AUDITORY ALARM DESIGN FOR NASA CEV APPLICATIONS

Durand R. Begault[1]
Martine Godfroy[1]
Aniko Sandor[2]
Kritina Holden[3]

[1]Human Systems Integration Division (Code TH)
NASA Ames Research Center
Moffett Field CA 94035
[2]LZ Technology [3]Lockheed Martin Corporation
Usability Testing and Analysis Facility (Code SF3)
NASA Johnson Space Center
Houston TX 77058
Durand.R.Begault@NASA.gov

ABSTRACT

This monograph reviews current knowledge in the design of auditory caution and warning signals, and sets criteria for development of 'best practices' for designing new signals for NASA's Crew Exploration Vehicle (CEV) and other future spacecraft, as well as for extra-vehicular operations. A design approach is presented that is based upon cross-disciplinary examination of psychoacoustic research, human factors experience, aerospace practices, and acoustical engineering requirements. Existing alarms currently in use with the NASA Space Shuttle flight deck are analyzed and then alternative designs are proposed that are compliant with ISO 7731, "Danger signals for work places – Auditory Danger Signals", and that correspond to suggested methods in the literature to insure discrimination and audibility. Future development of auditory "sonification" techniques into the design of alarms will allow auditory signals to be extremely subtle, yet extremely useful for indicating trends or root causes of failures. A summary of 'best practice' engineering guidelines is given, followed by results of an experiment involving subjective classification of alarms by ten subjects.

1. AUDITORY ALARMS IN THE CONTEXT OF THE SHUTTLE CAUTION AND WARNING SYSTEM

An 'auditory alarm' for purposes of this report refers to any audio signal used for alerting or warning a user within a human-machine interface, while an 'alarm' refers generically to either audio or visual cues. The use of auditory alarms in current shuttle applications is reviewed in technical documents ("Shuttle Crew Operations Manual- SCOM-Section 2.2 Caution and Warning"- available at http://www.shuttlepresskit.com; and "Space Flight Operations Contract. Caution and Warning C&W 21002, USA 006019, October 1, 2004"). Auditory alarms are part of the collective caution and warning (c/w) system that consists primarily of visual cues (either illuminated light displays and switches, an illuminated message on an dedicated matrix panel, or a text message on a CRT).

There are four classes of alarms used on shuttle, which can be prioritized in ascending order as follows. A "class 0" alarm visually indicates up and down arrows on the CRT display next to a specific parameter, indicating that it has exceeded its predefined upper or lower boundary limits. There is no auditory component for a class 0 alarm. A "class 3" alarm is technically an "alert" and generates a steady tone of 512 Hz for approximately 1 second (this can be changed by the crew to longer durations, up to 99 seconds), along with an illuminated button and fault message on the CRT. A "class 2" alarm generates an illuminated text message on a dedicated matrix panel (panel number F7), and illuminates parameter lights on another panel (number R13U). The alarm consists of an alternating tone between 375 and 1000 Hz. It is silenced ("killed") by pressing a master alarm switch.

There are two types of class 1 "emergency" alarms that are highest priority: (1) smoke detection and (2) rapid cabin depressurization. The smoke detection alarm consists of a "siren" sound, i.e., a tone varied from 666 to 1,460 Hz and then back to 666 Hz over a 5 second interval. Smoke detection lights are indicated on a dedicated panel (number L1). The cabin depressurization alarm is indicated via a "klaxon" sound, consisting of two tones at 270 and 2500 Hz that are periodically iterated. Pressing the master alarm switch also silences these alarms. Under the current design, it is possible for all of the auditory alarms to sound simultaneously.

These auditory alarms have three primary functions. First they indicate that a specific condition exists that did not occur previously in time, and that now requires attention. This may include the corollary function of waking a sleeping crewmember. Second, they have a rudimentary function of stating: "*look over here at this specific visual display*". This is a form of "directed attentional shift" that is significant in the larger context of the cognitive challenge of fault management [1]. Third, their function is to relate the relative urgency of the alarm through the semantic content contained in the alarm type. The type of alarm indicates: "*where in the hierarchy of possible auditory alerts does this new alarm lie?*" and "*how quickly do I need to attend to this problem?*"

The class 1-3 auditory alarms used in Shuttle are useful reference points from which to discuss best practices in the development of future alarms for CEV (Crew Exploration Vehicle) and EVA (Extra-Vehicular Activity). They are illustrative of a coherent, useful approach to alarms that nevertheless can be improved upon, given subsequent human factors research and the possibility of implementing superior alarm generation hardware.

