How to Tackle Auditory Interface Aesthetics? Discussion and Case Study

Grégory Leplâtre and Iain McGregor
Napier University
School of Computing
10 Colinton Road, Edinburgh EH10 5DT
{g.leplatre,i.mcgregor}@napier.ac.uk

ABSTRACT

This paper discusses the importance of auditory interface aesthetics and presents an empirical investigation of sound aesthetics in context. The theoretical discussion examines the relationship between sound aesthetics and user satisfaction and concludes that, despite the creation of numerous auditory design methods and guidelines, none are dedicated to achieving aesthetically pleasing designs. In a case study, an empirical investigation is conducted to evaluate the relationship between the functional and aesthetic value of an auditory interface. By investigating two different tasks, this study demonstrates that the nature of the tasks allocated to subjects has a significant impact on the aesthetic judgments made by the subjects. Consequently, functional and aesthetic properties of auditory cannot be dealt with independently.

1. INTRODUCTION

In the last twenty years the developing Auditory Display (AD) community has demonstrated that using audio at the human-computer interface can improve its usability [1]. Yet, although Gaver presented the *SonicFinder* almost 20 years ago, one must acknowledge that computers around us have remained mostly silent. One of the reasons why the potential of non-speech sound has not materialised, except in game and multimedia design, is because the sounds created and investigated by the AD community notoriously lack *design quality*. Gaver cites Cohen and Brook's OutToLunch [2] as a rare counter-example [3]. The design of non-speech sounds largely focuses on improving the performance of tasks, but the aesthetics of the sounds used is often disregarded. In 1997, Gaver stated that one of the three major lines of endeavour that should help the field of Auditory Display really mature is that: "we focus our attention on sounds that are aesthetically controlled, as subtle and beautiful as those as those we hear in the orchestra hall, or on a walk though the woods" [3]. Last year, Thornton and colleagues recognized that despite putting the emphasis on sound aesthetics in their design, "there remains a need for more sustained, ongoing, inquiry into the role of aesthetics in auditory interfaces" [4]. This suggests that the AD community has still to shift its focus from the dry representation of information through audio to an approach that takes into account the aesthetics and expressiveness of the medium. There is currently no design guideline or method dealing with the aesthetic qualities of auditory interfaces.

The audio medium is not alien to aestheticians. Music has long been an area of interest to them. Yet there are two chief reasons why their work is not sufficient for the AD community's needs. Firstly, the sounds used in auditory interfaces are not necessarily musical. *Earcons*, which are brief structured non-speech sounds can be regarded as musical, but *auditory icons*, "everyday sounds mapped to computer events by analogy with everyday sound-producing events" [3], are not. Secondly, even for musical sounds, the aesthetics of a sound depends on the context in which the sound is heard/listened. The only relevant context for music aestheticians is a musical context. For an auditory interface designer, the relevant context is the interaction in which the sound is used and the context in which the interaction takes place. It is arguable that the relationship between the context in which a sound is used *i.e.*, interaction with the interface, and the sound itself contributes largely to the perceived aesthetic qualities of the sound, of the interface and hence to user satisfaction. We propose to start tackling this question by investigating the aspects of auditory interface interaction that contribute to user satisfaction. This involves the intrinsic aesthetical qualities of the sounds themselves and the relationship between sounds and the interaction in which they are used.

2. AESTHETICS AND USER SATISFACTION

In 1992, Peterson and Wilson estimated that more than 15,000 academic and commercial articles appearing in the previous two decades deal with customer satisfaction [5]. User satisfaction, on the other hand, has far from received the same level of interest in HCI. The very title of a recent study by Lindgaard and Dudek ("What is this evasive beast we call user satisfaction") is symptomatic of our lack of understanding of this notion [6]. Yet satisfaction is widely regarded as one of the three pillars of usability [7]. There are obvious reasons why effectiveness and efficiency remain the focus of the HCI hardcore. In particular, effectiveness and efficiency are easily quantifiable notions that cognitive ergonomics methods have under control. This situation is however unfortunate because it reveals a gap between the practice of the computer and media industry on the one hand, and the foci of the HCI community on the other hand [8].

