Using Non-Speech Sound to Overcome Information Overload

Stephen A. Brewster

Glasgow Interactive Systems Group
Department of Computing Science
The University of Glasgow
Glasgow, G12 8QQ, UK
Tel: +44 (0)141 330 4966
Fax: +44 (0)141 330 4913
stephen@dcs.gla.ac.uk
http://www.dcs.gla.ac.uk/~stephen/

With ever increasing amounts of visual information to take in when interacting with computers, users can become overloaded. One reason is that computers communicate solely by graphical output. This paper suggests the use of non-speech sound output to enhance the graphical display of information can overcome overload. The question is how to integrate the display of sound and graphics to capitalise on the advantages each offer. The approach described here is to integrate sound into the basic components of the human-computer interface. Two experiments are described where non-speech sounds were added to buttons and scrollbars. Results showed sound improved usability by increasing performance and reducing time to recover from errors. Subjective workload measures also showed a significant reduction. Results from this work show that the integrated display of graphical and auditory information can overcome information overload.

Keywords: Sonically-enhanced widgets, auditory interfaces, sonification, buttons, scrollbars.

INTRODUCTION

With ever increasing amounts of visual information to take in when interacting with computers, users can become overloaded. What causes this problem? In our everyday lives we are able to deal with an enormous amount of complex information of many different types without difficulty. One reason for the problem is that computers communicate solely by graphical output, putting a heavy burden on our visual sense which may become overloaded. In the real world we have five senses and the combination of these avoids any one sense becoming overloaded. The next step forward in display design is to allow the use of these other senses when interacting with a computer. Such *multimodal* interfaces would allow a greater and more natural communication between the computer and the user. They also allow the user to

1

employ appropriate sensory modalities to solve a problem, rather than just using one modality (usually vision) to solve all problems.

This paper suggests the use of non-speech sound output to enhance the graphical display of information at the human-computer interface. There is a growing body of research which indicates that the addition of non-speech sounds to human-computer interfaces can improve performance and increase usability [4, 6, 15]. Our visual and auditory senses work well together: The visual sense gives us detailed data about a small area of focus whereas the auditory provides data from all around. Users can be informed of important events even if they are not looking at the right position on the display (or even not looking at the display at all). This is particularly important for large-screen, high-resolution, multiple monitor interfaces. The question is how to integrate the display of sound and graphics to capitalise on the advantages each offer.

The motivation for this research is that users' eyes cannot do everything. As mentioned, the visual sense has a small area of high acuity. In highly complex graphical displays users must concentrate on one part of the display to perceive the graphical feedback, so that feedback from another part may be missed. This becomes very important in situations where users must notice and deal with large amounts of dynamic data. For example, imagine you are working on your computer writing a report and are monitoring several on-going tasks such as a compilation, a print job and downloading files over the Internet. The word-processing task will take up all of your visual attention because you must concentrate on what you are writing. In order to check when your printout is done, the compilation has finished or the files have downloaded you must move your visual attention away from the report and look at these other tasks. This causes the interface to intrude into the task you are trying to perform. It is suggested here that some information should be presented in sound. This would allow you to continue looking at the report but to hear information on the other tasks that would otherwise not be seen (or would not be seen unless you moved your visual attention away from the area of interest, so interrupting the task you are trying to perform). Sound and graphics will be used together to exploit the advantages of each. In the above example, you could be looking at the report you are typing but hear progress information on the other tasks in sound. To find out how the file download was progressing you could just listen to the download sound without moving your visual attention from the writing task.

Current interfaces depend heavily on graphical output. One reason for this is that when current

interaction techniques (such as buttons, scrollbars, etc.) were developed, visual output was the only communication medium available. However, technology has progressed and now almost all computer manufacturers include sophisticated sound hardware in their systems. This hardware is unused in daily interactions with machines (the sounds are really only used to any extent in computer games). This research will take advantage of this available hardware and make it a central part of users' everyday interactions to improve usability.

Even though sound has benefits to offer it is not clear how best to use it in combination with graphical output. The use of sound in computer displays is still in its infancy, there is little research to show the best ways of combining these different media. This means sounds are sometimes added in *ad hoc* and ineffective ways by individual designers [1, 21]. The approach described here is to integrate sound in a structured way into the basic components of the interface to improve the display of information from the bottom up. This paper describes two experiments where non-speech sounds were added to buttons and scrollbars to correct usability errors that are due to the system requiring the user to look at more than one place at a time.

SOUNDS USED

The non-speech sounds used for this investigation were based around structured audio messages called *Earcons* [5, 6, 24]. Earcons are abstract, synthetic tones that can be used in structured combinations to create sound messages to represent parts of an interface. Detailed investigations of earcons by Brewster, Wright & Edwards [9] showed that they are an effective means of communicating information in sound. The sounds were designed using earcon construction guidelines proposed by Brewster *et al.* [11].

Earcons are constructed from motives. These are short rhythmic sequences that can be combined in different ways. The simplest method of combination is concatenation to produce *compound earcons*. By using more complex manipulations of the parameters of sound *hierarchical earcons* can be created [5] which allow the representation of hierarchical structures.

All the sounds used in the experiments were played on a Roland D110 multi-timbral sound synthesiser. The sounds were controlled by an Apple Macintosh via MIDI through a Yamaha DMP 11 digital mixer and presented to participants by loudspeakers. A web

demo of all of the earcons described in the paper will be provided.

TESTING FRAMEWORK

In order to test the sonically-enhanced widgets an experimental testing framework was created. This allowed the testing of the widgets in a simple and consistent manner. The same types of measures and designs would be used for each.

