2018-10-01

Getting alarm sounds into a global standard: A cases study with reflections

Edworthy, Judy

10.1177/1064804618763268
Ergonomics in Design
SAGE Publications

Getting alarm sounds into a global standard: a case study with reflections

Judy Reed Edworthy[1], Richard R McNeer[2], Christopher L Bennett[3], Roman Dudaryk[2], Siné J P McDougall[4], Joseph J Schlesinger[5], Matthew L Bolton[6], Jonathan D Reed Edworthy[1], Elif Ožcan Vieira[7], Andrew D Boyd[8], Scott K J Reid[1], Michael F Rayo[9], Melanie C Wright[10], David Osborn[11]

[1] Cognition Institute, Plymouth University, Plymouth, Devon, UK

[2] Miller School of Medicine, University of Miami, Miami, FL , USA

[3] Frost School of Music, University of Miami, Miami, FL, USA

[4] School of Psychology, Bournemouth University, Bournemouth, Dorset, UK

[5] Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN

[6] Department of Industrial and Systems Engineering University at Buffalo, SUNY, Buffalo, NY

[7] Faculty of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands

[8] Department of Biomedical and Health Information Sciences College of Applied Health Sciences University of Illinois at Chicago, Chicago, IL

[9] Department of Quality and Patient Safety, Wexner Medical Center, Ohio State University, Columbus, OH

[10] Saint Alphonsus Regional Medical Center, Boise, ID

[11] Philips, Salem, NH

The reserved set of audible alarm signals embodied within the global medical alarm system standard, IEC 60601-1-8, are known to be problematic and in need of updating. The current alarm signals are not only suboptimal, there is also little evidence beyond learnability (which is known to be poor) demonstrating their performance in realistic and representative clinical environments. In this paper we describe the process of first designing, and then testing, potential replacement audible alarm signals for IEC 60601-1-8, starting with the design of several sets of candidate sounds, initial tests on learnability and localizability, followed by testing in simulated clinical environments. We demonstrate that in all tests the alarm signals selected for further development outperform the current alarm signals (by a long way), and we describe the process of collecting considerably more data on the performance of the new sounds than we have for the current sounds, which will ultimately be of use to end-users. We also reflect on the process and practice of working with the relevant committees and other practical issues beyond the science which also need constant attention if the alarms we have developed are to be included successfully in an updated version of the standard.

Keywords: auditory alarms; standards; alarm learnability; clinical alarms; safety

1. Introduction

Audible alarm signals are very important across high-workload industries and their use in those environments is not always driven by the best science, but by other factors such as customer reaction, budget, lack of expertise in design and application of knowledge, inflexible and/or conservative approaches to known problems, and so on. Therefore there are many examples of high-workload, safety-critical environments where the audible alarm signals leave much to be desired in terms of both their implementation and design, though increasingly there are also many examples of thoughtful and designed implementations. In clinical environments the problem of bad alarm system implementation has reached colossal proportions, where patient deaths have been attributed to 'alarm fatigue' (Drew et al, 2014; Sendelbach & Funk, 2013). Until a national summit in the US in 2011 (http://www.aami.org/events/eventdetail.aspx?ItemNumber=1153&loggedOut=True), little was being done about the general problem of alarm condition over-use but now there are well-documented and successful attempts to reduce the problem of over-alarming in general (Cvach, 2012; Welch et al, 2011; Whalen et al, 2014). The audible alarm signals that annunciate the hazards have traditionally also left a lot to be desired from the point of view of design, but now that the broader alarm system problems are slowly being resolved, the time is right to improve upon the audible alarm signals as well. In this paper we describe a project intended to upgrade and update the audible alarm signals in a global medical equipment standard.

The challenge of carrying out what is essentially an applied, customer-based problem while maintaining the best scientific approach one can muster is a challenge. This issue is

highlighted by Morrow and Durso (2011) in their editorial on a special issue of JEP:Applied on Cognitive Issues in Healthcare. They introduce their papers thus:

'…we focus on the need for research that is sufficiently comprehensive to identify threats to patient safety, yet specific enough to explain how provider and patient factors interact with task and health context to engender these threats. Such research should be theory-based, yet also problem-driven; exert experimental control over theoretically relevant variables, yet also involve participants, tasks, and contexts that represent the problems of interest. A tension exists between theory-based, experimentally controlled research on the one hand, and problem-driven research with representative situations on the other' (p.191)

The challenge in terms of audible alarm signal design is to bring the scientific evidence to bear on the problem, but also to commit at some point during the process to a specific set or set of sounds so that a research database can be built around them.

The evidence base for auditory alarm signal design is considerably more advanced than the typical sorts of alarm signals that are used in practice might suggest. Bridging the 'valley of death' between theory and application is always a problem, made more acute in auditory work given the difficulty of talking to non-experts (often the client) about sound in any abstract way, and given the predisposition that clients have to like or dislike a sound designed for a specific application. Reactions to alarms can sometimes be colored by the existing, often adverse, alarm environment. For example, nurses are typically already overwhelmed with alarm signals (Honan et al, 2015) so anything which looks like an addition to the alarm system environment (such as a new set of audible alarm sounds) needs to be presented within the context of a transition which will ultimately be of benefit to those working with those alarms on a day-in, day-out basis.

2. The standard: IEC 60601-1-8

IEC 60601 is a set of standards concerned with the safety of medical electrical equipment (so covers almost all medical equipment). Parts 1-8 specifies the basic safety and essential performance requirements and tests for the alarm systems contained within that equipment. Thus this standard governs almost all medical equipment across the globe. It was published first in 2006, then updated in 2012, had something of an update in 2015 and is due for another, major update by the end of 2019. The key feature of the standard in terms of audible alarm signals is that it specifies the acoustic and structural elements of the audible alarm signals that should accompany specific clinical hazards or categories (IEC, 2012).

