Matthew L; Zheng, Xi; Li, Meng; Edworthy, Judy Reed; Boyd, Andrew D. An Experimental Validation of Masking in IEC 60601-1-8:2006–compliant Alarm Sounds. *Human Factors*; Available from: <https://doi.org/10.1177/0018720819862911>. In Press
4. Bolton, Matthew L; Edworthy, Judy Reed; Boyd, Andrew D. Masking Between Reserved Alarm Sounds of the IEC 60601-1-8 International Medical Alarm Standard: A Systematic, Formal Analysis. *Human Factors*. ND; Under Review

CONFERENCE PAPERS, ABSTRACTS, AND TALKS. 3 total.

1. Bolton, Matthew L; Edworthy, Judy Reed; Boyd, Andrew D. A Formal Analysis of Masking Between Reserved Alarm Sounds of the IEC 60601-1-8 International Medical Alarm Standard. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Los Angeles: SAGE; 2018. p. 523–527; Available from <https://doi.org/10.1177%2F1541931218621119>

2. Bolton, Matthew L; Zheng, Xi; Li, Meng; Edworthy, Judy Reed; Boyd, Andrew D. An Experimental Validation of Masking Predictions in IEC 60601-1-8–compliant Alarm Sounds. Chicago, IL; 2019. Presented at the 2019 Human Factors and Ergonomics in Health Care Symposium
3. Bolton, Matthew L; Darget, Thomas; Bilttekoff, Elliot; Edworthy, Judy Reed; Boyd, Andrew D. Medical Alarm Audibility System Checker (MAASC): A Computational Tool for Checking Medical Alarm Configurations for Simultaneous Masking. In: *Proceedings of International Symposium on Human Factors and Ergonomics in Health Care*. Los Angeles: Sage; Under Review

MEDIA COVERAGE. 7 total.

1. UIC Today article (the online magazine of the University of Illinois at Chicago) on October 22, 2019: <https://today.uic.edu/medical-alarms-may-be-inaudible-to-hospital-staff>
2. Profiled in the Buffalo Engineering 2018 magazine
3. WIVB, the local Buffalo CBS affiliate, did a television news story that aired on March 28, 2017, at 5:41 pm: <https://goo.gl/rzWj2F>
4. Featured cover story on UBNOW (the online magazine at the University at Buffalo) on March 20, 2017: <https://goo.gl/hEuemM>
5. UB issued a press release: <https://goo.gl/cuUpLG>
6. Covered in the M&A Malpractice blog: <http://goo.gl/qVnpJw> and <http://goo.gl/Jjk7sr>
7. Article in NetworkWorld's Mobile Cloud Blog: <https://goo.gl/UJtnVH>

SOFTWARE TOOLS. 2 total.

1. MAASC (Medical Alarm Audibility System Checker): <http://fhsl.eng.buffalo.edu/MAASC/>
2. Signal Detection Theory software. This project created an application for running auditory signal detection theory experiments as an Excel plugin. This has been used in three sections of IE 500 (Programming for Human Factors Engineering) at the University at Buffalo to teach students (30 in total) about computer programming and signal detection theory.

REFERENCES

- [1] Ambikairajah, E.; Davis, A.G.; Wong, W.T.K. Auditory masking and MPEG-1 audio compression. *Electronics & Communication Engineering Journal*. 1997; **9**(4):165–175.
- [2] Bolton, M.L.; Darget, T.; Bilttekoff, E.; Edworthy, J.R.; Boyd, A.D. Medical Alarm Audibility System Checker (MAASC): A Computational Tool for Checking Medical Alarm Configurations for Simultaneous Masking. In: *Proceedings of International Symposium on Human Factors and Ergonomics in Health Care*. Los Angeles: Sage; Under Review.
- [3] Bolton, M.L.; Zheng, X.; Li, M.; Edworthy, J.R.; Boyd, A.D. An Experimental Validation of Masking in IEC 60601-1-8:2006–compliant Alarm Sounds. *Human Factors*; Available from: <https://doi.org/10.1177/0018720819862911>. In Press.
- [4] Bolton, M.L.; Hasanain, B.; Boyd, A.D.; Edworthy, J. Using model checking to detect masking in IEC 60601-1-8–compliant alarm configurations. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Los Angeles: SAGE Publications; 2016. p. 636–640.
- [5] Bolton, M.L.; Edworthy, J.; Boyd, A.D. A computationally efficient formal method for discovering simultaneous masking in medical alarms. *Applied Acoustics*. 2018; **141**:403–415.
- [6] Bolton, M.L.; Edworthy, J.R.; Boyd, A.D. A Formal Analysis of Masking Between Reserved Alarm Sounds of the IEC 60601-1-8 International Medical Alarm Standard. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Los Angeles: SAGE; 2018. p. 523–527.

