Alarms and human behaviour: implications for medical alarms

J. Edworthy* and E. Hellier

School of Psychology, University of Plymouth, Devon PL4 8AA, UK
*Corresponding author. E-mail: jedworthy@plymouth.ac.uk

Alarms are used in many clinical applications, but they are often less than optimal because the design and implementation of alarms has not always taken the cognitive capacity and processing mechanisms of the user into account. As a result alarms are frequently too loud, irritating, confusing, badly designed, and too numerous, resulting in them often being turned off and hindering, rather than enhancing, task performance. This paper reviews some of the main areas where it is essential to take account of the cognitive system of the user and behavioural processes more generally. Five central areas of concern are discussed: the number of alarms and ways that this might be reduced; false alarm rates and their impact on human responses; the design of alarms and the application of research into auditory cognition on design; intelligent alarm systems; and the proposals for alarm design set out in a recent worldwide medical alarms standard. In each area some background is given and the implications for alarm design and implementation outlined. The conclusion is that there are some indications that alarm design and implementation takes account of relevant research data, but that there is still some way to go before these findings are fully integrated and the situation is improved upon further.

Br J Anaesth 2006; 97: 12–17

Keywords: equipment, alarm; monitoring, anaesthetist activity; psychological responses

Alarm sounds are to be found everywhere; not just in the clinical environment but throughout industry, transport and other settings. Over the last 20 yr their use has become even more widespread because of a greater interest in and attention to safety, and also to the litigation which may follow when adverse incidents occur. In practice, alarms perform considerably below their optimum level for many reasons. The main problems found with alarms are that they can often be too loud and shrill (because they are usually installed on a ‘better-safe-than-sorry’ logic) with the risk that they are turned off because they are so irritating, there are usually too many of them, they are often difficult to tell apart from one another and there are usually too many false alarms for the system to be trusted by the user.12 All of these factors contribute to alarms often being seen as getting in the way of a task rather than improving performance on that task. All of these problems are found in the clinical world. In many ways the problem is greater here than in other high-workload safety-critical environments because of the sheer pace of technological change and development in medical equipment, and also because of the transient nature of most clinical applications in which alarms are used. Other safety-critical areas such as nuclear power plants have considerable monitoring and alarming issues to deal with, but usually the equipment used to do this is upgraded or changed slowly and the subject of the monitoring (the core and its associated functions) stays fixed from day to day. There has also been considerable effort by the nuclear industry to apply what is known about alarm design and alarm handling appropriately, and this has generally not been the case in clinical care. In medicine, the equipment used can vary for every patient and every procedure and thus the demands placed on alarm users are considerably greater. Concern has been expressed about the quality of medical alarms by a large number of authors over recent years72 43 34 4 and in the current climate of patient safety the inefficiency of alarms has become something of an issue as patients can die or become injured because of alarm problems. The Joint Commission on the Accreditation of Healthcare Organizations, a US body which evaluates health care and patient safety in more than 15 000 US medical organizations, made clinical alarm safety one of its patient safety goals in 2003; thus emphasizing the importance of proper attention to alarms and alarm issues in clinical care.

This article will review some of the main areas of concern and highlight the problems as well as showing where there have been improvements and recent developments. The areas covered here will be the number of alarms typically used, the design of the alarm sounds themselves, false alarm rates, the use of intelligent alarm systems, and a new and...
than to equipment functions, and then to standardize alarm signals to physiological functions rather than to equipment functions, and then to standardize upon this so that manufacturers are obliged to use the alarms appropriate to the physiological function(s) being monitored as these will not change. This was an idea first put forward 20 yr ago and has been implemented in both older standards and a new alarms standard which is likely to become influential, IEC 60601, to be discussed later.

Deaths have been attributed to confusions between pieces of equipment with similar alarms and in part this is because of the sheer number of alarms often found in clinical settings. One of the most cited pieces of psychological research demonstrates that our ability to remember pieces of unrelated information is limited to 7, plus or minus 2. This applies to remembering numbers, letters, words, sounds and many other items. The implication of this is that it should be difficult to learn and retain the large numbers of alarms typical in many clinical settings. For example, it has been demonstrated that of the nearly 50 alarms in the operating room and intensive care unit of a large Canadian hospital, fewer than half could be recognized by the clinical staff, even when they worked in this same area on a daily basis. Alarms were recognized correctly 39% of the time by nurses and 40% of the time by anaesthetists and operating-room technicians. This study also demonstrated that the acoustic urgency of the alarms was not always linked to the clinical urgency of the situation being signalled. Thus if the person did not recognize the function of the alarm (which was true in the majority of cases) they could not use clues in the alarm to try to narrow down the function which it might be signalling. The type of alarms used in this setting were largely abstract alarms such as beeps and buzzers where there is no intuitive or obvious link between the sound and its function other than one which would have to be learned, making the retention of the alarms even harder, as discussed in the next section.

