2017-02

Classifying Alarms: Seeking Durability, Credibility, Consistency, and Simplicity

Edworthy, Judy

http://hdl.handle.net/10026.1/9552

10.2345/0899-8205-51.s2.50
Biomedical Instrumentation & Technology
Association for the Advancement of Medical Instrumentation (AAMI)

How should we classify alarms? Seeking durability, credibility, consistency, and simplicity

Judy Reed Edworthy[1], Joseph J Schlesinger[2], Richard R McNeer[3], Michael Sonne Kristensen[1] & Christopher L Bennett[4]

1 Cognition Institute, Plymouth University, Plymouth, Devon, UK PL4 8AA

2 Vanderbilt University Medical Center, Nashville, TN 37203

3 Miller School of Medicine, University of Miami, Miami, FL 33136

4 Frost School of Music, University of Miami, Miami, FL 33146

Submitted to 'Horizons' special issue on clinical alarms. Do not share and do not quote

Corresponding author: Judy Edworthy jedworthy@plymouth.ac.uk

Accepted for publication 27th January 2017

Alongside the development and testing of new audible alarms intended to support IEC 60601-1-8 - a global standard concerned with alarm safety - the categories of risk that the standard denotes require further thought and possible updating. In this article we revisit the origins of the categories currently in the standard, which are based on the ways in which tissue damage can be caused. We consider these categories from the varied professional perspectives of the authors: from that of human factors; semiotics; clinical practice; and patient/family (layperson). We conclude that while the categories possess many clinically applicable and defensible features from our range of perspectives, the advances in alarm design now available may allow a more flexible approach. We present a three-tier system with superordinate, basic and subordinate levels that fit both within the thinking embodied in the current standard and possible new developments.

Background

Work is currently underway to update the audible alarms associated with an important global medical device standard, IEC 60601-1-8[1]. The standard, which is concerned with the safety of medical devices, specifies the audible alarms that should accompany the risk categories set out in the standard. The audible alarms themselves have been demonstrated to be less that optimal[2-5]. Four sets of prototype updates have been developed and are in the process of being benchmarked[6]. The alarms will then be made available for further testing. An important issue emerging from this work is that it may not simply be the audible alarms, but also the categories of risk specified in the standard, that require updating. In this paper we revisit the categories from our multidisciplinary perspective, in an attempt to open up discussion of the categories, and suggest how updating them might be approached. The writing team consists of a human factors and auditory alarms specialist; a semiotician; two anesthesiologists; and, for the purposes of this article, a layperson who might be a patient, or a member of a patient's family. This paper is not a systematic, or even a narrative, review, but a collection of viewpoints aimed at stimulating debate. We also provide an updated proposal in an attempt to stimulate the debate further.

The categories

The standard specifies eight risk categories. The six central ones come from Kerr [7] The thinking behind Kerr's paper was influential in the development of the principles still embodied in IEC 60601-1-8, and the categories suggested by Kerr are still present in the standard. In Kerr's paper there is some discussion about different ways in which alarms might be categorized – always bearing in mind that an important aim was to keep the numbers of alarms to a manageable number – after which Kerr comes down in favour of what he refers to as a 'risk-and-response-based' approach. The basis of the approach is that it captures all of the ways in which tissue damage can occur, and the response required to ensure that damage does not occur. The categories are as follows: *Hypoxia (H* - failure to deliver oxygen (check oxygen supply)*; Ventilation problems (V)* - including disconnection and high airway pressure (check airway and ventilator); *Cardiovascular problems (C* -check circulatory status*; Interruption to artificial perfusion (P)* - check perfusion machine e.g. haemodialysis*; Drug administration error (D)* - check syringe/infusion pump*; Thermal risk (T)* - check heating and cooling devices.

One of the consequences of Kerr's categories is that a single piece of equipment may need to produce more than one alarm. For example a ventilator might need to produce H, V and T alarms. Anecdotally, this is one of the aspects of the categories that sometimes attracts criticism as many of those who interact with equipment on a day-to-day basis might reasonable argue that the most obvious way to classify alarms is on the basis of the piece of equipment for which the alarm is relevant. The traditional objection to an equipment-based approach is that it would firstly lead to proliferation of alarms and secondly be difficult to standardize for, as new equipment constantly comes on to the market.

