Designing Audio and Tactile Crossmodal Icons for Mobile Devices

Eve Hoggan and Stephen Brewster

Glasgow Interactive Systems Group, Department of Computing Science

University of Glasgow, Glasgow, G12 8QQ, UK
+44 141 330 4966

{eve, stephen}@dcs.gla.ac.uk   www.tactons.org

ABSTRACT

This paper reports an experiment into the design of crossmodal icons which can provide an alternative form of output for mobile devices using audio and tactile modalities to communicate information. A complete set of crossmodal icons was created by encoding three dimensions of information in three crossmodal auditory/tactile parameters. Earcons were used for the audio and Tactons for the tactile crossmodal icons. The experiment investigated absolute identification of audio and tactile crossmodal icons when a user is trained in one modality and tested in the other (and given no training in the other modality) to see if knowledge could be transferred between modalities. We also compared performance when users were static and mobile to see any effects that mobility might have on recognition of the cues. The results showed that if participants were trained in sound with Earcons and then tested with the same messages presented via Tactons they could recognize 85% of messages when stationary and 79% when mobile. When trained with Tactons and tested with Earcons participants could accurately recognize 76.5% of messages when stationary and 78% of messages when mobile. These results suggest that participants can recognize and understand a message in a different modality very effectively. These results will aid designers of mobile displays in creating effective crossmodal cues which require minimal training for users and can provide alternative presentation modalities through which information may be presented if the context requires.

Categories and Subject Descriptors

H.5.2. [**User Interfaces**]: Haptic I/O, Auditory (non-speech feedback).

General Terms

Human Factors.

Keywords

Tactons (tactile icons), Earcons, crossmodal interaction, mobile interaction.

1. INTRODUCTION

Providing non-visual information to mobile device users is becoming an important area of research in multimodal and crossmodal interaction. We spend a great deal of our lives using our mobile devices. Whether it is in our bag or we are in a meeting, at a party, or listening to music, we still want to be able to interact with our mobile device. In these situations, visual feedback is not always appropriate. Although a user's eyes may be busy focusing on the primary task, many activities do not otherwise restrict users from attending to information using their remaining available senses. This is when multimodal interaction is of benefit so that, for instance, messages can be presented through the audio modality and alerts can be presented through the tactile.

The manufacturers of mobile devices already include audio and vibrotactile feedback in products like PDAs and mobile phones, allowing feedback to be designed for our senses of touch and hearing. Unfortunately, when the device is in a bag or pocket, tactile feedback can go unnoticed. When a user is in noisy environments like parties or listening to music, audio feedback can be ineffective. For example, Sam is on her way to a business meeting walking along a busy street with her mobile phone in her bag when she receives an important calendar reminder. As her phone is not in contact with her body, a tactile alert would probably go unnoticed so the reminder would be best presented in audio. Next, Sam gets on a train to continue her journey with her phone in her pocket. As the train leaves the station, Sam starts downloading some music for her phone. Given that the train is noisy and she has placed her phone back in her pocket so she can read the newspaper, audio alerts alone would be insufficient to inform her of her completed download. At the same time, tactile alerts would be slightly masked as the phone is in her pocket. At this time, a combination of audio and tactile feedback could let her know when her song has been downloaded. Finally, Sam arrives at her business meeting. As the boss makes a presentation, Sam receives an urgent email from her husband. Everyone in the meeting room is listening to the presentation and it would be rude for Sam to disrupt the meeting with audio feedback informing her of the incoming email. In this case, a tactile cue would be much more subtle and more socially acceptable. This scenario is an example of the need for mobile devices to provide alternative presentation modalities through which information may be presented if the context requires. As the context changes, so should the feedback modality.

As mentioned, multimodal feedback is often used to reduce the visual load on mobile device users. There has been a large body of

research into mobile multimodal interaction with each individual modality [8, 16, 17]. However, as this scenario has demonstrated, users need to be able to switch effortlessly between different modalities depending on the situation. They also need the option of several different modalities. Much of the research so far does not give the user a choice of modalities but simply provides one modality, resulting in unimodal interaction.

The approach used in this research to combat the problems mentioned above involves crossmodal audio and tactile feedback. Unlike multimodal interaction, crossmodal interaction uses the different senses to provide the same information. This is much like sensory substitution where one sensory modality is used to supply information normally gathered by another [12]. Sensory substitution systems have proven to be an effective means of communicating information to people with sensory impairments [12] so could provide an alternative method through which information can be presented to mobile device users. By employing concepts from sensory substitution, mobile devices could translate information into an auditory or tactile form so that it can be presented in the most appropriate modality to suit the context. For example, alerts providing information to the user about incoming messages (e.g. SMS, MMS, or phone call) could be crossmodally encoded in both the audio and tactile modalities. By making this information available to both the auditory and tactile senses, users can receive the information in the most suitable way, without having to abandon their primary task to look at the device.