The recent emergence of "affective computing" has come as a rare but remarkable effort to breath *humanity* in the almost mechanistic HCI hardcore [9-13]. However, Lindgaard and Dudek point out that such studies tend to focus more on changing the behaviour of a system based on perceived user emotions, rather than on understanding what elements of the interaction make the experience "exciting, fun, or boring" [6]. Lindgaard and Dudek's study reveals that: "the concepts of aesthetics, emotion, expectation, likeability and usability play a major role in shaping the user experience" [6].

The problem of aesthetics is that it is a notion that is hardly less evasive than user satisfaction. Lindgaard and Dudek have identified that there are at least five distinct meanings for it [14]. Beyond philosophical questions regarding the definition of beauty, it is however recognised that aesthetics plays a role in the sense of satisfaction experienced by the user of a system.

As Karvonen points out, aesthetic notions are commonly used to assess the quality of an interface [15]. For instance, "design quality" is often used to describe the aesthetic quality of an interface [16]. Similar aesthetic judgement can also be made with the term "pleasantness", in the sense that a high "design quality" results in a pleasant experience [17]. Karvonen argues that "simplicity" is also used to describe the aesthetic qualities of an artefact. Simplicity is a popular topic in HCI studies to the point of being designated by Nielsen a key factor in the design of usable systems. However, in Nielsen's work and most of the HCI literature, simplicity does not have any aesthetic connotation.

A few studies have investigated the impact of aesthetic features of graphical interfaces on their usability. Tractinsky *et al.*, have shown that there is a strong correlation between the perceived aesthetics quality of a system and the perceived usability of the system [8]. These conclusions are based on a study conducted on a computer simulation of an automated teller machine. In a study of the first impressions of web pages, Schenkman and Jonsson observed that the overall rating of a page was correlated to the "beauty" rating of the page [1]. In addition, there is evidence of cross-culture differences between the perceived aesthetics of interfaces [18].

3. DESIGNING AESTHETICALLY PLEASING AUDITORY INTERFACES

3.1. Presentation of the Problem

Among the AD community, the importance of user satisfaction with auditory interfaces is recognised to be an important issue, but no research has been carried out to address it. The relationship between user satisfaction and the aesthetics of the sounds used have been restricted to individuals' intuitions. Despite the evident emotional potential of non-speech sounds, unparalleled in graphical user interfaces, the aesthetics of auditory interfaces is widely overlooked by the AD community. Most studies investigating sound quality focus on effectiveness and efficiency of sounds in conveying information (see for example two recent studies on the perception of train and car horn sounds [19, 20]). One has to deplore that the division between the cognitive bound HCI hardcore and graphics designers and artists also exists between the AD community and sound designers.

In most media, such as film and television, sound is often regarded as a mere companion to the more dominant visual medium. This is largely due to the widespread use of printed materials as a means to convey information since the invention of printing [21, 22]. However, Somers points out that it is easier for a blind person to follow a play or a film without additional aural content than for a deaf person to follow the same media without additional visual captions [23]. In 1973, a BBC program called "Ways of Seeing" demonstrated how strong the impact of a soundtrack is on the perception of a visual scene (as pointed out by Somers [23]): the program clearly showed how the interpretation of an oil painting could be drastically undermined by the nature of the sound or music played to the viewer. Audio in a user interface has the same effect. The potentially dramatic impact of sound on an interface in terms of its aesthetics, user satisfaction and as a result, usability, pleads for investigating means to keep the aesthetics of the (auditory) interface under control. Many questions must be addressed: what is a pleasant sound? How can we design them? Should the emotional potential of music be used in auditory interface design? Can we design *emotionally neutral* sounds?

One of the main reasons why the AD community is sensitive to the design qualities of auditory interfaces is because a lack of these causes annoyance. Gaver and Mandler have observed that: "people tend to find the repetition of tunes annoying" [24]. Gaver also noted that musical phrases may be hard to integrate in a working environment [3], which confirms the important of context in the use of non-speech sounds. Conscious of the potential annoyance of non-speech sound in user interfaces, Brewster added an "annoyance" measure to the NASA TLX workload test [25] in the evaluation procedure [26]. The research that needs to be conducted in this area must provide the community with much more than ways to design auditory interfaces that are not annoying.