The reserved set of alarm signals was designed with the best of intentions (Block et al, 2000), based on some aspects of what was known about alarm signal design at the time (but not all). The sounds embodied important acoustic features that would increase their resistance to masking (compared at least with single harmonics), and improve their general acceptability over the earlier beeps, buzzers, and bells. The structure of the alarm signals and their categories is shown in Table 1. There are eight categories of risk specified. Each has a high- and a medium-priority form. In our studies only the high-priority version was tested, though generic medium- and low-priority alarm signals were also tested for this update.

Table 1: IEC 60601-1-8 High Priority Alarm Signal Characteristics (from Edworthy et al, 2017a)

Function of Alarm	**Alarm Signal Characteristics**
General	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: c c c – c c
Power down	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: C c c – C c
Cardiovascular	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms, followed by a burst of two regularly spaced pulses in the following pattern: c e g – g C
Perfusion	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: c f# c – c f#
Drug Administration	A burst of three regularly spaced pulses (each pule ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: C d g – C d
Oxygen	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: C b a – g f
Ventilation	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: c a f – a f
Temperature	A burst of three regularly spaced pulses (each pulse ranging between 100ms – 300ms), followed by a burst of two regularly spaced pulses in the following pattern: c d e – f g

A key problem with the design was that the alarm signals, which sound like short, tonal melodies, all possess the same number of pulses and the same rhythm, making them very hard to distinguish between (Lacherez et al, 2007; Sanderson et al, 2006; Wee & Sanderson, 2008). The lack of diversity between the sounds is a major contributor to the known problems with learning and recognizing these alarm signals, and the finding is no surprise given that our ability to distinguish between stimuli depends on the number of dimensions along which they vary (Miller, 1956). A shared rhythm is also a key component of a listener's confusion between sounds (Patterson 1982). Calls to update and improve the sounds have been numerous, with the designer of the sounds himself issuing an apologia for the current sounds (Block, 2008).

It has become clear that almost anything would be better than the current alarm signals, which presents its own problem. Atyeo and Sanderson (2015) demonstrated that a similar set of alarm signals designed prior to the 2006 version of IEC 60601-1-8 (designed for an earlier version of the standard, Patterson et al 1986) outperform the current alarm signals, and other evidence shows that a random set of audible sounds, with no association to the meanings or functions of the alarm conditions, was easier to learn than the current alarm signals (Edworthy et al, 2014). The earlier (1986) set of sounds was rejected on a non-empirical basis which allowed interested parties to call into a telephone line and listen to the alarm sounds, and then to voice an opinion. However, that was the 1980s and patently, replacing the current alarm signals with sounds that simply perform better than the current alarm signals – even those designed in the 80s which turned out to be better than the alarm signals in the standard - is not enough.

Commentary 1

Despite knowing of the existence of the 'IEC 60601-1-8 alarm problem' for years prior to the start of the project, it was important to conduct the project with the endorsement of the body charged with updating the standard rather than conducting the work in isolation and then presenting it to that body, and to wait for a head of steam to build up over any potential replacement. The bodies in this case are the IEC 60601-1-8 and AAMI 60601-1-8 standards committees, through an IEC alarm systems joint working group, which have a common core and some cross-over membership. Access to this group was made possible by the Association for the Advancement of Medical Instrumentation (AAMI) having an open policy on membership of its own parallel IEC 60601-1-8 committee, AAMI 60601-1-8. The first author joined and began attending meetings. AAMI later made a grant to the first author to carry out the initial development work.

Changing and updating standards is akin to the proverbial changing of the course of a ship using a teaspoon. The process of bringing about change in standards is very slow, and requires sustained attention. The challenges and demands of achieving global standardization in our increasingly technological world is well-documented across several spheres such as finance and medical devices (Abbot & Snidal, 2001; Cheng, 2003; Hallstrom, 2004; Mattli & Büthe, 2003). Achieving standardization even of the relatively straightforward and contained issue of medical device alarms inevitably involves stakeholders with many different vested interests, most of which are market- and financially-driven.

The fate of earlier work heightens our awareness of non-empircally-based criticisms and potential scuppering, which is best met with empirically-based answers. Thus a key

element of our strategy is to create a published and accessible database at every point in the process.

