Are Melodic Medical Equipment Alarms Easily Learned?

Alexandra N. Wee, BIT(Hons)*

Penelope M. Sanderson, PhD, FASSA†

BACKGROUND: We tested melodic auditory alarms recommended in the IEC 60601-1-8 standard for medical electrical equipment for ease of learning and discrimination, and for effectiveness during a timeshared task.

METHODS: Twenty-two critical care nurses learned the IEC 60601-1-8 melodic alarms over two training sessions more than a week apart, with or without mnemonics suggested in the standard. Subsequently, the nurses identified alarms arriving at quasi random intervals while performing a timeshared arithmetic task.

RESULTS: Only one nurse (4.5%) identified the alarms with 100% accuracy after two training sessions. Mnemonics did not improve overall alarm identification accuracy (mnemonic = 56%, nonmnemonic = 55%) but led to a narrower range of confusions between alarms. Nurses responded faster ($P < 0.0001$) and more accurately ($P = 0.032$) to medium priority than high priority alarms, despite rating high priority alarms as sounding more urgent ($P < 0.0001$). Nurses with at least 1 yr of formal musical training identified the alarms much more accurately (musical training = 73%, no musical training = 38%, $P < 0.0001$), perceived a greater distinction between high and medium priority alarms ($P = 0.002$), and found identifying the alarms easier overall ($P = 0.023$). During the timeshared task, nurses' responses were slower ($P = 0.002$) and became less accurate ($P = 0.02$).

CONCLUSIONS: The slow rate of learning and persistent confusions suggest that the IEC 60601-1-8 melodic alarms should be redesigned before they are adopted for clinical practice.

(Anesth Analg 2008;106:501–8)

The design and use of auditory alarms in medical electrical equipment has recently received attention from many organizations. The Joint Commission on Accreditation of Healthcare Organizations included alarm safety on a list of national patient safety goals in 2004 after a Sentinel Event report motivated by concerns with auditory alarms.[1] The American Society of Anesthesiologists[2] and American Association of Nurse Anesthetists[3] have adopted the Anesthesia Patient Safety Foundation's 2004 alarm summit recommendations that the variable-pitch pulse oximeter tone and capnography auditory alarm should always be on and audible.[4,5] In 2005, an international standard, IEC 60601-1-8[6] also appeared that guides the design of visual and auditory alarm systems in medical electrical equipment and that will appear as a collateral standard in many countries.

IEC 60601-1-8 reflects the input of many anesthesiologists. It includes many good recommendations for alarm design and is far more detailed than its predecessor, ISO 9703-2.[7] Included in IEC 60601-1-8, however, is a well-intentioned recommendation for a set of melodic alarms, motivated by a controversial earlier design[8], that distinguishes the physical or physiological system each alarm represents[9,10] rather than having one alarm for all systems (Table 1). A mnemonic rationale, which explains the mapping of the melody, accompanies each alarm (e.g., "drops of an infusion falling and splashing back up" for the up-and-down infusion alarm melody). Elsewhere, lyrics have been proposed to aid learning (e.g., "IN-FU-SION A-LARM").[9]

The melodies in IEC 60601-1-8 were designed under the challenging constraint of conforming to fixed rhythmic patterns already established for medium and high priority alarms in ISO 9703-2[7] (IEC 60601-1-8, p. 87). Unfortunately, there was no formal test of the effectiveness of the melodic alarms with representative users before they were included in IEC 60601-1-8. Manufacturer experience at implementing the alarms is mixed. Independent research suggests that people's ability to identify and distinguish the melodic alarms is low, with persistent confusions between similar-sounding alarms.[11–13] There are further concerns which are rooted in auditory design theory.[14,15]

Each empirical study has shortcomings in either training, target population, or the range of alarms

---

From the *School of Information Technology Electrical Engineering, and †School of Medicine and ARC Key Centre for Human Factors, The University of Queensland, St. Lucia, Queensland, Australia.

Accepted for Publication August 3, 2007.

Supported by a UQ postgraduate research scholarship to Alexandra Wee.

Reprints will not be available from the author.

Address correspondence to Penelope M. Sanderson, PhD, FASSA, ARC Key Centre for Human Factors, The University of Queensland, St. Lucia, QLD 4072, Australia. Address e-mail to psanderson@itee.uq.edu.au.