A two-condition, within-subjects design was used to test both of the widgets. In one of the conditions the standard graphical widget was tested, in the other condition the sonically-enhanced widget. The order of presentation was counterbalanced to evenly distribute learning effects from Condition 1 to Condition 2. Table 1 shows the format of the experiment (progressing from left to right). After the test of each condition participants were presented with workload charts which they had to fill-in (this is described in detail below). Instructions were read from prepared scripts.

Participants	Condition 1		Condition 2	
Six Participants Æ	Sonically-enhanced Widget Train & Test	Workload Test	Visual Widget Train & Test	Workload Test
Six Participants Æ	Visual Widget Train & Test		Sonically-enhanced Widget Train & Test	

Table 1: Format of the experiments.

Measures

In order to get a full range of quantitative and qualitative results time, error rates and subjective workload measures (see below) were used as part of the framework. Time and error rate reductions would show quantitative improvements and workload differences would show qualitative differences. This gives a balanced view of the usability of a system [3].

Hart and Wickens ([17], p 258) define workload "...as the effort invested by the human operator into task performance; workload arises from the interaction between a particular and task and the performer". The NASA Human Performance Research Group [20] analysed workload into six different factors: Mental demand, physical demand, time pressure, effort expended, performance level achieved and frustration experienced. The NASA-Task Load Index (TLX) [16] is a set of six rating scales and was used for estimating these subjective workload factors in the experiments described here.

The basic six factors were used as described but a seventh factor was added: Annoyance. This is often cited as a reason for not using sound for display as it is argued that continued presentation of sound would be an annoyance for the user. So, by adding this as a specific factor in the usability assessment it would be possible to find out if participants felt that sonic

feedback was an annoyance. Participants were also asked to indicate overall preference: Which of the two interfaces they felt made the task easiest. Participants had to fill in paper-based workload charts for all of the eight factors after both conditions of the experiment.

Participants

For each of the tests twelve participants were used. They were postgraduate students from the Department of Computer Science at the University of York, UK. All had more than three years experience of graphical interfaces and buttons. Expert participants were used because the type of errors studied here are action slips (which are typical of experts working fast, see below).

SONICALLY-ENHANCED BUTTONS

One of the most fundamental widgets in all graphical human-computer interfaces is the graphical button (to avoid confusion, *graphical button* will here be used to refer to the button on the computer display and *mouse button* will be used to refer to the button on the mouse). Although these are very common they are not without problems [8, 12, 13]. One of the main difficulties is that the user may think the graphical button has been pressed when it has not. This can happen because the user moves off the graphical button before the mouse button is released. This is caused by a problem with the feedback from the
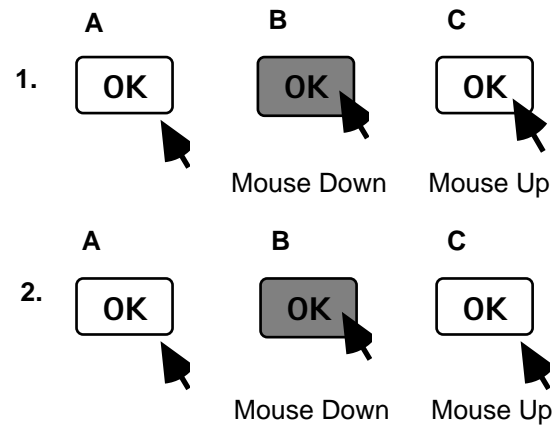

Figure 1: Feedback from pressing and releasing a graphical button. (1) shows a correct button selection, (2) shows a slip-off.

graphical button (see Figure 1). Both correct and incorrect presses start in the same way (1A and 2A). In the correct case, the user presses the graphical button and it becomes highlighted (1B), the mouse button is then released with the mouse still over the graphical button, it becomes un-highlighted (1C) and the operation requested takes place. The button slip-off starts in the same way. The user presses the mouse button over the graphical button (2B), then moves (or slips) off the graphical button and releases the mouse button (2C), the graphical button becomes un-highlighted (as before) but no action takes place. The feedback from these two different situations is *identical*. This problem occurs infrequently but, as the error may not be noticed for a considerable time, the effects can be serious. With a one-step undo facility users must notice before the next action takes place otherwise they may not easily be able to correct the mistake.

The identical feedback would not be a problem if the user was looking at the graphical button to see the slip-off, but this is not the case [13]. Dix & Brewster suggest there are three conditions necessary for such slip-off errors to occur:

i) The user reaches *closure* after the mouse button is depressed and the graphical button has highlighted.

ii) The visual focus of the next action is at some distance from the graphical button.

iii) The cursor is required at the new focus.

Closure occurs when a user perceives a task as being completed, which in this case is when the graphical button is highlighted (the mouse button is down). In reality, the task does not end until the mouse button is released. Because closure (i) is reached before this, the user starts the next task (mouse movement, iii) in parallel with the mouse up and a slip-off occurs. The user's attention is no longer at the graphical button (ii) so the feedback indicating a slip-off is not noticed. The problem occurs with expert users who perform many operations automatically and do not explicitly monitor the feedback from each interaction. This type of error is an *action slip* [22], as Lee [18] describes (p 73): "…as a skill develops, performance shifts from 'closed loop' control to 'open-loop' control, or from monitored mode to an automatic, unmonitored mode of processing." As

users become familiar with a task they no longer monitor the feedback so closely but continue to look at the information in which they are interested. If they must look at the widget then it forces the interface to intrude upon the task they are trying to perform.

These problems occur in graphical buttons that allow a 'back-out' option: Where the user can move off the graphical button to stop the action. If the action is invoked when the mouse button is pressed down on the graphical button (instead of when it is released) then these problems do not occur as the user cannot slip off. These buttons are less common because they are more dangerous as users cannot change their minds.