The research presented here investigates the design of crossmodal auditory and tactile messages, called crossmodal icons [9], which are abstract icons that can be instantiated in one of two equivalent forms (auditory or tactile). These can be used in interfaces as a means on non-visual output.

2. CROSSMODAL ICONS
Crossmodal icons enable mobile devices to output the same information interchangeably via different modalities. They can be automatically instantiated as either an Earcon or Tacton, such that the resultant Earcons or Tactons are equivalent and can be compared as such [9].

The auditory cues used in crossmodal icons are Earcons which are a common type of non-speech auditory display, which Blattner *et al.* defines as "non-verbal audio messages that are used in the computer/user interface to provide information to the user about some computer object, operation or interaction" [2]. Brewster [4] has conducted detailed investigations of Earcons, which have shown that they are an effective means of communicating information in sound.

Tactons are used as the vibrotactile counterparts of Earcons in the design of crossmodal icons. These are structured vibrotactile messages which can be used to communicate information non-visually [3]. They are the tactile equivalent of Earcons and visual icons, and could be used for communication in situations where vision is overloaded, restricted or unavailable [3]. Tactons are created by manipulating the parameters of cutaneous perception to encode information. For example, Brown *et al.* [5] encoded three pieces of information using rhythm, roughness and spatial location into a Tacton to create messages for mobile telephones.

Any attribute that can specify similar information across modalities is considered to be *amodal* in nature [13]. Thus, the crossmodal parameters used in auditory and tactile icons to encode the same information are the amodal attributes available in those two senses. Auditory and tactile displays were chosen because they are ideal candidates for crossmodal combination in view of the fact that both modalities share temporal and spatial properties. The amodal attributes between our senses of hearing and touch include intensity, rate, rhythmic structure and spatial location [13].

Several dimensions of information can be represented in crossmodal icons by encoding each dimension in a different amodal parameter (e.g. a parameter available in both modalities). To develop a set of Earcons/Tactons as crossmodal icons, the information represented must be able to be encoded in both modalities. In other words, an email alert encoded using a specific melody in the audio domain could not be crossmodal as there is no tactile equivalent to melody. Whereas, an email alert encoded using a particular spatial location in audio (e.g. using the cardinal points in a 3D audio soundscape) could be used as a crossmodal alert as there is a tactile equivalent available (e.g. presenting the tactile cue with vibrotactile transducers placed in a circle on the body).

Previous research has identified rhythm, texture and spatial location as suitable crossmodal parameters or amodal attributes for use with auditory and vibrotactile cues [9]. Furthermore, it has shown that roughness can be mapped between modalities using amplitude modulation in the vibrotactile cues and differing timbres in the audio domain [9]. Spatial location can be perceived as equivalent in the audio and tactile modalities when tactile body positions around the waist are mapped to audio positions in a 3D soundscape [10] around the head.

Although there are now three possible parameters which allow easy mappings between the auditory and tactile modalities, no complete set of crossmodal icons which use a combination of these parameters has been created to test whether the concept works and whether users can transfer knowledge of messages between senses. Therefore, in this research we develop a complete set of crossmodal icons and assess learning and the extent to which this learning transfers between the two modalities before testing recognition rates during an absolute identification and absolute matching experiment for the resulting crossmodal icons.

3. DESIGN OF 3-DIMENSIONAL ICONS
In this study, crossmodal icons were created to represent alerts which might occur on a mobile phone to inform the user of incoming messages. Three pieces of information were encoded in each crossmodal icon using the parameters identified earlier: the type of message was encoded in the rhythm, the urgency of the message was encoded in the roughness, and the sender of the message was encoded in spatial location. These types of information were chosen as they are common alerts provided through the visual modality on current mobile devices and would be familiar to participants. The type of message had three possible values: text, email, or voicemail, the urgency of the message had two possible values: urgent or not urgent, and the sender of the message had three possible values: work, personal, or junk. This resulted in a set of 18 crossmodal icons. Therefore, there were 18 Earcons representing the message alerts, and 18 Tactons representing the same message alerts.

3.1 Type of Message

Three different rhythms were used to represent the three types of message: text, email, and voicemail. These rhythms have already been used successfully in tactile experiments [6]. Each rhythm was made up of a different number of beats, with the text rhythm consisting of one short beat and one long beat, the email rhythm consisting of two long beats and two short beats, and the voicemail rhythm consisting of one long beat, three short beats, and two long beats. Using a different number of beats in each rhythm helps to make the rhythms distinguishable [6]. These rhythms are presented in Figure 1 using standard musical notation.

Figure 1. Text rhythm, email rhythm, and voicemail rhythm (from [6]).