DOI: 10.1213/01.ane.0000286148.58823.6c

Table 1. IEC 60601-1-8 Alarms with Melody, Mnemonic Lyric, and the Rationale for the Mapping of the Melody to the Alarm

	Melody* and mnemonic lyric		Rationale mnemonic (other information in support of mapping)
Alarm	Medium priority	High priority	
General	C4-C4-C4	C4-C4-C4—C4-C4 (repeated)	Fixed pitch, traditional (usual) ISO 9703 sound
Oxygen	C5-B4-A4 "OX-Y-GEN"	C5-B4-A4—G4-F4 (repeated) "OX-Y-GEN A-LARM"	Slowly falling pitches; top of a major scale; falling pitch of an oximeter
Ventilation	C4-A4-F4 "VEN-TI-LATE" "RISE-AND-FALL"	C4-A4-F4—A4-F4 (repeated) "VEN-TI-LATE A-LARM" "RISE-AND-FALL AND-FALL"	Old "NBC chime;" inverted major chord; rise and fall of the lungs
Cardiovascular	C4-E4-G4 "CAR-DI-AC"	C4-E4-G4—G4-C5 (repeated) "CAR-DI-AC A-LARM"	Trumpet call; call to arms; major chord
Temperature (or delivery of energy)	C4-D4-E4 "TEM-P'RA-TURE"	C4-D4-E4—F4-G4 (repeated) "TEM-P'RA-TURE A-LARM"	Slowly rising pitches; bottom of a major scale; related to slow increase in energy or (usually) temperature
Infusion (drug delivery)	C5-D4-G4 "IN-FU-SION"	C5-D4-G4—C5-D4 (repeated) "IN-FU-SION A-LARM"	Jazz chord (inverted 9th); drops of an infusion falling and "splashing" back up
Perfusion (artificial perfusion)	C4-F#4-C4 "PER-FU-SION"	C4-F#4-C4—C4-F#4 (repeated) "PER-FU-SION A-LARM"	Artificial sound; tri-tone; similar to "yo-ee-oh" of the Munchkins in "The Wizard of Oz"
Power failure	C5-C4-C4 "POW-ER FAIL" "GO-ING DOWN"	C5-C4-C4—C5-C4 (repeated) "POW-ER GO-ING DOWN"	Falling pitch as when the power has run down on an old Victrola

Medium priority alarms sound once to make a pattern of three tone pulses, designated as a "3" pattern. High priority alarms are repeated to make a pattern of 10 tone pulses, designated as a "3-2, 3-2" pattern. Alarms can be heard at http://www.itee.uq.edu.au/cerg/alarms.htm.

* C4 is middle C; C5 is the C above middle C.

tested. Williams and Beatty[12] trained 21 nonclinicians to identify the melodic alarms using the mnemonic lyrics supplied by Block et al.[9] but did not include the IEC 60601-1-8 mnemonic notes giving the rationale for the mapping. After two sessions of practice, alarm identification accuracy was only between 10% and 61% and participants often confused alarms. Sanderson et al.[11] tested 33 nonclinicians who learned either with or without the mnemonic lyrics and rationale, and found no difference between conditions. Participants often confused perceptually similar alarms. Finally, Lacherez et al.[13] examined 14 critical care nurses' ability to learn the high priority alarms unaided by mnemonic lyrics or rationale, as might occur in busy clinical contexts. Despite nurses' domain expertise, alarm identification was still poor and confusions between alarms were common.

The mnemonic lyrics and rationale should improve learning when participants have clinical monitoring experience. To determine this, first we tested whether critical care nurses can learn the IEC 60601-1-8 melodic alarms to 100% accuracy when supported by mnemonic lyrics and notes. Second, we hypothesized that learning will be better with mnemonics than without. Third, we tested whether learning is different for high priority versus medium priority alarms. Fourth, we hypothesized that alarm identification will be worse when participants are distracted, particularly when the alarms have been learned without mnemonics.

METHODS

Participants

The study was approved by Human Research Ethics Committees of The University of Queensland and of the participating hospitals. Participation was voluntary and participants provided written informed consent. They received AUD$40 and a small gift for their participation.

The participants were 22 registered critical care nurses between 21 and 58 yr-of-age (average 35) working in major hospitals in Brisbane. They were allocated randomly to learning condition. Six were male and 16 female with between 1 and 36 yr of nursing experience (average 13 yr). Participants were considered musically trained if they had at least 1 yr of formal training on a musical instrument. No participants had previous experience with the IEC 60601-1-8 melodic alarm sounds. A power analysis based on previous data[16] indicated that with 22 participants a 25% improvement in performance with mnemonics would be detected with statistical power of 0.71 and a 25% improvement in performance with musical training would be detected with statistical power of more than 0.9.

Apparatus and Stimuli

Alarms were created with the IEC 60601-1-8 "3-2, 3-2" pattern of tone pulses for high priority alarms and the "3" pattern for medium priority alarms (Table 1). Total duration of high priority alarms was 2924 ms.

