HOW MANY AUDITORY ICONS, IN A CONTROL ROOM ENVIRONMENT, CAN YOU LEARN?

Ronja Frimalm

Interactive Institute Swedish ICT
Acusticum 4, SWE-94128 Piteå, Sweden
ronja.frimalm@tii.se

Johan Fagerlönn

Interactive Institute Swedish ICT
Acusticum 4, SWE-94128 Piteå, Sweden
johan.fagerlonn@tii.se

Stefan Lindberg

Interactive Institute Swedish ICT
Acusticum 4, SWE-94128 Piteå, Sweden
stefan.lindberg@tii.se

Anna Sirkka

Interactive Institute Swedish ICT
Acusticum 4, SWE-94128 Piteå, Sweden
anna.sirkka@tii.se

ABSTRACT

Previous research has shown that auditory icons can be effective warnings. The aim of this study was to determine the number of auditory icons that can be learned, in a control room context. The participants in the study consisted of 14 control room operators and 15 people who were not control room operators. The participants were divided into three groups. Prior to the testing the three groups practiced on 10, 20 and 30 different sounds. Each group was tested using the sounds that they had practiced. The results support the potential for learning and recalling a large number of auditory icons, as many as 30. The results also show that sounds with similar characteristics are easily confused.

1. INTRODUCTION

In numerous user environments, such as vehicles, control rooms, and various health care institutions, sounds frequently assume the form of auditory warning signals. The use of sound as an information carrier is an effective way to present information, especially when there is a simultaneous demand for visual attention (e.g. operator monitoring a process). In contrast with visual information sounds can be perceived without direct attention to a specific location.

In a control room environment sound information can be especially beneficial. Operators are required to keep track of a considerable amount of information, which is presented on numerous screens and stations. By designing sounds that an operator can easily associate with various parts of the process, operators can be guided to problematic areas and relevant monitors. However, this approach is not a typical method for the selection of sound alarms in this type of environment. Sounds frequently serve as abstract warnings with an arbitrary connection to the alarming event and the same sound is frequently employed for several different alarms. This makes it difficult for the operators to indentify the problematic area and the appropriate intervention. Simplification of the identification process for an alarm may increase the effectiveness of control room operators. Improvements may also provide the opportunity to employ additional sounds in an environment without increasing the risk of confusion.

Several different approaches exist for designing sounds that are used as information carriers. One method involves the use of abstract sounds that are arbitrarily connected to the object/event they are intended to represent. This approach is a common method for creating sounds to be applied in different types of products. Many household products e.g. dishwashers and micro wave ovens emit various beeping sounds to call attention to the completion of an operation or cycle. These sounds are arbitrarily connected to the meaning "finished". Learning and recalling abstract warning sounds can be challenging. Patterson and Mayfield [1] describe a study that concluded that four to six arbitrary warnings can be rapidly learned. The rate of acquisition decreases for a larger number of warnings: thus further effort is required. Patterson and Mayfield also suggested that the application of more than 10 warnings in one context, an aircraft, is excessive.

Earcons are defined by Blattner et al. [2] as "nonverbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation or interaction". Garzonis et al. [3] suggest that earcons, which are typically described as abstract musical sounds lack meaningful relationships with their referents. As a result the user is required to learn and memorize these relationships.

Speech warnings can also be used to convey information. Speech is a beneficial method of presenting information in several different conditions. Because speech is a direct form of communication the relationship between the spoken word and the object/event it represents does not have to be learned.

The research presented in this paper was funded by EU:s Structural Funds, the County Administrative Board of Norrbotten, Piteå Municipality, Skellefteå Municipality, and RISE.

Speech is also suitable for the presentation of complex information. However, speech warnings have several disadvantages. They are easily masked by background noise, they interfere with other verbal communication and the language of different operators varies among different countries.

