# The perceptual and cognitive aspects of warning sound design

J. Edworthy

School of Psychology, University of Plymouth, Drake Circus, Plymouth, Devon, United Kingdom PL4 8AA, jedworthy@plymouth.ac.uk

## INTRODUCTION

Ensuring that an auditory warning is appropriately loud for its environment (i.e. it is neither too loud nor too quiet), localisable, and resistant to masking is a demanding set of challenges. However, even if these issues are successfully resolved, it is still the case that warnings and alarms will not necessarily work simply by virtue of being appropriately loud. In this paper I will outline some warning design issues which are the domain of cognitive psychology, rather than that of acoustics and psychoacoustics. I will demonstrate with some examples.

## LEARNABILITY

In some environments, particularly in healthcare, there are a great many alarms and so the issue of learnability has become important. While there might be an expectation that alarms might require some learning (for example, alarms used in places where there are patients and relatives may require some level of coding), alarms that are too difficult to learn will become an unnecessary burden to staff. There is some evidence to show that alarms taken from different classes of sound are differentially difficult to learn and retain (Leung et al. 1997; Ulfvengren 2003; Keller & Stevens 2004). The factor which seems to underpin the ease with which people can learn sounds is the degree of signal-referent relationship, which is the degree to which the sound (the warning or alarm) is associated with the object or situation which it represents. Petocz et al. (2008) have clarified that this association can come from either physically determined relationships between sounds and objects (for example, the sounds actually made by an object when it is doing the thing one wishes to signal, which in semiotic terms is called a sign) or from learned relationships. An example of the former would be tyres skidding when we are braking violently, and an example of the latter might be a school bell which is associated (by most of us) with the end of school lesson. The best example of learned associations between sounds and objects is, of course, speech. The available studies which are underpinned by the strength of the signal-referent relationship demonstrate that both modern and more traditional abstract alarms are difficult to learn, taking many trials. 'Auditory icons', which are a large group of sounds where there is usually some kind of metaphorical relationship between the sound and its referent (for example, a monkey screech representing 'attack') are much easier to learn, and speech is the easiest of all to learn.

Thus, different types of sounds are easier and more difficult to learn, and this might be an important tool in warning sound design. The UK's Rail Safety and Standards Board (RSSB) has an alarms and alerts evaluation tool with a sound library demonstrating different classes of sounds, what their advantages and disadvantages might be, with suggestions as to how they might be used (http://www.rssb.co.uk/sitecollectiondocuments/pdf/research-toolkits/T326/ index.html).

IEC 60601-1-8 (General requirements, tests and guidance for alarm systems in medical electrical systems, 2006) is an important standard for medical alarms and there has been much debate about the efficacy of the alarms designated in the current standard. One issue pertinent to these alarms (there are more which I will come to later) is that the alarms are tonal and abstract. The evidence would suggest that, as a class, tonal alarms would be difficult to learn and research studies looking specifically at these alarms confirms this to be the case (Wee & Sanderson 2008; Sanderson et al. 2011).

## DIFFERENTIATION BETWEEN ALARMS

Many sets of alarms in specific work contexts (such as medicine, the rail industry, vehicles) suffer from being too similar to one another. Many alarms are still very shrill and high-pitched and do not use the auditory space available for design. There are a number of reasons for this: responders to alarms often have a view as to what an alarm should sound like, and therefore expect alarms to be of a particular design; the process of standardisation tends to force the people who design and specify the alarms into a design niche as uniformity is often either a goal, or a consequence, of the process of standardisation; and the methods of producing and signalling alarms, even in an electronic age, tend to be rather conservative. Thus within a particular alarm set typically only one or two classes of sound might be used from the available range. The range available includes traditional alarm sounds, modern electronic abstract alarms, auditory icons (which can in theory be any type of sound but in all cases there is at least a metaphorical link between the sound and the event it is representing), earcons (Brewster et al. 1992) which are tonal cues capable of representing hierarchical organisation through sound, and speech. In addition to alarm sets usually representing only one of the available classes of sound, many are even more restricting as they often do not explore the range of alarms possible even within a particular design niche or class of sounds.

