UNIVERSITY OF LEEDS

This is a repository copy of *Semantic congruency of auditory warnings*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/105180/

Version: Accepted Version

Article:

Isherwood, SJ orcid.org/0000-0002-8022-3110 and McKeown, JD (2017) Semantic congruency of auditory warnings. Ergonomics, 60 (7). pp. 1014-1023. ISSN 0014-0139

https://doi.org/10.1080/00140139.2016.1237677

© 2016 Informa UK Limited, trading as Taylor & Francis group. This is an Accepted Manuscript of an article published by Taylor & Francis in Ergonomics on 21st September 2016, available online: http://www.tandfonline.com/10.1080/00140139.2016.1237677.

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Semantic congruency of auditory warnings

Sarah J. Isherwood[a], Denis McKeown[b]

[a]School of Medicine, University of Leeds, Leeds, UK

[b] School of Psychology, University of Leeds, Leeds, UK

Corresponding author: Sarah J. Isherwood, School of Medicine, University of Leeds, Leeds, LS2 9JT. E-mail: s.j.isherwood@leeds.ac.uk

Semantic congruency of auditory warnings

The aim of this study was to explore operator experience and performance for semantically congruent and incongruent auditory icons and abstract alarm sounds. It was expected that performance advantages for congruent sounds would be present initially but would reduce over time for both alarm types. Twenty-four participants (12M/12F) were placed into auditory icon or abstract alarm groupings. For each group both congruent and incongruent alarms were used to represent different driving task scenarios. Once sounded, participants were required to respond to each alarm by selecting a corresponding driving scenario. User performance for all sound types improved over time, however even with experience a decrement in speed of response remained for the incongruent iconic sounds and in accuracy of performance for the abstract warning sounds when compared to the congruent auditory icons. Semantic congruency was found to be of more importance for auditory icons than for abstract sounds.

Keywords: auditory warnings, auditory icons, human-computer interaction, user experience, auditory displays

Practitioner summary. Alarms are used in many operating systems as emergency, alerting, or continuous monitoring signals for instance. This study found that the type and representativeness of an auditory warning will influence operator performance over time. Semantically congruent iconic sounds produced performance advantages over both incongruent iconic sounds and abstract warnings.

Introduction

Auditory display is the term used to encapsulate a range of intentionally produced auditory signals and possibly also the context in which they are used and the manner with which they are transmitted, such as through headphones or loudspeakers (Hermann, 2008; Walker and Kramer, 2004). Such sound signals can include alarms, auditory warnings, alerts, auditory icons, earcons, auditory graphs, spearcons and data audification for instance (Nees and Walker, 2009). Auditory warnings form a subcategory of auditory display sounds. The purpose of an auditory warning is to attract attention in order to provide support or information for a listener facing a potentially dangerous situation (Edworthy and Hellier, 2006). Auditory warnings tend to convey binary information about the occurrence of an event, signalling only when a particular threshold value is reached for the event, and they will frequently require a prompt response from a listener (Edworthy and Hellier, 2006; Nees and Walker, 2009).

Both speech and non-speech sounds can be used as auditory warning signals. Speech warnings offer the advantage of being able to convey complex messages without prior user experience of the warning. Non-speech sounds however may be chosen over speech sounds where privacy is a concern, where competing speech (natural or artificial) may mask the speech warning, or where a warning that is not language specific is required. Many traditional non-speech auditory warnings have made use of abstract sounds, consisting of simple tones, pulses, bells, sirens or buzzes for instance (Graham, 1999, Patterson 1982, 1989). Auditory icons can also be utilised as warning sounds these are natural, or environmental, sounds that have been employed to act as signifiers for human-machine interfaces (see Belz, Robinson and Casali, 1999; Gaver 1986). Whereas the acoustic features of a signal are crucial in the design of abstract alarms, auditory icons rely on associations with their real-world referents in order to enable a user to infer the meaning of the alarm signal, and are often classified according to how closely they ‘map’ to their referents.

Abstract warning signals

The acoustic features of abstract auditory warnings may be manipulated to convey urgency and localisation information for instance but for many such alarms the associations between the warnings and their real world referents have to be learnt. To accomplish this, a listener has to extract the features of a sound and associate those features in memory with a particular referent (Perry et al, 2007). Patterson (1982) proposed that people could learn up to four to six auditory signals relatively quickly but that much more effort was required to learn additional signals. The simultaneous presentation of abstract alarms may also create difficulties in distinguishing and identifying individual alarms.