The discussion that follows indicates criteria for a design approach based upon a cross-disciplinary examination of psychoacoustic research, human factors experience, aerospace practices, and acoustical engineering requirements. Not considered here is the use of 'sonification' — the use of sound to continually monitor the status of a system, as opposed to only using audio to signal that a limit has been exceeded (e.g., [2]). Some research has been concerned with manipulating a single type of sound to convey differing levels of urgency. For example, a tone might get faster in its repetition cycle akin to a Geiger counter measuring an increasing level of radioactivity. Sonification has been shown in surgical applications to allow anesthesiologists to maintain high situational awareness while performing other tasks more effectively, compared to visual-only displays [3].

2. VISUAL VERSUS AUDITORY ALARMS

When is it appropriate to use an auditory, as opposed to a visual, alarm? Perhaps the most obvious function of the auditory alarm is to alert a person to inspect a visual display. Less obvious is the relationship between the alarm's informational content and the preferred modality for communicating to a user.

Table I summarizes some important differences between visual and auditory alarms from the standpoint of human factors and multimodal perception capabilities [4]. Auditory alarms are pervasive and independent of where the listener is in the environment, as long as the level of the alarm is audible. Visual alarms require the user to be looking at the specific alert (hence the use of auditory alarms to guide attention to the visual alarm message). Auditory alarms are far faster for conveying a specific message than a visual alarm, particularly one connected with a text-based display. It is also possible to immediately convey an urgent versus a non-urgent meaning regarding the alarm. On the other hand, there is a trade-off that the semantic content of the alarm cannot be overly complex; the order of messages is far easier to retain from visual information, compared to auditory information. Finally, the magnitude of the noise environment (distractors) in a specific perceptual modality can be considered more or less irrelevant to the message in a different modality.

Table I. *Visual versus auditory alarms: perceptual factors*

	Auditory	*Visual*
Reception	No directional search	Requires attention, selection
Speed	Fastest	Slowest
Message order, complexity	Low retention	High retention
Urgency	Easy to convey	Difficult to convey
Noise	Independent of visual noise	Independent of auditory noise

From these perceptual performance differences, Table II can be derived in order to determine guidelines for when to use either auditory or visual alarms.

Table II. *Guidelines for using auditory versus visual alarms.*

Auditory alarm preferred for:	*Visual alarm preferred for:*
Simple message	Complex message
Short message	Long message
Not referred to later	Referred to later
Requires immediate action	Does not require immediate action
When visual system overloaded	When auditory system overloaded
Moving persons	Stationary persons

3. AUDIBILITY: SOUND PRESSURE LEVEL AND FREQUENCY CONTENT

Sound Pressure Level. An auditory alarm must be audible with a very high degree of reliability. This typically requires the sound pressure level of the alarm to 'penetrate' background noise. At the same time, alarms need to be conducive to effective fault management, and not merely audible. Hence, the challenge for designing a good alarm in terms of its level can be expressed simply as "not too loud, not too soft- but just right"! This becomes a challenging matter when listeners are at varying distances from loudspeakers, or when wearing hearing protection devices without headset delivery of sound.

In terms of Signal Detection Theory (SDT), the presence of an auditory alarm in the expected noise environment should have a 100% "hit rate" in terms of audibility. This is usually a matter of calculating a signal-noise ratio based on prior research into auditory signal detection. Most auditory alarm engineering guidelines "err" towards making the level higher than might be predicted by auditory masking experiments [5]. However, if the alarm is too loud or too pervasive, negative effects on human performance can occur; from the perspective of effective fault management, a startle effect requires time for recovery [6]. Many alarms have a startling, excessively high level that is counter-productive from the perspective of human factors research; simultaneous alarms can exacerbate the problem.

Overall levels of 15–20 dB(A) are commonly cited target signal-to-noise ratios for alarms, but this disregards the frequency content (spectra) of the alarm or the noise. International Standard ISO 7731, "Danger signals for work places-auditory danger signals" examines the role of the spectral components of alarms with regards to masking in concurrent and adjacent spectral bands. Figure 1 is an example of a one-third-octave band analysis from ISO 7731 that calculates the signal-noise ratio in greater spectral detail. The requirement states that the signal must be >= 13 dB relative to the *masked threshold* in one or more octave bands [7]. The masked threshold is modeled by allowing a contribution towards the masking level from adjacent bands of noise, as well as the frequency band that is concurrent with the alarm.

Note that an alarm can certainly be audible but not "heard." It is well known that certain repetitive or irritating sound sources, including alarms, can be ignored or dismissed from memory, independent of level, through a process of habituation. Conversely, certain other sounds, such as a child's voice heard by their parent, can be very effective at harnessing attention when in competition with other sound sources. The ability to hear through the auditory 'scene' of multiple sound sources is a process known as 'auditory streaming' [8].