The issue of number is therefore a very important one. People cannot remember large numbers of alarm sounds for good psychological reasons and the medical profession is no exception. Given that the selection of both the alarms themselves and the situations deemed to give rise to alarms is to some extent up to individual manufacturers (though some of this is controlled by standards), then we can expect the introduction of a new alarm almost every time a new piece of equipment appears. It is more parsimonious to allocate alarm signals to physiological functions rather than to equipment functions, and then to standardize

### Number of alarms

Deaths have been attributed to confusions between pieces of equipment with similar alarms and in part this is because of the sheer number of alarms often found in clinical settings. One of the most cited pieces of psychological research demonstrates that our ability to remember pieces of unrelated information is limited to 7, plus or minus 2. This applies to remembering numbers, letters, words, sounds and many other items. The implication of this is that it should be difficult to learn and retain the large numbers of alarms typical in many clinical settings. For example, it has been demonstrated that of the nearly 50 alarms in the operating room and intensive care unit of a large Canadian hospital, fewer than half could be recognized by the clinical staff, even when they worked in this same area on a daily basis. Alarms were recognized correctly 39% of the time by nurses and 40% of the time by anaesthetists and operating-room technicians. This study also demonstrated that the acoustic urgency of the alarms was not always linked to the clinical urgency of the situation being signalled. Thus if the person did not recognize the function of the alarm (which was true in the majority of cases) they could not use clues in the alarm to try to narrow down the function which it might be signalling. The type of alarms used in this setting were largely abstract alarms such as beeps and buzzers where there is no intuitive or obvious link between the sound and its function other than one which would have to be learned, making the retention of the alarms even harder, as discussed in the next section.

The issue of number is therefore a very important one. People cannot remember large numbers of alarm sounds for good psychological reasons and the medical profession is no exception. Given that the selection of both the alarms themselves and the situations deemed to give rise to alarms is to some extent up to individual manufacturers (though some of this is controlled by standards), then we can expect the introduction of a new alarm almost every time a new piece of equipment appears. It is more parsimonious to allocate alarm signals to physiological functions rather than to equipment functions, and then to standardize

### Table 1 The characteristics of an ideal alarm sound and knowledge which suggests how those characteristics can be achieved

<table>
<thead>
<tr>
<th>Characteristics of ideal alarm sound</th>
<th>Relevant findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to localize</td>
<td>The ear uses two mechanisms for localizing sound, one at high frequencies and one at low frequencies. Neither functions well in the mid-to-high normal band of frequencies of normal hearing</td>
</tr>
<tr>
<td>Resistant to masking by other sounds</td>
<td>Sounds that are acoustically ‘rich’, that is contain a number of harmonics, are more resistant to masking</td>
</tr>
<tr>
<td>Allow communication</td>
<td>Continuous sounds are more likely to be irritating and interfere with communication</td>
</tr>
<tr>
<td>Easy to learn and retain</td>
<td>People are poor at retaining the absolute pitch of a tone and find it difficult to distinguish sounds that vary only in pitch unless they are heard in close temporal proximity. In addition, abstract sounds are harder to learn and retain than environmental sounds or auditory icons.</td>
</tr>
</tbody>
</table>

The design of alarm sounds

If a medical alarms designer wanted to design the best possible alarm sound for a given environment, there is a substantial amount of evidence which suggests how to do so (summarized in Table 1). The resultant alarm would be easy to localize, resistant to masking by other sounds and therefore not easily missed, would not interfere with communication at the very time when it was absolutely necessary, would be easy to distinguish from other alarms, and easy to learn and retain. This ‘idealized’ alarm is unfortunately very different from the ubiquitous medical alarm of a few years ago, the continuous tone, which still to a large extent remains the ‘preferred’ alarm design on many pieces of medical equipment.