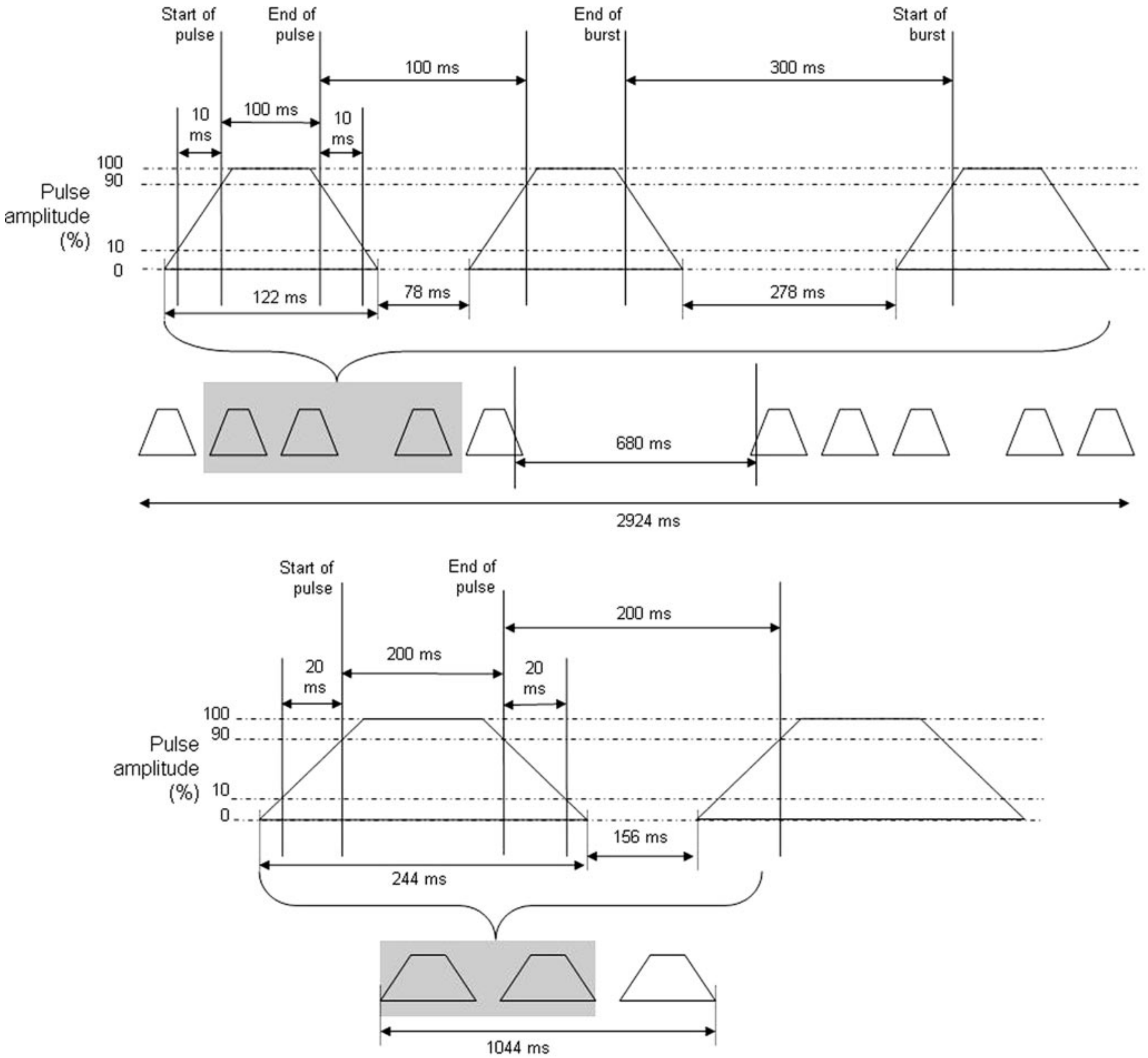

Figure 1. Temporal profile of the high priority alarms (top) and the medium priority alarms (bottom) used in the experiment, constructed to conform to general constraints provided in IEC 60601-1-8 (proportions in figures are not to scale).

Pulse durations and spacing followed constraints specified in IEC 60601-1-8 as follows (Fig. 1).

- Pulse rise time (time for amplitude to rise from 10% to 90% of maximum pulse amplitude) and pulse fall time (time for amplitude to fall from 90% to 10% of maximum pulse amplitude) for each tone was 10 ms.
- Pulse duration was 100 ms, consisting of 98 ms at maximum amplitude, plus 1 ms each for the period from 90% to 100% and 100% to 90% amplitude at the start and end of the pulse.
- Pulse spacing was 100 ms, measured from the end of one pulse (defined as when its amplitude fell to 90% of maximum) to the start of the next pulse (defined as when its amplitude rose to 90% of maximum).
- Spacing between the end of burst (end of third pulse) and start of next burst (start of fourth pulse) was 300 ms (pulse start and end defined as above).
- Spacing between the fifth and sixth pulses was 680 ms (pulse start and end defined as above).