A key and heavily-cited piece of psychological research is that by Miller (1957) which demonstrates that people are able to learn between five and nine items. Beyond this. remembering new material is a problem for short-term memory. However, this paper also demonstrated that the number of dimensions by which stimuli differ from one another also affects memory for those stimuli. Stimuli which differ over more dimensions are, all other things being equal, likely to be more resistant to forgetting and also easier to learn. This has direct application to the design of alarms, as it suggests that a useful design principle would be to design heterogeneity into alarm sets.

A demonstration of the usefulness of heterogeneity in alarm sets is seen in Edworthy et al. (2011). In this study we took a set of already-designed alarms which were considered in some quarters to be a little problematic, partly because of their number (17 in all) and partly because of their design. We increased the heterogeneity of the alarm set by making small changes to some alarms (for example, within the set there were a trio of alarms with related functions, which were very similar to one another, which we modified to sound more different from one another but were still clearly related) and replacing some abstract alarms with auditory icons. We carried out two studies each for the old set and for our redesigned set of alarms. In the first study in each case, we asked participants to rate the difference between all pairs of alarms and derived a hierarchical tree structure (dendrogram) demonstrating the closeness of the links between each of the alarms in the set. These studies showed that the original set was in general closer in similarity than our redesigned set. There were also some undesirable similarities between alarms with different functions in the old set which were rated as very similar to one another. In the second of the two experiments for each set of alarms, we carried out learning trials. We found that our redesigned set was easier to learn. Thus we demonstrated that increased heterogeneity (which is confirmed by the difference data) led to, or at least is correlated with, ease of learning. This is a useful principle for auditory warning design in general. The two examples below demonstrate how designing heterogeneity into auditory warning sets may be used in practice.

## Train protection warning system

We recently undertook a design project to design an auditory warning to accompany a new standard for Train Protection Warning Systems in the UK (Railway Group Standard GE/RT 8083 Issue 3). The standard is to come into operation in April 2012 and supports a new standard for the visual display associated with this function. One of the key requirements of the project was to design an auditory warning which would not be confused with other alerts and alarms in the train cab. Thus a major task within the project was to carry out a review of alarms and alerts already in the train cab. The review showed that many of the alarms were either high-pitched or extremely high-pitched (above 3 kHz), with short, repetitive pulses; one or two were tonal, with variation in pitch (for example, a C-E-G sequence); there were several 2- or 4-pulse repeating alarms; there were a few telephone-like rings (as both external and internal telephones are used in rail cabs); there were a few modern alarms which appeared to be designed along the principles set out by Patterson (1982); and there were one or two traditional-style warnings (a bell and a horn).

The review of the current alarms therefore showed what kinds of designs were already in use, and where a design niche might lie. The decision was taken to design a 3-pulse unit, relatively low in pitch, with a frequency-modulated pulse. This was because there were no frequency-modulated alarms within the set, there were no 3-pulse units in the alarms surveyed, and there were relatively few low-pitched alarms (though good design practice would suggest that low-pitched alarms are better for localisation, irritation and so on). The entire design process was achieved by iterating the design through meetings with a steering committee, who approved the actual design as well as the design remit.

Although the project involved design rather than extensive testing, our research and design projects allow us to predict with confidence that the TPWS alarm will be easily recognisable and stand out from the set of alarms in the train cab. Of course, if there were several frequency-modulated alarms already within the set typically used in the cab, this would be a bad design decision. Thus taking this approach to alarm design has to be case-driven and design decisions depend on what alarms are already used in that environment. Of course, there are some absolute principles of good auditory warning design (e.g. Patterson 1982; Edworthy et al. 1991; Watson & Sanderson 2007) which should be adhered to regardless of the set of alarms as a whole. However, the specific alarm context provided by other alarms in the same environment needs to be taken into account. Heterogeneity within an alarm set is a useful goal.