3. The process

Figure 1 shows the process we have adopted in developing the alarm sounds

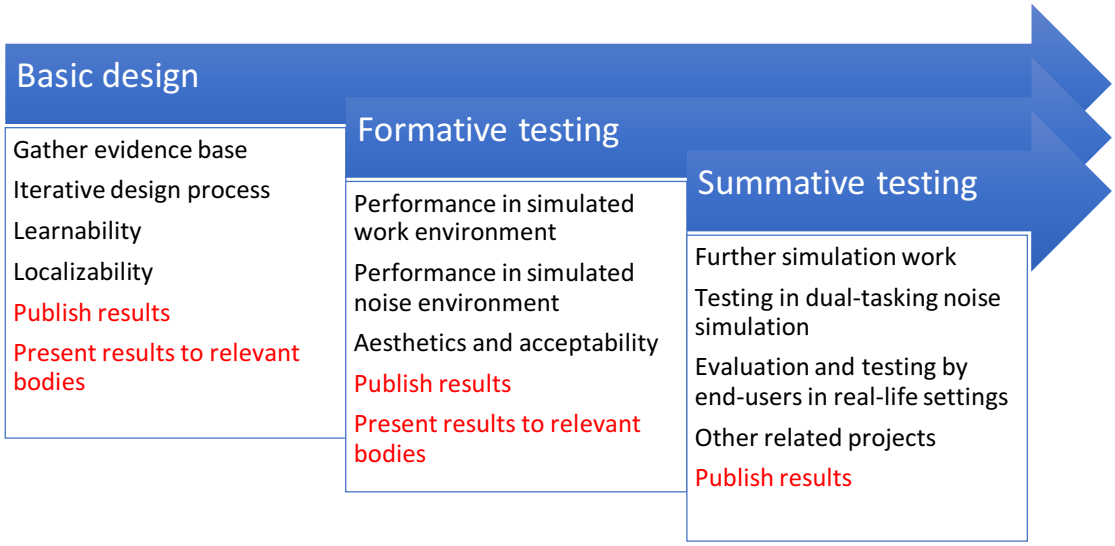


Figure 1: The stages of updating the audible alarm signals for IEC 60601-1-8

Whereas medical equipment audible alarm signals have traditionally been produced by poor quality sounding devices, many medical devices are now equipped with good quality speakers. Sound storage and reproduction is also much cheaper, all of which means that, potentially, almost any sound can be used as an alarm signal, and the sound reproduction can be of high quality. This doesn't make the work of the designer any easier, indeed it focuses the effort required to demonstrate that any new alarm signals are not only 'better', but 'the best', or among the best, possible. A key question is what constitutes 'best'. Here, we have to start with learnability (whether or not it is important, though it

probably is) as learnability is the only data we have on the current alarm sounds and comparisons are a good starting point, indeed essential in making the preliminary arguments for adoption of any new sounds.

3.1. Basic design

There is ample evidence to show that the concrete-abstract continuum plays a big role in the learnabiity of sounds, and there are many published examples of 'auditory icon' alarm signal designs which outperform abstract sounds (Belz et al, 1999; Edworthy et al, 2014; Graham, 1999; Keller & Stevens, 2004; Leung et al, 1997; Perry et al, 2007; Petocz et al, 2008; Stephan et al, 2006; Ulfvengren, 2003). Auditory icons, which are typically real-world sounds with direct associations to their meanings, are obviously high in their concreteness though there are other ways in which metaphors can be achieved. Actual speech-based sounds (including speech itself) are also readily learnable, as demonstrated in encouraging findings for 'spearcons', speech-based alarm sounds in a clinical context (Li et al, 2017). For the standard itself it was felt that speech was not appropriate, but we did include a set of alarms for testing which were based on word rhythms and patterns.

We developed four sets of sounds which used different types of metaphors for the eight alarm conditions, and compared them to the current sounds, which have no, or very minimal, metaphors. We tested them in terms of learnabiity and localizability. In one ('word rhythms'), the eight words of the functions were imitated and stylized in terms of number of syllables, rhythm, and tonal structure. For example, the Cardiovascular alarm was a 6-tone, evenly-spaced sound ('Car-di-o-vas-cu-lar') with the first three pitches were the same, and the second three were the same, with the second three lower than the first. This set is

mapping (I.e. the current sounds, which did make some attempt at mnemonics). A second set ('Resilient') were designed with lower acoustic fidelity, aimed at devices which might have low sound reproduction quality. For these, half again used the word-rhythm association and half used simple metaphors – for example, for temperature the alarm sound was a tone glide upwards and for power down it was a tone glide downwards. We expected these metaphors to be relatively easy to learn, and the word rhythms to be approximately the same as for the 'word rhythm' set. The other two sets were both auditory icons, one set of which contained an abstract 'pointer' and one of which did not. The sets were identical other than this. For each of the eight alarm categories, a combination of focus groups, questionnaires and repeated discussions within the research group led to the identification of appropriate metaphors for each of the alarm conditions. For some alarm conditions the most appropriate metaphor was obvious (for example, a heartbeat sound for Cardiovascular) but for others the most appropriate metaphor was less obvious. Although we refined and tested three metaphors for each function in later testing (see later), subsequently it turned out that by and large we had selected the 'best' metaphors at this first attempt. We also took care to ensure that there was acoustic variability across the set of auditory icons, in order to minimise possible confusion.

The learnability data for the sound sets can be seen in Figure 2. All of our designs were more memorable than the existing set (all lines on the graph were significantly different from one another except the two at the top), but there was variation across our experimental sets also, with the auditory icon sets being the most memorable. The performance data suggests that we have covered the range of responses here, in that the performance for the auditory icons was almost at ceiling level from the start, and the

