

# Testing New Alarms for Medical Electrical Equipment



**Abstract.** This paper reports on the first of several tests of new auditory alarms proposed by Block et al. [1] and formalized in IEC 60101-1-8 for use in medical electrical equipment. We test whether participants who are supplied with the IEC-recommended mnemonics while learning label-alarm associations can more accurately recall what they have learnt in the short term. Results for 18 participants strongly indicate that there is a mutual confusability between certain alarm pairs in both learning conditions, but that mnemonics may strengthen rather than diminish certain key confusions.

## 1. Introduction

It has been a concern for many years that alarms in the ICU—and in other critical care areas in the hospital, such as the operating room—are too numerous and are often inappropriate and confusing [2]. It is reported that even in low stress conditions humans do not readily learn and retain more than eight sounds [3].

Block et al. [1] have proposed new alarm sounds that comply with existing standards from the International Organization for Standardization (ISO), European Committee for Standardization (CEN) and the American Society for Testing and Materials (ASTM). To help distinguish the alarm sounds, each alarm sound has been given a melody and assigned a mnemonic, which is a small phrase that can be “sung” to the alarm melody and that might suggest the nature of the device. Each alarm conforms to the same rhythmic pattern, as defined by stringent existing equipment standards, and varies only in its melodic pattern. The alarm sounds also have to be identifiable within the first few notes, and high and medium priority alarms must be distinguishable.

To date, there appears to be no indication that the alarms as proposed in these standards have been formally tested for discernability and memorability. This would seem to be a considerable oversight given that the proposed alarms are part of an international standard for a safety critical domain. In our study we aim to provide this information. In our study two key issues were investigated: (1) whether the varying melody was sufficient for users to readily discern the differences between one alarm and another and (2) if the 16 sounds, even if distinguished into 8 separate categories, were too many to be readily learned.

## 2. Method

**Participants.** Participants were first year undergraduate students from [REDACTED], recruited through the School of Psychology Research Sign-Up System. Although no formal hearing tests were administered, participants were asked prior to the experiment whether they had any known hearing problems.

**Apparatus and Stimuli.** Based on criteria outlined in the recently published IEC standard for medical equipment (IEC 60601-1-8) 16 auditory warnings were created using Csound. Sounds were processed on a Pentium® 4 1.9 GHz PC compatible with Soundblaster live 5.1 digital soundcard. The sounds were presented via Harmon/Kardon HK695 speakers at a mean amplitude of 60dB SPL(A), located about one meter from the participant. Alarm labels and mnemonics were displayed on a 19 inch touch screen monitor, which also recorded the participants' responses.

**Procedure.** The experiment required the participant to learn the 16 different alarm sounds (8 medium priority and 8 high priority) proposed by Block et al (2000) for auditory alarms in anesthesia machines. Learning proceeded in a series of learn-test cycles until the learning criterion was satisfied.

In the learning phase the experimenter introduced the sound and meaning of each alarm. A screen was then presented that displayed the alarm labels categorized by priority. Participants learned the alarms by touching the label of the alarm they wished to hear. They were allowed to listen to each alarm only once and could choose the order in which they listened to the alarms. (Participants who had already mastered certain alarm sounds could elect not to hear them again during a subsequent learning session.) Half the participants were randomly assigned to a group who learned with the help of mnemonics given in [1] [4] whereas the other half did not use mnemonics.

After the participant had listened to the alarms (all 16 or fewer if desired) a testing phase followed. Each test consisted of a random alarm being played. The participant then indicated which label, out of a list of 16 provided, was correct for that sound.

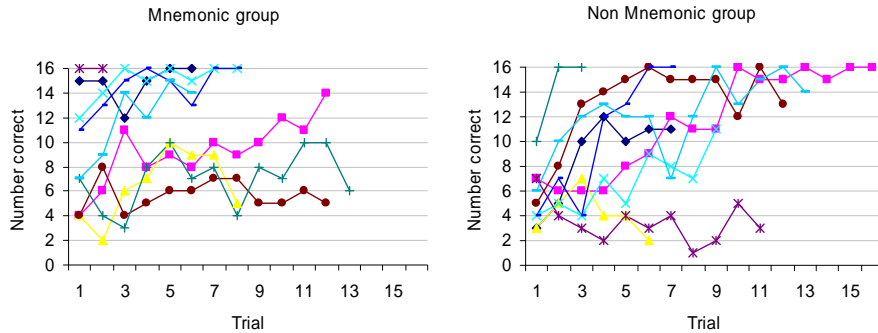
The learn-test cycle continued until the participant achieved two consecutive sets of correct tests or until 35 minutes had passed—whichever came first. After each testing phase the participant returned to the learning phase screen to review any alarms they were unsure about before attempting another testing phase.