## IEC 60601-1-8

The international standard on medical equipment alarms was mentioned earlier. This is currently under review, and part of the review concerns the design of the actual alarms in use. The standard precedes an earlier standard, for which there was considerable objection to the specified alarms, which were designed according to Patterson's principles (1982, demonstrated in Patterson et al. 1986). Both the current standard and the earlier one used what are effectively tonal alarms. The Patterson et al. alarms were in fact intended to portray Patterson's design principles based on the construction of a pulse containing acoustic information necessary for localisation, the construction of a short burst of sound akin to a melody, and the construction of complete warnings through repetition of bursts at different levels of loudness, speed and pitch to signify different levels of urgency. However, they were thought of by receivers as melodies. The newer alarms, the ones currently supporting the standard, are considerably more homogeneous than the Patterson et al. alarms, in that they all have the same number of pulses and each conform to a regular and standardised temporal pattern. Specifically, the medium urgency version of the alarms has three pulses and the urgent version consists in each case of the medium priority alarm plus the addition of two pulses. We would predict that this more homogenous set of alarms should be harder to learn than the previous set simply because there is less for the learner to use in order to differentiate between them. A recent study by Sanderson et al. (2011) which compares the earlier set with the current set, as well as a revised new set, shows that all three sets are relatively difficult to learn. This is not surprising, as the research literature shows that tonal alarms are difficult to learn. The results also demonstrate that the newer alarms are more difficult to learn by non-musicians, but that the Patterson et al. alarms do not disadvantage non-musician learners. We can assume that because there is so little variation in the newer set of alarms, only those with special knowledge and ability to attend to small variation in melodic structure (i.e. musicians) can adequately use the cues that differentiate between the alarms. Thus, a narrow design envelope seems to have compromised the learnability of the alarms associated with IEC 60601-1-8. A broader design envelope would seem to be a partial solution to this problem when the alarm sets are redesigned.

Although the broadening of the design envelope seems to help in alarm differentiation and learnability, there are many other issues which need to be borne in mind if this regime followed. For example, listeners expect alarms to have particular attention-getting qualities and there may be some sounds which simply do not work in this way for acoustic or other reasons. The degree to which sounds can function as alarms might interact with the nature of the signal-referent association. For example, alarms with low signal-referent associations may need to sound 'alarm-like', but those with high signal-referent associations may not need to sound 'alarm-like'. We respond to sounds in our environment all of the time in (usually) an appropriate manner, so there is no reason to suppose that we would respond any differently to alarm sounds with strong signal-referent relationships. Sounds with strong signal-referent relations are, almost by definition, either environmental sounds or abstract sounds which are very familiar to us.

## FALSE ALARMS

The issue of false alarms is probably the most prevailing current problem with alarms, particularly in the medical arena. This is particularly true in areas where complex information is being presented and sometimes monitored by the apparatus which presents the alarms. In the medical arena, false alarms sound all the time and there are many different causes such as setting alarm thresholds too low or conservatively; the equipment or the patient falsely triggering an alarm because of some state which cannot be avoided, but which is not important in that context; and medical and other interventions which trigger the alarms (Imhoff & Kuhls 2006). There is consequently a burgeoning of practical protocols aimed at reducing the number of alarms in several medical spheres (e.g. Keller et al. 2011) which have proved successful.

One important source of false alarms is the way the parameters which are being monitored relate to the alarm and its triggering. Ultimately, the success of an alarm system in this type of application can only be as successful as the algorithm underlying the information flow from the patient, to the monitor, and to the recipient of the alarm. Imhoff & Kuhls (2006) describe a number of algorithms which may be used, but report that the use of these algorithms is rather spasmodic in practice. Herein lies an important issue for the alarm designer. No matter how good the design, the efficacy of the alarm will always be undermined by a high false alarm rate, as high false alarm rates leads to high scepticism about the reliability of the underlying alarm system (Bliss et al. 1995; Bliss & Dunn 2000). For example, there is ample research literature demonstrating how the perceived urgency of alarms can be varied, in order to allow mapping between the urgency of the alarm and the urgency of the situation being signalled (e.g. Edworthy et al. 1991; Hellier et al. 1993). The extent to which this can be achieved effectively is necessarily restricted by the efficacy of the underlying alarm algorithms.