In this situation sound has potential because the user's eyes are occupied. Moving the mouse to the location of the next action requires visual attention so that the mouse can be positioned correctly. Therefore, the user cannot look at the button to see feedback indicating a slip-off. It would be very difficult to correct this problem using visual feedback. Buttons could be changed so that they indicated a difference between a successful and unsuccessful click. For example, the button could flash in a different way. This would not work because users will not be looking at the button but at the location of the next action. The area of visual focus is too small to allow

them to see the feedback. Perhaps feedback could be given at the mouse location but again we cannot be sure that the user will be looking there either. Sound would allow us to present the information to the user without knowing where he/she was looking.

Experimental hypotheses

The workload felt by participants should be reduced as the extra feedback would provide information that the participants needed. Participants should have to expend less effort recovering from errors. There should be no increased frustration or annoyance due to the addition of sound as the auditory feedback will provide information that the participants need.

The extra feedback provided by the earcons should make it easier for participants to recover from errors. They will notice that the errors have occurred more quickly than in the visual condition. This should result in faster error recovery times in the auditory condition. More data (codes) should be typed in the time available due to less time being spent on error recovery.

Experimental task

Figure 2 shows the interface to the task. Participants were required to enter five digit codes via the on-screen keypad. The task was designed to be simple so that the participants could easily learn it and reach a level of automaticity in the task where slip-off errors would occur [6]. The codes were displayed in the 'Code to type' field. The numbers entered appeared in the 'Code' field above the keypad. When the code had been typed and the 'OK' button pressed, the 'Next' button was used to display the next code. This maximised the number of button presses and mouse movements that the participants had to make. In the visual condition the buttons acted like normal Macintosh buttons. In the auditory condition the buttons made the sounds described below.

Experimental design and procedure

The design of the experiment was as described above. Each condition lasted 15 minutes and the participants had to type in as many codes as possible in that time.
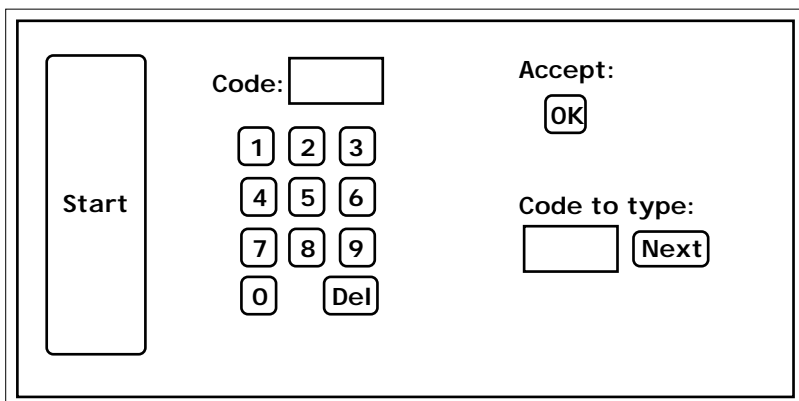
Total time taken, time taken to recover from errors and the total number of codes typed were recorded, as well as workload data.

The design of the sonically-enhanced buttons

Three earcons were needed to overcome the usability problems

Figure 2: The button testing program (reduced in size).

discussed above: One to indicate to the user when the mouse was over a graphical button; one to be the auditory equivalent of the graphical highlight when the mouse button was pressed down on the graphical button; the other to indicate when a button was pressed correctly or when a slip-off occurred.

A base earcon was created for when the mouse was moved over a screen button. This was a continuous tone at pitch C_4 (130Hz)[1]. The volume of this was kept to just above the threshold level. This sound was played for as long as the mouse was over a graphical button. When the mouse button was pressed down over a graphical button a continuous sound at pitch C_3 (261Hz) was played. This continued for as long as the mouse button was down and the mouse was over the graphical button. If the mouse was moved off the graphical button the sound stopped. If the mouse was released over the graphical button then a success sound was played. This consisted of two notes, played consecutively, at C_1 (1046Hz) each with a duration of 40 msec. This success sound had to be kept short so that users would not get confused as to which button the feedback was coming from. The mouse button down and success sounds differentiated a successful and unsuccessful mouse click.

---

[1] Classification system described by Scholes [23].

The earcons used a combination of pitch, duration and intensity to get the listener's attention. This meant that a lower level of intensity could be used, making the sounds less annoying for the primary user and others working nearby [2]. It is important to note that intensity is not the only way to get the user's attention. As Edworthy *et al.* [14] showed, attention-grabbing sounds can be created by varying other sound parameters, such as those used here.

Workload results

Figure 3 shows the average score for each of the workload categories. They were all scored in the range 0-20 on rating scales. Paired T-tests were carried out on the auditory versus visual conditions for each of the workload categories. An analysis of the individual scores showed that none were significantly different between conditions. However, the sonically-enhanced buttons were given a significantly higher overall preference rating (T(11)=5.14, p=0.0003). This strongly significant result indicates that the participants found the task easier with the sonically-enhanced buttons but this did not affect the workload required for the task.

There was no significant difference in terms of annoyance. Four participants rated the auditory condition more annoying than the visual but five participants rated the visual more annoying than the
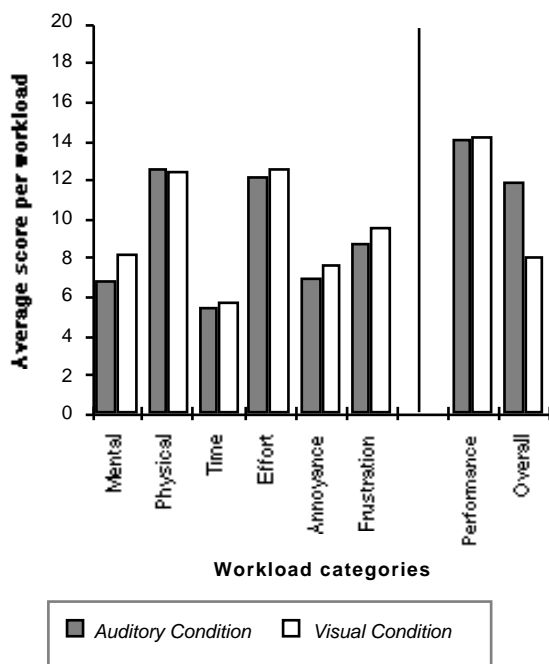
Figure 3: Average TLX workload scores for the auditory and visual conditions of the buttons experiment. In the first six categories higher scores mean higher workload. The final two categories, performance and overall preference, are separated because higher scores mean less workload.

auditory. This indicated that the participants did not find the sound feedback annoying to use.