There are certain instances where conventional acoustic alarms have become very successful in being clearly associated with specific events. Such alarms include for example, fire and smoke alarms, emergency-vehicle sirens and telephone bells. This leads to the question of whether the success of these alarms occurs entirely through learning the meaning of the alarm, and its association to a real-world counterpart, or through the acoustic characteristics of the sound. It is possible that the level of a listener's familiarity with the signal being presented may have a critical role in determining the signal's compatibility with an event.

Environmental sounds and auditory icons

Although the term environmental sound encompasses a broad category of sounds with enormous variation in spectral-temporal variation (Gygi, Kidd and Watson, 2004) research has shown that listeners are able to very accurately identify numerous commonly heard sounds (Marcell et al, 2000; Shafiro, 2008). This ability to recognise environmental sound remains robust even with degraded acoustical properties (for instance, see Gygi et al, 2004). Gygi et al (2004) propose that identification of environmental sounds is facilitated by naturally occurring limitations in the possible spectral-temporal structures and the probabilities of sequences of such sounds, and by their complex spectral-temporal composition. This results in a number of acoustic cues being available for both bottom-up and top-down processing. Over time, listeners will have developed meaningful associations with various acoustic sequences and the geneses of these sequences; resulting in a large corpus of sounds becoming well learnt and familiar. Leung et al (1997) posit that the large memory capacity for environmental sounds may partly result from the fact that they are easily assigned verbal labels.

The act of signification, developed for the purpose of human-machine communication, differentiates auditory icons from environmental sounds. It is the listener's familiarity with a sound in the real-world that is exploited in order to facilitate their understanding of the intended meaning of the auditory icon. Gaver (1986) is credited with developing the concept of auditory icons. Guided by Gibson's (1979) ecological approach to perception, Gaver (1986, 1993) theorised that reliable causal relationships between sounds and sound-generating events enabled listeners to focus on the dimensions and properties of the distal sound-events rather than singularly focussing on the acoustic features of proximal stimuli. This notion of event oriented listening has been explored in a number of studies (for instance see Ballas and Howard, 1987; Carello, Anderson and Kunkler-Peck, 1998; Grassi, 2005).

The inherent meaningfulness of well-designed auditory icons, in comparison to abstract warnings, could be particularly useful for alarms attempting to convey meaning in infrequently occurring emergency situations and where learning or training opportunities with the warnings sounds are limited. The similarity between auditory icons and environmental sounds should facilitate the recognisability of numerous auditory icons, although as noted by Brazil, Fernström and Bowers (2009) a performance decrement can occur if more than three to six auditory icons are presented simultaneously. Research investigating how well people respond to different types of warning signal has often revealed that auditory icons result in faster and more accurate responses than abstract sounds in both laboratory tasks (McKeown and Isherwood, 2007; Perry et al, 2007; Stephan et al, 2006) and in driving simulator tasks (Belz, Robinson and Casali, 1999; Graham, 1999; McKeown, Isherwood and Conway, 2010).

The question arises as to whether the superior responses elicited by auditory icons, in comparison to more conventional abstract alarms, occur due to the potentially meaningful signal-referent mappings afforded by these types of alarms or perhaps as a result of some other feature of the alarm such as its novelty, memorability or distinctiveness. Studies exploring this issue have generally shown that well-mapped auditory icons elicit performance advantages when compared to more arbitrarily, or distantly, mapped auditory icons (Keller and Stevens, 2004; McKeown and Isherwood, 2007; McKeown, Isherwood and Conway, 2010; Stephan et al, 2006). Thus providing support for the importance of close semantic links between the alarm and its referent. Such arguments have parallels in visual icon research.

Comparison with visual icon studies

Research into visual icons has revealed how key icon characteristics are important in determining user performance with visual displays. These characteristics include concreteness, visual complexity, semantic distance and a user's familiarity with an icon or function for instance (Byrne, 1993; Forsythe, Mulhern, and Sawey, 2008; Goonetilleke et al, 2001; Green and Barnard, 1990; Isherwood, McDougall and Curry, 2007; McDougall and Isherwood, 2009; McDougall, de Bruijn and Curry, 2000; Rogers and Oborne, 1987; Scott, 1993; Stotts, 1998). Early research often focussed on icon pictorialness, or concreteness, as being an icon's most important property however later studies have shown that other icon characteristics transcend the concreteness of an icon in terms of predicting user performance. These characteristics include: semantic distance, which is an index of the closeness-of-fit of icon function relationships irrespective of whether icons are abstract or concrete, and icon familiarity (McDougall, Curry, and de Bruijn, 2001; Isherwood, McDougall and Curry, 2007; McDougall and Isherwood, 2009). However such work has also shown that the importance of such predictors is not necessarily straightforward as these characteristics tend to be interconnected and their influence on performance may change over time. For instance initial performance advantages for representational and closely-mapped icons can diminish with user experience (Green and Barnard, 1990; Isherwood, McDougall and Curry, 2007; McDougall, de Bruijn and Curry, 2000).