In IEC 60601-1-8 itself, these categories have been translated as Oxygen, Ventilation, Cardiac, Artificial perfusion, Drug or fluid delivery, and Thermal risk. A 'Power down' category was subsequently added, as was a general alarm, intended to substitute for any of the others and to be a superordinate alarm. In 1986 a set of audible alarms was designed for these eight categories, closely following Kerr's recommendations, which are now usually referred to as the Patterson-Edworthy sounds[8]. Though they pre-date the alarms currently supporting IEC 60601-1-8, they have recently been shown to outperform the current IEC alarms[9] .

The Human Factors perspective

Human factors seeks to design around humans and their limitations, rather than imposing systems upon them. Human factors approaches are typically user-centred, and employ a variety of techniques aimed at eliciting and understanding users' needs. In the case of clinical alarms, a user-centred approach to alarm categorization might attempt to understand the way that information is represented and organised at a cognitive level by those for whom the alarms are relevant. Work on visual displays focusing on ecological interface design[10] , which usually begins with a technique known as cognitive work analysis, can and does provide detailed information about both what is to be acted upon, and what is subsequently done. Work on integrated visual displays using these techniques provide a useful research basis for understanding how clinicians group and understand patient monitoring information, and how their understanding might be enhanced or compromised by how that information is presented[11-13]. These techniques and findings are relevant to understand users' needs in terms of alarm categories and is a topic which should be explored in future research.

Human Factors approaches also favour standardization wherever possible, and this is certainly relevant in this case as standardization will reduce the burden of learning new alarms and possibly new categories when moving from one workplace to another. Standardization also helps to minimize other, residual problems associated with auditory alarms. Key among these is the risk of masking (where one sound conceals the sounding of another), and irritation to the user, the risk of both of which will be reduced if the number of alarms is kept relatively small. Other goals would be some level of stimulus-stimulus compatibility (in that there are links between sound and situation, possibly in terms of meaning and/or urgency). From this perspective, Kerr's system represents good human factors and ergonomics as the number of categories is small, and (if designed well) the alarms associated with the categories can convey the meanings intended. Whether or not the categories are meaningful and relevant to the user is currently not well understood, and is worthy of further investigation.

Kerr was very keen to restrict the numbers of different alarms to 6-8 (with a maximum of 10), as evidence at the time suggested that this was the limit to how many alarms could reasonably be learned[14]. Indeed, his proposed classification system was partly driven by the fact that the number of categories could not be increased (because the causes of tissue damage were unlikely to change) and would thus render the system future-proof and self-limiting. However, while the alarms currently specified in IEC 60601-1-8 are difficult to learn and retain, some types of alarm are not difficult to learn[15-17] and some of the suggested updates are also easy to learn[5,6]. This means that we may have greater flexibility in thinking about future classification systems, though avoiding proliferation of alarms should always be a core aim.

The semiotician's perspective

Semiotics refers to the study of signs and symbols, with a wide application across science, the arts, and social science. As we are talking about alarms and the categories that are annunciated by those alarms, semiotics is highly relevant and revealing and gives us a useful viewpoint for looking at the categories. The categorization of objects and events is a fundamental human activity and the way in which we categorize, and the level of granularity that we apply to our classifications, is directly connected to the issue described in this paper. While Human Factors methods can inform what we alarm about, and when, the broader and perhaps more theoretical issue as to how we build a categorization system with the appropriate balance of generality and granularity remains. The study of category formation can help in this.

In her seminal work on category formation Eleanor Rosch proposed a three-tiered taxonomy (‘superordinate’, ‘basic-level’, and ‘subordinate’) to describe how humans categorize objects in the world, and showed that basic-level categories have the highest degree of cue validity[18]. For instance, the concepts *furniture (superordinate)*, *chair (basic level)*, and *armchair (subordinate)* are very closely related semantically but differ in their levels of informativeness. ‘Chair’ is a much more tangible concept than ‘furniture’ (as its features in the physical world can be perceived and represented) and therefore considered more meaningful. And despite being *less* specific than ‘armchair’ it is a much more commonly used word in people’s vocabulary cross-culturally. In other words, there seems to be a golden mean when it comes to finding an appropriate level of abstraction for things in the world that we want to refer to by means of some sign. This will be true of the alarm categories of IEC 60601-1-8 and indeed will be true of any categorization system where something (a sound, picture, icon etc) represents something else.