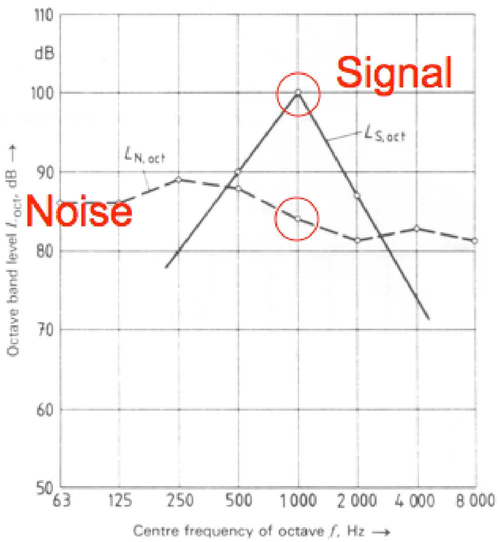

Figure 1. *An alarm signal with a spectral component 15 dB higher than the noise in the 1 kHz one-third-octave band.*

While it might be considered naïve at first from a psychoacoustic perspective to design a very loud, startling auditory alarm, there are certain useful applications. For example, building evacuation alarms are designed to be annoying to compel persons to leave the vicinity as quickly as possible. In this case, the message is to "leave immediately", and the listener does so because the alarm is loud and irritating. These types of auditory alarms are of course counter-productive to environments such as flight decks, but the design of many older alarms (e.g., as used on the Boeing 707) inherited the 'hue and cry' design of older paradigms such as the familiar siren and fire bell (see Figure 2).

Figure 2. *The sound of mechanical alarms (here, a fire bell and siren from the 1930s) continues to represent familiar typologies in the formation of electronic sirens and bells.*

Frequency Content. The frequency content of an auditory alarm is as important as its level to ensure audibility. Several standards specify that auditory alarms contain frequency components in the region of relative maximal hearing sensitivity, 0.2–4 kHz, which is the primary region of acoustical energy for speech sounds.

ISO 7731 specifies that alarm signals contain frequency components between 0.3–3 kHz, and with sufficient energy between 0.3–1.5 kHz to accommodate high frequency hearing loss or those wearing hearing protection devices (HPDs). Patterson [6] recommends having four or more spectral components that are harmonically related, to allow "fusion" of spectral components. He also specifies that the fundamental frequency (the first spectral component) be between 0.15–1 kHz. Military standard 1472C "Human Engineering Design Criteria For Military System, Equipment, and Facilities" (1981) specifies frequencies between 0.2–5 kHz but with an upper limit of 1 kHz for distances greater than 300 m to account for acoustical 'shadow zones'. This is because shorter wavelengths that correspond to higher frequencies can be blocked by solid objects, while lower frequencies can 'bend' around them. Other standards are similar in terms of specified frequency content; for example, the Society of Automotive Engineers Standard SAE J994b, "Performance, Test, and Application Criteria for Electrically Operated Backup Alarm Devices" (1974) indicates spectral components between 0.7–2.8 kHz.

These frequency specifications are simplistic and most applicable to synthetic tones. The time-varying spectral content of many candidate alarms sounds, as well as the role of brief noise or broad-band spectral components are not typically addressed by these standards. For example, the timbre (tone color) of the fire bell shown in Figure 2 consists of a fast attack transient rich in harmonic and inharmonic partials caused by the clapper making contact with the bell. This is followed by the decay of the bell resonance, which is comparably harmonic in nature. The spectral complexity of the sound is important, but would be unrealizable with, e.g., only four spectral components.

4. DISCRIMINABILITY: TYPOLOGY AND TEMPORAL PATTERN

What should an alarm 'sound like' for a given context? People identify alarms based on their frequency content and temporal pattern. This infers that it be easily discriminated from background noise or other types of signals. Section 4 discussed frequency requirements in general, but for a specific context, it may be important to determine to what degree non-alarm auditory signals overlap with a potential alarm. For instance, if there is a constant harmonic tone from a fan that has significant energy at 1 kHz, the use of a constant tone may not be an ideal type of alarm, even if its frequency components conform to ISO 7731 by being >= 13 dB over that tone. Equally as important, in an auditory display using multiple alarms, the alarms themselves should be easily discriminable.

The *typology* of an auditory alarm refers to the temporal aspects and frequency content of the sound 'object' that gives it a specific semantic content. The typology of an auditory alert can be identified by using ecological reference for determining an auditory alarm's meaning and level of urgency. For example, the sound of a siren is a learned cultural reference that differs from region to region, but is easily identified. Research has indicated that the ease of recognizing an auditory alarm is driven partly by learned associations as to what the sounds represent [4, 9]. The design of an auditory alarm system can take advantage of 'known' alarm typologies. It is also useful in some cases to associate the typology with a specific action. For instance, the rattling sound and vibration of a stick shaker in an aircraft during a stall alert is caused by the same object that must be attended to.

One researcher classifies alarms in four categories, "Siren – klaxon- horn- electronic". Klaxon Signals LTD., a company specializing in alarms, indicates categories of "electronic sounders, sirens, buzzers, hooters, fire alarms and beacons" (http://www.klaxonsignals.com). The sound of alarms can be considered a subset of what Gaver [9] has termed 'auditory icons', which can range from abstract to literal representations of sounds (e.g., a fire alarm being represented by the sound of something burning).