Gaver [4] defined auditory icons as “sounds that provide information about an event that represents desired data“ and further describe them as caricatures of naturally occurring sounds. Auditory icons have several advantages over both speech and abstract sounds. They are easier to learn, they require fewer training trials and they yield a lower error rate than abstract warnings. The response times for auditory icons can be faster than the response times of both abstract sounds and speech warnings and the risk of being masked by background communication is reduced compared with speech warnings [5, 6, 7, 8]. In the study by Lucas [8], in which six sounds were tested, it is suggested that a larger sample of sounds can more accurately describe the communicative ability of auditory icons. Auditory icons also seem to be less likely to be forgotten compared with more conventional auditory warning signals [9]. An explanation of why auditory icons are appropriate for use as warning signals is provided by Fagerlönn and Alm [10], who found that the use of sounds which are meaningful in the context in which they are applied can reduce the number of required cognitive resources compared with the use of arbitrarily mapped sounds.

Numerous studies have shown that auditory icons are more appropriate warnings compared with abstract sounds and other types of warnings. However, few studies focus on the number of warnings for an industrial context such as a control room.

Therefore the aim of this study was to investigate the maximum number of auditory icons that can be learned and recalled in a control room environment. The selection of an industry as context was based on previous studies about this issue. A larger project that examines alarm signals in an industrial control room for several months has been conducted; this study is an extension of this project [11]. Based on the problems described previously and the complex information flow in a control room, this type of auditory icon was determined to be beneficial for this type of environment.

2. METHOD

2.1. Participants

The study included a total of 29 participants: 14 women and 15 men. Fourteen of the participants worked as control room operators at a local paper mill (11 men, and 3 women). To obtain a larger group, 15 additional participants who were not control room operators, were also recruited (4 men, and 11 women). The age of all participants ranged from 20 to 58 years (mean age: 35). The mean age of the control room operators was 43 years. The mean age of the remaining participants was 28 years. None of the participants reported any hearing impairments. The participants were compensated with a movie ticket for their contributions to the study.

2.2. Apparatus

The test was performed using an Apple MacBook laptop. The training and test application were developed in Java. The participants wore headphones (Koss PortaPro) during the test.

2.2.1. Sounds

Thirty different sounds were created for the study. All sounds are auditory icons that were supposed to represent the type of events and objects that may occur in an industry setting. All sounds ranged from 0.5 and 5.0 seconds. The sounds had a sample rate of 44.1 kHz. The aim was to design sounds with a direct relationship to the object/event that they were supposed to represent, because direct relations create rapid and immediate recognition of the target [12]. For some sounds, these direct relationships were difficult to find; therefore, an indirect relationship was employed. According to Stephan et al. [13] indirect associations are an excellent alternative when direct associations are not feasible. In this case a direct relationship can be defined as an event that causes a sound, which is also the referent of the sound. For example, the sound of a burning fire is associated with “fire”. It is an indirect relationship when the referent of a sound is not the event, but can be associated with the event in another way. An example of an indirect association is the sound of a running train, which is associated with “stacker”.

2.3. Procedure

The participants were divided into three test groups. The first group listened to 10 different sounds, the second group listened to 20 different sounds and the third group listened to 30 different sounds. There was an equal distribution of participants from the two participant groups among the different sound groups.

The control room operators performed the test in two control rooms at the paper mill. The group of non-operators performed the test in an office environment. Figure 1 provides an example of the sound levels in one of the control rooms and the office environment.

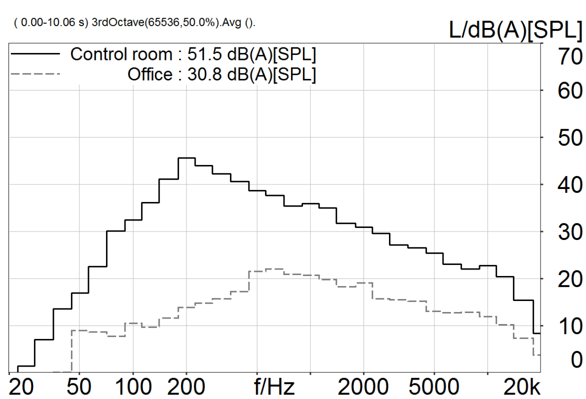

Figure 1: A-weighted SPL in 1/3-octave bands of the background noise of one of the control rooms and the office.