Aims

This study was carried out to explore the relationship between levels of user experience with congruent auditory icons (in terms of situation represented), and abstract warning sounds (in terms of urgency), and auditory icons and abstract alarms with conflicting situational and urgency mappings. It was expected that, as has been found with visual icon research, performance advantages for congruently-mapped sounds would be present at first but would reduce over time for both types of alarm. Participants were required to rate the perceived urgency and pleasantness at the beginning and end of the study, it was predicted that ratings would differ based on alarm-type and congruency of sound-situation pairing.

Methods

Participants

Twenty-four participants, twelve men and twelve women, took part in this experiment, they were between 19 and 42 years of age (mean = 26.38 years). All reported no known hearing problems and had either normal, or corrected-to-normal, vision. They received payment for participation.

Materials and stimuli

Sound stimuli

Tables, listing the eight sounds employed for the experiment, were used to record the participants’ perceived ratings for the pleasantness and urgency of the sound stimuli. Two sound groups were used: ‘auditory icons’ and ‘abstract’ sounds (see table 1). The study utilised a driving scenario in which each sound resembled a situation that may occur whilst driving a vehicle. The driving situations and accompanying sounds were chosen from situations utilised in a previous experiment (McKeown and Isherwood, 2007). The abstract sound category included sounds that were not considered to convey any type of meaning to listeners but had previously been matched to the perceived urgency of each of eight driving scenarios. The auditory icons were selected because they were thought to produce a realistic representative sound that one would expect to hear if they were in each of the driving situations. The sounds used for this experiment therefore had, necessarily, considerable acoustic variation. Sound durations ranged from 2.11 to 8.09 s, however in practice listeners did not wait until the end of sounds before responding. Sounds were presented at a comfortable listening level, and did not exceed 75 dB SPL, through BBC design LS3/5A loudspeakers in stereo configuration in a sound proofed booth.

[table 1 near here]

Computer task

Black and white pictures, approximately 27 mm by 27 mm, were used to represent the eight driving situations. They were displayed equidistantly in a semi-circle. At the bottom-centre of the semi-circle was a small red circle, diameter 15 mm, present at the beginning of each trial. The action of the participant moving the mouse cursor to the centre circle would simultaneously activate the 8 pictures on screen and start one of the sound stimuli. The pictures appeared in a random order on the screen for each trial. Participants' response time measurements were recorded from the start of the sound and ended when the participant clicked on one of the eight pictures.

Each of the two sound groups (abstract sounds and auditory icons) was further split into two sub-groups. One group had congruent pairings between the sounds and situations for half of the scenarios and incongruent for the other half. The second group had the reverse pairings to the first group. For the abstract group appropriate pairings were based on the perceived urgency of the sounds and situations. For the auditory icon group appropriate pairings were based on how well-mapped, due to representativeness, that the sounds were to the situations (see table 2).

[table 2 near here]

Procedure

The experiment was conducted over four sessions. For the first session participants listened to each of the eight sounds and rated them twice for pleasantness and for urgency. For all of the ratings participants were told to concentrate on sound rather than meaning. Participants were then informed of the sound-situation pairings (according to their group and sub-group) and advised to remember these pairings.

For the computerised task participants initially carried out a practice session where they heard each sound twice in random order (two blocks of eight trials). The correct picture corresponding to each sound would always flash after the participant had clicked on a picture regardless of whether their response was correct or not. The main experimental task involved each of the sound-picture pairings being presented four times in a random order. Participants received no feedback regarding the correct sound-picture pairings during these experimental trials.

The second session was carried out the following day. The participants' recall of the sound-picture pairings was tested using one block of eight sound-picture trials. Participants received no feedback regarding their performance during the recall test. Participants were then given a training session to complete; this consisted of one block of eight trials in which feedback was again provided so participants could refresh their memory for each of the pairings. Finally participants completed four blocks of eight experimental trials as had been done at the end of the first session. The third and fourth sessions were carried out one week later; they followed the same format as the second session. At the end of the fourth session participants were again required to perform two ratings of both the pleasantness and urgency of each sound.

Results

Computer task

Two measures were made of participants' performance whilst completing the computer task, accuracy and response time, to determine not only how quickly and correctly participants were responding to the alarm sounds but also to examine whether there had been any trade-off between these two measures. Greenhouse-Geisser corrections were used where appropriate in all ANOVA, however uncorrected degrees of freedom are reported.