3.2 Urgency of Message

Two levels of roughness were used to represent urgent (very rough) and not urgent (smooth) messages. Brown *et al.* used amplitude modulation to create different levels of roughness [6], the ones used here were based on those: an unmodulated 250Hz sine wave (smooth) and a 250Hz sine wave modulated by 30Hz sine wave (rough). The Earcons used differing timbres as levels of roughness based on previous experiments on crossmodal parameters [9]: a piano was used for smooth whilst a vibraphone was used for rough.

3.3 Message Sender

Three locations on the user's waist were used to encode information about the sender in the tactile crossmodal icons – three vibrotactile actuators were placed on a Velcro belt on the left hand side, the front centre, and the right hand side of the waist (Figure 4). A previous study showed that these body locations can be effectively mapped to 3D audio locations in both a mobile and stationary environment, and that the waist was the most effective location [10]. The audio crossmodal icons used three locations in a 3D audio soundscape to encode the information about the sender of the message – sounds were placed on a horizontal plane around the users head. A vibration or sound to the left hand side indicated that the message was from 'work', the centre indicated that the message was 'personal', and the right hand side represented 'junk' (Figure 2).

As an example, an urgent email from work in a tactile form would be the email rhythm with a rough texture to the left hand side of the user's waist, and the audio version would present the email rhythm played by a vibraphone to the left hand side of the 3D audio soundscape.

Figure 2. 'Junk' message indicated by audio panned to the right (Earcon) and tactile pulse on the right of the waist.

4. EXPERIMENT 1 – LAB BASED STUDY OF CROSSMODAL DISPLAYS

An experiment was conducted to investigate absolute identification of crossmodal icons encoding three dimensions of information to see if users would be able to use them and transfer knowledge of messages learned in one modality to the other. Half of the participants were trained and tested in different modalities: One quarter of the participants was trained to identify the crossmodal Earcons and then tested with crossmodal Tactons, another quarter was trained with Tactons and tested with Earcons. As a control, the other half of the participants were trained and tested in the same modality (Table 1). Data were recorded on the identification of the three parameters – type, urgency, and sender. In addition, participants were informally interviewed about their experiences after the experiment.

Participant Group	Training	Testing
1	Audio	Tactile
2	Tactile	Audio
3	Audio	Audio
4	Tactile	Tactile

Table 1. Experiment conditions

4.1 Aim and Hypotheses

The aim of this experiment was to investigate whether, if users are trained to understand alerts in one modality, they can then identify them in the other. The hypotheses were as follows:

1. If trained to identify the information encoded in audio crossmodal icons, participants will be able to identify the same information in the corresponding tactile crossmodal icons.

2. If trained to identify the information encoded in tactile crossmodal icons, the participant will be able to identify the same information in the corresponding audio crossmodal icons.

3. The rate of identification in the crossmodal training will be the same as that for participants trained and tested in the same modality.

4.2 Experiment Set Up

The C2 Tactor from EAI (Figure 3) is a small wearable linear vibrotactile actuator, which was designed specifically to provide a lightweight equivalent to large laboratory-based linear actuators [15]. The contactor in the C2 is the moving mass itself, which is mounted above the housing and pre-loaded against the skin. This

helps to provide localized feedback as only the contact point vibrates instead of the whole surrounding. The C2 is resonant at 250Hz but is also designed to be able to produce a wide range of frequencies unlike many current mobile phone actuators which have limited frequency ranges [15].

Figure 3. A C2 Tactor from Engineering Acoustics Inc.

When being tested or trained in the tactile modality, three C2 EAI Tactors (www.eai.com) were attached to the participant's waist using a belt lined with Velcro (Figure 4). The participant also wore headphones to eliminate any inadvertent audio feedback from the actuators. Tactile sensitivity can vary across the waist therefore the vibrations could feel very different in intensity at different points on the waist [7]. To counteract this, each participant was asked to set the levels of the transducers so that they all felt of equivalent intensity at the start of the experiment.

When being tested or trained in the audio modality, the participants again wore headphones attached to a soundcard on a PC through which the audio alerts were played. The audio cues used in this experiment were created using the AM:3D (www.am3d.org) audio engine and were placed on a plane around the user's head at the height of the ears to avoid problems related to elevation perception. The sounds were located in front of the nose (0°) and ±90° to the left and right at each ear. Participants were asked to set the volume levels of the audio to a comfortable level at the start of the experiment.
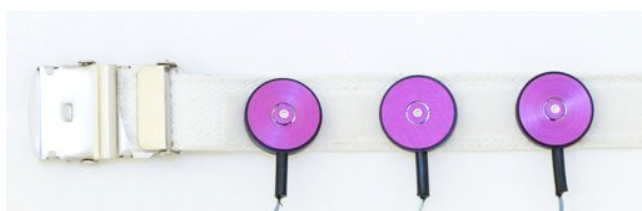

Figure 4. Belt lined with velcro used in experiment with 3 C2 Tactors attached.