The total duration of medium priority alarms was 1044 ms. Pulse rise times, duration, and spacing were twice as long as for high priority alarms and consistent with ranges specified in the standard.

The lowest alarm note was middle C (C4: 278.4 Hz) and the highest was an octave higher (C5: 556.8 Hz) (Table 1). The alarms were created in Csound and processed on a Pentium 4 1.9 GHz Acer Laptop with integrated soundcard. They were presented via AKG K 240DF Studio Monitor earphones with WAV on

Table 2. Phases of the Experiment

Day 1
Demographic questionnaire
Familiarization with alarms (Mn group introduced to mnemonics)
Learn-test cycles (learn alarms then self-test)
Questionnaire
Day 2 (approximately 1 wk after Day 1)
Part 1—Long-term memory (LTM)
LTM test without relearning alarms
Part 2—Relearning
Relearning in learn-test cycles (learn alarms then self-test)
Questionnaire
Part 3—Transfer
Timeshared transfer task (arithmetic expression true-false judgments and identify alarms)
Questionnaire

maximum and main volume control at level three from the lowest. The experimenter checked that the sound level was clear and comfortable for the participant. The alarm labels and mnemonics were displayed on a 17 in. flat screen display at 1280 × 1024 screen resolution by a Java program, which ran the experiment and recorded participants' responses.

Procedure

Participants were randomly allocated into either the mnemonic (Mn) ($n = 11$) or the nonmnemonic (NMn) group ($n = 11$). Each participant completed the experiment in two sessions on separate days (Table 2).

Day 1

Participants worked in a quiet room. They completed a questionnaire that asked about their age, any known hearing problems, time in the nursing profession, current specialty, and any formal musical training. During the familiarization phase, participants listened to the 16 alarms and were told the identity of each alarm (e.g., "high-priority oxygenation" or "medium-priority perfusion"). Participants in the Mn group were introduced to the mnemonics; the experimenter sang each alarm melody using the lyric and explained the rationale for its mapping to the alarm sound (Table 1). Participants in the NMn group simply learned the alarm label without being informed of the lyric or the rationale for each mapping.

After familiarization, participants learned the alarms in several learn-test cycles. The alarms were presented as 16 large labeled buttons on a computer screen. In each learning phase, participants could listen to each alarm as often as they needed and in any order until they were ready for the test phase. In the Mn condition, when each alarm button was pressed, the alarm's mnemonics (both the lyric and the rationale) appeared at the bottom of the screen.

In each test phase, participants clicked a button to hear an unidentified alarm and then identified the alarm by clicking one of 16 possible alarm buttons. After all alarms were tested, the screen displayed the actual alarms played, in the order in which they had been presented, along with the participant's response to the alarm and whether the response was correct. Participants then returned to the learning phase.

The learn-test cycles continued until either the participant reached the learning criterion of two consecutive sets of perfect test scores or 45 min had elapsed.

Day 2

Participants returned for a second session 6–11 days after the first session. Part 1 of Day 2 was a long-term memory (LTM) test in which the participants performed two tests similar to those in Day 1, but without first relearning the alarms.

In Part 2 of Day 2, participants performed learn-test cycles, which continued until the participants achieved two consecutive perfect scores, 45 min had elapsed, or the learn-test cycle had been completed eight times.

In Part 3 of Day 2 participants identified the alarms while performing a timeshared task. Participants saw a simple equation such as "$4 - 2 = 12$" and responded with the "P" key if the equation were true and the "Q" if false. The keys and their mappings were clearly indicated on the screen. At quasi-random moments, between 21 and 55 s apart, a melodic alarm would sound. The participant clicked the button on the screen representing the alarm they heard. The alarms were presented in random order with each alarm occurring once in each of three trials.

At the end of Day 1, after Part 2 of Day 2 (Relearning), and after Part 3 of Day 2 (Transfer), participants answered questionnaires that asked them to rate on 7-point scales the ease of learning the alarms, the ease of associating alarm meanings with their sounds, and the relative urgency of the medium and high priority alarms.

Statistical Analysis

Data were tested for significance with Statistica 7 using two-way tests with $\alpha = 0.05$. Responses were scored correct if the participant correctly identified both the alarm and its priority level. Latency was the time from cessation of the last pulse of the alarm to the onset of the participant's response. Accuracy and latency of identifying alarms were tested in mixed design ANOVAs with between-subject factors of learning condition and level of musical training and within-subjects factors of alarm priority and experiment phase. Differences between means were tested with *post hoc* Tukey HSD tests with $\alpha = 0.05$. Responses to questionnaires were tested using a mixed design ANOVA with between-subjects factors of learning condition and musical training, and the within-subjects factors of phase and, where appropriate, alarm priority. The range of types of confusions between alarms was assessed with interaction effects