There are, however, indicators here and there of improvement in the design of medical alarms in that generic research evidence on alarm design and response is filtering through to the design process. Recent research evidence suggests that abstract sounds are much harder to learn than other types of sound which might be used. One study compared the learnability of abstract alarms, auditory icons and speech. Speech has the advantage that it can clearly signify a problem without any learning being necessary. An auditory icon is a sound which bears some relationship to its function and could, for example, be an everyday or familiar sound. The results showed that the speech alarms were learned almost perfectly from the outset, the auditory icons took a little longer to learn and the abstract alarms took considerably longer. This result was repeated for both initial learning and subsequent recognition of the sounds. A similar finding was shown by another study which compared speech, auditory icons, abstract alarm sounds of the type recommended by
research work,\textsuperscript{40} traditional alarm sounds such as beeps, bells and buzzers, and animal sounds. Again, the speech alarms required little or no learning, the auditory icons took a little longer, the animal sounds took considerably longer, as did both sets of abstract alarm sounds. Other studies\textsuperscript{22,23,25,26,35} have also shown not only that people learn and remember auditory icons better than abstract sounds but also respond more quickly to them. Thus the research data suggest that abstract alarms, the type of alarms traditionally and typically used in the clinical environment, are harder to learn than most other types of sound. Speech appears to be much better, as might be expected, although the applicability of speech messages in clinical settings, where the transparency of the message has the potential to alarm conscious patients and their families would need to be carefully considered. The broad category of sounds classed as ‘auditory icons’ appears to be worth developing in clinical applications. Although the term ‘auditory icon’ may be a loose and rather meaningless label,\textsuperscript{14,17} there is considerable evidence that sounds which bear a closer relationship with their referent are easier to learn than ones which do not.\textsuperscript{29} In the clinical setting this might translate into using a breathing-type sound as a ventilator alarm, for example. Using sounds in this way in clinical settings has been proposed\textsuperscript{8} and their use would be supported in principle by what research in cognitive psychology can tell us about how people understand and learn sounds. The extent to which other issues might mitigate against the use of such sounds is still unclear; for example, an alternative alarm implementation strategy would be to use a general alarm sound that directs the clinician to a ‘master caution’ panel that describes the situation being signalled on a visual display, thereby avoiding the need for clinicians to learn the meaning of individual alarm sounds. Alarm implementation strategies which combine the ‘master caution’ approach for low frequency alarms with the operator learning the meaning of individual high frequency alarms are usually recommended.\textsuperscript{13,16} One issue which has received considerable attention and is relevant to the design of abstract alarm sounds (which still dominate clinical settings) is that of perceived urgency. There is considerable research evidence showing how to construct alarm sounds that are reliably different in their urgency and where the urgency order can be predicted.\textsuperscript{19,22,23,25,26,35} These data can be used to design abstract alarms varying in their degree or urgency so that the alarms can be mapped to the urgency of the situation being signalled. The principles can to some extent be applied to other sound types too including speech in particular,\textsuperscript{18,27} and could be applied to some types of auditory icons also. The advantage of urgency mapping means that, even if the hearer does not know the function of the alarm immediately, the degree of urgency with which the problem should be attended can be conveyed directly through the alarm.\textsuperscript{1,16,37,46}

One further issue with the design of alarm sounds is worthy of mention and that is the burgeoning research area of sonification. Sonification is the science of turning data into sound.\textsuperscript{16,37,46} In sonification, each dimension which the operator wishes to monitor is assigned to a separate acoustic parameter (such as pitch, loudness, speed, harmonic content and so on) and a composite auditory signal is generated. The signal sounds all of the time, and the user is required to listen to it and monitor continually. As there are changes in the parameter in question, so the acoustic dimension which portrays it varies and if there is change in more than one parameter at any one time, then the commensurate number of dimensions varies as a consequence. In clinical settings, sonification has already been applied to the pulse-oximeter, so that a downward trend is indicated by a decrease in the pitch of the audible tone. It is possible that in future clinical applications, a range of different parameters that might need to be monitored (such as heart rate, blood pressure, temperature and so on) could be assigned to different acoustic parameters and variation in one or more of these would indicate a change in the value of the clinical parameter(s) assigned to those acoustic dimensions. Initial research into the usefulness of sonification in clinical settings is encouraging,\textsuperscript{37} but there are issues that remain to be explored. For example, the extent to which clinicians could successfully monitor simultaneous changes in more than one audible parameter remains to be tested empirically. In addition, the practical issue, that alarms may no longer be necessary because the patient is being constantly monitored auditorily, and thus the decision as to when intervention is necessary is one left to the clinician rather than the equipment, remains to be considered. A hybrid system of alarms and trend-monitoring is a further system that has been suggested.\textsuperscript{13,16}