3.2. Basic Aesthetic Principles

Although we have discussed that there is a lack of work dedicated to auditory interface aesthetics, a few results can be found in the literature. This section does not intend to review them all, but to present a limited number of ideas introduced in previous work by Leplâtre [27]. The design recommendations described below will serve as a basis for a practical investigation of the effect of a limited number of design parameters on the perceived aesthetic qualities of sounds:

- Homogeneity of the design – Designers often need to maximise the differences between sounds to make them more easily distinguishable. This compromises the homogeneity of the auditory interface and hence its overall aesthetics. For example, the designer should ensure that, if the auditory interface is composed of individual musical sounds that can be played in different orders depending on the interaction, all the possible sequences are melodically and harmonically sound.
- Temporal envelope – Sounds used in auditory interfaces must often be brief and interruptible. In this case, the information conveyed by the sound should preferably be located in the onset of the sound and fade-ins and fade-outs should be used to soften the transition between the sounds.
- Sonic density – In brief, sonic density refers to the perceived density of a sound. The contributing factors are duration, intensity, spectrum, number of instruments, etc. For example, in the case of navigation in a sonified mobile phone menu hierarchy (see study by Leplâtre and Brewster [28]), in which users navigate quickly in the menu and consequently in which fast sequences of sounds are played, annoyance is best avoided by limiting the density of the sounds.

4. CASE STUDY: AN AUDITORY EMAIL NOTIFICATION SYSTEM

The impact of various design parameters (such as those mentioned in Section 3.1) on the design's aesthetic qualities was investigated in a case study. For this case study, the design problem chosen was the creation of auditory email preview. Audio email previews can be regarded as sophisticated email notification messages. This design problem is not particularly new nor original. It has been addressed in the past, in particular by Hudson and Smith [29]. Nevertheless there are several reasons why it has been chosen: Firstly, it is a problem that everyone understands, including sound designers. One of the issues with working with sound designers is that they are not necessarily familiar with the domain that pertains to the design (sonification of scientific data, for instance). This issue did not arise in present study. Secondly, this problem is challenging enough to lead to the design of interesting sonifications.

4.1. The sound design problem

The challenge of this design is to create a sonification that conveys a relevant subset of the information presented in this section in an *acceptable* time interval. Indeed, one of the main constraint of the design is to keep the duration of the sounds as short as possible. The information potentially conveyed in sound is listed below (note that the presentation is written from the mailbox user's point of view. Therefore, in what follows, "me" refers to the user, not the author):

1. New message received
    a. From **X**
        i. **X** belongs to Group **G**
        ii. **X** emailed me **n** minutes ago
        iii. I have **m** unread messages from **X**
    b. Sent to email address **E**
    c. Sent to **Y:**
        i. **Y** is me only
        ii. **Y** belongs to Group **H**
2. Number of unread messages **p** in my mailbox
3. Number of message **q** in my mailbox
4. Contains attachment **A**
    a. No
    b. Of type **T**

Where,
G = {friends&family, colleague, junk, unknown},
H = {University, department, research group, unknown},
T = {.doc, .pdf, .ppt, .xls, .gif, .jpg, .exe, unknown}.

In its current description, a potentially large number of sounds would need to be used to convey the requested information. Yet, it is intended that the resulting sonification should not be significantly more invasive than common email notifications. Hence the designer was provided with a priority list (compulsory feedback at the top, dispensable at the bottom):

1. New message.
2. From group G.
3. To me only | to mailing list.
4. Contains an attachment | doesn't.
5. Number of unread messages in mailbox.