Timing and error results

Figure 4 shows the results of error recovery. The time to recover from each slip-off was calculated. This was taken as the time from when the slip-off occurred until the user pressed the mouse button down on the correct graphical button again. It was found that participants in the auditory condition recovered from slip-off errors significantly faster than in the visual condition (T(12)=3.51, p=0.004). Average recovery times ranged from 2.00 seconds in the auditory condition to 4.21 seconds in the visual condition. The number of mouse button downs and button ups taken to recover from slip off errors was also significantly

reduced in the auditory condition (T(12)=4.40, p=0.0008). The average number of clicks to recovery was 1.5 in the auditory condition and 5.89 in the visual. In the auditory condition the participants recognised an error had occurred and often fixed it by the next mouse button down. In the visual condition it took nearly six button ups and downs before the participants recovered from an error. These results confirmed the hypothesis that sound can help users recover from slip-off errors more quickly.

The auditory condition had an average of 6.6 slip-off errors per participant and the visual condition 3 per participant. There was no significant difference between these scores (T(11)=2.03, p=0.067). There was no significant difference in the total number of codes typed in the two conditions (T(11)=0.401, p=0.696). The average number of codes typed per participant in the auditory group was 64.5 and in the visual 65.5.

Discussion

The workload analysis showed that there were no significant differences between the conditions on any of the factors. This showed that the sonic enhancements did not reduce the workload of the task. However, the participants very strongly preferred the sonically-enhanced buttons to the standard ones. This may have been because the auditory buttons allowed participants to recover from errors more
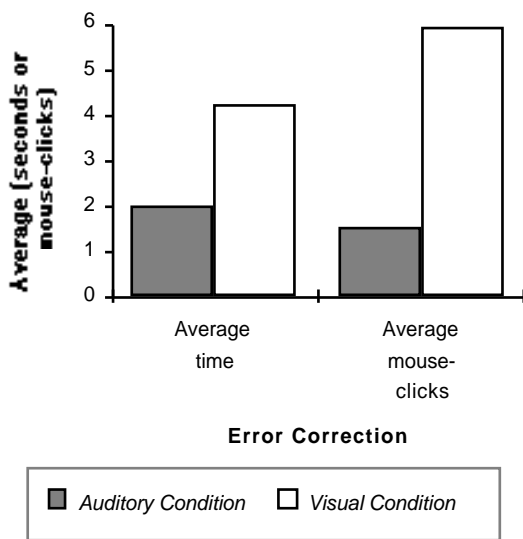
Figure 4: Error recovery in the buttons experiment. The graph shows average error recovery times and the average number of mouse clicks needed for recovery.

quickly. It is unclear why this was not reflected in the workload scores. It may be that recovering from errors was seen as a separate activity from the main task and therefore did not figure in workload estimates but might have affected preference ratings

The sonically-enhanced buttons did not increase the annoyance or frustration felt by the participants. This gives strong evidence to suggest that if sounds provide useful information they will not be perceived as annoying to the main user of the computer. The sounds used here were kept at a low intensity so that others working nearby would not be annoyed by them [11].

The main hypothesis, that the addition of sound would speed up error recovery, was proved correct by the experiment. Time to recover from errors and the number of keystrokes needed were both significantly reduced. These results indicate that, if the simple sound-enhancements suggested here are used, slip-off problems can be dealt with very effectively. This improvement is due to the sharing of the task across the visual and auditory senses. The positioning of the mouse required precise hand-eye coordination. This meant that the participants could not look at the location of the previous button press to see if a slip-off had occurred. Therefore, in the visual condition, errors were not corrected until the whole 5-digit code had been typed. In the auditory condition the participants could hear when a slip-off occurred whilst still performing the mouse positioning task. They could then correct it immediately.

There was no significant difference between the total number of codes typed in either condition. Even though the participants in the auditory condition recovered from errors more rapidly they did not type more codes. The participants in the auditory condition made slightly more slip-off errors than in the visual (although this difference was not significant). The fact that the participants made slightly more errors in the auditory condition wiped-out the advantage gained in recovering from errors more quickly. It was as if participants became more careless with their clicking because they knew that they could recover from errors with little cost. However, recall that the difference in total errors was not significant so we cannot make any strong conclusions about the number of errors that would be observed in a real interface. It is hoped that when sonically-enhanced buttons are used in real interfaces users will make the same number of errors and recover more quickly.

SONICALLY-ENHANCED SCROLLBARS

There are two main usability problems with scrollbars [10]: Position awareness in documents and scrollbar 'kangarooing'.