- [7] Bolton, M.L.; Zheng, X.; Li, M.; Edworthy, J.R.; Boyd, A.D. An Experimental Validation of Masking Predictions in IEC 60601-1-8-compliant Alarm Sounds. Chicago, IL; 2019. Presented at the 2019 Human Factors and Ergonomics in Health Care Symposium.
- [8] Bolton, M.L.; Edworthy, J.R.; Boyd, A.D. Masking Between Reserved Alarm Sounds of the IEC 60601-1-8 International Medical Alarm Standard: A Systematic, Formal Analysis. *Human Factors*. ND; Under Review.
- [9] Bosi, M.; Goldberg, R.E. *Introduction to Digital Audio Coding and Standards*. New York: Springer; 2003.
- [10] Brandenburg, K.; Stoll, G. ISO/MPEG-1 audio: A generic standard for coding of high-quality digital audio. *Journal of the Audio Engineering Society*. 1994; **42**(10):780–792.
- [11] Clarke, E.M.; Grumberg, O.; Peled, D.A. *Model checking*. Cambridge: MIT Press; 1999.
- [12] Zwicker, E.; Feldtkeller, R. *Das Ohr als Nachrichtenempfänger*. Stuttgart: Hirzel Verlag; 1967.
- [13] ECRI Institute; ISMP. Connecting remote cardiac monitoring issues with care areas. *Pennsylvania Patient Safety Authority*. 2009; **6**(3):79–83. Available from: [http://patientsafetyauthority.org/ADVISORIES/AdvisoryLibrary/2009/Sep6\(3\)/Pages/79.aspx](http://patientsafetyauthority.org/ADVISORIES/AdvisoryLibrary/2009/Sep6(3)/Pages/79.aspx).
- [14] ECRI Institute. 2019 Top 10 Health Technology Hazards. *Health Devices*. 2018; **October**. Available from: https://www.ecri.org/Resources/Whitepapers_and_reports/Haz_19.pdf.
- [15] Edworthy, J.; Hellier, E. Fewer but better auditory alarms will improve patient safety. *Quality and Safety in Health Care*. 2005; **14**(3):212–215.
- [16] Edworthy, J.; Hellier, E. Alarms and human behaviour: Implications for medical alarms. *British Journal of Anaesthesia*. 2006; **97**(1):12–17.
- [17] Edworthy, J.; Meredith, C.S. Cognitive psychology and the design of alarm sounds. *Medical Engineering & Physics*. 1994; **16**(6):445–449.
- [18] Edworthy, J.R.; McNeer, R.R.; Bennett, C.L.; Dudaryk, R.; McDougall, S.J.P.; Schlesinger, J. J.; Bolton, M.L.; Edworthy, J.D.; Vieira, E.O.; Boyd, A.D.; et al. Getting alarm sounds into a global standard. *Ergonomics in Design*. 2018; **26**(4):4–13.
- [19] Emerson, E.A. Temporal and modal logic. In: van Leeuwen, J.; Meyer, Albert R.; Nivat, M.; Paterson, M., Perrin, D., eds.. *Handbook of Theoretical Computer Science*. Cambridge: MIT Press; 1990. Chapter 16; p. 995–1072.
- [20] Fastl, H.; Zwicker, E. *Psychoacoustics: Facts and models*. vol. 22. Springer; 2006.
- [21] Hasanain, B.; Boyd, A.D.; Bolton, M.L. An approach to model checking the perceptual interactions of medical alarms. In: *Proceedings of the 2014 International Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica: HFES; 2014. p. 822–826.
- [22] Hasanain, B.; Boyd, A.D.; Bolton, M.L. Using Model Checking to Detect Simultaneous Masking in Medical Alarms. *IEEE Transactions on Human-Machine Systems*. 2016; **46**(2):174–185.
- [23] Hasanain, B.; Boyd, A.D.; Edworthy, J.; Bolton, M.L. A formal approach to discovering simultaneous additive masking between auditory medical alarms. *Applied Ergonomics*. 2017; **58**:500–514.
- [24] IEC 60601-1-8:2006+AMD1:2012. *Medical Electrical Equipment: Part 1–8: General requirements for basic safety and essential performance – Collateral standard: General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems*. Geneva: International Electrotechnical Commission; 2012.
- [25] Konkani, A.; Oakley, B.; Bauld, T.J. Reducing Hospital Noise: A Review of Medical Device Alarm Management. *Biomedical Instrumentation & Technology*. 2012; **46**(6):478–487.
- [26] Lacherez, P.; Seah, E.L.; Sanderson, P. Overlapping melodic alarms are almost indiscriminable. *Human Factors*. 2007; **49**(4):637–645.
- [27] Lutfi, R.A. Additivity of simultaneous masking. *The Journal of the Acoustical Society of America*. 1983; **73**(1):262–267.