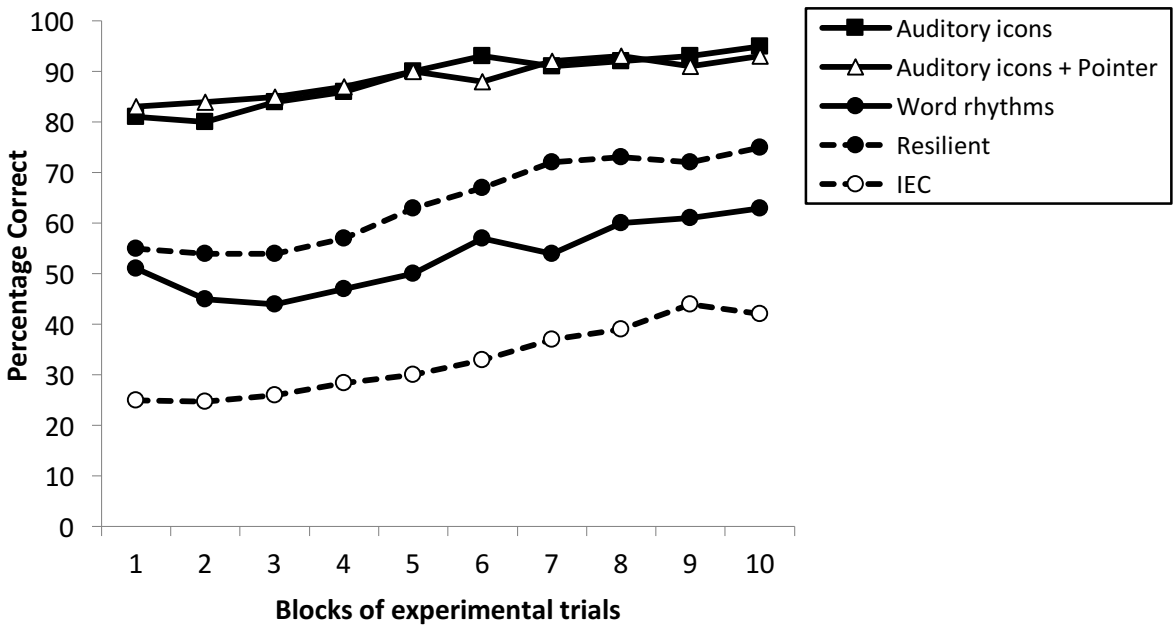


Figure 2: Percentage correct responses for each set of alarm signals, across ten trials (from Edworthy et al 2017a)

The candidate alarm signals were also varied in their harmonic complexity and denseness, as by and large more harmonically dense sounds are easier to localize. Very few tests of alarm signal localizability have been conducted (Alali, 2011; Catchpole et al, 2004; Vaillancourt et al, 2014), though localizability is often a pertinent issue in clinical care (for example in a multibed ICU). Our results confirmed that the more harmonically dense alarm signals were easier to localize, and that the least complex, the current alarm signals, were weakest in localizability (Edworthy et al, 2017a).

Reflection 2

The findings from the basic study (Edworthy et al 2017a) were presented to the standards alarm system joint working group in April 2016. They were also presented to the AAMI 60601-1-8 committee in June 2016, and to a meeting of the AAMI alarms coalition in

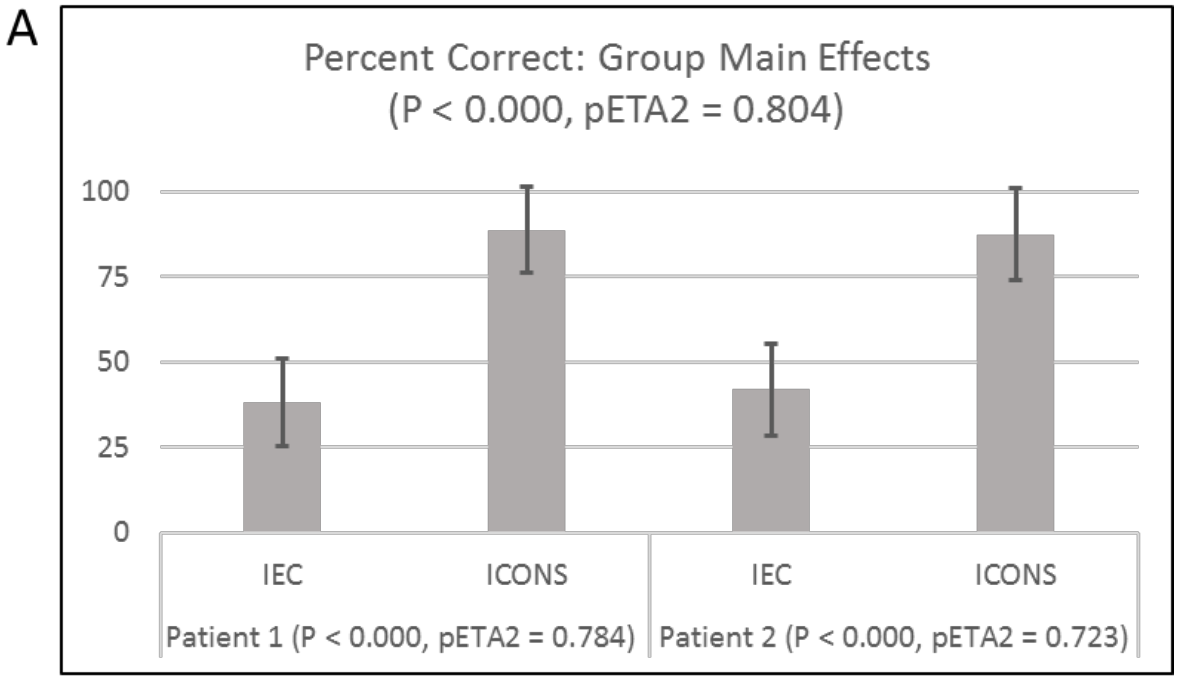
July 2016. The empirical evidence was presented along with the sounds As a consequence of this, the alarm system joint working group decided they wished to go ahead with the auditory icons plus pointer design, and supplied a list of activities, some formative and some summative, they would like to see undertaken prior to the committee recommending the adoption of the alarm signals into the standard. A further grant from AAMI to the first author was negotiated on this basis.