4.3 Methodology

Sixteen people took part in the experiment, aged between 22 – 38 years, 9 female and 7 male, all members of staff or students at the University. The experimental method used a between groups design where each participant was trained in either audio or tactile and tested in either audio or tactile (see Table 1). At the beginning of the session participants were presented with a tutorial to introduce them to the concept of crossmodal icons, roughness, rhythm, etc., they were then allowed to experiment with either the crossmodal Earcons or Tactons (depending on the group to which they belonged). After familiarizing themselves with either the Earcons or Tactons, the participants began training using a custom training/testing application we developed (Figure 5).

The application is a purpose built experimentation system that can present audio and tactile cues of different types in multiple locations. The system presents the participant with either a tactile or audio cue at the beginning of each task. Then the participant can press the 'replay' button to have the cue presented again. Once participants have identified the information in the cue, they can select the corresponding button and submit their answer (using the tick button). After submitting the answer, a button appears which the participants press when they are ready to move on to the next task. The system records the participant's responses, the time taken to respond, and also the number of times a cue was replayed. Participants were allowed to play each cue up to 4 times per task. Replaying the cues was allowed because the expected usage of these icons is in mobile devices where standard cues such as ringtones for incoming calls are commonly presented several times.
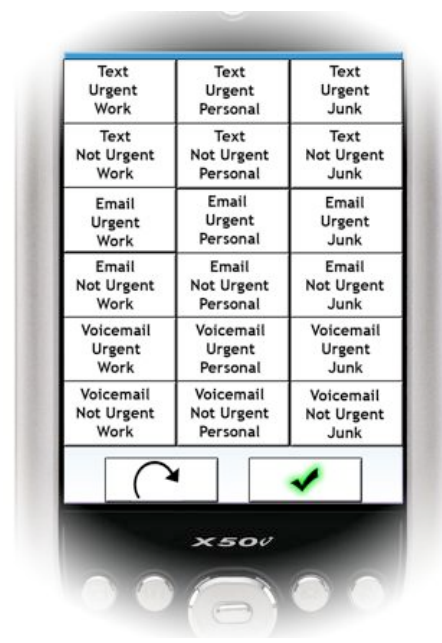

Figure 5. Screenshot of training and testing application.

4.4 Training

For training and testing, the standard Absolute Identification (AI) paradigm with trial-by-trial correct-answer feedback was employed. The AI paradigm used involves a set of k stimuli, a set of k responses, and a one-to-one mapping between the stimuli and responses. The stimuli are presented one at a time in random order and the subject is instructed to respond to each stimulus presentation with the response defined by the one-to-one mapping, i.e., to identify which of the k stimuli was presented.

Originally, for the purposes of this experiment, training was used purely to ensure that all participants reached an appropriate level of understanding. However, we also became interested in how long it would take the participants to learn the sets of Earcons and Tactons as there is little data on how long it takes to learn such cues and if the learning required by the modalities is different. This would also allow us to compare the results of crossmodal

training to training within the same modality to see if there were differences. The set of stimuli used to train the participants was identical to the set on which they would be later tested except that during the training phase the stimuli were presented in the same modality. The application shown in Figure 6 was used to record participants' answers.

The participants had to identify the information in the cue they heard or felt and then choose the appropriate button on the display shown in Figure 6. Each stimulus alternative was applied twice during each training run, resulting in a total of 36 tasks per run. During training the participants were required to repeat experimental runs (in audio or tactile) until a run with >= 90% correct identification was achieved so that we could measure how long it took for them to reach a good level of performance. If a participant did not reach 90% at the end of a training run, he/she received further training before being given another training run.

4.4.1 Training Results

During the training and the experiment itself data were collected on the number of correct responses to the complete crossmodal icons. The learning curves for each participant and each stimulus set during training are shown in Figures 6 and 7. The amount of time to reach the performance criterion varied across participants. These results show that, on average, it takes 2 training sessions for participants to be able to identify Earcons with recognition rates of 90% or higher. They also show that, on average, it takes 3 training sessions for participants to identify Tactons with recognition rates of 90% or above.
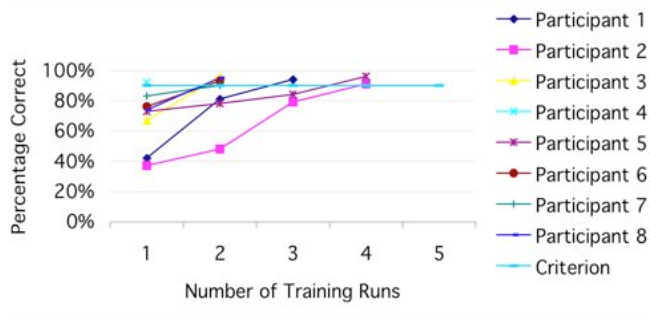

Figure 6. Learning curve for audio training.