Table 3. Average Accuracy (Percentage Correct) and Latency (s) Under Mnemonic and Nonmnemonic Learning Conditions for Day 1 Start, Day 1 End, Day 2 LTM, Day 2 End, and Transfer

	Phase				
	Day 1 start	Day 1 end	Day 2 LTM	Day 2 end	Transfer
Accuracy (%)					
Mnemonic					
High-priority	37	54	46	72	58
95% CI	26–48	42–67	37–56	60–84	44–73
Medium-priority	50	63	55	69	66
95% CI	40–60	53–73	45–65	55–83	55–77
Nonmnemonic					
High-priority	36	59	52	66	58
95% CI	25–47	46–72	43–62	54–78	44–73
Medium-priority	46	60	55	66	65
95% CI	36–56	50–71	45–64	52–80	54–76
Overall average	42	59	52	68	62
Latency (s)					
Mnemonic					
High-priority	9.0	5.6	6.5	5.2	6.3
95% CI	6.6–11.3	4.3–6.8	5.0–7.9	4.3–6.0	4.8–7.7
Medium-priority	6.1	5.4	5.1	5.0	5.6
95% CI	4.7–7.5	3.6–7.0	4.0–6.1	3.9–5.9	4.3–6.7
Nonmnemonic					
High-priority	7.7	6.2	7.0	5.1	6.8
95% CI	5.3–10.0	4.9–7.4	5.6–8.4	4.2–6.0	5.3–8.2
Medium-priority	6.3	5.6	5.8	5.0	5.7
95% CI	4.9–7.7	3.9–7.3	4.7–6.8	3.9–6.0	4.5–6.9
Overall average	7.3	5.7	6.1	5.1	6.1

Results are given for medium and high priority alarms.

CI = confidence interval.

in a log-linear model to test whether learning condition was a significant predictor of the range of types of confusions seen of the maximum possible of 56.

RESULTS

Figure 1 shows average accuracy and latency under Mn and NMn learning conditions for the first two test trials on Day 1 (termed "Day 1 start" below), the last two test trials on Day 1 ("Day 1 end"), the Day 2 LTM test ("Day 2 LTM"), the last two trials in Part 2 of Day 2 ("Day 2 end"), and for Part 3 of Day 2 during the timeshared task ("Transfer").

Response Accuracy During Learning

Average accuracy across alarms other than the general alarm at the end of Day 2 was poor (range 37%–82%; average 57%) but was good for the general alarm (average 89% correct). Two participants (9%) achieved 100% correct identification on one test cycle only on Day 1 and six participants (27%) did so on Day 2. However, no participant achieved the learning criterion on Day 1 (100% correct identification in two consecutive test cycles) and only one participant (4.5%) did so on Day 2.

There was no significant effect of learning condition on accuracy (Mn = 56%, NMn = 55%, $P = 0.88$) but participants with musical training identified alarms far more accurately than those without (musical = 73%, non musical = 38%, $P < 0.0001$).

Accuracy changed over test phases (Day 1 start = 42%, Day 1 end = 59%, Day 2 LTM = 52%, Day 2 end = 68%, $P < 0.0001$) (Table 3). Improvements in accuracy from start to end of Day 1, from LTM test to Day 2 end and from Day 1 start to Day 2 end were significant at $P < 0.05$. Musical participants gained accuracy over phases faster than the nonmusical participants, leading to an interaction of musical training with phase ($P = 0.008$). Participants identified medium priority alarms more accurately than high priority alarms (medium = 58%, high = 53%, $P = 0.032$).

In Figure 2, participants' accuracy at identifying each of the eight alarms is shown for the Mn and NMn learning conditions and for high and medium priority alarms for all test trials in Part 2 of Day 2. Large-font percentages adjacent to alarm names indicate the average identification accuracy for each alarm. Smaller-font percentages on arrows indicate the percentage of participants who misidentified the sound at the origin of the arrow as the alarm at the end on more than 25% of their test trials ("persistent confusions"). There were further persistent confusions unique to a single participant (not shown on figures). Although there was no difference in overall accuracy between the Mn and NMn conditions, Mn participants showed a relatively narrow range of persistent confusions (high = 17, medium = 20) whereas NMn participants showed a broader range of persistent confusions (high = 31, medium = 29). The interaction of learning condition