## 3. Results

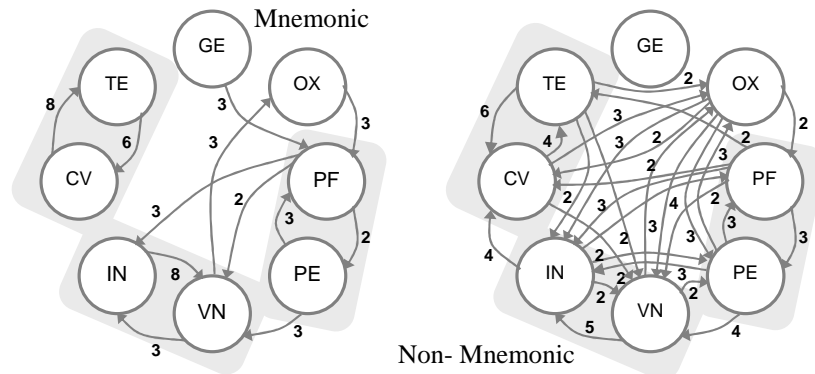
**Learnability.** Mean accuracy across the different alarms was poor. The exception was the general alarm, which had a mean accuracy across all conditions of 0.91. Participants' accuracy at identifying alarms was tested in a mixed-design ANOVA with factors of learning condition (mnemonic vs. no mnemonic) and alarm type (excluding the general alarm due to its near-perfect recognisability). There was a significant effect of alarm type,  $F(6, 96) = 3.025$ ,  $p = .009$ . Some alarms were learned more easily; accuracy and RT were better for them compared with other alarms. A significant interaction of alarm type and learning condition was also found,  $F(6, 96) =$

2.843,  $p = .014$ . The basis of this interaction still needs to be explored. Reaction time results are consistent with accuracy and do not show any speed-accuracy tradeoffs.

In the 35-minute learning period, only 9 out of 18 participants were able to complete at least one test with 100% accuracy, and only 7 of those 9 could fulfill the requirement of achieving 100% in two successive trials. Some participants found the alarm sounds very easy to learn whereas others were having difficulty in learning them, even showing very little signs of improvement over trials (see Figure 1).



**Fig. 1.** Successive test results for participants showing test trial and number correct in each test trial. Each line represents a different participant.



**Fig. 2.** Pattern of confusions between alarms. See text for description.

**Confusions.** We examined each participant’s confusions between different alarms. We noted the strongest confusion and second strongest (in cases where confusions between first and second were close) for each participant. Figure 2 shows the eight alarms. Links indicate that the alarm at the origin of the arrow is confused for the alarm at the end of the arrow. Numbers on the links indicate the number of participants for whom the confusion was the strongest or near-strongest confusion. Links were included only when two or more participants were confused. Participants in the mnemonic condition appear to be confused about a relatively small number of parameters (principally CV/TE and IN/VN) but these confusions are often shared by many participants. In contrast, participants in the non-mnemonic condition have a broader and more idiosyncratic pattern of confusions.

## 4. Discussion

To date, our results suggest that if no mnemonic is provided when participants are learning the IEC 60601-1-8 alarms, then they develop their own strategies. This leads to a wider variety of confusions between alarms across the group. For the mnemonic group, the confusion between some pairs of alarms seems more consistent. The mnemonics seem to have an effect of restricting confusions to a smaller number, but these confusions are then strong. This could result in more people stumbling over those few pairs and never learning the full set. Finally, some pairs of alarms were more easily confused with each other whether or not mnemonics were used. Overall, these results are unexpected. In ongoing work, 22 additional participants will be tested to increase power. Long-term retention and alarm discrimination while performing dual tasks will be assessed through follow up testing sessions.

The results also raise the issue of how much training is required for the majority of participants to attain 100% learning. In the present results we can see almost three distinct groups forming: (1) participants who quickly learn all alarms, (2) participants who take a while to learn and (3) participants who do not seem to learn at all. It is important to focus on the latter groups to discover the most effective ways in which we can help them learn the alarms faster and more accurately.

If they are robust, these results indicate there may be problems when healthcare providers use equipment with the new alarm sounds. Therefore it is important to test the alarm sounds with healthcare professionals in a follow-up study that we are planning.

## Acknowledgements

We thank Philippe Lacherez for preparing the alarm stimuli and his input on the project. Thanks also to members of our research group for their input on the manuscript.

## References

1. Patterson R.D. & Milroy R. (1980) *Auditory Warnings on Civil Aircraft: The learning and Retention of Warnings*. CAA paper 7D/S/0142. Civil Aviation Authority, London.
2. Meredith C. & Edworthy J. (1995) Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of Advanced Nursing* 21, 15-20.
3. Block, F. E., Rouse, J. D., Hakala, M., & Thompson, C. L. (2000). A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. *Journal of Clinical Monitoring*, 16, 541-546.
4. IEC 60601-1-8 (2003-08-14). *Medical electrical equipment - Part 1-8*. Geneva, Switzerland: International Electrotechnical Commission.