Another unexpected consequence is that there appears to be a substantial amount of dissent over the categories of risk themselves. We have approached this by writing a paper to open out discussion of the categories themselves (Edworthy et al, 2017b). AAMI has made a grant available to Dr Wright to carry out research on this issue.

3. Formative testing

Mindful of Morrow & Durso's call for the use of contexts, tasks, and participants of relevance (2011), the formative testing involves more realistic tasks, using clinically-trained participants. Using a range of already-developed and published techniques (Bennett & NcNeer, 2012; Bennett et al 2015; McNeer et al 2016), a paradigm was developed whereby trained anesthesiologists carried out a short clinical simulation task. They were required to monitor two patients and respond to alarm signals by indicating the nature of the alarm condition (its category), and their reaction times also were measured. Prior to this they were given a brief exposure to either the auditory icon plus pointer alarm signals, or the current IEC alarm signals. Results indicated very early on that the auditory icons produced faster and more accurate responses than the current IEC alarm signals (McNeer et al, 2017a, b). Results of the early trials can be seen in Figure 3. Secondary workload and fatigue

icons are also less frustrating and impede performance less than the current alarm signals. Here, we may be tapping into 'alarm fatigue'. This is important, because though the concept of alarm fatigue is generally accepted, and there certainly is a clinical alarm problem, the details of its manifestation and dimensions are somewhat sketchy (Deb & Claudio, 2015; Kristensen et al, 2016; Rayo & Moffat-Bruce, 2015).



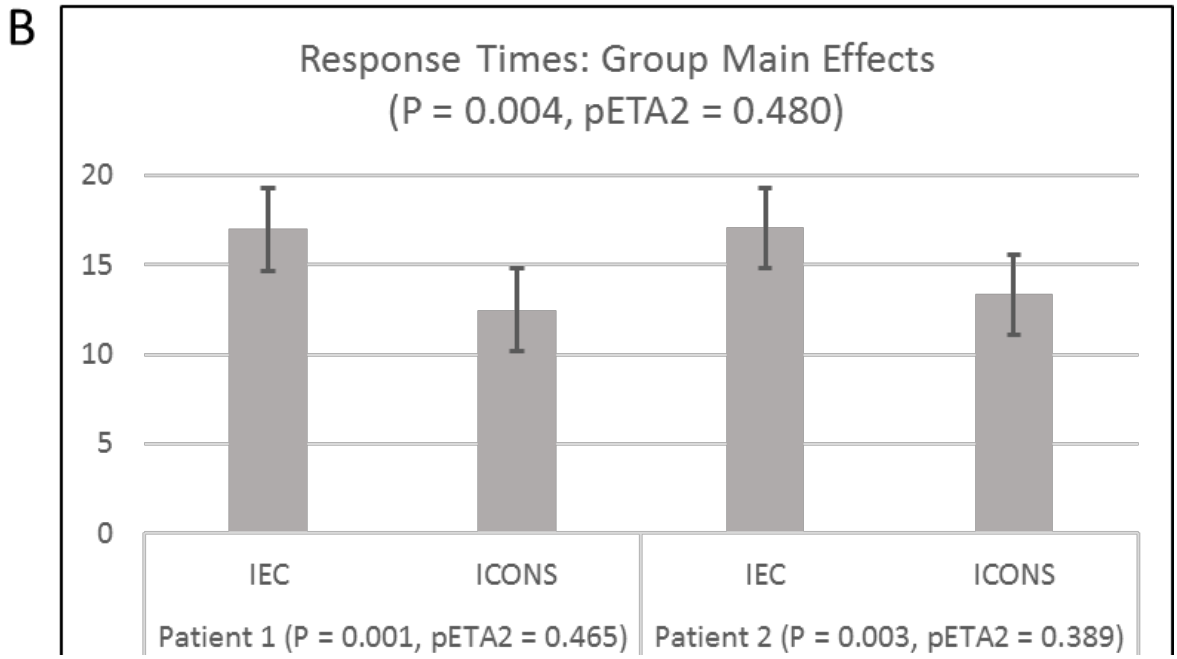


Figure 3: Mean percentage correct identification and mean reaction times to the new alarm signals ('Icons') or the current IEC alarm signals ('IEC') (from McNeer et al, 2017a). x axis = percent correct, y axis = IEC audible alarms ('IEC') or auditory icons ('ICONS'). 'Patient 1' and 'Patient 2' refer to the two simulated patients being monitored by participants. A shows percent correct and B shows reaction time.

The final phase of the formative testing in simulation was to test three versions of each auditory icon. Three different auditory icons were generated for each function (we also added two further functions, 'brain activity' and 'monitor error' see comments about the categories later) and tested each of them in the simulation paradigm. We derived a 'dream' and a 'nightmare' set dependent on performance. The compound results for both reaction time and accuracy in identification is shown in Figure 4, which has undergone a transformation so that for both measures, higher scores are better. Here we see that the 'dream' team outperforms the 'nightmare' team (statistically significantly) and that lower reaction times are associated with more accurate recognition. Thus, some auditory icons simply work better than others.