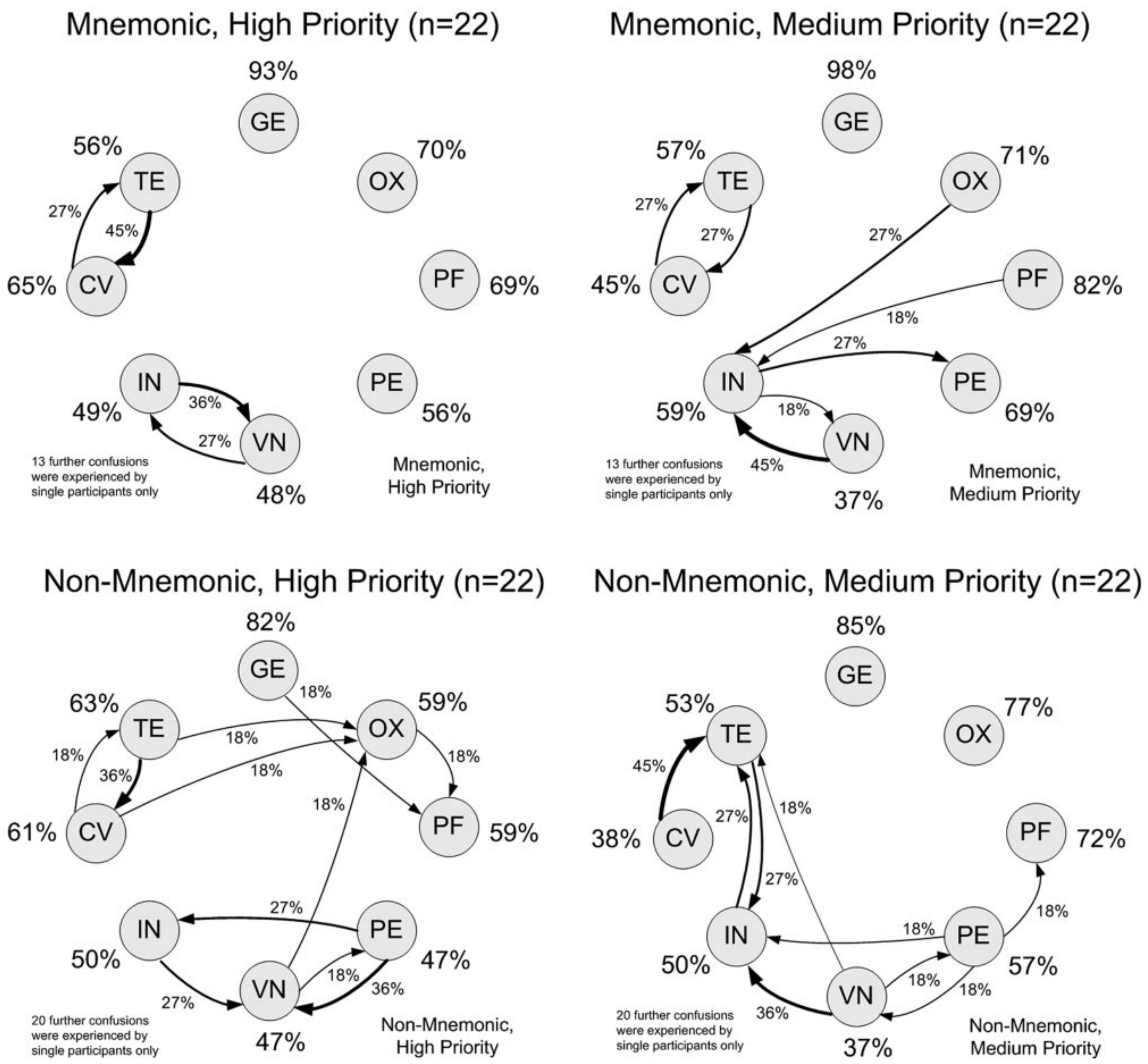

Figure 2. Average percentage accuracy across Day 2 Part 2 for each alarm (large numerals) for Mnemonic high priority condition, Mnemonic medium priority condition, Nonmnemonic high priority condition, and Nonmnemonic medium priority condition. Percentages on links (small numerals) represent the percentage of participants who misidentified the sound at the origin of the arrow as the alarm at the end of the arrow on more than 25% of trials. The number of further persistent confusions confined to single participants only is noted on each graph. Alarms are GE = general, OX = oxygenation, PF = power failure, PE = perfusion, VN = ventilation, IN = infusion, CV = cardiovascular, TE = temperature.

with the presence or absence of persistent confusions absorbed the log-linear model with a significant partial association ($\chi^2(1) = 9.52$, $P = 0.002$).

Response Latency During Learning

There was no effect on latency of learning condition (Mn = 6.0 s, NMn = 6.1 s, $P = 0.9$) or of musical training (no musical training = 6.5 s, musical training = 5.6 s, $P = 0.23$). Latency changed over test phases (Day 1 start = 7.3 s, Day 1 end = 5.7 s, Day 2 LTM = 6.1 s, Day 2 end = 5.1 s, $P < 0.0001$) with the start of Day 1 slower than the end of Day 1 or than either part of Day 2 ($P = 0.0006$, 0.012, and 0.0002, respectively). Participants responded significantly faster to medium priority alarms than to high priority alarms (medium = 5.5 s, high = 6.6 s, $P < 0.0001$), even though responding was measured from the end of the last pulse of each alarm.