Recall sessions

Recall tests were carried out at the beginning of the second, third and fourth experimental sessions. Three factors were examined: sound-type, congruency and experience (for sessions two, three and four) for both accuracy and speed of response. The first three-way ANOVA for accuracy revealed a significant effect of time, showing accuracy performance to improve for each session, $F(2, 44) = 6.89$, $p < .01$, $\eta_p^2 = .24$; a significant interaction between congruency and time, F(2, 44) = 5.50, $p < .01$, $\eta_p^2 = .20$; and a significant effect of sound-type, with auditory icons producing more accurate responses than abstract sounds F(1, 22) = 7.23, $p < .05$, $\eta_p^2 = .25$ (see figure 1).

[figure 1 near here]

A three-way ANOVA examining the factors of sound-type, congruency and experience for participants' speed of response showed a significant effect of congruency, F(1, 20) = 7.46, $p < .05$, $\eta_p^2 = .27$; of sound-type, F(1, 20) = 6.07, $p < .05$, $\eta_p^2 = .23$; and a significant interaction between congruency and sound-type, F(1, 20) = 6.55, $p < .05$, $\eta_p^2 = .25$. Bonferroni-corrected post hoc comparisons showed a significant effect of congruency for speed of response to auditory icons for each of the three recall sessions, $p < .01$. There was no effect of congruency for the abstract sounds. Finally the three-way ANOVA again showed a significant effect of experience, F(2, 40) = 10.19, $p < .001$, $\eta_p^2 = .34$, participants' speed of response improved during the three recall sessions (see figure 2)

[figure 2 near here]

Experimental sessions

Participants completed four experimental sessions, with each session consisting of four repetitions of each trial presented in a random order. A three-way ANOVA examining sound-type, congruency and experience for participants' accuracy of response showed a significant effect of sound-type, F(1, 22) = 7.26, $p < .05$, $\eta_p^2 = .25$, with auditory icons producing more accurate responses than abstract sounds, this finding also holds even for the incongruently matched auditory icons (see figure 3). The ANOVA also showed a significant effect of time, F(3, 66) = 12.78, $p < .001$, $\eta_p^2 = .37$; and a significant interaction between congruency and time, F(3, 66) = 3.93, $p < .05$, $\eta_p^2 = .15$.

[figure 3 near here]

For participants' speed of response a three-way ANOVA revealed a significant effect of time, $F(3, 66) = 27.01$, $p < .001$, $\eta_p^2 = .55$, as participants' responses improved for each experimental session. There was a also a significant effect of congruency, $F(1, 22) = 7.85$, $p < .05$, $\eta_p^2 = .26$; and a significant interaction between congruency and sound-type, $F(1, 22) = 8.54$, $p < .01$, $\eta_p^2 = .28$. The interaction revealed congruency to be an important contributing factor for speed of response to auditory icons but not for abstract sounds, this was significant at $p < .01$ according to Bonferroni-corrected post hoc comparisons (see figure 4).

[figure 4 near here]

Participant ratings

Figures 5 and 6 show participants' initial ratings for the pleasantness and urgency of the auditory icons and abstract sounds. These ratings were made prior to the sounds being associated with any of the driving scenarios.

[figure 5 near here]

[figure 6 near here]

A Wilcoxon Signed-ranks test was carried out to determine whether participants rated the auditory icons and abstract icons differently for urgency and pleasantness at the first and final rating sessions. Results from these tests showed significant differences, with medium to large effect sizes, occurring between the ratings for the auditory icons but not for the abstract sounds.

For pleasantness the auditory icon for 'door open' showed a significant difference in participants ratings, $z = -2.11$, $p < .05$, $r = -.43$ (Mdn for session 1 = 4.75, for session 4 = 5.50). For the urgency ratings the auditory icon for ‘low tyre pressure’ was rated as being significantly less urgent between the first (Mdn = 6.38) and final rating sessions (Mdn = 5.25) z = -2.29, p < .05, r = -.47, this was similarly the case for ‘low oil’ z = -2.36, p < .05, r = -.48 (Mdn for session 1 = 3.50, session 4 = 2.00), ‘drifting off road’ z = -2.20, p < .05, r = -.45 (Mdn for session 1 = 6.00, session 4 = 4.50), ‘car in blind-spot’ z = -2.23, p < .05, r = -.46 (Mdn for session 1 = 9.00, session 4 = 8.25) and ‘headway quickly closing’ z = -2.11, p < .05, r = -.43 (Mdn for session 1 = 8.50, session 4 = 7.88).