There have been no other such studies into the training and learning of Earcons and Tactons. These results are promising for using audio and tactile interchangeably and would seem to indicate that there is no significant difference in the time taken to learn these crossmodal cues in either modality. Further studies will look at the effectiveness of explicit versus implicit learning in crossmodal interaction to reduce the amount of training time needed.
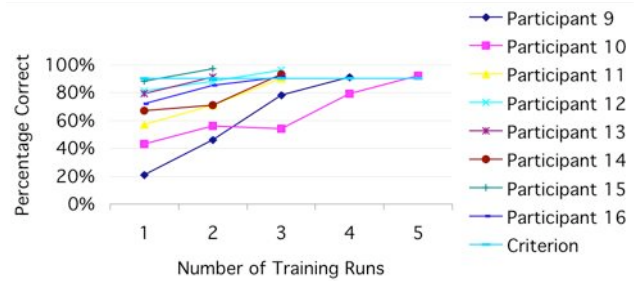

Figure 7. Learning curve for tactile training.

4.5 Testing in Alternative Modality

Once the participants in Groups 1 and 2 in Table 1 had achieved the correct level of training, they completed the absolute identification test using the same online system and tasks but with cues presented in the other modality. Participants in the control groups (Groups 3 and 4) continued through the absolute identification test using the same tasks in the same modality after training.

In total there were 36 tasks in the experiment, with all 18 crossmodal icons (either audio or tactile) presented twice during the experiment. The order in which the crossmodal icons were presented was random for each participant. In each task the participant was presented with a crossmodal icon which he/she could replay up to 4 times. The participants had to identify the corresponding alert and then select the corresponding button in the dialogue box (Figure 5).

4.6 Results

The results from the control group in comparison to the crossmodal testing group are shown in Figure 8.
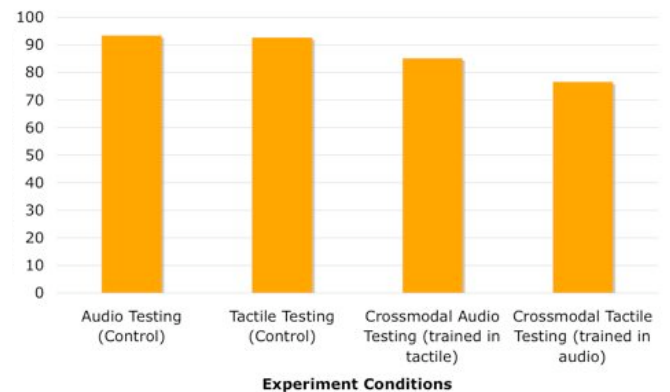

Figure 8. Average percentage correct responses during testing.

The results for overall Earcon recognition when trained with Tactons showed an average recognition rate of 85.1%. The alert 'personal urgent text' achieved the highest recognition rate of 94% while the alert 'work not urgent voicemail' resulted in the lowest recognition rate of 61%. The results for overall Tacton recognition when trained with Earcons showed an average recognition rate of 76.5%. The alert 'personal not urgent text' achieved the highest recognition rate of 83% and once again the alert 'work not urgent

voicemail' resulted in the lowest recognition rate of 56%. Thus hypothesis 2 can be accepted.

Having examined the data in depth, there does not seem to be any clear reason for the low scores produced by the 'work not urgent voicemail' cue. All of the individual parameters performed well in general (Figure 12) and there was no apparent misunderstanding by the participants. Further analysis will be done in the future to investigate this to ensure that it is an anomaly and not an issue with the design of the cues.

An ANOVA showed that there was no significant difference in the recognition rates between the results of the four different Groups (training in audio / tested in tactile, training in tactile / tested in audio, training and testing in tactile, training and testing in audio) with $F_{(3,60)} = 2.1$, $p=0.1$). With the standard deviations in each condition varying only slightly from 8.9 to 9.9 and the mean scores very close, the analysis suggests that information learnt in one modality can be recovered in the alternative modality in a way which is comparable with recognition of the same information in the trained modality. Thus hypothesis 1 can be accepted

The results suggest that if a user is taught to understand alerts provided by crossmodal Tactons, they could be expected to understand crossmodal Earcons with no audio training with approximately 85 % accuracy and if a user is taught to understand alerts provided by crossmodal Earcons, they could be expected to understand crossmodal Tactons with no tactile training with approximately 76.5% accuracy. These results are comparable to previous research in 3-dimensional Earcons where McGookin's results [14] showed recognition rates of around 70% for identification of complete 3-dimensional messages in audio. They are also comparable with previous Tactons research which produced recognition rates of 81% for identification of complete 3-dimensional messages in tactile icons [5].