When scrolling through a document it can be hard to maintain a sense of position. The text can scroll too fast to see (and the thumb wheel only gives an approximate position). The user really wants to look
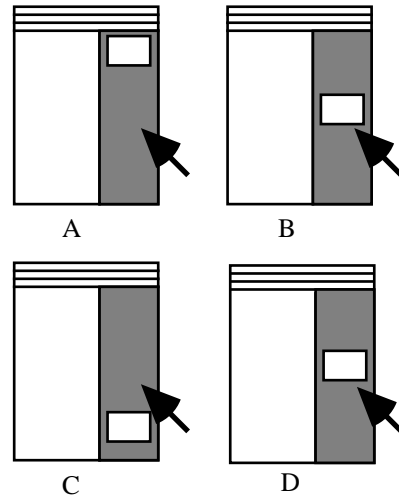

Figure 5: Scrollbar 'kangarooing'.

at the information in the window, not the scrollbar. However, in order to get location information he/she must look at the scrollbar, forcing visual attention away from the document which is what is really of interest and causing the interface to intrude into the task being performed. As before, the user cannot be looking in two places at once. Some systems put a page count at the bottom of the screen but this is too far from the centre of visual focus - again the user must move his/her visual focus to the page counter. Sound can be used to present this location information so that the user can perceive it wherever he/she is looking.

The second problem is scrollbar 'kangarooing'. Repeatedly clicking in the scroll area above or below the thumb wheel scrolls by a window-sized step. Clicking below the thumb scrolls down in the document and clicking above scrolls up. Figure 5

shows an example of kangarooing. In A the user begins clicking to scroll down towards the mouse pointer. In B the thumb wheel is just above the pointer. In C the user has clicked and the thumb has scrolled below the pointer. In D the user clicked again and this time the thumb scrolled back above the pointer and kangarooing occurred. Unless he/she is looking at the thumb it can be hard to recognise that this has happened. If the user continues clicking the thumb wheel will bounce above and below the mouse pointer location. Again, users are forced to look at two places at once: They really want to look at the document they are working on but they must look at the scrollbar to avoid kangarooing. Information about kangarooing could be presented in sound so that the user could perceive errors even if he/she was looking at the document.

Experimental hypotheses

The workload felt by participants should be reduced as the extra feedback provided information that they needed. Participants would have to expend less effort recovering from errors and remembering their whereabouts in the document. There would be no increased frustration or annoyance due to the addition of sound as the auditory feedback provided information that the participants needed.

The extra auditory feedback should make it easier for participants to recover from kangarooing errors. They will receive auditory feedback telling them when an error has occurred so that they will be able to correct it immediately. This should result in faster recovery from such errors. This should also result in a reduction in the total time taken in the tasks.

Participants should better be able to maintain their sense of position in the document with more page feedback and therefore give fewer wrong page answers. If participants lost their sense of position the time cost was high. For example, they would have to go back to the top of the data file and work out their position from there and this would take much time. Therefore, if they did not lose their sense of position, time to complete tasks should be reduced. The demanding audio feedback should make it easier for participants to perceive page boundaries and so make fewer wrong page errors.

Experimental tasks

A simple document browser was created (see Figure 6). This browser allowed participants to navigate around a document using a scrollbar and indicated page boundaries with a dotted line, in a similar way to many wordprocessors. The scrollbar used in the browser only allowed clicking in the grey scroll area above or below the thumb wheel to scroll by a window of data either way. It did not provide a page counter.
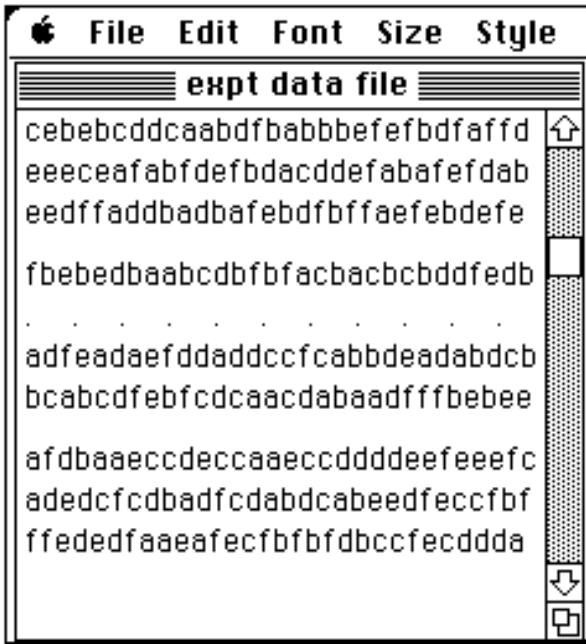
Figure 6: The scrollbar experiment browser program (reduced in size). It shows example data and a page boundary marked by dots.

The data files used in the browser were made up of groups of three lines of 30 randomly generated 'a' to 'f' characters separated by a blank line. This data was used so that the participants would not be able to learn and remember their way around the files, they would have to search for what they were looking for.

Participants were given two types of task. The first, which will be called the *Search Tasks*, involved the participants visually searching through a file of data to find significant features. These features were such things as whole line of 'a's together. When the target was found the participants had to say which page it occurred on. The other tasks, which will be called the *Navigate Tasks*, involved participants being given instructions to go to a specific point on a specific page and read the first six characters of the data that was there. These types of tasks covered the main ways users interact with scrollbars. They might be searching through a document to find something or they might be looking for a specific page to find the data they want. The data were described to the participants as 'experimental results data'. The rationale given to the participants for the tasks was that they were searching through the data to find significant features for analysis.

Experimental design and procedure

The experiment used the testing framework described above. In the visual condition, participants used an ordinary graphical Macintosh scrollbar (but restricted as described above). In the auditory condition the sonically-enhanced scrollbar described above was used. Training was given in both types of task before the main test was started. Each participant was given the search task questions first and then the navigate ones. They had to complete the tasks as fast as possible.

Two types of errors data were collected: The number of times kangarooing occurred (kangaroo errors) and the number of times the wrong page was chosen (wrong page errors). Total time taken, time taken to recover from errors, the total number of codes typed and workload data were also recorded.