4.2. First Design

A professional sound designer produced a design based on the problem description and requirements presented in Section 4.1. The design is presented here in decreasing order of importance of the feedback, according to the priority list mentioned earlier:

1. **New message**: the notification of a new message is implicit when a sound is played.
2. **From group G**:
    a. Friends & Family: children's voices announcing sender's name.
    b. Colleague: adult voices announcing sender's name.
    c. Junk: thud of heavy metal object hitting other metal objects. This emulates the sounds from a metal scrap yard.
    d. Unknown: knocking sound on wood simulating door knock..
3. **Sent to:**
    a. **Me only**: my name spoken softly by my own voice.
    b. **To group H:**
        i. University: Large clock/church bell
        ii. Department: hand bell.
        iii. Research group: small tinckle (miniature bell).
4. **Attachment(s):**
    a. .doc: Typewriter song.
    b. .pdf: Magazine page turning.
    c. .ppt: Slide projector changing slides.
    d. .xls: Old fashioned adding machine.
    e. .gif: Scraping sound similar to an artist using a palette knife.
    f. .jpg: Camera shutter and motor drive.
    g. .exe: Windows truncated startup sound.
5. **Number of unread messages in mailbox**: Simple beeps with increasing/decreasing pitch (high pitch = high number of messages). Higher frequency demands a greater degree of attention.

The remaining information that did not feature in the priority list was also allocated a feedback sound:

1. Number of attachments: Repetition of samples of the corresponding type.
2. Number of unread messages from sender: the pitch of the sample used for sender increases with the number of unread messages.
3. Number of messages in mailbox: series of simple beeps (same as those used for the number of unread messages). Emulates the number of messages stored on an answering machine, or the number of files to be accessed before space can be found for the new message.

The sequence in which the samples should be played completed the design description:

1. Number of messages in mailbox (q).
2. Sent to me or group (Y or H).
3. From (X).
4. Group (G).
5. Number of unread messages from X (m).
6. Attachment type and number (T).

There is scope for discussions about the design and its effectiveness, but this study is primarily interested in the aesthetic qualities of the design in different contexts. However, the durations of the email notifications produced with this design were too long. Therefore it was decided to review and modify the initial design.

4.3. Second Design

The following modifications were made to the attributes listed in the previous section:

1. **New Message** – no change
2. **From Group G**:
    a. Friends & Family: Name of the sender spoken by the sender. The rationale is that one can envisage getting their friends and family to record their own names, rather than using the voice of a child, or worse, a synthesized child voice.
    b. Colleague: Name of sender spoken by a speech synthesizer.
    c. Junk – no change
    d. Unknown – no change
3. **Sent to:**
    a. **Me only:** standard MS Windows email notification.
    b. **To group H:**
        i. Research Group: Same as above with pitch decreased by 6 tones
        ii. Department: Same as above with pitch decreased by 6 tones
        i. University: Same as above with pitch decreased by 6 tones.

The rationale is that the lower the pitch, the larger the group of recipients.

4. **Attachments:** only the .exe sound was modified. Instead of using a MS windows related sound, we decided to use a sequence of beeps simulating an old computer executing a program.

In addition to these changes, we also decided to represent the number or unread messages in the mailbox with reverberation: the more reverb, the more unread messages in the mailbox. With these changes, the duration of the sounds used in our experiment was reduced to a more acceptable standard.

4.4. Experiment Presentation

For this study, three groups of twelve users were recruited to make comparative judgments between email previews designed as descried above, and chosen transformations of these previews. The transformations chosen were related to rhythm and tempo. This choice was made for two reasons: Firstly, we wanted to tackle an aspect of the design that has rarely been studied. Secondly, the sequential aspect of the sounds designed for this case study was calling for rhythmic patterns to be investigated. Lastly, but most importantly, this promised to give us an indication of the relationship between the user's task and their perception of the sound. Indeed, rhythm has two distinct functions in the design. The first function is aesthetic: the user will like the rhythm of the sequence to various degrees. The second function is informative: if the rhythm is such that the duration of an element of the sequence that the user wants to monitor is long, the user may appreciate that rhythm for its informative value. Of course, a rhythm might be aesthetically good and have no informative value. This is an illustration of the typical tensions between designer's designs and interaction designer's designs. The experiment has been devised as follows:

4.5. Sounds used

Three notifications were designed according to the principles described earlier:

- Sound 1 corresponded to the notification of an email sent "to me" by "Fiona", who is a friend, containing an attached executable file. There was a large number of unread messages in the mailbox, therefore a large amount of reverb was applied to the sound. According to the design, the first part of the sound was the standard MS Windows email notification, the second part of the sound was the name "Fiona" spoken by a female English speaker and the last part of the sound was a short sequence of beeps.
- Sound 2 was the notification of an email sent to the "department", by "John", who is a colleague, with three attached "pdf" documents. No unread messages were present in the mailbox. The first part of the sound was the standard MS Windows email notification transposed to a low pitch, the second part of the sound was the name "John" generated by a speech synthesizer and the last part of the sound were three repetitions of a typewriter sound.
- Sound 3 was the notification of an email sent "to me" by "Iain" who is a friend. This email didn't contain any attachment, therefore this notification only contains 2 parts, whereas the previous 2 contain three parts. There was only a few unread messages in the mailbox hence, little reverb was used. The first part of the sound was the standard MS Windows email notification and the second part of the sound was the name "Iain" spoken by a male English speaker.

For each sound, rhythmic transformations were applied. Five different variations were used for Sounds 1 and 2 whereas three were used for Sound 3. For sounds 1 and 2, the variations were structured in the following way (These 2 sounds were composed of three different parts for recipient, sender and attachments):

- V1: long, long, long.
- V2: long, short, short.
- V3: very short, long, short.
- V4: short, short, short.
- V5: very short, very short, very short.

For Sound 3, less options were available as the sound was only made of two parts. The three variations for that sound were:

- V1: long, long.
- V2: short, short.
- V3: very short, very short

In the remainder of this paper, we will refer to these three groups of sounds as Sound groups 1, 2 and 3.

4.6. Experimental protocol

The main aim of the experiment was to investigate the relationship between functional and aesthetic values of sounds. In order to measure meaningful functional aspects of the sounds, different tasks were allocated to the participants of the experiment: Group 1 were asked to pay attention to the recipient of the message. Group 2 were asked to pay attention to the sender of the message. Group 3 were given no such instructions. All the participants were presented the sounds by pair (78 distinct pairs).

In order to ensure that the members of the first two groups did pay attention to the recipient/sender of the messages, they were asked, for each sound of each pair, who was the sender/recipient.

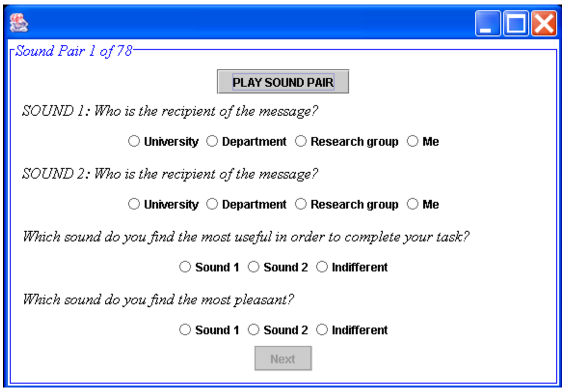

Figure 1. *Software used for the experiment with Group 2.*

For each pair, Group 1 and Group 2 were asked to express their preference on two different aspects of the interface:

1. "which sound do you find the most useful in order to carry out your task?" (functional)
2. "Which one do you find the most pleasant?" (aesthetic)

Participants were given three options to answer these questions: sound 1, sound 2, or indifferent. There was a risk in giving users the option to be indifferent because it was feared that they may take advantage of that option and not make an effort to make a judgment. However, removing the "indifferent" option and forcing them to always choose one sound may have led them to select random sounds when they were truly indifferent and thus bias the results.

Figure 1 shows the software used for the experiment. The screenshot represents a typical sound comparison window as used for Group 2. According to the experimental process, the first two questions were different for Group 1. Note that, to prevent errors, the *Next* button was only enabled once all the questions had been answered.

Group 3 were only asked the second question ("Which sound do you find the most useful?").

Given the exploratory nature of the experiment, no hypothesis regarding the results was made. Comparisons between the choices made by the three groups would indicate whether the judgments were made on functional or aesthetic grounds.

4.7. Results

The first indicator of the participants' performance came from *indifference*. Because of the nature of the differences between the sounds used, it was interesting to identify whether the participants did express any preference at all for any of the sounds or if they found them all equally useful/useless. This was assessed by measuring the number of occurrences of the *indifferent* option being selected, for both questions and across the groups.