**False alarm rates**

A study of alarm occurrence\textsuperscript{39} showed that of 1455 alarm soundings, only 8 represented critical and potentially life-threatening risk to the patient. Another also suggested that alarms only indicated actual patient risk for 3% of the time that they went off, and in 75% of cases the alarms were in any case spurious. This was shown over a survey of 50 separate operations.\textsuperscript{31} A further study suggested that anaesthetists often turn alarms off, the main reason for this being the unacceptably high false alarm rate.\textsuperscript{8} Thus it is clear that false alarm rates make alarm systems unwieldy and of little value. Research in the psychological domain shows that people adjust their behaviour according to the perceived false alarm rate. If a system is perceived to be 90% reliable then they will respond practically every time an alarm is sounded, whereas if it is perceived to be only 10% reliable then they will respond only infrequently.\textsuperscript{5,6} Thus a medical alarm which is perceived to have high false alarm rates will probably not be responded to or at least the situation which it purports to be signalling will be disregarded until further evidence is collected to confirm that there really is some patient risk.
The problem of false alarms in clinical settings is one which is almost bound to occur and one for which the solutions are not straightforward. For example, even if an anaesthetist is able to set the alarms him- or herself for each individual patient there may be a natural tendency to set them conservatively, with the consequence that the alarms will sound even when there is no immediate risk. In this example, the alarm is really being used as a monitoring sound, suggesting that a risk might develop but it has not yet reached that point. Being able to set multiple levels with possibly different levels of alarm urgency would be a more ergonomic way of dealing with this problem.

Another issue is that alarms often sound when a piece of equipment is turned on, for example when the equipment is self-checking or if it is being checked by the anaesthetist or medical physicist. Under these circumstances the tester knows that the alarms are effectively ‘false’, but this does not necessarily mean that the information is not treated as a false alarm by the cognitive system and thus influences the perceived reliability of the system. Of course, alarms do need to be checked so there is little that can be done about this particular problem.

Thus false alarm rates are typically very high with clinical equipment and much of the problem stems from normal, unavoidable practice. However, the addition of intelligent support typically reduces the false alarm rate significantly, as will be discussed in the next section.

**Intelligent alarm systems**

One of the more recent developments in alarm technology is to have some type of intelligent or expert system underpinning the alarm system. An intelligent system could be supported by a decision-making model using IF–THEN rules or fuzzy logic, by monitoring and integrating the outputs of several parameters at once or, in principle, by a neural part. Part of the burden for interpreting data from the patient falls on the system rather than the clinician, the intelligent system can make decisions about whether or not to alarm, or can be used more generally as decision support for the clinician. There have been successful attempts to develop such systems and one of their key advantages is their low false alarm rate. In some cases the alarm system is graded in some way so that events that are deemed to be more serious are signalled by more urgent or strident alarms. In general, however, the quality of the expert system is not mirrored by the sophistication of the alarm system which is often quite rudimentary, aside from simply prioritizing alarms.

**IEC 60601**

An important and potentially influential IEC standard now governs many aspects of electrical medical equipment, and includes detailed specification of the supporting alarm systems. IEC 60601-1-8 has been adopted as the regulatory framework in the majority of countries including the USA, Japan and the European Union. The standard advocates the assignment of particular alarms to particular physiological functions, as was originally advocated 20 yr ago. 30 IEC 60601 is similar to an earlier standard, IEC 9703; both standards assign alarms to physiological functions, both limit the number of alarms to eight and both use a cautionary and an emergency version of each alarm. Some of the principles of designing perceived urgency into sound have been applied to these signals.

The most important feature of the standards is that they standardize particular sounds for particular physiological functions, which should ensure that the number of alarms is contained, instead of varying across different manufacturers. Also, standards can ensure indirectly that alarms are similar across environments so that once a clinician had learned the meaning of the alarms there would be no need to re-learn different sounds depending on the introduction of new equipment or a move to a new clinical environment.