All participants completed a training session prior to the start of the test session in which they were given written instructions in Swedish on how to perform the training. An audio player (used to trigger the sounds), a picture of the intended association and a word that explained the intended association were presented to the participants on a laptop. The participants' task was to learn the item-sound pairs. During the training session, the participants were asked to learn the intended mapping for each sound and the event that each sound represented. The participants in each group listened to all sounds they were subsequently tested on e.g. the group with 10 sounds practiced on the 10 sounds and was subsequently tested on these sounds.

During the training session the participants were presented with the auditory icons through their headphones. The sounds were presented in random order. Throughout the training session, the participants were allowed to listen to each sound as many times as needed. However, they could not listen to a sound again after advancing to the next sound. Figure 2 provides an example of the training application.

Figure 2: Training application.

The participants completed the test session immediately after completing the training session. The participants received another set of written instructions. The test session included all the sounds that each participant had previously practiced on. Similar to the training session, the sounds were presented in random order and the participants were presented with an audio player. However, the intended associated word was not provided. Below the audio player, pictures of all 30 sounds were displayed (for all groups); and the participants' task was to select the picture that was associated/ paired with each sound by clicking on the correct box. The participants could not change previously selected answers (refer to figure 3).

The performance measure of the study was the number of correct answers.

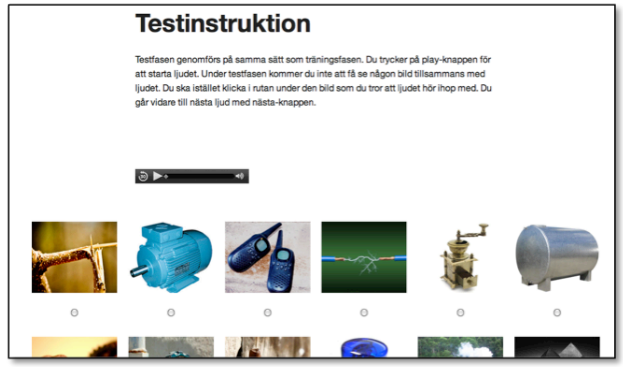

Figure 3: Test application.

3. RESULTS

The percentages of correct answers for the three groups are presented in figure 4.

Because a normal distribution was not assumed a Kruskal-Wallis test was performed to analyze the differences among the three groups. A significant difference among the groups was observed ($H(2) = 9.001$, $p = 0.011$). A Mann-Whitney U test with a Bonferroni correction for multiple comparisons (alpha = 0.05/3 = 0.017) showed a significant difference among the groups with regard to ten and thirty sounds. No other significant differences were observed.

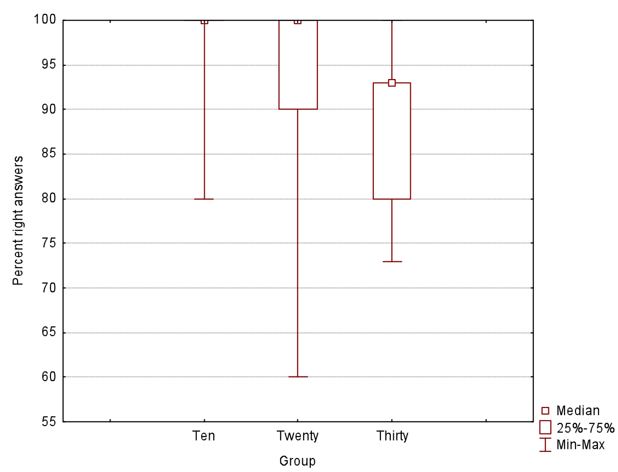

Figure 4: Percentage of correct answers for the three groups.

Figure 4 also reveals that only one participant in the group with 10 sounds did not receive 100 % correct answers. In the group with 20 sounds, the majority of the participants received 100 % correct answers. The majority of the participants in the group with 30 sounds received 93 % correct answers. One participant in the group with 30 sounds received 100 % correct answers.

A two-tailed Mann-Whitney U test was performed to assess the difference between the group of operators and the group of non-operators. The test results indicated a significant difference between the groups ($U = 48.5$, $n1 = 15$, $n2 = 14$, $p = 0.012$).