Performance of Timeshared Task

To test whether alarm identification performance worsened during the timeshared task, average performance for all of Part 2 on Day 2 (Relearning) was compared with average performance after transfer to the timeshared task in Part 3 of Day 2 (Transfer) (Table 3). Accuracy decreased significantly (Day 2 Relearning = 66%, Transfer = 62%, $P = 0.02$) with high priority alarms showing the greater decrease ($P = 0.043$). There were no main effects or interactions with learning condition. Latency slowed significantly (Day

2 Relearning = 5.2 s, Transfer = 6.1 s, $P = 0.002$) with high priority alarms slowing down more than medium priority alarms ($P = 0.038$). Again, there were no further effects.

Questionnaires

Perceived ease of distinguishing alarms was moderately low to moderate on the 7-point scale and did not change significantly with phase (Day 1 = 2.8, Day 2 Relearning = 3.6, Transfer = 3.5, $P = 0.08$). Participants with musical training rated it easier to distinguish alarms than did nonmusically trained participants (musical = 3.7, nonmusical = 2.9, $P = 0.03$). There was no effect of learning condition (Mn = 3.3, NMn = 3.3, $P = 0.84$).

Perceived ease of associating alarms with their meanings was moderately low to moderate and did not change significantly during relearning (Day 1 = 3.4, Day 2 Relearning = 4.0, Transfer = 3.3, $P = 0.07$) but participants with musical training rated it easier to make the associations than those without musical training (musical = 4.2, nonmusical = 3.0, $P = 0.009$).

Participants rated the high priority alarms as sounding significantly more urgent than the medium priority alarms (high = 5.1, medium = 3.3, $P < 0.0001$). There were no main effects of phase, learning condition, or musical training. However, musically trained participants differentiated the high and medium alarm more markedly than did participants without musical training leading to an interaction (for musical, high = 5.6 and medium = 3.1; for nonmusical, high = 4.5 and medium = 3.4; $P = 0.002$).

DISCUSSION

This is the first study to evaluate the effects of mnemonics and timeshared tasks on clinicians' performance in identifying the IEC 60601-1-8 melodic alarms. First, after two learning sessions, only 1 of the 22 nurses could identify the alarms with 100% accuracy in two successive tests. The average percentage correct on Day 2 was only 66% (including the general alarm).

Second, there was no significant effect of mnemonics on speed and accuracy of identifying the alarms or on retention. This is surprising because the mnemonics were intended to help clinicians remember the mappings of melodies to systems. Memory research indicates, however, that mnemonics are most useful when people generate their own associations[15] whereas IEC 60601-1-8 suggests a specific set of mnemonics.

Third, performance was always faster and more accurate for the medium priority alarms than for the high priority alarms, even though the high priority alarms sounded more urgent to participants. It would be better for clinicians to respond faster and more accurately to the high priority alarms, which may indicate more urgent situations, than to the low priority alarms.

Fourth, the timeshared task slowed alarm identification from 5.2 s to 6.1 s and led to a small but significant decrease in accuracy from 66% to 62%, indicating that identification was resource-demanding. With the timeshared task, performance worsened more markedly for the more difficult high priority alarms and the mnemonics did not preserve participants' ability to identify the alarms.

This study reinforces the conclusions of previous investigations[11–13] into the IEC 60601-1-8 melodic alarms. The mapping of melodies to alarms through clinical associations may be no more effective than if melodies had been assigned at random to the alarms. Acoustically differentiated alarms may be confusing in clinical practice if misidentified alarms cause inappropriate initial visual checks.

Confusion between sounds occurs if sounds have too many common characteristics.[17] The IEC 60601-1-8 melodic alarms share a consistent rhythmic structure that was established in the previous ISO 9703-2 standard and retained for IEC 60601-1-8. Moreover, the notes are all between C4 and C5 in the key of C major and it is implied, even if not stated, that they should share the same timbre, making them harder to distinguish.

The IEC 60601-1-8 melodic alarms share many similarities with earcons, which are short musical motifs used to represent members of a set of objects or events such as types of software applications in a computer operating system.[18,19] Earcons that have similar rhythm, timbre, and key are difficult to identify when heard singly and are exceptionally difficult to discriminate when overlapping.[11,13,19]

Given the above, it is no surprise that musically trained participants identify the alarms much more accurately than participants without musical training. However, a successful design should be one in which all participants do well, with far less (and possibly no) effect of musical training, as demonstrated by Brewster et al.[18]

There are several limitations to our study. First, the melodic alarms were tested for auditory discriminability only in the absence of redundant visual alarms. Second, the alarms were not tested under clinical conditions in the context of an episode of care. Third, the standard does not require all alarms to be differentiated, so the result might apply to extreme conditions only. Some confusions between alarms were so persistent for some participants, however, that they might have been experienced even if the set of melodic alarms had been smaller.