IEC 60601 specifies in some detail the precise nature of the alarm sounds to be used. These alarm sounds are all 5-tone (emergency version) or 3-tone (cautionary version) melodies. The cautionary versions of the alarms are shown in Table 2 and can be listened to at: http://www.usyd.edu.au/anaes/alarms/. These sounds have been tested and there is some indication that they may be effective.8 The earlier standard, IEC 9703, was also supported by a set of demonstration sounds which were more differentiated than the IEC 60601 sounds. While both sets are effectively tonal (i.e. ‘melodies’), the IEC 9703 sounds varied in their temporal patterns, number of tones in each alarm and so on, whereas the IEC 60601 alarms are all completely regular in rhythm, and two share the same melodic contour. The IEC 9703 alarms should therefore be easier to discriminate between than the IEC 60601 alarms as cognitive psychology tells us that the more dimensions available to us for discrimination between stimuli, the more able we are to differentiate between them. One of the key factors in our ability to discriminate between alarms of this type is their rhythm. Sounds which are different in other ways but share the same rhythm are easily confused. Also, melodic contour has been shown to be a much more salient feature of short melodies.

**Table 2** Characteristics of the IEC 60601-1-8 alarms (cautionary mode). C4 is middle C; C5 is an octave above middle C and C3 is an octave below

<table>
<thead>
<tr>
<th>Alarm</th>
<th>Melody and mnemonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>C4 C4 C4*</td>
</tr>
<tr>
<td>Oxygen</td>
<td>C5-B4-A4 (OX-Y-GEN)</td>
</tr>
<tr>
<td>Ventilation</td>
<td>C4-A4-F4 (VEN-TI-LATE)</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>C4-E4-G4 (CAR-DI-AC)</td>
</tr>
<tr>
<td>Temperature</td>
<td>C4-E4-D4 (TEMP-PRA-TURE)</td>
</tr>
<tr>
<td>Infusion</td>
<td>C5-D4-G4 (IN-FU-SION)</td>
</tr>
<tr>
<td>Perfusion</td>
<td>C4-F4-C4 (PER-FU-SION)</td>
</tr>
<tr>
<td>Power failure</td>
<td>C5-C4-C4 (GO-ING-DOWN/ POW-ER-FAIL)</td>
</tr>
</tbody>
</table>

Perfusion C4-F#4-C4 (PER-FU-SION)

Infusion C5-D4-G4 (IN-FU-SION)

Temperature C4-E4-D4 (TEMP-PRA-TURE)

Cardiovascular C4-E4-G4 (CAR-DI-AC)

Ventilation C4-A4-F4 (VEN-TI-LATE)

Oxygen C5-B4-A4 (OX-Y-GEN)

General C4 C4 C4*
than precise pitch values and thus we would predict that on the basis of their shared rhythms and contours, the new IEC alarm set will be difficult to learn and distinguish between. Research has recently demonstrated that fewer than 30% of anaesthetists were able to identify the new IEC alarms with 100% accuracy after training, particularly if the mnemonic system proposed for their learning was used. This study also found that the participants responded more quickly and accurately to the cautionary than to the emergency version of the alarms even though they rated the urgency of the emergency alarms as being higher than that of the cautionary alarms. Anaesthetists with some musical training also performed better than those without. As research also shows that tonal alarms are difficult to learn and retain (see earlier) this would suggest that the alarms supporting IEC 60601 are considerably less than optimal. However, the effect of standardization on cutting down the total number of alarms to be learned should ease the overall problem considerably.

References

7 Block FE. Do we monitor enough? We don’t monitor enough. J Clin Monit 1986; 2: 267–69
9 Cooper JB, Couvillon LA. Accidental breathing system disconnections. Interim report to the Food and Drug Administration, Cambridge, 1983
15 Edworthy J, Hellier EJ. Fewer but better auditory alarms will improve patient safety. Qual Saf Health Care 2005; 14: 212–15
26 Hellier E, Edworthy J, Dennis I. Improving auditory warning design: quantifying and predicting the effects of different warning parameters on perceived urgency. Hum Factors 1993; 35: 693–706
34 Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 1957; 63: 81–96
35 Momtahan KL. Mapping of psychoacoustic parameters to the perceived urgency of auditory warning signals. Unpublished
Master’s thesis, Carleton University, Ottawa, Ontario, Canada, 1990


39 O’Carroll TM. Survey of alarms in an intensive therapy unit. *Anaesthesia* 1986; 41: 742–4


47 Watson M, Sanderson P. Sonification helps eyes-free respiratory monitoring and task timesharing. *Hum Factors* 2004; 46: 497–517