4.7.1. Indifference

It was found that the level of indifference was significantly higher for Question 1 than for Question 2, for Group 1 (t_{12}=0.045). No significant difference was found for indifference between Question 1 and Question 2 for Group 2, although indifference was higher for Question 1 than for Question 2. This suggests that the task that the subjects were asked to carry out had an impact on their approach to both questions. No difference was found concerning indifference between Group1 and Group 2, for either Question 1 or 2. It was also found that indifference for Group 3 was very low. These results are summarized in Table 1:

	Question 1	Question 2
Group 1	52.2%	25.8%
Group 2	40.2%	31.8%
Group 3	NA	17.1%

Table 1. *Measure of indifference for both questions and three groups.*

The most probable explanation of the differences between groups may be linked to the level of difficulty of the tasks involved. The hardest task was arguably that allocated to Group 2. The subjects had to identify the recipient of the message, which could only be achieved by interpreting the pitch of the first part of the mail notification. On the other hand, Group 1, who were asked to identify the sender, only had to recognize whether the second part of the notification was synthesized or recorded speech. Given the (medium/low) quality of the speech synthesizer used, this was not a difficult task. Finally, Group 3 were not allocated any task to perform at all, which means that they had the easiest of them all.

4.7.2. Agreement

In order to further assess the significance of the users' response, we tried to answer the following question: how much did respondents agree with each other? Practically, we looked at the number of pairs for which more than *n* subjects made the same judgment, for different values of *n*, for each group and for each question. For example, *n=6*, for a group containing 12 subjects, means that a majority of subjects made the same judgment. Figure 2 shows the agreement distribution for various values of *n*, for Question 2. The graph shows that, for high values of *n* (the most interesting ones), Group 3 tended to make more consistent judgments. However, for very high values of *n* (9, 10 and 11), the trend disappeared. A closer look at the data revealed that agreement with large values of *n*, in Group 1 and 2 was very often agreement to be indifferent. This study was therefore complemented by an examination of non-indifferent judgments: all the "indifferent" choices were removed from the following analysis. Figure 3 shows the non-indifferent agreement pattern for Question 2. Two main differences can be identified between these two graphs. Firstly, the differences between Groups increased in Figure 2. Secondly, strong agreement (high values of *n*) is clearly higher for Group 3 than other groups. Overall, for Question 2, there is an interesting difference between the three groups (See Figure 3). Strong agreement is also noticeably high for Group 3. For instance, Figures 2 and 3 show that, for almost half of the sound pairs presented to the participants, 8 or more subjects out of 12 made a consistent judgment.

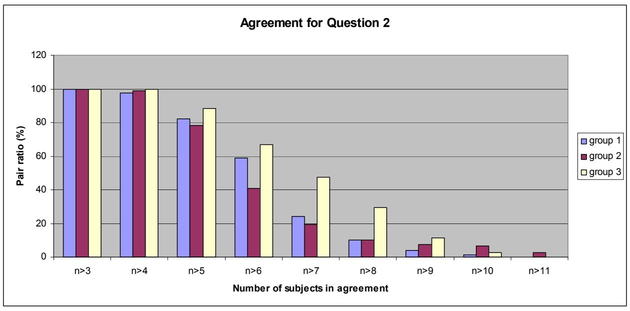

Figure 2. *Agreement for Question 2*

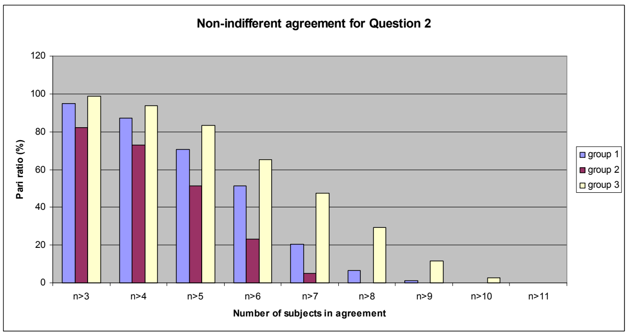

Figure 3. *Non-indifferent agreement for Question 2.*

Looking at agreement for both questions revealed an interesting fact, as Figure 4 shows: There was a noticeable difference between questions within each group (when applicable, *e.g.*, for Groups 1 and 2). This difference was minor in Group 2, but very significant in Group 1. These figures confirm the pattern that emerged in Section 4.7.1 about indifference:

- There is a more noticeable difference between Question 1 and 2 in Group 1 than in Group 2.