5. EXPERIMENT 2 – MOBILE STUDY OF CROSSMODAL DISPLAYS

As discussed at the start of the paper, crossmodal icons are being developed for users of mobile devices. Such users are often in motion when they use their devices so any alerts provided by the mobile device must be designed to be discernible in these situations too and not just when the user is stationary. There are many ways in which motion could affect perception of crossmodal output: mobile environments tend to change frequently with light, volume and vibration levels changing often. Consequently, another experiment in crossmodal identification was conducted which investigated the effects of motion on the results and assessed whether the good results observed in the laboratory would carry over to a more real world situation. The overall experiment involved 16 new participants who were either trained in audio or in tactile and then tested in audio or tactile whilst walking. Both the methodology and the crossmodal icons used in the experiment were the same as before to allow result comparisons.

The setup of this experiment was identical to the stationary one above in every respect except that participants were asked to walk on a treadmill during the experiment as opposed to sitting in a chair (Figure 9).
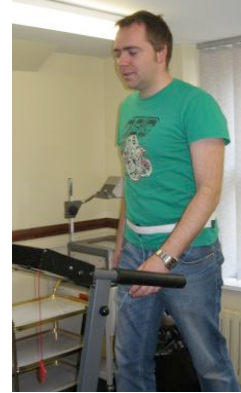

Figure 9. Mobile condition experimental set up.

This mobile experiment used a treadmill set up in a usability lab to simulate mobility because the actuators used to present the tactile cues were controlled from a PC and therefore we could not test in a real mobile environment. Studies show that using treadmills to simulate motion is good for mimicking workload [11] when performance measures are of key interest and is a more controllable environment [1]. Furthermore, using a treadmill permitted us to set a standard walking speed for all participants (in this case, all participants walked at a constant speed of 5km/hr during the experiment).

The hypothesis in this experiment was:

4. Being mobile will increase errors produced during crossmodal icon identification and matching between modalities as compared to being stationary.

5.1 Results

The average number of errors for audio and tactile identification is shown in Figure 10. As before, the average recognition rate for both the audio and tactile groups was calculated but this time for the mobile condition as well.
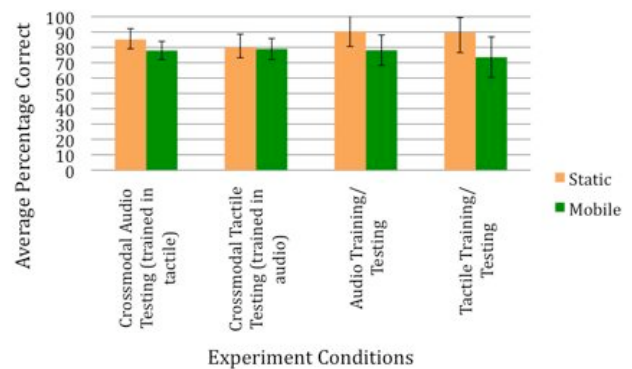

Figure 10. Average correct responses in stationary and mobile conditions.

The results for overall Earcon recognition when mobile and trained with Tactons showed an average recognition rate of 78%. The results for overall Tacton recognition when mobile and when trained with Earcons showed an average recognition rate of 79%. To establish whether there is a significant difference between the

mobile and stationary results, a 2 factor ANOVA was applied using the two training conditions (audio or tactile) and stationary/mobile as the two factors. The ANOVA showed that there was no significant difference in the recognition rates between the results from the mobile and stationary conditions ($F(1,15) = 3.4$, $p > 0.01$) or when trained in audio or tactile with ($F(1,30) = 0.7$, $p>0.01$).

These results show that training with crossmodal Tactons achieves slightly better instant recognition in crossmodal Earcons when mobile than *vice versa* but there are no actual significant differences. Therefore if a user is taught to understand alerts provided by crossmodal Tactons, they could be expected to understand crossmodal Earcons with no training when mobile with about 78 % accuracy and if a user is taught to understand alerts provided by crossmodal Earcons, they could be expected to understand crossmodal Tactons with no training when mobile with approximately 79% accuracy.