The most important aspect of alarm typology is that it allow for ease of discrimination amongst a set of alarms that would be used in a human interface. The alarms must have an inherent means of conveying level of urgency and be easy to learn. Hence, the use of alarms that are already familiar to a user makes categorization and learning easy. For example, astronauts familiar with the caution and warning signals described in section 1 will make an easy transition to similar alarms that maintain the same typology of 'siren, klaxon, electronic tone'. On the other hand, the ability to learn and remember a set of abstract alarms is severely limited; Patterson [6] set a limit of four alarms for easy acquisition, while learning with up to three additional alarms is far more difficult.

In aviation flight decks, the 'attention getting' component of some alarms is followed by and distinguished through the use of an "added" synthesized speech message. While an extended discussion on the use of speech messages is beyond the scope of the current monograph, it should be noted that speech messages take longer to comprehend than an auditory alarm, and are more easily masked by background noise. Although speech can convey complex ideas that cannot be conveyed by a non-speech auditory alarm, the chances for misidentification is far greater compared to non-speech alarms.

Electronic (synthesized) tones have a far richer potential for differentiation than they once did because of the ease and economy of using *sound sampling* techniques. This involves the use of PROM (programmable read-only memory) technology to store virtually any type of sound, and to allow proper software "hooks" to post-signal processing algorithms for changing the sound dynamically. Prior to around 1990, it was far less expensive to use digital or analog oscillator chips, which were efficient but severely limited in terms of timbral differentiation. This difference is audible in everyday computers or video games of the current time, compared to the personal computer 'beep' tones familiar from the 1980s. Hence, it is potentially far easier to create candidate alarms having far more 'discriminable' acoustic features.

The use of distinct temporal patterns has been proposed as a means of conveying urgency and for aiding discrimination between multiple alarms [6]. This is because temporal pattern- the sequence of 'on' and 'off' iterations of the alarm sound- is easily heard and discriminated by a listener.

American National Standard ANSI S3.41 "Audible Emergency Evacuation Signal" recommends a specific temporal pattern of three on pulses, each with a one second period, followed by 1.5 s of silence. International Standard ISO 9703-2:1994 "Anesthesia and respiratory care alarm signals" indicates two specific patterns, as shown in Figure 3. An alarm 'burst' is formed by multiple pulses with a silent interval in-between of 0.15- 0.5 s, depending on the ranking of the alarm. Between each burst is a silent period, here termed the 'inter-burst interval.' Such silent intervals allow time to think, verbally communicate, and take action in a constructive manner, compared to the counter-productive use of a constant alarm. It remains a question for each specific application how long the inter-burst interval could be increased such that awareness of whether or not the alarm was still active would be relevant.

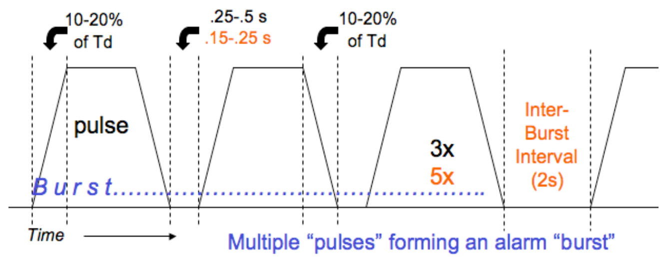

Figure 3. *Pulses forming an alarm 'burst' in ISO 9703-2:1994, "Anesthesia and respiratory care alarm signals".*

A final note regarding discrimination is based on an informal report that class 2 and class 3 alarms as heard on the International Space Station could be confused. The class 2 auditory alarm consists of an alternating low and high tone. When the class 2 tone is heard at a sufficient distance, e.g. within a different module, the upper tone is masked due to shadowing effects, thereby causing only the low tone to be heard. This can be confused with the sound of the class 3 alarms, since they are not differentiated by temporal interval and the 375 Hz tone is not perceivably different in isolation from the 512 Hz tone. The use of distinct temporal patterns may help mitigate this problem.

5. MINIMIZING STARTLE EFFECT

Humans possess a startle reflex that is involuntarily activated by objects abruptly entering the visual space or by loud noises. Physiological responses include anxiety, arousal, and tightening of muscles. At the most basic level this is likely a hard-wired evolutionary adaptive mechanism which helps protect us from potential dangers in our environment. From a human factors standpoint, the effect of startle is counter-productive to effective fault management.

There are two factors responsible for startle: 1) overall level; and 2) temporal transition of the amplitude envelope from zero state to a maximum. The overall level can be mitigated as described in ISO 7731. An additional concept is the use of a 'precursor' alert. The level of the alert is played -6 to -10 dB lower on its initial presentation, compared to successive presentations. This mimics the effect of hearing a gradually approaching emergency vehicle; sirens are far more startling when standing near a vehicle that initiates the alarm, versus when the vehicle is heard approaching from a distance.