Another two-tailed Mann-Whitney U test was performed to determine the potential effect of age. The participants were divided into two age groups; participants younger than 30 and participants older than 40. No significant difference between the different age groups was observed ($U = 43.0$, $n1 = 13$, $n2 = 11$, $p = 0.106$). Pearson product-moment correlation coefficient showed no correlation between age and result within the group of operators ($r = -0.039$, $p = 0.895$) or within the group of non-operators ($r = -0.366$, $p = 0.18$).

To determine any gender differences a Mann-Whitney U test was performed. No significant difference between the genders was observed ($U = 93.0$, $n1 = 15$, $n2 = 14$, $p = 0.621$).

Note that the results from the Mann-Whitney U tests should be carefully interpreted due to the incidence of several ties between scores.

The total number of incorrect answers was 33. Table 1 shows that among the 33 incorrect answers some sound pairs were incorrectly paired more than one time.

Two sound pairs were incorrectly paired at least four times. Sounds 4 and 8, were incorrectly paired four times. Sounds 15 and 22 were paired a total of six times (sound 15 was mistaken for sound 22, four times, and sound 22 was mistaken for sound 15, two times).

Table 1: Sounds with the highest number of incorrect parings
* Sounds 15 and 22 were paired a total of six times.

Number of incorrect sound pairings

Sounds	Number of parings
1-25	2
2-3	3
4-6	2
4-8	4
5-13	3
15-22*	4
19-27	2
21-26	3
22-15*	2

4. DISCUSSION

The findings of this study indicate that it is possible to learn and recall a large number of sounds in a control room environment if the sounds are characterized as auditory icons. Based on previous studies, Patterson and Mayhem [1] suggested that 10 or more alarm sounds in an aircraft is probably too many. The difficulty with learning arbitrarily mapped (abstract) sounds is supported by other studies [6, 10]. The results from this study indicate that it is possible to learn and recall a significant number of sounds if they are designed differently, that is, if they are designed as auditory icons with a relationship to the referent. The results demonstrate the feasibility of learning and recalling up to 30 different sounds. In this study the participants did not perform any typical operator tasks during the test. To gain more ecological validity future research should focus on further investigating sounds in the actual control room environment, with operators performing typical operator tasks simultaneously as discriminating between sounds.

Because the majority of the participants received a high percentage of correct answers, it seems like the majority of people can learn many more sounds. The results are not surprising considering the large number of sounds recognizable in everyday life. The results suggest that if designers apply existing associations between sounds and events, auditory displays can be implemented that inform operators in a verity of critical events. These displays may significantly improve the work efficiency as well as the overall working environment.

Although these results were obtained during a brief period of practice only a few participants received less than 90 % correct answers. These results support the simple learning of auditory icons. The low percentage of correct answers for some of the participants may be explained by the brief period of time allotted to learn the sounds. Providing additional practice occasions may improve the results. In a real-life context the operator may be given more time and opportunity to practice the sounds.

Follow-up sessions could be conducted in future studies. The test in this study was performed immediately after the training session. A longer period of time between the training session and the actual test phase may influence the number of recalled sounds.

A statistically significant difference between the operators and non-operators was observed. We examined any differences in age that may have caused different results. The group of control room operators exhibited a higher mean age (43) than the group of non-operators (28). Although no significant difference between the age groups was observed, a tendency toward a decline in the results with age was evident. However, no significant correlations between age and scores were found for either operators or non-operators, indicating that age probably had a minor effect.

There are other possible factors that could have contributed to the observed difference. First, the control room participants performed the test in their workplace, at the paper mill; this noisy environment may explain the poorer results of the operators. The background noise in the control room was louder than the background noise in the office (refer to figure 1). Secondly, the operators participated in the study during their working hours, which might have been a bit stressful and non-motivating. The non-operators participated during their free time. Perhaps they felt less stressed and more motivated to learn the sounds and reply correctly.

We also searched for any gender differences. The majority of the participants in the group of control room operators consisted of men (11 men and 3 women). The majority of the participants in the group of non-operators consisted of women (11 women, and 4 men). However, no gender differences were observed that explained the differences between the two groups.