Discussion

Ratings results

For both types of alarm the sounds chosen to represent the more urgent driving scenarios were generally rated as such by participants, and vice versa for the less urgent situations. Pleasantness ratings often show an inversed pattern to the urgency ratings. Participant ratings for the auditory icons did not show consistency across the first and final rating sessions, particularly when assessing the urgency of the sound. Participants had been advised to make judgements on the sound characteristics rather than the sound-situation mappings. However the fact that this measure was not stable could perhaps reflect the difficulty of assessing the urgency of auditory icons or confusions occurring as a result of the sound pairings with incongruent situations. Given Gaver’s (1986, 1993) assertion that auditory information is mapped to events rather than to sounds it may be that the listeners, when completing the ratings, became confused between the proximal and distal stimuli information being presented by the auditory icon particularly when they were paired with the incongruent driving situations.

Computer task: Accuracy of responses

Accuracy of response was noticeably better for participants responding to auditory icons than to abstract sounds. Participants’ responses to the auditory icons suggest that listeners have a strong mental representation of the environmental sounds which may help to facilitate their coupling of the sound with its intended referent, resulting in improved accuracy responses for these stimuli. It is possible that these ‘familiar’ sounds require minimal processing of their acoustic characteristics, as the sound only needs to be processed until it can be identified in long-term memory, once the sound has been identified and a mental representation is activated further ‘exhaustive’ processing of the sound parameters can be discontinued (Guillaume et al, 2003). Guillaume et al (2003) posited that such a ‘representational’ processing mode would result in immediate identification of a sound, be fast and automatic, and require fewer attentional demands. Following the arguments presented by Gaver (1986, 1993), perhaps it was the case that listeners were also able to utilise cues relating to the reliable relationships between the sounds and their sources in order to more directly identify the sound-producing events, rather than exclusively analysing on the acoustic features of the sound.

The listener must then associate this identified sound with its signified meaning. This will be achieved with the signal-referent relationships falling along a continuum from ‘closely-mapped’ auditory icons, providing intuitive links with their referents requiring minimal learning, to more ‘distantly-mapped’ signal-referent associations requiring greater levels of learning. The fact that the auditory icons were familiar sounds for listeners (albeit at times counterintuitive with their intended referents) resulted in stronger associations being built between the icon sound and its referent than was the case for the abstract sounds.

Conversely, for abstract sounds, where no specific prior experiential association existed, the listener would have to carefully process the acoustic features of the sound and form a new association in memory between the sound and its referent. As noted by Perry and colleagues (2007) and Guillaume et al (2003) learning the association between an abstract sound and its referent can be an effortful process requiring greater levels of attention.

The performance advantage found for auditory icons in comparison to abstract sounds may also result from such sounds being more likely to enable both verbal and imaginal encoding in memory than abstract sounds. Studies have found that verbal labels can facilitate how well sounds can be remembered (Bartlett, 1977; Paivio, Philipchalk, and Rowe, 1975; Philipchalk and Rowe, 1971).

With regard to the congruency between sound and situation pairings for performance accuracy in the present study, this was significantly important during participants' first exposure to the auditory icons but only a trend for congruency effects was observed during the following trials; congruency was less important for abstract warnings. Performance advantages for well-mapped auditory icons over more distantly mapped sounds have been reported in a number of studies (for instance see Keller and Stevens, 2004; McKeown and Isherwood, 2007; McKeown, Isherwood and Conway, 2010). The importance of congruency for novice, but not for experienced, listeners of the auditory icons also mirrors visual icon research in which the representativeness and closeness of icon-referent pairings has been found to be of most importance for inexperienced users of icon-sets (Green and Barnard, 1990; Isherwood, McDougall and Curry, 2007; McDougall, de Bruijn and Curry, 2000). Although not statistically significant, it is notable that the incongruent auditory icons did not reach the same level of accuracy of performance as their congruent counterparts at any time during the experiment.

Computer task: Speed of responses

For both experimental and recall tasks congruent auditory icons resulted in a consistently good speed of performance in comparison to the other sound types. Incongruent auditory icons generally resulted in the slowest responses. Both congruent and incongruent sounds did improve over time but for auditory icons there is a consistent and noticeable difference in the speed of response for well-mapped sounds and those with conflicting mappings. This progresses the argument that auditory displays, especially if using representational sounds in a time critical setting must be well designed to 'map' as directly as possible with their referents.

The congruent and incongruent abstract sounds produced similar results in the experimental phase of the experiment, generally falling between participants' responses for congruent and incongruent auditory icons; they also showed the most improvement over time. For the recall task congruent abstract sounds elicited generally good responses with the incongruent abstract sound responses being more variable.