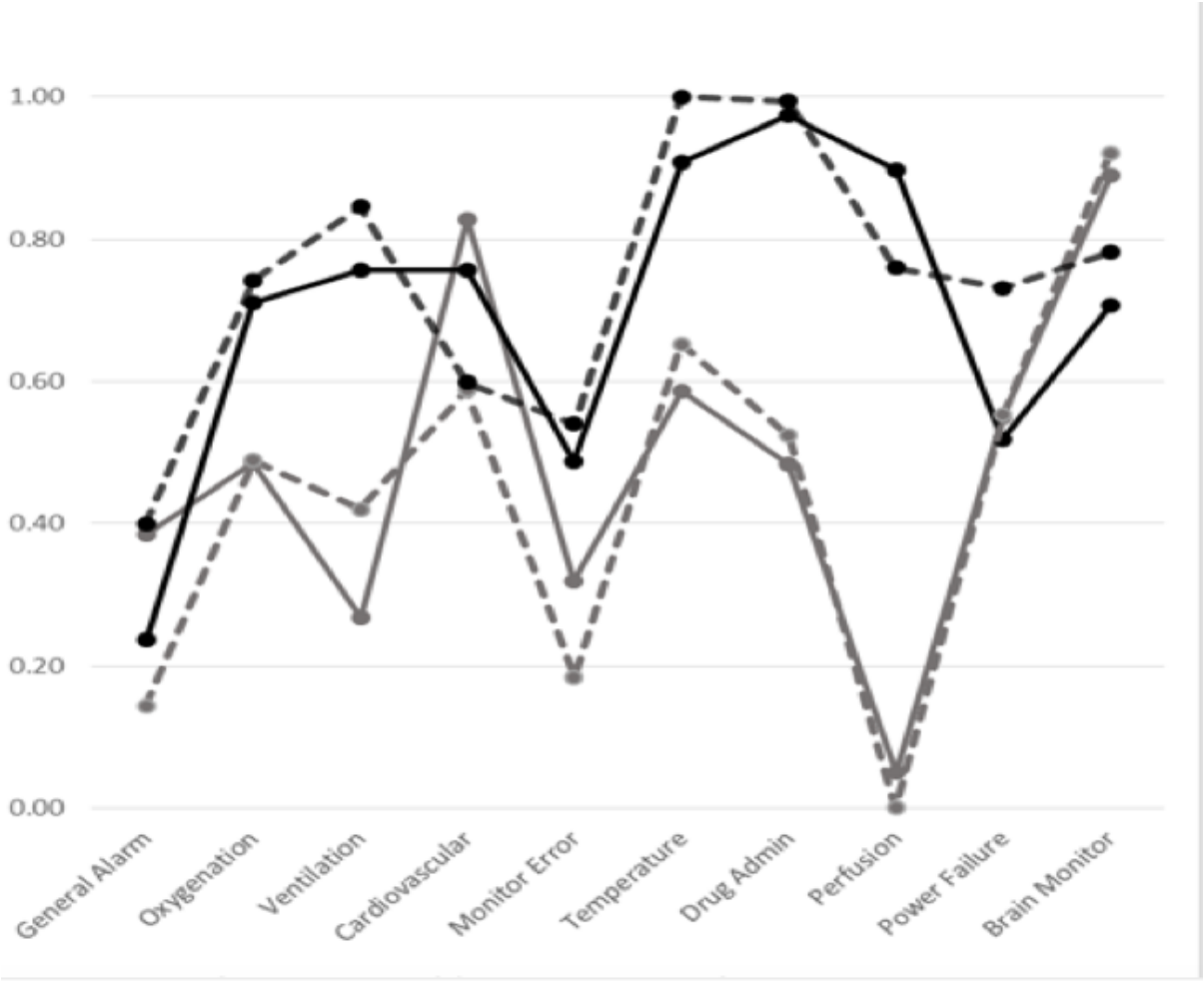


Figure 4: Binary response (a transformed composite of reaction time and accuracy) for 'dream' and 'nightmare' alarm sets. x axis = accuracy/time index relative to best performing sound for RT (Temperature); y axis = 10 alarm sound categories

Other studies currently being carried out as part of the formative (and more recently summative) testing include the audibility of the alarm signals in realistic listening conditions; findings thus far indicate that the sounds work well in relatively low signal-to-noise ratios (a finding being demonstrated for alarm signals more generally in other studies (Schlesinger et al, 2014; Stevenson et al, 2013) and that the presence of the pointer enhances audibility. The pointer in particular was found to be audible in noise that was four times louder.

Reflection 3

Because the alarm signals are intended for the update of the standard and therefore access to them will be of commercial advantage, the final sounds will be released to medical instrumentation companies via a website, through AAMI (the final details of this process are yet to be decided). Several companies are keen to do their own testing on the sounds once they are released. Another aspect of updating the standard is to update and enhance the guidance given to stakeholders, particularly sound designers, human factors engineers working with clinical device safety, medical instrument companies and test houses, among others.

3.3. Summative testing and other work

The summative testing will follow the broad protocols of the formative testing, with additional researchers testing the sounds in a range of clinical environments using protocols yet to be developed, as well as accepted and published protocols (Schlesinger et al, 2014; Stevenson et al, 2013). There is also other, related work being conducted. Dr Bolton is currently leading an AHRQ-funded project grant looking at the issue of masking of auditory alarm signals with specific reference to IEC 60601-1-8 (Hasanain et al, 2017). This research will fill a large gap in terms of understanding where and when auditory masking will occur, which is somewhat beyond the scope of the immediate project described but is very important in general terms in understanding audible alarms from a human factors perspective. This model checking approach uses formal methods, which are computing methods used for the specification, verification, and modelling of systems. It works as

propositions of the system. If the properties hold, the model can confirm this, and if it doesn't hold (i.e. it throws up a counterexample) then the specific set of values which gives rise to the counterexample can be checked. Thus it is an efficient way of assessing a system which could otherwise not be achieved. It has often been used to assess automated systems in human factors but not for auditory masking specifically. The model uses several sub-models including a clock submodel, an alarms submodel, and a masking computation submodel. Using actual audible alarms as input, the model can predict whether alarms will mask one another or not under specific conditions (for example, the onset of the timing of one alarm relative to one or more others). The model is still in the process of refinement and testing with human participants. Naturally this project is aware of both the current alarm signals and the projected new alarm signals, which will help ensure its validity as a practical instrument and also pushes the functionality of the software to more complex masking tasks. We are also carrying out more theoretical studies on the contributions of strength of metaphorical link and auditory diversity in sound set learning, as these two dimensions are thought to be large contributors to the effectiveness of any set of alarm signals.