Some of the sounds employed in the study were incorrectly paired several times. Specifically, sounds 4 and 8, and sounds 15 and 22 were incorrectly paired four times and six times respectively (refer to table 1). Sounds 4 and 8 are both squeaking noises. Sounds 15 and 22 sound like a snap/slam. In both cases, the correct answer and the incorrect answer sounded quite similar. The incorrect answers may be explained by the similarity in the characteristics of these sounds, which may have confused some of the participants. It seems important to keep this in mind when designing sounds. Similar sounds can be confusing. If all sounds can be easily distinguished, a greater number of sounds may be possible to be learned and recalled.

Another way to increase the number of sounds that can be learned and recalled is to involve the users in the design process. This approach can ensure a strong association between the sound and the referent. The users are experts on their own working environment and should know the appropriate types of associations. A stronger association may contribute to an even higher recall rate.

This study focused on alarm sounds in an industrial context. However, in addition to control room environments, auditory signals are used in other environments in which auditory icons can make considerable difference. For example, numerous different warning/alarm sounds are employed in hospital environments.

The use of auditory icons instead of abstract sounds creates the opportunity to incorporate sound in interface design to a greater extent. Here, it was tested in the context of a control room environment. The results indicate the potential for the use of a large number of sounds, if it is desirable.

5. REFERENCES

[1] R. D. Patterson, T. F. Mayfield, “Auditory Warning Sounds in the Work Environment [and Discussion],” *Philosophical Transactions of the Royal Society of London. B, Biological Sciences,* vol. 327, no. 1241, pp. 485-492, 1990.

[2] M. M. Blattner, D. A. Suikawa, and R. M. Greenberg, “Earcons and Icons: Their Structure and Common Design Principles,” *Human-computer interaction,* vol. 4, pp. 11-44, 1989.

[3] S. Garzonis, S. Jones, T. Jay, E. O’Neill, "Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 1513-1522.

[4] W. W. Gaver, “Auditory Icons: Using Sound in Computer Interfaces,” *Human-Computer Interaction,* vol. 2, pp. 167-177, 1986.

[5] N. C. Perry, C. J. Stevens, M. W. Wiggins, C. E. Howell “Cough Once for Danger: Icons Versus Abstract Warnings as Informative Alerts in Civil Aviation,” *Human Factors,* vol. 49, no. 6, pp. 1061-1071, 2007.

[6] Y. K. Leung, S. Smith, S. Parker, R. Martin, ”Learning and Retention of Auditory Warnings,” in *Proc. of the 4th Int. Conf. ICAD*, Palo Alto, CA, USA, 1997.

[7] R. Graham, “Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application,” *Ergonomics,* vol. 42, no. 9, pp. 1233-1248, 1999.

[8] P. A. Lucas, “An Evaluation of the Communicative Ability of Auditory Icons and Earcons,” in *Proc. of the 2nd Int. Conf. ICAD,* Santa Fe, NM, USA, 1994, pp. 121-128.

[9] S. M. Belz, G. S. Robinson, J. G. Casali, “A New Class of Auditory Warning Signals for Complex Systems: Auditory Icons,” *Human Factors,* vol. 41, no. 4, pp. 608-618, 1999.

[10] J. Fagerlönn, H. Alm, “Auditory signs to support traffic awareness,” *IET Intelligent Transport Systems,* vol. 4, no. 4, pp. 262-269, 2010.

[11] A. Sirkka, J. Fagerlönn, K. Delsing, S. Lindberg, ”The design of an auditory alarm concept for a paper mill control room,” *Manuscript in preparation,* 2014.

[12] P. Keller, C. Stevens, “Meaning From Environmental Sounds: Types of Signal–Referent Relations and Their Effect on Recognizing Auditory Icons,” *Journal of Experimental Psychology: Applied,* vol. 10, no. 1, pp. 3-12, 2004.

[13] K. L. Stephan, S. E. Smith, R. L. Martin, S. P. A Parker, K. I. McAnally, “Learning and Retention of Associations Between Auditory Icons and Denotative Referents: Implications for the Design of Auditory Warnings,” *Human factors,* vol. 48, no. 2, pp. 288-299, 2006.