As well as the level of granularity at which the categories are set, the level of categorical consistency is also important. Categorical consistency has two elements: vertical and horizontal consistency. Horizontal taxonomic consistency designates the level of variability *between* categories. The pertinent question here is to what extent different alarm categories should work according to the same semiotic principles. As discussed in more detail below (the perspective of the anaesthesia provider) a prerequisite for optimizing the semiotic power (i.e. strength of representational value) of a *set* of alarm sounds is to assign priority to the most significant elements of each category. For standardization purposes this has the important consequence that the optimal alarm philosophy might involve a (severely) skewed distribution of sound-interpretant mappings between the different alarm categories – for example, towards having several sounds for different cardiovascular functions.

By vertical taxonomic consistency is meant the degree of variability in the level of abstraction within the alarm categories. Vertical consistency clearly differs from one category to another in the current approach. For example the cardiovascular system consists of various components that have perceivable manifestations in the physical domain: You can see and touch the heart, the blood, and the vessels. But the cardiovascular system can also be attributed with a quasi-perceivable property such as 'pumping' (from the movement of the heart one can infer that the heart is pumping but one cannot really *perceive* the pumping, only the heart). And it can be attributed with more abstract properties like *metabolism*, *circulation*, and *transportation*. Other categories, for example drug delivery, are associated largely with infusion devices so is much more straightforward than cardiovascular and some of the other categories. Thus the categories fall short of ideal in terms of both vertical and horizontal consistency. One solution might lay in developing subordinate categories for some of the risks.

In the sections which follow we discuss whether the category 'cardiovascular', for example, is the appropriate level of granularity or whether subordinate categories might be more meaningful to the clinician in certain contexts, for example the operating room.

The clinician's perspective

It is important to note that our perspective here is that of the anesthesia provider only. We have not included other views and in particular we have not canvassed the view of nurses. It will be important to canvass the nurses' view as they are often the people who interact most, along with the anesthesia provider, with clinical alarms. From a clinician's perspective one of the key aims is to keep the number of alarms to a minimum, as much as anything to keep the noise levels down (these can become very high, particularly in the operating room). Thus, categorization of alarms at a meaningful and quite general level may be useful. Also, indicating an appropriate level of urgency is helpful as this will give a first indication of the speed with which a response should be made. The categories of alarm however have to be directly useful to the clinician, and the current categories may be somewhat suboptimal because of their lack of consistency and practical relevance, as described in the previous section. There is also scope for tweaking both the categories and any subordinate levels within the categories as a function of the type of activities that go on in different areas. Here we highlight the ICU and the OR. Ideally, the issue as to what to alarm about and when to do it is best approached by using known knowledge elicitation techniques and building on existing knowledge, as described earlier.

The ICU

The intensive care unit (ICU) is rife with a myriad of alarms – some true, some false, some indicating minor physiologic abnormality, and some indicating patient decompensation. The clinician must be able to discern the alarms and prioritize the auditory signal to provide safe and effective patient care.[19] The anesthesiologist-intensivist leads a team comprised of nurse practitioners, medical students, interns, residents, a pharmacist, a nutritionist, and other allied health students. This multidisciplinary team must take care of critically ill patients while working with bedside nurses and interacting with families with confidence, skill, and grace. The unique practice environment of the ICU, compared to the operating room (OR), is the increase in patient census and the presence of families. With the high occurrence of alarms and a large percentage of them false, there is a high degree of mistrust in the information provided by alarms.[20-22] An apathetic attitude towards alarms imparts a lackadaisical air of dedication to patient care to the observant families. Decreasing the number of alarms, admixed with improving the information transfer of the auditory signal can serve to improve the complex and interesting multidisciplinary approach to patient care in the ICU. The anesthesiologist-intensivist's perspective to this problem centers on differing practice locations with different equipment, availability of biomedical support, varying demographics of patient pathophysiology in specific ICUs, urgency of information, central versus peripheral alarms, and patient exposure to alarms.