Reflection 4

The work is on track to be completed to the satisfaction of the IEC alarm system joint working group well before the updated standard is published in 2019. By this time, there will be several published papers documenting the performance of the alarm signals from basic testing to their performance in simulated environments, their performance in noise and in other, increasingly realistic, tasks. Of course, the project will not have reached a satisfactory

conclusion until the alarm signals and the relevant advice is embodied within the standard and there is still a way to go and other possible unknown threats along the way.

We anticipate that our work will improve patient safety and clinical work performance, as well as contributing to the science of alarm design and implementation.

References

Abbott, K. W., & Snidal, D. (2001). International'standards' and international governance. *Journal of European Public Policy*, *8*(3), 345-370.

Alali KA. (2011) Azimuthal Localization and Detection of Vehicular Backup Alarms Under Electronic and Non-electronic Hearing Protection Devices in Noisy and Quiet Environments. Ph.D. thesis. Blacksburg, Virginia, USA: Virginia Polytechnic Institute and State University; 2011.

Atyeo, J., & Sanderson, P. M. (2015). Comparison of the identification and ease of use of two alarm sound sets by critical and acute care nurses with little or no music training: a laboratory study. *Anaesthesia*, *70*(7), 818-827.

Belz, S. M., Robinson, G. S., & Casali, J. G. (1999). A new class of auditory warning signals for complex systems: Auditory icons. Human Factors: The Journal of the Human Factors and Ergonomics Society, 41(4), 608-618.

Bennett, C. L., & McNeer, R. R. (2012). PT-SAFE: A software tool for development and annunciation of medical audible alarms. Anesthesia & Analgesia, 114(3), 576-583.

Block Jr, F. E. (2008). “For if the trumpet give an uncertain sound, who shall prepare himself to the battle?”(I Corinthians 14: 8, KJV). Anesthesia & Analgesia, 106(2), 357-359.

Block FE, Rouse JD, Hakala M, Thompson CL. (2000) A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. J Clin Monit Comp, 16:541– 6

Catchpole, K. R., McKeown, J. D., & Withington, D. J. (2004). Localizable auditory warning pulses. Ergonomics, 47(7), 748-771.

Cheng, M. (2003). *Medical device regulations: global overview and guiding principles*. World Health Organization.

Cvach, M. (2012). Monitor alarm fatigue: an integrative review. *Biomedical instrumentation & technology*, *46*(4), 268-277.

Deb, S., & Claudio, D. (2015). Alarm fatigue and its influence on staff performance. *IIE Transactions on Healthcare Systems Engineering*, *5*(3), 183-196.

Drew, B. J., Harris, P., Zègre-Hemsey, J. K., Mammone, T., Schindler, D., Salas-Boni, R., ... & Hu, X. (2014). Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit

Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., ... & Pye, E. (2017a). The Recognizability and Localizability of Auditory Alarms: Setting Global Medical Device Standards. *Human Factors*, 0018720817712004.

Edworthy, J. R., Schlesinger, J. J., McNeer, R. R., Kristensen, M. S., & Bennett, C. L. (2017b). Classifying alarms: Seeking durability, credibility, consistency, and simplicity. *Biomedical Instrumentation & Technology*, *51*(s2), 50-57.

Edworthy, J., Hellier, E., Titchener, K., Naweed, A., & Roels, R. (2011). Heterogeneity in auditory alarm sets makes them easier to learn. International Journal of Industrial Ergonomics, 41(2), 136-146.

Edworthy, J., Page, R., Hibbard, A., Kyle, S., Ratnage, P., & Claydon, S. (2014). Learning three sets of alarms for the same medical functions: A perspective on the difficulty of learning alarms specified in an international standard. *Applied ergonomics*, *45*(5), 1291-1296.

Graham, R. (1999). Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. Ergonomics, 42(9), 1233-1248.

Hallstrom, K. T. (2004). Organizing international standardization: ISO and the IASC in quest of authority.

Hasanain, B., Boyd, A. D., Edworthy, J., & Bolton, M. L. (2017). A formal approach to discovering simultaneous additive masking between auditory medical alarms. Applied Ergonomics, 58, 500-514.

Honan, L., Funk, M., Maynard, M., Fahs, D., Clark, J. T., & David, Y. (2015). Nurses' perspectives on clinical alarms. *American Journal of Critical Care*, *24*(5), 387-395.

International Electrotechnical Commission. (2006, 2012). IEC 60601-1-8: Medical electrical equipment–General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems.

Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: types of signal-referent relations and their effect on recognizing auditory icons. Journal of experimental psychology: Applied, 10(1), 3.

Kristensen, M , Edworthy, J & Ozcan Vieira, E (2017) In search of acoustical signs of alarm fatigue. *Sound Effects*, special issue on *Sound and Listening in Healthcare and Therapy,* 6(1), 88-104

Lacherez, P., Seah, E. L., & Sanderson, P. (2007). Overlapping melodic alarms are almost indiscriminable. Human Factors: The Journal of the Human Factors and Ergonomics Society, 49(4), 637-645.