5.1.1 Individual Parameter Results and Discussion
To establish the performance of each of the crossmodal parameters used, further analysis was performed on the data produced by both the audio and tactile versions of each parameter. The average percentage of correct responses in each audio parameter and each tactile parameter are shown in Figures 11 and 12.
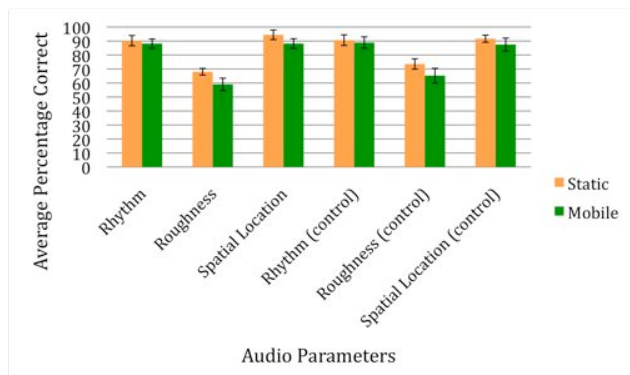

Figure 11. Average percentage of correct responses in each audio condition.
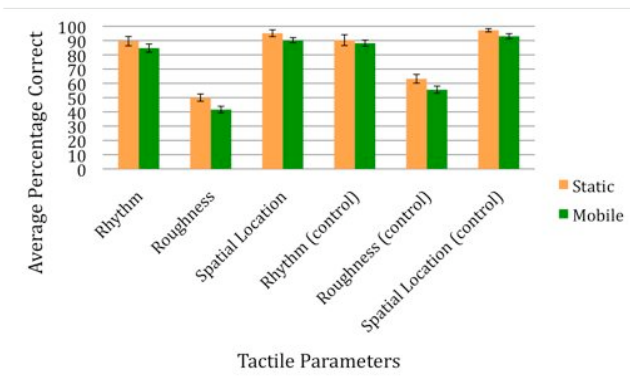

Figure 12. Average percentage of correct responses in each tactile condition.

An ANOVA showed no significant differences between audio rhythm and tactile rhythm (stationary or mobile) or between audio spatial location and tactile spatial location (stationary or mobile), however there was a significant difference between audio roughness ($F(5,18) = 4.01$, $p = 0.09$) and tactile roughness ($F(5,18) = 6.76$ $p = 0.04$) with both producing significantly poorer results than the other parameters in stationary and mobile environments.

These results suggest two different issues: firstly, overall the crossmodal roughness parameter is not as effective as rhythm and spatial location indicating that a different parameter may need to be used; secondly, when trained to identify roughness in one modality, participants struggle to then identify it in another modality. Although results in the stationary and mobile experiments show no significant difference in performance with crossmodal icons using rhythm and spatial location, audio and tactile roughness recognition rates are significantly lower when mobile.

The mobile results are comparable to the results of the stationary conditions. Overall, the results indicate that, if a user is trained in one modality, the accuracy achieved when they are asked to identify the same information in the other modality is comparable even when they are placed in a mobile situation.

These results indicate that crossmodal icons could be effective in mobile displays. Although the mobile environment used in this experiment was much more controlled than a real world environment, these results give an indication of the sorts of effects that may be seen when a user is in motion. Future experiments will be conducted in real world situations such as walking and traveling on a train or bus for example.

6. CONCLUSIONS
This paper presented an experiment which investigated the crossmodal transfer of information between the auditory and tactile modalities. Previous research had investigated identification of information in Tactons [5] and Earcons [4] showing that both could effectively encode information in three dimensions. Previous research had also established that information encoded in single parameters in the auditory and tactile modalities can be perceived as equivalent [9, 10] if the appropriate crossmodal parameters are used (rhythm, spatial location – locations in a 3D audio soundscape matched with body locations around the waist, and texture – roughness levels created with amplitude modulation and differing audio timbres e.g. smooth piano, rough vibraphone. This research investigated whether, if trained to understand multidimensional audio alerts, a user can then also understand the corresponding tactile alerts with no additional training and *vice versa*. Our results suggest that this is possible. The experiments described here are the first studies to be conducted investigating training and the transfer of training to other modalities in multimodal interaction. The overall findings from the experiment can be summarized as follows:

- Users in a stationary environment can accurately recognize 85% of messages presented by Earcons, if they have been trained to recognize the same alerts presented by Tactons.

- Users in a mobile environment can accurately recognize 78% of messages presented by Earcons, if they have been trained to recognize the same alerts presented by Tactons.

- Users in a stationary environment can accurately recognize 76.5% of messages presented by Tactons, if they have been trained to recognize the same alerts presented by Earcons.

- Users in a mobile environment can accurately recognize 79% of messages presented by Tactons, if they have been trained to recognize the same alerts presented by Earcons.

The results of this research indicate that it may not be necessary to train users to understand icons in all the modalities a system might use. If crossmodal icons are used to present information, training is only required in one modality as results show that users will then be able to understand the same messages in the other modality. Using crossmodal icons to communicate information to mobile device users could therefore reduce the learning time for the user and also increase the number of modalities through which this information may be transmitted.

