Pilot study: Designing Medical Alarm Sounds For Semantic Association And Ease Of Learning

G Kanagachandran, P Beatty

Citation

G Kanagachandran, P Beatty. *Pilot study: Designing Medical Alarm Sounds For Semantic Association And Ease Of Learning*. The Internet Journal of Anesthesiology. 2006 Volume 12 Number 2.

Abstract

The performance of a new set of alarm sounds, generated from recordings of words descriptive of the situation relevant to the alarm and based on the principle of phonic abbreviation, were tested. Twenty-one volunteers aged from 24 to 54 were recruited for the study. Before being trained to recognise the sounds, they correctly identified 16.57% (range 0.00% to 76.19%) of the descriptive words used. The mean correct identification rates for the sounds after training was 56.35% (range 17.78% to 93.02%). Limited urgency mapping testing showed that the intrinsic urgencies of the low, medium and high priority forms of the sounds were statistically significantly separate ($p<0.01$). These results suggest that it may well be possible to construct a set of alarm sounds, for use on physiological monitoring instruments, using the concept of phonic abbreviation, which may have superior learnability to other systems without being unacceptably easy for casual hearers to understand.

INTRODUCTION

Following the work of Block et al (2002), the current IEC 60601-1-8 (IEC, 2005) standard for medical equipment alarms offers equipment manufacturers the option to create melodic alarms that distinguish the physical or physiological system that each alarm represents. The logic of such alarms is that their meanings should be easier to remember than abstract sounds, and therefore they might improve the speed and accuracy of clinician response. However, two studies by Williams and Beatty (2005), and Sanderson, Wee and Lacherez (2006) have shown that there are difficulties in learning and memorising the meanings of the melodic alarms in the IEC standard.

The notion that melodic alarms might be easier to learn and remember, stems from a study by Block (1992) that found that anaesthetists rapidly achieved good learning of the meaning of a set of such alarms. However, unlike those in the IEC standard, the alarms Block tested had names associating the alarm source with a simple phrase usually taken from the title of a popular song (e.g. oxygenation and ‘Love is Blue'). Could the difference in reported performance be attributed to the effect of this clear associative labelling rather than something to do with the design of the alarm sounds themselves?

In a survey of the preferences of anaesthetists for the type of sounds that might be used to form the basis of new alarms an overwhelming majority preferred using a synthesised voice to produce a verbal alarm. However, there are practical objections to using such verbal alarms. In critical care, alarms on equipment are audible to patients, who may be conscious, and to visitors. Both groups might be disturbed by the implications of such alarms. Even where a ‘public' alarm of this type might only be heard by medical staff, as in an operating theatre, it may still not be advisable for everyone to be informed about the detailed meaning of an alarm intended for action primarily by the anaesthetist.

This paper reports the results of a pilot study to test whether a set of alarms based on vocalisations of key-words explicitly describing the meaning of an alarm could be used to create an abstract sound that would preserve sufficient inherent semantic association to be easier to learn and memorise, but would still be sufficiently abstract to be unintelligible to a casual untrained hearer. We have termed these sorts of alarms ‘phonic abbreviations'.

METHOD

CONSTRUCTION OF THE NEW ALARM SOUNDS

The sounds would be tested within the same overall framework as the IEC melodic alarms. In this framework there are 8 physical or physiological systems referred to (ventilation, perfusion, infusion of drugs or fluids, cardiac performance, oxygenation, temperature, power failure and general alarm) each with its own basic alarm sound. For each of these systems there are modifications to the basic alarm sound according to the urgency to be indicated. All the alarms have a medium and high priority form and the general alarm also has a low priority form. Thus 17 different alarm sounds needed to be created.

The creation of the sounds began with live recording of a female voice (26 years) saying key words or phrases that could describe the systems to be referred to (ventilation, perfusion, drug infusion, cardiac, oxygen, power, temperature, and alarm). These sounds were then edited using audio editing software (Roxio's Sound Editor, Adobe Audition version 5.0) to create the basic phonic abbreviations for the key-words. To create different levels of urgency these basic phonic abbreviations were repeated different numbers of times: 3 for the single low priority sound, 5 for the medium priority sounds and 10 for the high priority sounds. The exact length of the final sounds generated depended on the key-word used but the low priority sound was 4.16s in length, the mean medium priority sound 4.21s and the mean high priority 6.15s. Copies of the sounds are available from http://www.medicine.manchester.ac.uk/staff/987.

TESTING OF THE NEW ALARM SOUNDS.