It is interesting to note that participants were responding faster to the abstract sounds (whether congruently matched or not) than to the incongruent auditory icons. This shows, perhaps, that the processing speed for learning new signal-referent pairings (including the more exhaustive acoustic analysis of these initially unfamiliar sounds) was faster than for responding to familiar but conflicting iconic sounds. Or, perhaps, speeded responses to abstract sounds could to some degree have resulted from participants simply guessing, or quickly responding to sounds that they did not know the correct pairing for, as more abstract warnings were incorrectly responded to than was the case for the iconic sounds.

Computer task: Experience

The amount of experience a participant gained with the auditory alarms was a significant factor in each of the recall and experimental conditions for both accuracy and speed of response, performance improved for all types of auditory warning over time. The importance of congruency for icon sounds changed over time however this was more so for the accuracy performance than for speed of performance. With experience listeners were able to respond almost as accurately for incongruent icon pairings as they did for the congruent pairings but for speed of performance incongruent icons retained a notable performance decrement. There was a steady improvement for the abstract sounds, perhaps with further experimental trials the accuracy of response for the abstract sounds may reach that of the icon sounds. However this experiment does represent fairly significant repeated exposure to the sounds, a luxury of learning that may not always be possible for an alarm set in the real world. The abstract sounds in the experimental trials also showed a notable improvement over time for speed of response. Consequently this study shows that abstract sounds will particularly suffer from an accuracy decrement, and poorly mapped auditory icons from a response time decrement, that appears to be quite long-lasting.

User acceptance, confusability and alarm design

Even though auditory icons have been found in a number of studies to have potential performance advantages over abstract warning sounds they have not to date been widely employed for human-machine communication. One reason for this may be due to the lack of user acceptance for auditory icons (for instance see Stanton and Edworthy, 1998; Belz, Robinson and Casali, 1999). Perhaps more research should consider situations in which auditory icons may be more acceptable for instance hearing a bird briefly tweet to inform a computer-user that they have a message on Twitter may be considered more relevant and appropriate than hearing a horn blast to inform of a potential vehicle collision. Consequently the context, expected response and importance of an alarm and the sound of the auditory icon itself, will all play a part in listener acceptability for these warnings.

As noted by Stanton and Edworthy (1998) there may be resistance to radical departures in alarm design. It may be the case that abstract sounds have become very strongly associated with alarms and warning signals in general whereas auditory icons simply have not. In order to consider auditory icons for this role people would have to change their opinion of what an alarm signal sounds like and would have to learn to associate representative sounds with such signals. The question becomes one of whether this would be possible and whether operators would encounter issues with these alarms that research has not yet considered.

One such problem could be the potential for confusion between alarm sounds and real world sounds. An example of the type of confusion that may occur for a representational auditory icon could be a driver mistaking car horn blasts being used as auditory icons with an event outside of the vehicle or alternatively incorrectly associating a sound originating from outside of their vehicle with an auditory vehicle alert; potentially resulting in the driver making an erroneous response to the warning sound or the real-world sound. It may be the case that certain representative alarms are more likely to lead to confusions than others, for instance a car horn sounding is a realistic sound which could originate from inside or outside of the vehicle.

Perhaps consideration should be given to ensuring that these alarms are both discriminable from other auditory icons and also, in certain situations, from other environmental sounds. As noted by Graham (1999) this could be done through using representative sounds in conjunction with other more traditional abstract warnings or earcons, so that a listener can identify the warning sound as such. This would therefore utilise a listener’s well-learned associations with abstract sounds and alarm situations and would also enable them to benefit from the performance advantages offered by iconic sounds.

As with many objects or events that are presented using visual icons there is also a problem for auditory icons that there is not always a representative display for a real-world referent, and so the auditory warning will necessarily have to be abstract in design. This does not mean that the warning would necessarily be distantly-mapped to its referent. Consider for instance a portable Geiger-Müller counter producing audible clicks to indicate amount of radiation in an area or a car signalling nearness of objects through increasing the alarm’s pulse rate and decibel level, both provide examples of situations, in which abstract alarms have been designed to provide intuitive cognitive links, or a close semantic mapping, with a real-world referent. In fact it is the close cognitive link rather than the representational nature of the alarm that has been found in visual icon studies to be of particular importance for effective operator performance (Isherwood, McDougall and Curry, 2007; McDougall, de Bruijn and Curry, 2000).

Conclusion

This study has shown that well-mapped, congruent, iconic sounds produce both accuracy and response time advantages in comparison to incongruent iconic sounds and abstract warnings. User performance for all sound types did improve over time, however even with listener experience a decrement in accuracy performance remained for the abstract warning sounds, and in speed of response for the incongruent iconic sounds. These findings have relevance for design practice. They suggest that designers should consider the association between the signal-referent relationships for different types of auditory warnings particularly if the warnings need to be understood immediately with little or no learning. Research is needed to investigate the situations and contexts in which differing types of alarms may be suitable and how to best design an alarm to semantically ‘map’ to its intended referent.