13

The design of the sonically-enhanced scrollbars

Two types of sounds were needed to solve the problems described above: One to give scrolling information to indicate when the thumbwheel reached the target location (to avoid kangarooing) and one to give location information.

The first sound was a fixed tone of duration 180 msec. and was used to indicate a window scroll event with the thumbwheel. The sound was kept short so that it could keep up with the interactions taking place. If the user scrolled towards the bottom of a document, the short tone was played at a low-pitch, C_4 (130Hz). When scrolling up a high-pitched note C_0 (2093Hz) was played. High pitch was used as up and low pitch as down because of the natural tendency to perceive higher pitch as higher spatial location [19]. If a user was clicking to scroll down towards the mouse pointer location he/she would hear the repeated low-pitched sound. If kangarooing occurred then the user would hear a demanding high-pitched tone when not expected and this would indicate the error.

A low intensity, continuous tone was used to give location information. This earcon changed in pitch when a page boundary was crossed; lower pitch when scrolling downwards and higher when scrolling upwards. To indicate a page boundary event the background tone was increased in volume for two tones of 180 msec. each to demand the listener's attention. It then decreased again to just above threshold level so that it could be habituated. The different number of notes differentiated this earcon from the previous window scroll sound. Again the sound was short so that it did not hold up the interaction. The notes played when scrolling towards the bottom of the document decreased in pitch from B_1 (1975Hz) to C_4 (130Hz) when a page boundary was crossed. The reverse occurred when scrolling up from the bottom of the document. When the scrollbar was clicked the thumb sound was played first followed by the page boundary sound after a 180 msec. delay (if a page boundary had been crossed).

Workload results

Figure 7 shows the average score for each of the workload categories. Paired T-tests were carried out on the auditory versus visual conditions for each of the categories. Mental demand showed a significant decrease in the auditory condition over the visual (T(11)=3.23, p=0.008). Nine of the twelve participants rated the auditory condition lower in effort than the visual but this failed to reach significance (T(11)=1.83, p=0.09). There were no significant differences in any of the other workload categories.
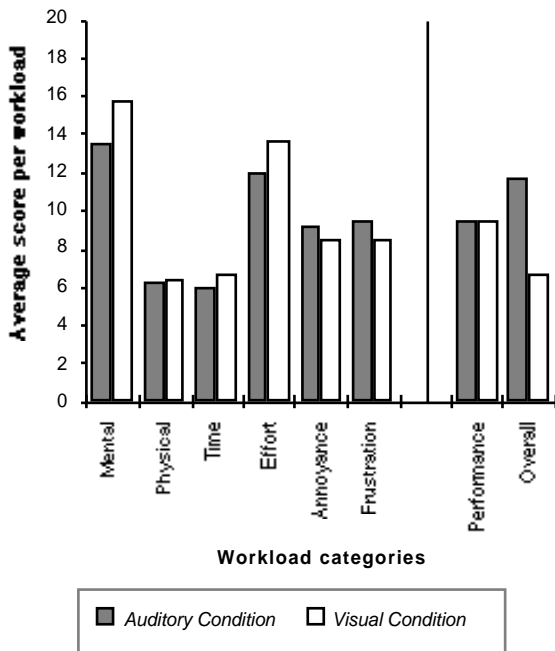
Figure 7: Average TLX workload scores for the auditory and visual conditions of the scrollbar experiment. In the first six categories higher scores mean higher workload. The final two categories, performance and overall preference, are separated because higher scores mean less workload.

taken by each of the participants in the two conditions for the search tasks. Nine of the twelve participants performed faster in the auditory condition but there was no significant difference in time scores at the 95% level (T(11)=1.846, p=0.09). However, an F-test between the auditory and visual conditions across participants showed a significant reduction in the variance in the auditory condition (F(11)=3.98, p=0.01).

Tasks/ Conditions	Search		Navigate	
	Wrong page	Kangaroos	Wrong page	Kangaroos
Auditory	13	5	40	4
Visual	11	3	51	8

Table 2: Totals of wrong page and kangaroo errors in both conditions of the scrollbar experiment.

The annoyance for the auditory condition was not significantly different to the visual condition (T(11)=0.516, p=0.615). Five participants rated the auditory condition more annoying than the visual and three rated the visual more annoying than the auditory. There was a difference in terms of overall preference. Here the auditory scrollbar was significantly better than the visual one (T(11)=2.55, p=0.02).

Timing and error results

Table 2 shows the total numbers of kangaroo and wrong-page errors. Figure 8 shows the total times

To find out if any underlying differences were hidden in the overall timing data a more detailed analysis was undertaken. The average time taken to answer a question where errors occurred was calculated for each question in both conditions of the search tasks (both types of errors were included in this analysis). There were no significant differences between the conditions in time taken to answer questions with errors.

Figure 8 also shows the total times for the two conditions in the navigate tasks. In these tasks there was a significant difference between the times taken. A paired T-test showed the auditory condition was significantly faster than the visual (T(11)=2.29, p=0.04). As before, there was also a significant reduction in the variance in the auditory condition (F(11)=6.36, p=0.001). To find whether the decrease in time taken for the auditory condition was due to faster recovery from errors, a more detailed analysis was undertaken. Recovery was significantly faster in the auditory than in the visual condition (T(9)=2.61, p=0.02). The average time taken to answer questions with no errors was also calculated and the auditory condition was again significantly faster than the visual (T(9)=4.18, p=0.002).

In both the search and navigate tasks there were no differences in the number of wrong page errors between conditions (see Table 2). In the navigate tasks there was a reduction from 51 to 40 in the auditory condition but this failed to reach significance.