All the tests were administered by computer to 21 non-medically trained volunteers, ten female and eleven male, aged between 24 and 54 (median 30 years). Eleven of the volunteers were students and staff within the Medical School and School of Biomedical Sciences at Manchester University, and 10 of them were non-academic volunteers based in London. All volunteers were recruited under conditions of informed consent. The study was given ethical approval by the COREC under reference number 05/MRE08/2. All volunteers had normal hearing. All had a good knowledge of English, though for 11 of them, English was not their first language.

Testing was divided into two sessions. In the first the rate of identification of the sounds was tested, prior to formal teaching of their meanings. In the second a limited urgency mapping test was performed before a formal test on how well the sounds were remembered.

Session 1 started with a test of how well the sounds could be understood by an untrained listener. The sounds were played to volunteers in a random order and they were asked, after hearing each sound, to identify the word they thought the sound was saying. The rest of the structure of session 1 followed the form used in Williams and Beatty (2005). To teach the meanings, sounds were presented to the volunteers in three batches: first the general alarm sounds, then the oxygenation, ventilation, cardiac and temperature sounds, and finally the power failure, drug infusion and perfusion sounds. In all cases the volunteers were allowed to click at will on ‘buttons' on the screen to hear the sounds in a given batch. At any time they were allowed to perform a ‘self-test' of how much they had learned on all the sounds heard up to that point. Volunteers were not allowed to progress to the next batch of sounds until they had scored 70% correct identification on a self-test. No limit on learning time was imposed on the volunteers.

Session 2 took place a few days later (median 9 days: range 7 to 24 days) and tested how well the volunteers could recall the meaning of the sounds. It started with an urgency mapping test in which the sounds were played in a random order and their urgencies scored on a scale of 1 (Not at all urgent) to 7 (Critically urgent) by the volunteers. A formal overall performance test was then administered where all the sounds were played twice in a random order and the volunteers were asked to click on buttons to identify which sound they had heard.

RESULTS

CORRECT IDENTIFICATION RATES

Table 1 shows the correct identification rates for the different alarms sounds at the end of session 2, which represent our best estimate of the expected performance of the alarm system in use. The table also shows the pre-training identification rates for the sounds, which is our estimate of how easy the sounds would be to identify for a casual hearer, such as a visitor in an ITU.

Figure 1

Table 1: Correct identification rates for the different sounds after training at the end of session 2 and before any training at the start of session 1. * indicates rates statistically above chance: $p < 0.05$.

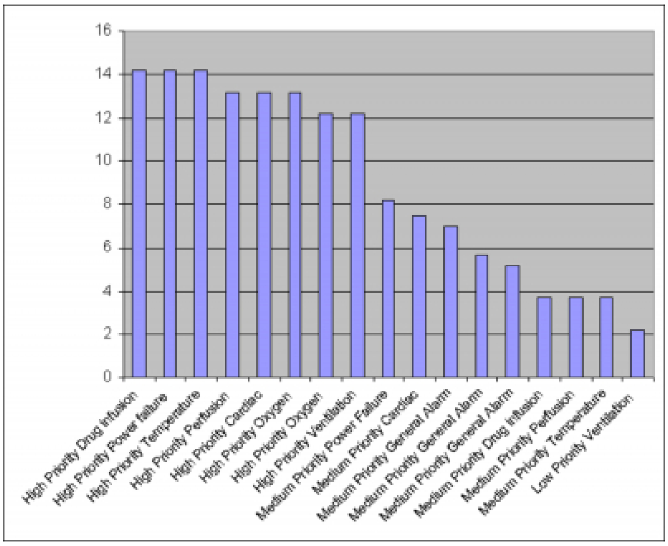
		Final Correct Identification Rate (%)	Correct Identification Rate Pre-training (%)
Medium Priority Alarms	Oxygen	93.02	76.19*
	Ventilation	30.00	26.92*
	Drug infusion	64.29	0.00
	Power failure	29.55	15.00
	Perfusion	71.11	29.55*
	Temperature	76.92	0.00
	Cardiac	44.19	0.00
High Priority Alarms	Oxygen	71.43	48.00*
	Ventilation	17.78	15.00
	Drug infusion	53.66	0.00
	Power Failure	42.86	4.35
	Perfusion	46.15	50.00*
	Temperature	57.14	3.70
	Cardiac	40.48	0.00
General Alarms	Low Priority	88.37	0.00
	Medium Priority	58.97	13.04
	High Priority	72.09	0.00
Overall correct identification rates		**56.35**	**16.57**

The mean overall correct identification rate found was 56.35% (range 17.78% to 93.02%). The mean overall correct identification rate pre-training on the phonically abbreviated alarm sounds was suitably low at 16.57%. However, this low mean figure disguises a large range of 0.00% to 76.19%. Five sounds had pre-training correct identification rates statistically significantly above random (medium priority oxygen, ventilation and perfusion: high priority oxygen and perfusion). Comparing pre-training and final correct identification rates suggest that the basic phonic abbreviations of oxygen and perfusion are too easy to identify without training. The performance of the basic phonic abbreviation for ventilation is also problematical, since it has a high pre-training identification in one of its forms (medium priority ventilation) with noticeably poor post-training identification rates at both levels of priority.