The crossmodal icons described in this paper were designed for a mobile phone notification application. Based on the positive results gained so far, there are many other potential applications that could benefit from the inclusion of crossmodal icons such as context aware navigation applications for the visually impaired and outdoor mobile games where it is dangerous for participants to concentrate visually on the device instead of their environment. Also, as mentioned earlier, mobile devices often have cluttered displays due to the lack of screen space. Crossmodal features could be added to buttons, scrollbars, and menus, etc. on touchscreen mobile devices so that information about those widgets can be presented non-visually. This would allow the widget size to be reduced (or even removed from the screen) and allow more information to be presented on the display.

Different tactile spatial locations do not necessarily have to be on the body but could be on the actual device. For instance, localized tactile feedback on touchscreen mobile devices can provide spatial information. Furthermore, audio spatial location using 3D audio without wearing headphones is now becoming available as mobile device manufacturers are beginning to incorporate stereo audio output.

Mobile technology incorporating audio and tactile output has now become widely available and our research has shown that feedback can be created which exploits users' abilities to transfer knowledge from one modality to another. By taking this into account and designing mobile applications with adaptive crossmodal icons, users will have the ability to interact with their devices even when their situation and surroundings are changing.

7. ACKNOWLEDGEMENTS

REFERENCES

[1] Barnard, L., Yi, J. S., Jacko, J. A. and Sears, A., An Empirical Comparison of Use-in-Motion Evaluation Scenarios for Mobile Computing Devices, in *International Journal of Human-Computer Studies* **62** (2005), 487 - 520.

[2] Blattner, M. M., Sumikawa, D. A. and Greenberg, R. M., Earcons and Icons: Their Structure and Common Design Principles, in *Human Computer Interaction* **4(1)** (1989), 11 - 44.

[3] Brewster, S. A. and Brown, L. M., Tactons: Structured Tactile Messages for Non-Visual Information Display, in *Proc AUI Conference 2004*, ACS (2004), 15 - 23.

[4] Brewster, S. A., Wright, P. C. and Edwards, A. D. N., An evaluation of Earcons for use in auditory human-computer interfaces, in *Proc InterCHI'93*, Amsterdam, ACM Press (1993), 222 - 227.

[5] Brown, L. M. and Brewster, S. A., Multidimensional Tactons for Non-Visual Information Display in Mobile Devices, in *Proc MobileHCI 2006*, ACM Press (2006), 231 - 238.

[6] Brown, L. M., Brewster, S. A. and Purchase, H. C., A First Investigation into the Effectiveness of Tactons, in *Proc WorldHaptics 2005*, IEEE (2005), 167 - 176.

[7] Cholewiak, R. W. and Craig, J. C., Vibrotactile Pattern Recognition and Discrimination at Several Body Sites, in *Perception and Psychophysics* **35** (1984), 503 - 514.

[8] Erp, J. B., Tactile Navigation Display, in *First international Workshop on Haptic Human - Computer Interaction, Lecture Notes in Computer Science* **2058** (2001), 165 - 173.

[9] Hoggan, E. and Brewster, S. A., Crossmodal Icons for Information Display, in *Proc ACM CHI '06 Extended Abstracts*, ACM Press (2006), 857 - 862.

[10] Hoggan, E. and Brewster, S. A., Crossmodal Spatial Location: Initial Experiments, in *Proc NordiCHI '06*, Norway, ACM Press (2006), 469 - 472.

[11] Kjeldskov, J. and Stage, J., New Techniques for Usability Evaluation of Mobile Systems in *International Journal of Human-Computer Studies* **60** (2004), 599 - 620.

[12] Lenay, C., Canu, S. and Villon, P., Technology and Perception: The Contribution of Sensory Substitution Systems, in *Proc ICCT*, IEEE (1997),

[13] Lewkowicz, D. J., The Development of Intersensory Temporal Perception: An Epigenetic Systems/Limitations View, in *Psychological Bulletin* **126** (2000), 281 - 308.

[14] McGookin, D. and Brewster, S. A., Understanding Concurrent Earcons: Applying Auditory Scene Analysis Principles to Concurrent Earcon Recognition, in *ACM Transactions on Applied Perception* **1(2)** (2004), 130 - 155.

[15] Mortimer, B., Zets, G. and Cholewiak, R. W., Vibrotactile Transduction, *submitted to Journal of the Acoustic Society of America* (2006).

[16] Sawhney, N. and Schmandt, C., Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments, in *Proc ACM Transactions on Computer Human Interaction*, ACM Press (2000), 353 - 383.

[17] Tan, H. Z. and Pentland, A., Tactual Displays for Wearable Computing, in *Proc the 1st IEEE International Symposium on Wearable Computers*, IEEE (1997), 225 - 230.