Discussion

The workload results indicated that the auditory scrollbar reduced the workload of the task. Mental demand (which dealt with how much mental and perceptual activity was required to perform the task) was significantly reduced. This was because it was easier for participants to hear page boundaries than it was to see them as the feedback was more demanding. Participants also got more feedback about kangaroo errors so making it less effort to recover from them. This confirmed the hypothesis that extra auditory feedback would lower workload. Although participants felt their performance was no better in the auditory condition than in the visual, they had an overall preference for the auditory scrollbar because it lowered mental demand and there was some decrease in effort expended. These factors indicated that an auditory enhanced scrollbar would be an effective addition to an interface and could lower the workload therefore freeing-up cognitive resources for other tasks.
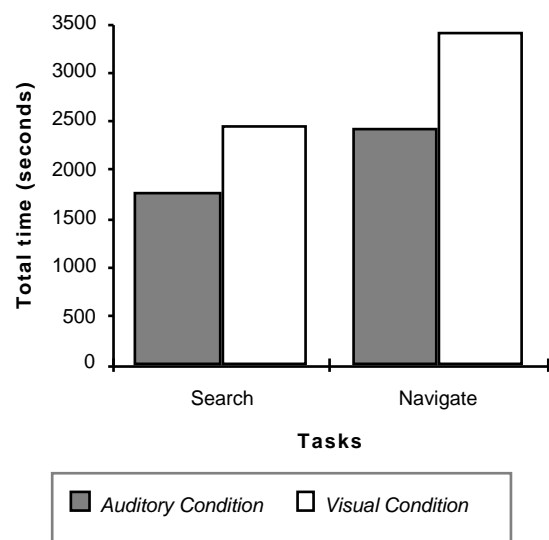
There was no significant difference in the annoyance

Figure 8: Total times for the search and navigate tasks in the scrollbar experiment.

or frustration felt by participants in the auditory condition. This indicated that auditory feedback, and especially constant auditory feedback, was not necessarily annoying when used at the interface. This confirmed the hypothesis that auditory feedback would not be annoying if it provided useful information to the user.

The significant reduction in time for the auditory condition in the navigate tasks indicated that the sonically-enhanced scrollbar improved performance. This is again evidence to suggest that auditory scrollbars are an effective extension to standard visual ones. The times for the search tasks were not significantly different. This may have been due to the nature of the task. A participant was required to visually search through the data file to find a target. The advantage conferred by sound may have been lost in the overall time to do the visual searching as this took up a large proportion of the time for this task. Position awareness within the document was bound up in this. The advantages due to sound were small and therefore lost in the large times for visual searching. In the navigate tasks, where the participants had to find a specific page, searching was based on page boundaries so there was a better comparison between the auditory and visual conditions.

There were no significant differences between conditions in the time taken to recover from errors in the search tasks. As described previously, the time to do the searching might have been the problem here. In the navigate tasks the auditory group was significantly faster overall. When this result was investigated in more detail the auditory group was found to be significantly faster at recovering from errors than the visual. The auditory group also performed better when there were no errors. It seems that the sounds helped increase general performance with the scrollbar.

One problem with the error analysis was that the frequency of kangaroo errors was too low to be a good measure. For example, in the search tasks there was fewer than one error per participant in each of the conditions. It turned out to be very difficult to generate many kangaroo type errors. It could be that, as the participants were experienced scrollbar users, they had developed strategies for avoiding kangarooing in their everyday work which they used in the experiment. However, two participants did say that the sounds helped them identify when a kangaroo error had taken place. These problems generating kangaroo errors meant that it was difficult to test the hypothesis that recovery from such errors would be quicker.

There were no differences between the conditions in the number of wrong-page errors. It may have been that participants counted the page boundaries whether they saw them or heard them, but it just took longer when they had to do it visually. This may have been one of the reasons for improved performance in the navigate tasks for the auditory condition. Further investigation of errors is therefore necessary.

It is noteworthy that there were significant differences between the auditory and visual conditions in terms of variance on both tasks. Eight of the twelve participants showed less variability in the auditory condition. However, a Sign test failed to reach significance. The data indicated that the best participants stayed the same when using the new scrollbar but the poorer ones improved. This may be because the best participants already had strategies to help them deal with scrollbar problems. However, the poorer ones did not so that the extra support provided by the sonic enhancement allowed them to perform better.

FUTURE WORK

The future of this work is to test and sonically enhance the rest of the standard interface widgets [7]. This will be done using the same techniques as described here. Once this has been done, the individual widgets will be combined into a complete toolkit. This toolkit will allow designers to include sounds that will improve usability and will be consistent across different applications. The toolkit will offer four advantages: 1) Simplification of the implementation of sonically-enhanced interfaces; 2) Allow designers who are not sound experts to create sonically-enhanced interfaces; 3) By experimental evaluation, ensure the sonically-enhanced widgets are effective and improve usability; 4) Make sure the sounds are used in a clear, coherent and consistent way across the human-computer interface.

CONCLUSIONS

This paper has described two experiments that have integrated sound into basic interface components. One problem with current displays is that they rely almost entirely on graphical output, no other senses are used to provide information to the user (as would occur in everyday life). The person using the system can then become overloaded with large amounts of visual information. One reason for this is that people can only look at one thing at once. However, when operating a computer users are often required to look at two (or more) things: The information they are interested in and also the interface to the computer, so can become overloaded. This paper suggested that sound should be used alongside graphics to avoid this problem. Sound and graphics work well together; our visual sense gives us detailed data about a small area

of focus whereas the auditory sense provides data from all around, even things that the listener cannot see. By using the advantages of sound and graphics together the display of information can be improved.