The temporal transition of the amplitude envelope can be made more gradual than an instantaneous onset via design of the alarm pulse. Figure 3 indicates the use in ISO 9703-2:1994 of an envelope rise and fall time equivalent to 10-20% of the overall duration of the pulse. The effect is to cause a 'fade-in' of the tone that helps to mitigate startle effect. It should be noted that several types of alarms, including the shuttle depressurization class 1 klaxon and a fire bell, cannot be faded in using this method without significantly altering the timbre (and therefore the recognition) of the signal.

Finally, ISO 7731 is the only standard that indicates that a human-in-the-loop test be conducted to insure that alarms are discriminable. A formal study using multi-dimensional scaling techniques would be the best means to determine the underlying perceptual scaling for differences between alarms.

6. CONSIDERATION OF LEVELS AT THE EAR, HPDS

To maintain levels within NASA standards, sound levels cannot exceed 85 dB(A) during orbit, and 105 dB(A) during launch or re-entry. To maintain levels above a changing background noise level, it is possible to integrate background noise monitoring systems that continually monitor the level and then adjust levels to the target signal-noise ratio. When wearing HPDs and helmets, alarms should be delivered via headsets, though not necessarily at as high a signal-noise ratio as from loudspeakers, since the position of the loudspeaker to the ear is predictable. A rough estimate would be 50% of the loudspeaker level (about 6 dB above the background noise level). Under circumstances where this level would exceed 105 dB(A), the use of non-auditory means of alerts (e.g., tactile-haptic actuators or only visual alerts) may be recommended to conserve hearing.

7. ADVANCED TECHNIQUES

To help facilitate fault management, it is possible to directionalize alarms to the source of a problem or to a visual panel that requires attention. For example, locating individual loudspeakers near each of the smoke detection sensors on a flight deck might allow faster determination of the source of a problem. It is also possible when wearing headsets to use a 3-D audio acoustic display to directionalize audio to a virtual source position [10]. Another advanced technique for increasing the detectability of an alarm is to use a technique termed 'spatial modulation', where the tone is moved laterally at a rate of 2-10 Hz. One study [11] has shown that the spatial movement allows the sound to be about 7 dB more detectible against a stationary background noise, compared to a stationary alarm.

8. SUMMARY: BEST PRACTICES FOR AUDITORY ALARM DESIGN

The preceding discussion can be summarized in terms of the following guidelines for best practices in forming auditory alarm design.

• Auditory alarms are preferred to visual alarms for short, simple messages requiring immediate action

• Auditory alarms should be designed to enhance, not hinder, effective fault management
by controlling level and factors causing a 'startle effect'

• ISO 7731 provides the best guidance for determining signal-noise ratio because it evaluates individual bands of frequencies with respect to the masked threshold

• Frequency content of the auditory alarm should correspond to maximal human sensitivity, i.e., between 200 Hz- 4 kHz.

• The fundamental frequency of the auditory alarm should be between 300 Hz – 1 kHz.

• There should be four or more harmonically-related spectral components to the auditory alarm

• Auditory alarms are best discriminated by taking into account listener association with specific typologies (e.g., sirens; bells; buzzer; etc.) rather than varying a single typology (e.g., different types of bells).

• Auditory alarms are best discriminated when a unique temporal pattern is associated with each one.

• The number of auditory alarms should be limited ideally to four, no more than seven.

• Speech alerts present challenges to design, for instance for the time to understand the message and for ensuring an effective signal-noise ratio.

9. RESULTS OF A STUDY EXAMINING SUBJECTIVE CLASSIFICATION OF ALARMS

9.1. Introduction

A pilot study was conducted to examine the relationship between predicted responses for alarm typologies and subjective responses from non-professional subjects. The results can indicate the degree of predictability based on level of urgency via 'post analyses of the sounds used. The study was also used to narrow down the number of stimuli to be evaluated in later studies involving crew or other domain specialists.

9.2. Method

Ten volunteers participated in the study (five male, five female). Some of them were already familiar with the definition of the alarms and two had previously heard the caution and warning sounds currently used on the International Space Station.

A set of 49 sounds was gathered for use as stimuli. The source of the sounds were from numerous sources, including (1) existing ISS caution and warning tones (2) synthesized variants of the tones (3) "novel" synthesized tones and (4) recordings of tones used in military and naval applications. They were categorized by one of the co-authors according to their coherence with a specific level of urgency, based on the criteria discussed previously in this document.