URGENCY MAPPING CHARACTERISTICS

Determining the urgency mapping characteristics of the new sounds was not a primary objective of this study but since an urgency mapping test was included in session 2 some urgency mapping results were available. Figure 1 shows a plot of the mean rank of reported urgency for the different sounds. There is no overlap between the high and low priority sounds in these results. The high priority mean rank range is from 14.17 (high priority perfusion) to 12.17 (high priority cardiac) and the medium priority range is from 8.17 (medium priority power failure) to 3.67 (medium priority oxygen). The sound with the lowest mean rank is the low priority general alarm (rank 2.17). The three main sound priority sub-groups were statistically significantly different from each other (p<0.01, Wilcoxon signed rank test).

Figure 2



DISCUSSION

DO THE PHONICALLY ABBREVIATED SOUNDS PERFORM BETTER THAN THE IEC MELODIC SOUNDS?

The form of this study, the structure of the alarm system tested, the number and age range of volunteers, and the method of test administration, are comparable with those in Williams and Beatty (2005). However, there are two differences between the studies that need to be considered before any comparisons of performance are made.

The mean lengths of the sounds in the different priority groups between the two studies were: low priority 0.22s (IEC) 4.16s (phonic); medium priority 1.02s (IEC), 4.21s (phonic) and high priority 4.18s (IEC), 6.15s (phonic). Length of sound is known to increase perceived urgency (Hellier and Edworthy, 1999) but there is no evidence that it increases learnability.

In this study 11 volunteers (52%) were not native English speakers compared with 8 (38%) in the study on the IEC melodic alarms, though all volunteers in both studies were fluent English speakers as assessed by the experimenters. It is unlikely that being a non-native English speaker would have affected the IEC study but in the case of this study it might have weakened the effect of the semantic associations of the sounds and thus made them harder to learn. Thus we might expect the results of this study to be depressed compared to a group with all native English speakers.

With these caveats in mind, we believe a limited comparison of the overall correct identification rates between the sounds tested here and the IEC melodic sounds is justified.

The mean overall correct identification rates found was 56.35% (range 17.78% to 93.02%) which compares favourably with similar figures for the IEC melodic alarms of 48.41% (range 10.30% to 90.00%), though the overall rates are not statistically significantly different. However, the correct identification rate for the medium priority alarms alone was 58.44% (range 29.55% to 93.02%) which was statistically significantly better than the comparable performance of the IEC melodic alarms ($p<0.05$).

CONCLUSION

The sounds tested in this paper were not designed in a rigorous or systematic way but as illustrations of what might be done using this sort of approach. The detailed reasons as to why some work well and others do not have not was not investigated but are anticipated to include such factors as the repletion of syllabic form between the original words and the abstracted sounds. However, the results of this pilot study suggest that design along these lines may well result in alarms of superior learnability to other systems, without it being unacceptably easy for casual hearers to understand their meaning. Determining the factors required for systematic construction of such sounds should form the basis of future research work.

CORRESPONDENCE TO

Dr Paul Beatty, Imaging Science and Biomedical Engineering (ISBE) The Stopford Building, The University of Manchester, Oxford Road, Manchester M13 9PT E-mail: paul.beatty@manchester.ac.uk Tel: 0161 275 5714

References

r-0. IEC. International Standard IEC 60601 2005 Medical electrical equipment - Part 1-8: General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems. Geneva: International electro-technical commission.
r-1. F. E. Block, J. D. Rouse, M. Hakala and C. L. Thompson 2000 A proposed new set of alarm sounds which satisfy standards and rationale to encode source information J. Clin. Monit. Comput.
16 541-546
r-2. S. Williams and P. Beatty 2005 Measuring the performance of audible alarms for anaesthesia Physiol. Meas.
26 571-581
r-3. P. M. Sanderson, A. Wee and P. Lacherez 2006 Learnability and discriminability of melodic medical equipment alarms Anaesthesia
61 142-147
r-4. F. E. Block. 1992 Evaluation of user's ability to recognize musical alarm tones. Journal of Clinical Monitoring 8 285-290
r-5. E. Hellier and J. Edworthy 1999 On psychophysical techniques to achieve urgency mapping in auditory warnings App. Ergon.
30 167-171

Author Information

Gasthurie Kanagachandran, M.Sc.
Imaging Science and Biomedical Engineering (ISBE), The University of Manchester

Paul Beatty, Ph.D.
Imaging Science and Biomedical Engineering (ISBE), The University of Manchester