At the most general level, two major bifurcations exist – what will cause immediate harm and requires immediate action, and who needs the alarm or alert information. Problems with ventilation, oxygenation, and hemodynamic stability are immediate alarms to which the entire care team should be exposed. Drug administration is primarily a nursing task (except in the operating room), and thermal risk is important, albeit typically less acute than the other categories of risk in the standard. This suggests that the categories have important differences between them in terms of their relevance to the whole or only parts of the team, as well as their maximum urgency levels. Thus the categories as they are currently proscribed do not necessarily possess all of the features that might be desirable in an optimal classification system.

The OR

Since 1985, several technologies anticipated by Kerr have become commonplace in the operating room arena, including the ‘alpha-numeric indicators’ or visual displays and to a limited extent the application of a ‘centralized (*though not yet smart*) alarm system’ (p.706), both of which are present in the modern anesthesia workstation. When attended to by the anesthesia provider who is often seated, the workstation is akin to a cockpit providing both continual and continuous patient- and ventilator-state updates through an audio-visual, non-standardized interface meant to facilitate and maintain situational awareness. In fact, delivery of general anesthesia can be categorized into three periods, each with an analogue to the airline industry: Induction (taking off); maintenance (cruising altitude); and emergence (landing). A single all-purpose alarm sound (one of Kerr’s suggestions) will fail to be useful during induction, emergence and emergencies, when visual redirection is often not possible and alarms may come from sources apart from the workstation. Almost every anesthesia provider has at one point experienced a ‘perfect storm’ scenario in which imminent harm to patient is heralded by a barrage of auditory signals, some coming from the surgical field or surgeons and others coming from various device- or patient-associated monitors. The anesthesia provider may expect that having an alarm system in place that consists of an easily learnt set of alarm sounds that conveys both meaning and urgency (potentially decreasing cognitive load or facilitating appropriate allocation of cognitive resources) would be useful for timely and effective crisis management.

From the perspective of the anesthesia provider, the alarm categories suggested by Kerr and currently part of the IEC 60601-1-8 standard may not be as useful in the OR relative to other patient care areas. Patients requiring hemodialysis usually receive this therapy either before or after surgery, and most providers will rarely provide anesthesia to patients requiring artificial perfusion for circulatory support (cardiac bypass) or oxygenation (extracorporeal membrane oxygenation). Therefore, the allocation of an audible alarm to this category may be seen as being superfluous and wasting cognitive resources that could be allocated to other more relevant categories – though if the category is never used, having an unused category may make little difference in practice. It may be beneficial to increase the number of categories related to cardiac status, or to subdivide the cardiac category. In semiotic terms, this means developing subordinate categories beneath the basic category of ‘cardiovascular’. Thus, instead of one ‘cardiac’ alarm being overburdened by being associated with blood pressure, heart rate, heart rhythm, and cardiac output, for example, it may be possible to design alarm sounds for each, thus increasing the amount of information conveyed through the auditory medium during times when attention is directed elsewhere.

The layperson’s perspective

It is important to say from the outset that audible alarms emanating from machinery surrounding the patient are there to inform and attract the medical staff, not the patient. A consequence of this is that they are also heard by the patient. Our introspections about patient and patient visitors/family lead us to suggest that important requirements may be that of understanding the level of urgency, and not being constantly bombarded by overly-urgent sounding alarms.

Just like most interactions that occur in clinical settings, the patient and bedside family members expect that the clinicians will serve as an “expert interpreter” of the medical devices and their various alarms. This emerges from the patient/care provider covenant in which the patients trust that the clinician will be both knowledgeable and transparent regarding their interpretation. The question as to what degree should the alarming information be apparent to the patients *without first having gone through the “filter” of the clinical staff* is key. The answer to this question bifurcates depending on the situation and the patient. In some instances, knowing the exact meaning of the alarm can assuage an anxious patient. For example, in a recovery room, it is typical for an IV alarm to sound when the IV bag is running low. However, to an unsuspecting layperson, this alarm sounds as ominous as a much more serious alarm. In this case, bypassing the clinician’s interpretation by knowing that the alarm is simply a reminder to swap out the IV bag would *provide a benefit to the patient*. Equally, if the alarms themselves demonstrated some level of urgency mapping then the patient would be able to interpret the urgency from the sound itself.