Leung, Y. K., Smith, S., Parker, S., & Martin, R. (1997). Learning and retention of auditory warnings. Proceedings of the International Community for Auditory Display, 1997. http://www.icad.org/Proceedings/1997/LeungSmith1997.pdf

Li, S. Y., Tang, T. L., Hickling, A., Yau, S., Brecknell, B., & Sanderson, P. M. (2017). Spearcons for Patient Monitoring: Laboratory Investigation Comparing Earcons and Spearcons. *Human Factors*, 0018720817697536.

Mattli, W., & Büthe, T. (2003). Setting international standards: technological rationality or primacy of power?. *World Politics*, *56*(1), 1-42.

McNeer, R. R., Bennett, C. L., & Dudaryk, R. (2016). Intraoperative noise increases perceived task load and fatigue in anesthesiology residents: a simulation-based study. Anesthesia & Analgesia, 122(6), 2068-2081.

McNeer, R, Bennett, C.L., Horn, D.B., Dudaryk, R & Edworthy, J (2017a) The Aggregate Performance of Novel Auditory 'Icon' Alarms in a Simulated Intensive Care Unit Environment. Poster presented at International Anesthesia Research Society ,Washington DC, May 2017

McNeer, R, Bennett, C.L., Horn, D.B., Dudaryk, R & Edworthy, J (2017b) Use of Ordinal Regression Analysis for Assessing and Comparing Individual Alarm-type Performance in a Simulated Intensive Care Unit Environment. Poster presented at International Anesthesia Research Society ,Washington DC, May 2017

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Morrow, D. G., & Durso, F. T. (2011). Patient safety research that works: Introduction to the special issue on human performance and health care. *Journal of Experimental Psychology: Applied*, *17*, 191-194.

Patterson, RD (1982) Guidelines for auditory warnings systems on civil aircraft. Civil Aviation Authority paper 82017: Civil Aviation Authority.

Patterson RD, Edworthy J, Shailer MJ, Lower MC, Wheeler PD (1986) Alarm sounds for medical equipment in intensive care areas and operating theatres. Report No. AC598: Institute of Sound and Vibration Research, Southampton, UK

Perry, N. C., Stevens, C. J., Wiggins, M. W., & Howell, C. E. (2007). Cough once for danger: Icons versus abstract warnings as informative alerts in civil aviation. Human Factors: The Journal of the Human Factors and Ergonomics Society, 49(6), 1061- 1071.

Petocz, A., Keller, P. E., & Stevens, C. J. (2008). Auditory warnings, signal-referent relations, and natural indicators: re-thinking theory and application. Journal of Experimental Psychology: Applied, 14(2), 165.

Rayo, M. F., & Moffatt-Bruce, S. D. (2015). Alarm system management: evidence-based guidance encouraging direct measurement of informativeness to improve alarm response. BMJ Quality & Safety, 24(4), 282–286

Sanderson, P. M., Wee, A., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. Anaesthesia, 61(2), 142-147.

Schlesinger, J. J., Stevenson, R. A., Shotwell, M. S., & Wallace, M. T. (2014). Improving pulse oximetry pitch perception with multisensory perceptual training. *Anesthesia & Analgesia*, *118*(6), 1249-1253.

Sendelbach, S., & Funk, M. (2013). Alarm Fatigue A Patient Safety Concern. *AACN advanced critical care*, *24*(4), 378-386.

Stephan, K. L., Smith, S. E., Martin, R. L., Parker, S. P., & McAnally, K. I. (2006). Learning and retention of associations between auditory icons and denotative referents: Implications for the design of auditory warnings. Human Factors, 48(2), 288-299.

Stevenson, R. A., Schlesinger, J. J., & Wallace, M. T. (2013). Effects of divided attention and operating room noise on perception of pulse oximeter pitch changes; A laboratory study. The Journal of the American Society of Anesthesiologists, 118(2), 376-381.

Ulfvengren, P. (2003). Associability: A comparison of sounds in a cognitive approach to auditory alert design. Human factors and aerospace safety, 3(4), 313-331.

Vaillancourt, V., Nélisse, H., Laroche, C., Giguère, C., Boutin, J., & Laferrière, P. (2013). Comparison of sound propagation and perception of three types of backup alarms with regards to worker safety. Noise and Health, 15(67), 420.

Wee, A. N., & Sanderson, P. M. (2008). Are melodic medical equipment alarms easily learned?. Anesthesia & Analgesia, 106(2), 501-508.

Welch, J. (2011). An evidence-based approach to reduce nuisance alarms and alarm fatigue. *Biomedical Instrumentation & Technology*, *45*(s1), 46-52.

Whalen, D. A., Covelle, P. M., Piepenbrink, J. C., Villanova, K. L., Cuneo, C. L., & Awtry, E. H. (2014). Novel approach to cardiac alarm management on telemetry units. *Journal of Cardiovascular Nursing*, *29*(5), E13-E22.

Acknowledgements

Research reported in this paper was supported by the Association for the Advancement of Medical Instrumentation in two grants to Dr Edworthy, 'Audible alarms in support of IEC 60601-1-8 Phase 1/Phase 2'.

Research reported in this paper was also supported by the Agency for Healthcare Research and Quality under award number R18HS024679 to Dr Matthew Bolton. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Dr. Michael Rayo's participation in this effort was supported by the Institute for the Design of Environments Aligned for Patient Safety (IDEA4PS) at The Ohio State University which is sponsored by the Agency for Healthcare Research & Quality (P30HS024379). The authors' views do not necessarily represent the views of AHRQ.