References

Ballas, J. A. and Howard, J. H. 1987. Interpreting the language of environmental sounds. Environment and Behavior, 19 (1): 91-114.

Bartlett J. C. 1977. Remembering Environmental Sounds: The Role of Verbalization at Input. Memory and Cognition 5 (4): 404-414. doi: 10.3758/BF03197379. Belz, S. M., Robinson, G.S., and Casali, J. G. 1999. A New Class of Auditory Warning Signals for Complex Systems: Auditory icons. Human Factors 41 (4): 608-618.

Belz, S. M., Robinson, G.S., and Casali, J. G. 1999. A New Class of Auditory Warning Signals for Complex Systems: Auditory icons. Human Factors 41(4): 608-618.

Brazil, E., Fernström, M. and Bowers, J. 2009. Exploring concurrent auditory icon recognition. In Proceedings of the 15th International Conference on Auditory Display (ICAD).

Byrne, M. D. 1993. Using Icons to Find Documents: Simplicity is Critical. Proceedings *of INTERCHI ’93*: 446–453.

Carello, C., Anderson, K. L. and Kunkler-Peck, A. J. 1998. Perception of object length by sound. Psychological science, 9 (3): 211-214.

Edworthy, J. and Hellier, E. J. 2006. Auditory warnings. In International Encyclopedia of Ergonomics and Human Factors. 2nd ed., edited by W. Karwowski. London: Taylor and Francis.

Forsythe, A., Mulhern, G. and Sawey, M. 2008. Confounds in Pictorial Sets: The Role of Complexity and Familiarity in Basic-Level Picture Processing. Behavior Research Methods 40 (1): 116-129.

Gaver, W. W. 1986. Auditory Icons: Using Sound in Computer Interfaces. Human-Computer Interaction 2 (2): 167-177. doi: 10.1207/s15327051hci0202_3.

Gaver, W.W. 1993. How do we hear in the world? Explorations in ecological acoustics. Ecological psychology, 5 (4): 285-313.

Gibson, J. J. 1979. The ecological approach to visual perception. New York: Houghton Mifflin.

Graham, R. 1999. Use of Auditory Icons as Emergency Warnings: Evaluation Within a Vehicle Collision Avoidance Application. Ergonomics 42 (9): 1233-1248. doi: 10.1080/001401399185108

Grassi, M. 2005. Do we hear size or sound? Balls dropped on plates. Perception and psychophysics, 67 (2): 274-284.

Goonetilleke, R. S., Shih, H. M., On, H. K. and Fritsch, J. 2001. Effects of training and representational characteristics in icon design. International Journal of Human-Computer Studies 55 (5): 741-760.

Green, A. J. K. and Barnard P. J. 1990. Iconic interfacing: The role of icon distinctiveness and fixed or variable screen locations. In Human-Computer Interaction - INTERACT '90, edited by D. Diaper, D. Gilmore, G. Cockton, and B. Shackel. Amsterdam, Netherlands: Elsevier Science Publishers.

Guillaume, A., Pellieux, L., Chastres, V. and Drake, C. 2003. Judging the Urgency of Nonvocal Auditory Warning Signals: Perceptual and Cognitive Processes. Journal of Experimental Psychology: Applied. 9 (3): 196-212

Gygi, B., Kidd, G. R. and Watson, C.S. 2004. Spectral-temporal factors in the identification of environmental sounds. The Journal of the Acoustical Society of America 115 (3): 1252–1265.

Hermann, T. 2008. Taxonomy and definitions for sonification and auditory display. Proceedings of the 14th International Conference on Auditory Display (ICAD).

Isherwood, S. J., McDougall, S. J. P., and Curry, M. B. 2007. Icon Identification in Context: The Changing Role of Icon Characteristics with User Experience. Human Factors 49 (3): 465–476. doi: 10.1518/001872007X200102.

Keller, P. and 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): 3-12.

Lazarus, H. and Höge, H. 1986. Industrial Safety: Acoustic Signals for Danger Situations in Factories. Applied Ergonomics 17 (1): 41-46. doi: 10.1016/0003-6870(86)90192-4.

Leung, Y. K., Smith. S., Parker. S., and Martin. R. 1997. Learning and Retention of Auditory Warnings. Proceedings of the 4th International Conference on Auditory Display (ICAD): 128-131.