Each participant was seated at a computer and the stimuli were presented sequentially from a loudspeaker at approximately the same level. A 3-alternative forced choice task was given to the subjects, where they had to categorize the sounds into one of three predefined groups according to its ‘best match’s. Prior to the experiment, the predefined groups were described as follows:

> Emergency (class 1): This is the most serious type of event. It is used in a life threatening condition that requires immediate action in order to protect the crew.
>
> Warning (class 2): This is less serious than emergency. It is used in a situation that requires immediate correction to avoid loss or a major impact to mission or potential loss of crew.
>
> Caution (class 3): This is a situation of a less time critical nature, but with a potential for further degradation if crew attention is not given.

After participants finished the categorizations, they were asked to sort the sounds they chose for class 1 into three sub-groups, for fire/smoke, rapid pressure change, and toxic atmosphere alarms. The study lasted on average 20 minutes.

9.3. Results

Figure 4 shows raw data for the categorization of each of the alarms by overall percentage of subjects, ordered by the categorization of the class 3 alarm. A clustering of results is clearly evident from the opposite trend lines of the class 1 alarm choices (red line) and the class 3 alarm choices (blue line). Furthermore, class 2 alarms appear to cluster towards the area where alarms are less obviously class 1 or class 3.

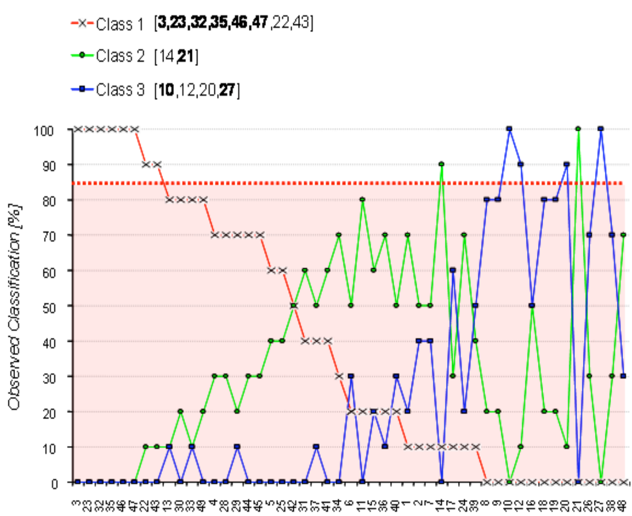

Figure 4. *Raw data showing percentage of subjects classifying each alarm as class 1, 2 or 3. The dashed line indicates an arbitrary criterion at 85% for choosing alarms based on consensus. The bold numerals indicate 100% of agreement between subjects.*

It is possible to set a criterion for calling an alarm ‘strongly identified’ for a particular class when agreement surpasses a threshold. Here, we have arbitrarily adopted an 85% criterion and above (indicated by the dashed line), where essentially 8 out of 10 subjects tested agreed as to the specific identification of the alarm. Based on this criterion, 12 alarms (3, 23, 32, 35, 46, 47 at 100%; 22 and 43 at 90 %) out of the 49 tested were ‘strongly identified’ as class 1 caution tones. Two alarms (14 and 21) were strongly identified as class 2 warnings. Four alarms out of the 49 were strongly identified as class 3 emergency alerts.

The primary physical attribute of these ‘strongly identified’ alarms for each category is rather clear between class 1 and class 3 types. The class 3 alarms had relatively longer inter-burst intervals and did not have pitch modulation within the alarm pulse. Furthermore, two of the alarms were characteristic of commercial airline ‘flight attendant’ chimes having two tones, akin to a doorbell ‘ding dong’. These chimes are most likely associated with previous experience and association with less urgent contexts. Contrasting this, the class 1 alarms were mostly of the ‘siren’ category, involving pitch modulation, and had far briefer inter-burst intervals or were constant. This is in line with the previous discussion regarding perceived urgency and with identification based on a recognizable typology.[1]

The class 2 alarms were more ambiguous, having some characteristic of either a class 1 or 3 alarms and without a recognizable typology. One had a rapid frequency modulation but a relatively long inter-burst interval. The other was a relatively slow frequency modulation over a very wide rage.

Table III indicates the percentage breakdown of class 1 alarms that exceeded the 85% criteria as a function of category - fire/smoke, pressure or toxic alarm. There is no strong consensus as to specific identification of the alarm by category.

Table III. *Categorization of class 1 alarm (one subject rated sound 22 as a class 2 alarm; total across columns therefore equals 90%).*

Sound	Pressure	FireSmoke	Toxic
3	30	50	20
22	30	30	30
23	60	20	20
32	50	50	0
35	60	20	20
43	20	40	30
46	30	50	20
47	30	50	20

Figure 5 indicates the results of the principal component (PCA) data clustering analysis. These analyses are based on input from all data as opposed to data based on the 85% consensus level, and establish subjective distances of the stimuli from one another in a graphic manner.

The correlations between variables can be explained in terms of underlying factors or latent variables. The correlation matrix shows the correlation (or relationship) of each variable with all the other variables. The factorial analysis allows regrouping the variables that correlate strongly with one another while dissociating with all the other variables weakly correlated. The eigenvalues represent the amount of variance that is accounted for by each factor. Each variable is attributed a weight in relation with each factor, showing how much a variable is correlated (or loads on) across different factors.