How should sound be included into the computer display? Previous to the work described here there was little research on how best to combine graphics and sound because this area is still in its infancy. This paper described one approach, suggesting that sound should be integrated into the basic components that make up a computer display, things such as buttons and scrollbars. Sonically-enhanced buttons were experimentally tested and shown to be effective at reducing the time taken to recover from slip-off errors. These errors occur because users cannot look at a button and the information they are working on at the same time. The number of keystrokes necessary to recover from these errors was also reduced. The sonically-enhanced buttons were strongly preferred by the participants over standard visual ones. These results suggest that the introduction of such buttons into human-computer interfaces would improve usability but not at a cost of making the interface more annoying to the user.

A sonically-enhanced scrollbar was also tested and found to significantly improve performance time on certain tasks. The problems with scrollbars ('kangarooing' and losing ones sense of position in a document) again resulted from the fact that users cannot look at the information they are concerned with and the scrollbar at the same time. Adding sound also significantly reduced the mental workload of using the scrollbar and the new scrollbar was rated with a significantly higher preference score than a standard visual one.

This research shows that the use of auditory and graphical feedback to create multimodal computer displays will provide more usable systems. Users will not be overloaded because information output will be shared across two different senses. As well as demonstrating that sound can improve performance it also gives a model for how it might be used in other systems. By integrating sound into a system from the bottom up, simple errors that occur when graphics alone is used can be overcome and usability improved.

ACKNOWLEDGEMENTS

REFERENCES

Any references by Brewster can be found at http://www.dcs.gla.ac.uk/~stephen/publications.html

1. Barfield, W., Rosenberg, C. and Levasseur, G. The use of icons, earcons and commands in the design of an online hierarchical menu. *IEEE Transactions on Professional Communication* 34, 2 (1991), 101-108.

2. Berglund, B., Preis, A. and Rankin, K. Relationship between loudness and annoyance for ten community sounds. *Environment International* 16 (1990), 523-531.

3. Bevan, N. and Macleod, M. Usability measurement in context. *International Journal of Man-Machine Studies* 13, 1 & 2 (1994), 123-145.

4. Blattner, M., Papp, A. and Glinert, E. Sonic enhancements of two-dimensional graphic displays. In *Proceedings of ICAD'92* (Santa Fe Institute, Santa Fe) Addison-Wesley, 1992, pp. 447-470.

5. Blattner, M., Sumikawa, D. and Greenberg, R. Earcons and icons: Their structure and common design principles. *Human Computer Interaction* 4, 1 (1989), 11-44.

6. Brewster, S.A. *Providing a structured method for integrating non-speech audio into human-computer interfaces*. PhD Thesis, University of York, UK, 1994.

7. Brewster, S.A. The development of a sonically-enhanced widget set. In *Proceedings of EWHCI'95* (Moscow, Russia) International Centre for Scientific and Technical Information, 1995, pp. 126-129.

8. Brewster, S.A., Wright, P.C., Dix, A.J. and Edwards, A.D.N. The sonic enhancement of graphical buttons. In *Proceedings of IFIP Interact'95* (Lillehammer, Norway) Chapman & Hall, 1995, pp. 43-48.

9. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of ACM/IFIP INTERCHI'93* (Amsterdam) ACM Press, Addison-Wesley, 1993, pp. 222-227.

10. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. The design and evaluation of an auditory-enhanced scrollbar. In *Proceedings of ACM CHI'94* (Boston, MA) ACM Press, Addison-Wesley, 1994, pp. 173-179.

11. Brewster, S.A., Wright, P.C. and Edwards, A.D.N. Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of BCS HCI'95* (Huddersfield, UK), 1995, pp. 155-159.

12. Dix, A., Finlay, J., Abowd, G. and Beale, R. Chapter 9.4 Status/Event Analysis. In *Human-Computer Interaction*(Ed.), Prentice-Hall, London, 1993, 325-334.

13. Dix, A.J. and Brewster, S.A. Causing trouble with buttons. In *Ancillary Proceedings of BCS HCI'94* (Glasgow, UK) Cambridge University Press, 1994.

14. Edworthy, J., Loxley, S., Geelhoed, E. and Dennis, I. The perceived urgency of auditory warnings. *Proceedings of the Institute of Acoustics* 11, 5 (1989), 73-80.

15. Gaver, W., Smith, R. and O'Shea, T. Effective sounds in complex systems: The ARKola simulation. In *Proceedings of ACM CHI'91* (New Orleans) ACM Press, Addison-Wesley, 1991, pp. 85-90.

16. Hart, S. and Staveland, L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human mental workload*, Hancock, P. and Meshkati, N. (Ed.), North Holland B.V., Amsterdam, 1988, 139-183.

17. Hart, S.G. and Wickens, C. Workload assessment and prediction. In *MANPRINT, an approach to systems integration*, Booher, H.R. (Ed.), Van Nostrand Reinhold, New York, 1990, 257-296.

18. Lee, W.O. The effects of skill development and feedback on action slips. In *Proceedings of HCI'92* (York, UK) Cambridge University Press, 1992, pp. 73-86.

19. Mansur, D.L., Blattner, M. and Joy, K. Sound-Graphs: A numerical data analysis method for the blind. *Journal of Medical Systems* 9 (1985), 163-174.

20. NASA Human Performance Research Group. *Task Load Index (NASA-TLX) v1.0 computerised version*. NASA Ames Research Centre, 1987, , .

21. Portigal, S. *Auralization of document structure*. MSc. Thesis, The University of Guelph, Canada, 1994.

22. Reason, J. *Human Error*. Cambridge University Press, Cambridge, UK, 1990.

23. Scholes, P.A. *The oxford companion to music*. Oxford University Press, Oxford, 1975.

24. Sumikawa, D., Blattner, M., Joy, K. and Greenberg, R. *Guidelines for the syntactic design of audio cues in computer interfaces*. Lawrence Livermore National Laboratory, 1986, Technical Report, UCRL 92925.