- [28] McNicol, D. *A primer of signal detection theory*. Mahwah: Lawrence Erlbaum; 2005.
- [29] Meredith, C.; Edworthy, J. Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of Advanced Nursing*. 1995; **21**(1):15–20.
- [30] Momtahan, K.; Hetu, R.; Tansley, B. Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics*. 1993; **36**(10):1159–1176.
- [31] Patterson, R.D.; Mayfield, T.F. Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*. 1990; **327**(1241):485–492.
- [32] Patterson, R.D. *Guidelines for Auditory Warning Systems on Civil Aircraft*. Civil Aviation Authority; 1982.
- [33] Sanderson, P.M.; Wee, A.; Lacherez, P. Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia*. 2006; **61**(2):142–147.
- [34] See, Judi E.; Warm, Joel S.; Dember, William N.; Howe, Steven R. Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors*. 1997; **39**(1):14–29.
- [35] Stead, Willam W.; Lin, Herbert S., eds. *Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions*. Atlanta: National Academies Press; 2009.
- [36] The Joint Commission. Medical device alarm safety in hospitals. *Sentinel Even Alert*. 2013 April; **50**.
- [37] Toor, O.; Ryan, T.; Richard, M. Auditory Masking Potential of Common Operating Room Sounds: A Psychoacoustic Analysis. In: *Anesthesiology*. Park Ridge: American Society of Anesthesiologists; 2008. p. A1207.

APPENDIX A: ENROLLMENT REPORTS

Experiment 1 Enrollment Report

Racial Categories	Ethnic Categories									Total
	Not Hispanic or Latino			Hispanic or Latino			Unknown			
	Female	Male	Unknown	Female	Male	Unknown	Female	Male	Unknown	
American Indian/Alaska Native	1									1
Asian	3	2								5
Native Hawaiian or Other Pacific Islander	1			1						2
Black or African American										0
White	14	4								18
More than One Race		1		1						2
Unknown or Not Reported										0
Total	19	7	0	2	0	0	0	0	0	28

Experiment 2 Enrollment Report

Racial Categories	Ethnic Categories									Total
	Not Hispanic or Latino			Hispanic or Latino			Unknown			
	Female	Male	Unknown	Female	Male	Unknown	Female	Male	Unknown	
American Indian/Alaska Native										0
Asian	8	1								9
Native Hawaiian or Other Pacific Islander										0
Black or African American	2									2
White	12	3		1						16
More than One Race				1						1
Unknown or Not Reported										0
Total	22	4	0	2	0	0	0	0	0	28

Total Enrollment Report (for both Experiments 1 and 2)

Racial Categories	Ethnic Categories									Total
	Not Hispanic or Latino			Hispanic or Latino			Unknown			
	Female	Male	Unknown	Female	Male	Unknown	Female	Male	Unknown	
American Indian/Alaska Native	1									1
Asian	11	3								14
Native Hawaiian or Other Pacific Islander	1			1						2
Black or African American	2									2
White	26	7		1						34
More than One Race		1		2						3
Unknown or Not Reported										0
Total	41	11	0	4	0	0	0	0	0	56