In other instances, however, alarms can be misleading outside of context. Patient monitors are typically pre-programmed with alarming thresholds, which may or may not be appropriate, depending on the patient – for example SpO_2 levels will be different for a fit young person in comparison to an ageing smoker, so similar (default) settings are not appropriate. In this case, the clinician’s interpretation is required to contextualize the alarm, and therefore bypassing the clinician would *provide <u>no</u> benefit to the patient*.

These examples highlight the need to appropriately “thread the needle”, so to speak, with the type and amount of information conveyed to the patient via audible medical alarms. “Self-diagnosis” is a double-edged sword that can either deteriorate or alleviate patient anxiety and even health status. From a patient’s perspective, having an appropriate sense of *urgency* is an important factor. In very few cases can the patient herself have an impact on the care provided, with the exception of calling for help. If urgency is well encoded within alarms and immediately apparent to the layperson or patient, then the difficult task of keeping the patient informed *just enough* may be achievable. And to whatever degree the alarms can sound less urgent than “life threatening” (which is unfortunately all too often the case), the less anxious the patients might be.

Conclusion

The idea of a self-limiting principle which sets the categories of risk at the basic level of categorization is appealing, ergonomic and useful for clinicians, but it is unclear as to what that principle should be at the point of writing. This should however be the topic of further research. For now, we present a summary and an example of how we could think more flexibly about alarm categories.

Our thinking has led us to the conclusion that basic level categories (whatever may drive them, be it equipment, risk level, risk category or something else) are useful but that the addition of subordinate categories might be added where needed. Let us assume that whatever system is developed, it will include a ‘general’ category and a ‘cardiovascular’ category at the very least. For illustration purposes we also refer to a ‘drug administration’ alarm in order to exemplify a less important category. In Figure 1 we set out a classification system where there is a general ‘alarm’ category (the superordinate category) which may not be meaningful at a clinical level, but helps to think about the issue of alarming. The basic level categories are small in number, and indeed could be as already proscribed in IEC

60601-1-8 (or with some modification, or driven by a different principle, depending on the outcome of future research). Below the categories lie the subcategories, which might have fewer or more categories themselves (including none) according to need. Thus in the case of the cardiovascular category there may be several subordinate alarms, and in the case of drug administration there may be none (assuming there is a drug administration category for the purposes of illustration). The level at which actual audible alarms should be proscribed is an issue still up for discussion – for example, the categories could simply be categories, with or without specific audible alarms to support them.

Finally, although we have talked about alarm categories rather than the audible alarms themselves, there are many issues surrounding the sounds themselves. The desire for urgency mapping is strong, as is the desire for alarm sounds which are easy to learn and are informative, rather than shrill and alarming. The work on the design of new alarms intended for IEC 60601-1-8 meets these requirements. Future work focusing on the categories themselves might well be fruitful if the benefits of the design work are to be optimized.