## FUTURE DIRECTIONS

Considerable research and knowledge on auditory processing and auditory cognition in relation to alarms and auditory warnings has been gained over the last few years. Little of this appears to have filtered through to practice and application, though there are examples which demonstrate the principles which flow from these findings. Alarms and auditory warnings problems are becoming increasingly salient to those in a position to change policy and practice, in particular in the medical arena. The increased interdisciplinary effort now being aimed at alarm and auditory warning practice should help the flow of expertise from one area to another, which is long overdue.

## REFERENCES

Bliss JP, Dunn MC (2000). Behavioural implications of alarm mistrust as a function of task workload. Ergonomics 43:1283-300.

Bliss JP, Gilson RD, Deaton JE (1995). Human probability matching behaviour in response to alarms of varying reliability. Ergonomics 38: 2300-2312.

Brewster S, Wright PC, Edwards A (1992). A detailed investigation into the effectiveness of earcons. In: Kramer G (ed): Proceedings of the First International Conference on Auditory Display. Santa Fe Institute, NM: Addison-Wesley.

Edworthy J, Loxley S, Dennis I (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. Hum Factors 33: 205-231.

Hellier E, Edworthy J, Dennis I (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. Hum Factors 35: 693-706.

Edworthy J, Hellier E, Titchener K et al. (2011). Heterogeneity in auditory alarm sets makes them easier to learn. Int J Industr Ergon 41: 136-146.

IEC 60601-1-8 (2006). International Standard: Medical electrical equipment – Parts 1-8: General requirements, tests and guidance for alarm systems in medical electrical equipment. Geneva: International Electrotechnical Commission.

Imhoff M, Kuhls S (2006). Alarm algorithms in critical care monitoring. Anesth Analg 102: 1525-1537.

Keller P, Stevens A (2004). Meaning from environmental sounds: Types of signal-referent relationship and their effect on recognising auditory icons. J Exp Psychol: Appl 10: 3-12.

Keller JP et al. (2011). Why clinical alarms are a 'top ten' hazard. Horizons 2011 (Spring): 17-23.

Leung YK, Smith S, Parker S et al (1997). Learning and retention of auditory warnings. In: Proceedings of the Third International Conference on Auditory Display. Palo Alto, USA.

Miller GA (1957). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 63: 81-96.

Patterson RD (1982). Guidelines for auditory warnings systems on civil aircraft. UK, Civil Aviation Authority contract report No. 82017. London: Civil Aviation Authority.

Patterson RD, Edworthy J, Shailer MJ et al (1986). Alarm sounds for medical equipment in intensive care areas and operating theatres. Southampton: Institute of Sound and Vibration, report no. AC598.

Petocz A, Keller P, Stevens C (2008). Auditory warnings: Signal-referent relations, and natural indicators: Re-thinking theory and application. J Exp Psychol: Appl 14: 165-178.

Sanderson PM et al (2011). Learning and remembering auditory equipment alarms: Patterson-Edworthy sounds help anaesthesiologists who have no musical training. Anesth Analg (in press).

Ulfvengren P (2003). Designing natural warning sounds in human-machine systems. Doctoral thesis, Stockholm Institute of Technology.

Watson MO, Sanderson PM (2007). Designing for attention with sound: Challenges and extensions to ecological interface design. Hum Factors 49: 331-346.

Wee AN, Sanderson PM (2008). Are melodic medical equipment alarms easily learned? Anesth Analg 106: 501-508.