[1] One subject commented, “Continuous sounds seem to be more alerting than discrete sounds, so I put more continuous sounds in the emergency category and more discrete sounds in the warning or caution category.”

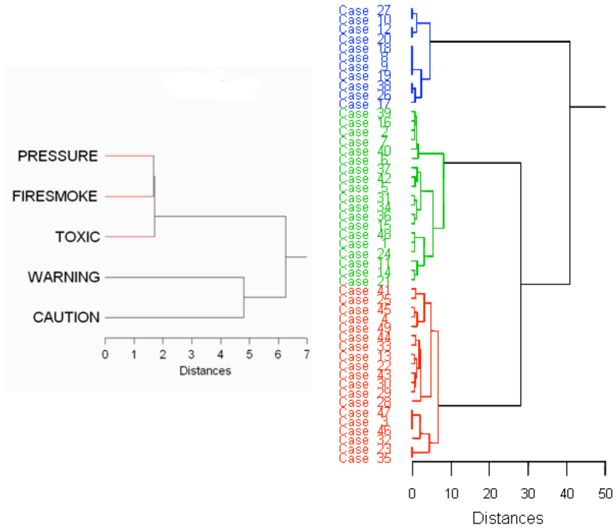

Figure 5. *Cluster tree analysis. The left dendogram shows the clustering of the alarms across the five categories. The right dendogram represents the relationships between sounds at a higher hierarchical level between classes (red= class 1, green=class 2 and blue=class 3)*

Clustering or classifying refers to the grouping of objects into sets based on their similarities and on differentiation between sets because of their differences. Similarity can be understood as some index of distance (between raw scores) for two or more stets of objects. Objects showing the highest similarity index are connected to each other by a line, and then form a new object. The shorter the lines, the more similar the objects are. The process continues until all the objects are joined.

As it can be seen from Figure 5, right, class 1 and class 2, as well as class 2 and class 3 show significant inverse correlations (class 1, class 2: r=-.493, p=.0002; class 2, class 3: r=-.724, p<.0001), while class 1 and 3 do not share any communality. At the level of the different categories within class 1, in Figure 5 left, we observe the highest correlation coefficient between the pressure and fire/smoke categories (r=.51, p=.0001), while all the other correlations remain significant.

These relationships are confirmed by the factorial analysis, showing that the variance is explained by two factors, the first factor representing 59% of the total variance explained by the model (see figure 6, left). Figure 6, right shows that the variance in class 1 is almost totally explained by Factor 1 (component loadings, class1: -.99) while the variance in class 3 is essentially explained by Factor 2 (-.92). Class 2 is somehow intermediate since its variance is explained by both factors (component loadings factor 1=.8, factor 2=.59).

Figure 7 illustrates the relationship between theoretical categorization of sounds and observed responses from all the subjects. The theoretical categorization was made by one of the co-authors (db) based on a subjective estimate of the physical characteristics of the sound with reference to the principles outlined earlier in this paper for conveying urgency (based on level of frequency modulation, pulse rate or inter-burst interval).

The left side plot shows the categorization in terms of the orange-colored intervals. The right side plot shows the observed deviation from these categories. The correlation between prediction and observation is significant for class 1 (r=.64, p<.0001) and class 3 (r=.45, p=.001) alarms, but not for class 2 alarms, as seen in Figures 7.

Overall, the results of this pilot study suggest that the best practice techniques for the design of new alarms outlined in this paper are salient to untrained listeners and can provide predictable results, for at least two extreme levels of alarms and possibly for an intermediate level of alarms.

10. ACKNOWLEDGEMENTS

This work was supported by a directed research project plan of NASA's Space Human Factors Engineering Project.

11. REFERENCES

[1] D. D. Woods "The alarm problem and directed attention in dynamic fault management" *Ergonomics,* vol. 38, 1995, pp. 2371-2393.

[2] S. Bly, "Presenting information in sound", in *Proc. CHI '82 Conf. on Human Factors in Computing Systems,* 1982, Association for Computing Machinery.

[3] M. Watson, and P. Sanderson. Sonification supports eyes-free respiratory monitoring nd task time-sharing. *Human Factors* 46:3, 497-517.

[4] N. A. Stanton, N. A. and Edworthy, J. (1999). "Auditory Warnings and Displays: An Overview". In *Human Factors in Auditory Warning Systems.* Aldershot, UK: Ashgate.

[5] G. S. Robinson, and J. G. Casali, "Speech Communications and Signal Detection in Noise", in *The Noise Manual* 5th edition (Burger, E. H., Royster, L. H., Driscoll, D. P. and Lyane, M., ed.s). Indianapolis, IN, American Industrial Hygiene Association, 2003.