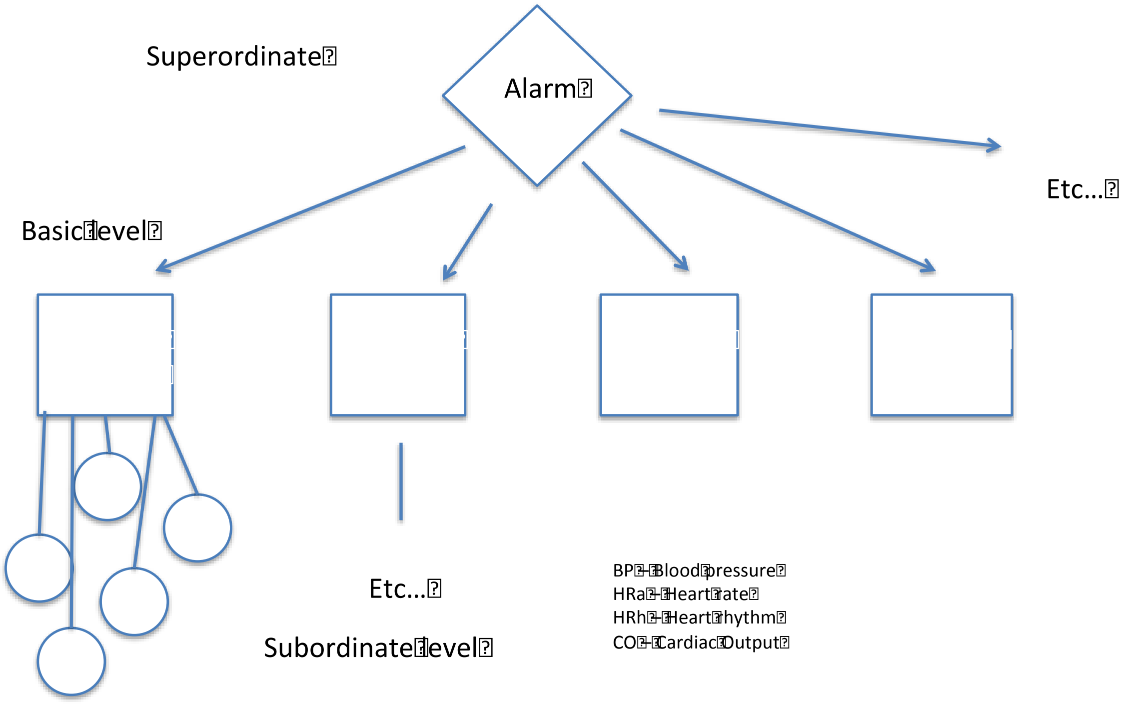

Figure 1: A framework for thinking about alarm categories

References

1. 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.
2. Wee, A. N., & Sanderson, P. M. (2008). Are melodic medical equipment alarms easily learned?. *Anesthesia & Analgesia*, *106*(2), 501-508.
3. 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.
4. Sanderson, P. M., Wee, A., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia*, *61*(2), 142-147.
5. 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.
6. Edworthy, J R, Reid, S, McDougall, S, Edworthy, J D, Hall, S, Bennett, D, Khan, J & Pye, E (under review) The learnability and localizability of auditory alarms: setting global medical device standards
7. Kerr JH. Warning devices. Br J Anaesth 1985;57:696 –708
8. 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.
9. 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.
10. K.J. Vicente, Ecological interface design: progress and challenges, *Human Factors: The Journal of the Human Factors and Ergonomics Society* 44 (2002) 62–78.
11. Miller, A., Scheinkestel, C., & Steele, C. (2009). The effects of clinical information presentation on physicians' and nurses' decision-making in ICUs *Applied ergonomics*, *40*(4), 753-761.
12. Koch, S. H., Weir, C., Westenskow, D., Gondan, M., Agutter, J., Haar, M., ... & Staggers, N. (2013). Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: A prospective randomized controlled study. *International journal of medical informatics*, *82*(8), 665-675.
13. Effken, J. A., Loeb, R. G., Kang, Y., & Lin, Z. C. (2008). Clinical information displays to improve ICU outcomes. *International journal of medical informatics*, *77*(11), 765-777.
14. Patterson, R D (1982) Guidelines for auditory warnings systems on civil aircraft, Civil Aviation Authority paper 82017: UK Civil Aviation Authority
15. 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.
16. 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

17. 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.
18. Roach, E., & Lloyd, B. B. (1978). *Cognition and categorization*. Hillsdale, New Jersey, 47.
19. Stevenson RA, Schlesinger JJ, Wallace MT. Effects of divided attention and operating room noise on perception of pulse oximeter pitch changes: a laboratory study. Anesthesiology. 2013;118(2):376-381.
20. 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 patients. *PloS one*, *9*(10), e110274.
21. Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behaviour in response to alarms of varying reliability. Ergonomics, 38(11), 2300-2312.
22. Lawless, S. T. (1994). Crying wolf: false alarms in a pediatric intensive care unit. Critical care medicine, 22(6), 981-985.