In summary, although the IEC 60601-1-8 recommendation for melodic alarms is well-intentioned, there is cause for concern that the ability of clinicians to learn the recommended alarms and to discriminate between them may not be adequate.[11–13,17,18] Overall, the design of the IEC 60601-1-8 melodic alarms appears to have been over-constrained. Interdisciplinary research teams should explore simple design changes to the melodic alarms such as modifications in rhythm, timbre, and pitch, and should test the results formally. Indeed, any design changes to existing medical alarms should meet the conditions proposed by the standard. In the meantime, medical electrical equipment manufacturers and the anesthesia and critical care communities should be aware of potential problems with the recommended IEC 60601-1-8 melodic alarms before introducing them into clinical settings. An important lesson learned is the need to test design innovations formally with representative users before proposing such designs in equipment standards.

ACKNOWLEDGMENTS

The authors thank Philippe Lacherez and Phil Cole for constructing the alarm stimuli and for programming aspects of the experimental software, Richard Morey for statistical advice, and Associate Professor Marcus Watson for help in earlier phases of the research. The authors thank Dr. Chris Thompson, Dr. Daryle Gardner-Bonneau, and Dr. Edward Israelski for insights into the development and/or implementation of the alarms, and also Professor Judy Edworthy plus two anonymous reviewers for valuable comments on the manuscript.

REFERENCES

1. JCAHO. 2004 National patient safety goals. Joint Commission and Health Care Safety and Quality. Retrieved from http://www.jointcommission.org/PatientSafety/NationalPatientSafetyGoals/04_npsgs.htm. Last accessed on January 28, 2007
2. ASA. Standards for basic anesthetic monitoring (Amended October 2005). American Society of Anesthesiologists. Retrieved from http://www.asahq.org/publicationsAndServices/standards/02.pdf. Last accessed on January 28, 2007
3. AANA. Scope and standards for nurse anesthesia practice. Retrieved from http://aana.com/resources.aspx?ucNavMenu_TSMenuTargetID=51&ucNavMenu_TSMenuTargetType=4&ucNavMenu_TSMenuID=6&id=783. Last accessed on January 30, 2007
4. Olympio M. APSF workshop recommends new standards. APSF Newsletter 2005:19
5. APSF. Clinical alarms. Retrieved from http://www.apsf.org/initiatives/clinical_alarms/ on Anesthesia Patient Safety Foundation website. Last accessed on January 30, 2007
6. IEC. International Standard IEC60601-1-8 (2005–08): 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. Geneva, Switzerland: International Electrotechnical Commission, 2005
7. ISO. International Standard ISO9703-2:1994: Anaesthesia and respiratory care alarm signals, Part 2: auditory alarm signals. Geneva, Switzerland: International Organisation for Standardization, 2006
8. Weinger M. Proposed new alarms standards may make a bad situation worse [letter]. Anesthesiology 1991;74:791–2
9. Block FE, Rouse JD, Hakala M, Thompson CL. A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. J Clin Monit Comp 2000;16:541–6
10. Kerr JH. Warning devices. Br J Anaesth 1985;57:696–708
11. Sanderson P, Wee A, Lacherez P. Learnability and discriminability of melodic medical equipment alarms. Anaesthesia 2006;61:142–7
12. Williams S, Beatty PCW. Measuring the performance of audible alarms for anaesthesia. Physiol Meas 2005;26:571–81
13. Lacherez P, Seah L, Sanderson P. Overlapping melodic alarms are almost indiscriminable. Hum Factors 2007;49:637–45
14. Edworthy J, Hellier E. Fewer but better auditory alarms will improve patient safety. Qual Saf Health Care 2005;14:212–5
15. Edworthy J, Hellier E. Alarms and human behaviour: implications for medical alarms. Br J Anaesth 2006;97:12–7
16. Sanderson P, Wee A, Seah L, Lacherez P. Auditory alarms, medical standards, and urgency. In: Proceedings of the 12th International Conference on Auditory Display. London: Queen Mary University of London, 2006
17. Meredith C, Edworthy J. Sources of confusion in intensive therapy unit alarms. In: Stanton N, ed. Human factors in alarm design. London: Taylor & Francis, 1994
18. Brewster S, Wright PC, Edwards A. A detailed investigation into the effectiveness of earcons. In: Kramer G, ed. Proceedings of the First International Conference on Auditory Display. Santa Fe Institute, Santa Fe, NM: Addison-Wesley, 1992
19. McGookin D, Brewster S. Understanding concurrent earcons: applying auditory scene analysis principles to concurrent earcon recognition. ACM Trans Appl Percept 2004;1:120–55