Marcell, M. M., Borella, D., Greene, M., Kerr, E. and Rogers, S. 2000. Confrontation Naming of Environmental Sounds. Journal of Clinical and Experimental Neuropsychology 22 (6): 830-864.

McDougall, S. J. P., Curry, M. B. and de Bruijn, O. 2001. The effects of Visual Information on Users' Mental Models: An Evaluation of Pathfinder Analysis as a Measure of Icon Usability. International Journal of Cognitive Ergonomics 5 (1): 59-84.

McDougall, S. J. P., de Bruijn, O. and Curry, M. B. 2000. Exploring the effects of Icon Characteristics on User Performance: The Role of Icon Concreteness, Complexity, and Distinctiveness. Journal of Experimental Psychology: Applied. 6: 291–306.

McDougall S, Isherwood S. 2009. What’s in a Name? The Role of Graphics, Functions and their Interrelationship in Icon Identification. Behavior Research Methods. 41 (2): 325–336. doi: 10.3758/BRM.41.2.325.

McKeown, D. and Isherwood, S. 2007. Mapping Candidate Within-Vehicle Auditory Displays to Their Referents. Human Factors 49 (3): 417-428. doi: 10.1518/001872007X200067.

McKeown, D., Isherwood, S. and Conway, G. 2010. Auditory Displays as Occasion Setters. Human Factors 52 (1): 54-62. doi: 10.1177/0018720810366861.

Nees, M. A. and Walker, B.N. 2009. Auditory Interfaces and Sonification. In The Universal Access Handbook, edited by C. Stephanidis, 507 – 521. New York: CRC Press.

Paivio, A., Philipchalk, R. and Rowe, E. J. 1975. Free and Serial Recall of Pictures, Sounds, and Words. Memory and Cognition 3 (6): 586-590.

Patterson, R. D. 1982. Guidelines for the auditory warnings systems on civil aircraft. (CAA paper 82017). London: Civil Aviation Authority.

Patterson, R. D. 1989. Guidelines for the design of auditory warning sound. Proceedings of the Institute of Acoustics 11 (5): 17–25.

Perry, N. C., Stevens, C. J., Wiggins, M. W. and Howell, C. E. 2007. Cough Once for Danger: Icons Versus Abstract Warnings as Informative Alerts in Civil Aviation. Human Factors 49 (6): 1061-1071.

Petocz, A., Keller, P., and Stevens, C. 2008. Auditory Warnings, Signal-Referent Relations and Natural Indicators: Re-thinking Theory and Application. Journal of Experimental Psychology: Applied 14: 165-178.

Philipchalk, R. P. and Rowe, E. J. 1971. Sequential and nonsequential memory for verbal and nonverbal auditory stimuli. Journal of Experimental Psychology 91 (2): 341-343.

Rogers, Y. and Oborne, D. J. 1987. Pictorial Communication of Abstract Verbs in Relation to Human-Computer Interaction. British Journal of Psychology 78 (1): 99-112. doi: 10.1111/j.2044-8295.1987.tb02229.x

Scott D. 1993. Visual Search in Modern Human-Computer Interfaces. Behaviour & Information Technology. 12: 174–189. doi: 10.1080/01449299308924378.

Shafiro, V. 2008. Development of a Large-Item Environmental Sound Test and the Effects of Short-Term Training with Spectrally-Degraded Stimuli. Ear and hearing 29 (5): 775-790.

Stanton, N. and Edworthy, J. 1998. Auditory Affordances in the Intensive Treatment Unit. Applied Ergonomics, 29 (5): 389-394. doi:10.1016/S0003-6870(97)00064-1.

Stephan, K. L., Smith, S. E., Martin, R. L., Parker, S. P. A. and 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. doi: 10.1518/001872006777724426

Stotts, D. B. 1998. The Usefulness of Icons on the Computer Interface: Effect of Graphical Abstraction and Functional Representation on Experienced and Novice Users. In: Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting: 453–457.

Walker, B.N. and Kramer, G. 2004. Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In Ecological Psychoacoustics, edited by J. G. Neuhoff, 150 – 175. London: Elsevier Academic Press.

Table 1. Driving situations with descriptions and waveforms of acoustic warnings

Table 2. Congruent and incongruent mappings for sub-groups 1 and 2

Figure 1. Recall session congruency and accuracy of response over time for auditory icons and abstract alarm sounds

Figure 2. Recall session congruency and speed of response over time for auditory icons and abstract alarm sounds

Figure 3. Experimental session accuracy of response for alarm type and congruency over time

Figure 4. Experimental session response times for alarm type and congruency over time

Figure 5. Auditory icon pleasantness and urgency ratings

Figure 6. Abstract sound pleasantness and urgency ratings