[6] R. Patterson, *Guidelines for auditory warning systems on civil aircraft* (report no. 82017). London, Civil Aviation Authority, 1982.

[7] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*. Springer-Verlag, Heidelberg, Germany, 1990.

[8] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA, MIT Press, 1990

[9] W. Gaver, "Auditory icons: using sound in computer interfaces," *Human-Computer Interaction,* vol. 2, 1986, pp. 167-177.

[10] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia.* Cambridge, MA, Academic Press, 1994.

[11] D. R. Begault, M. R. Anderson, and B. M. McClain, "Spatially-modulated auditory alerts for aviation," *J. Audio Eng. Soc.,* vol. 55, no. 4, 2006, pp. 276-282.

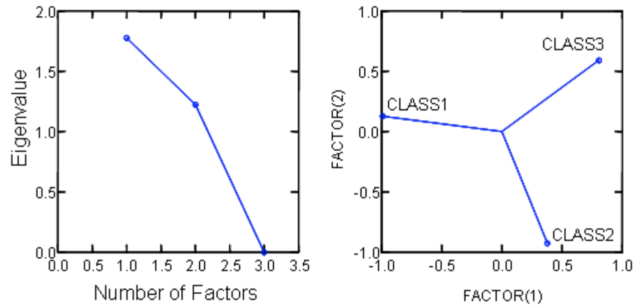

Figure 6. *Left: scree plot and Right: factor loading plot for the three classes of alarms. Two factors explain 59% of the total variance.*

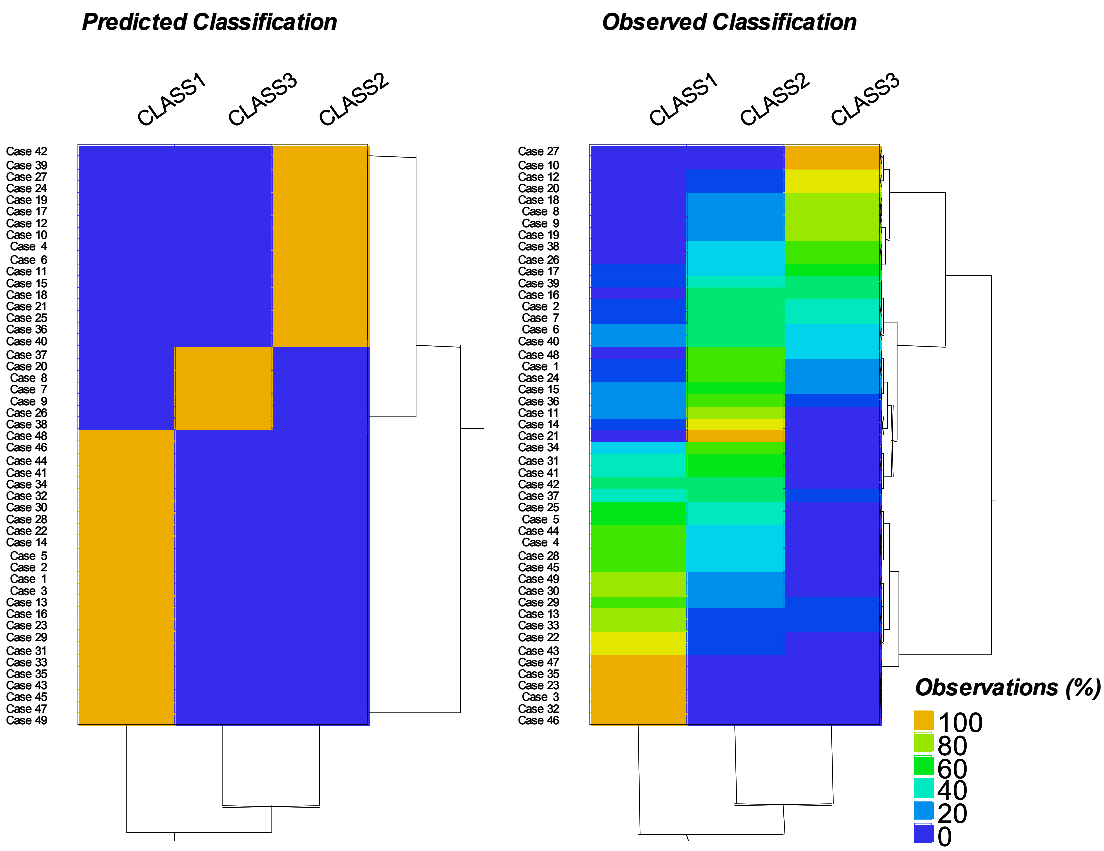

Figure 7. *Cluster trees (Distance metric is Euclidean distance, Ward minimum variance method) for the predicted classification (left) and the observed classification (right). The color label indicates frequency of the classification responses as a function of the Alarm (1 to 49). Note the permutation of the class 2 and class 3 between the predicted and the observed classification.*