- For question 2, Group 3 leads the trend, followed by Group 1 and then Group 2.

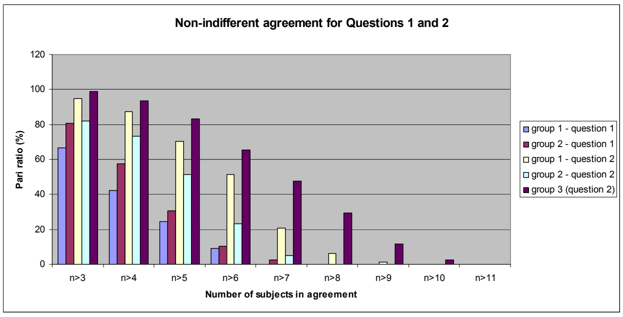

Figure 4. Non-indifferent agreement for Questions 1 and 2.

4.7.3. *Ratings of individual sounds*

Each of the 13 sounds studied were paired once with the 12 others. The number of occurrences when a sound was preferred to the other, or preferred to the *indifferent* option was measured for each sound. This number is comprised between 0 and 12 for each participant, question and sound.

Figures 5 and 6 show that it is difficult to identify sounds whose ratings stand out for Question 1. On the other hand, for Question 2, Figures 7, 8 and 9 show that sounds 1, 4, 11, 12 stand out noticeably, with ratings near or over 6 out of 12. This is apparent in Groups 1 and 3, but not in Group 2, which confirms the pattern observed so far.

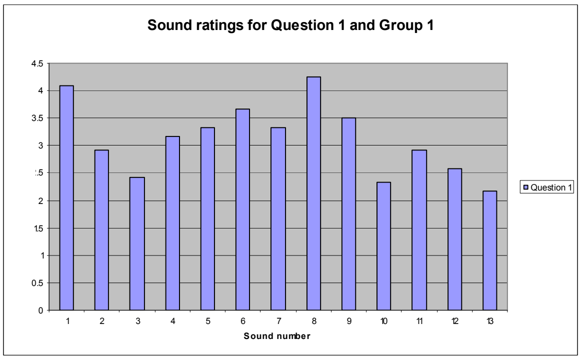

Figure 5. *Sound ratings for Question 1 and Group 1.*

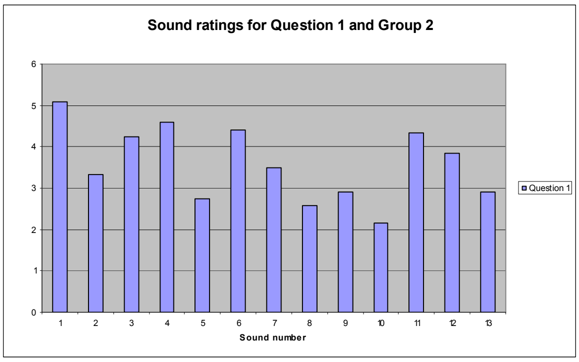

Figure 6. *Sound ratings for Question 1 and Group 2.*

A number of interesting pieces of information can be extracted from the figures observed:

- The three versions of the third sound group (Sounds 11, 12 and 13) were always rated consistently between each other *i.e.*, the rating for Sound 11 was always higher than that of Sound 12, which was itself higher than that of Sound 13.
- There were striking similarities between the contours of some of the graphs. The most noticeable of these is the fact that Figure 7 and Figure 9 show the same contour.
- The overall ratings for each sound group were significantly different in Group 3: Sound group 3 was rated significantly higher than Sound group 1 ($t_{12} = 0.04$) and Sound group 1 was rating significantly higher than Sound group 2 ($t_{12} = 0.02$). For Group 2, the same pattern was identified, but it was not as acute. Significant differences were only measured for this group on Question 2 between Sound group 2 and Sound Group 3 ($t_{12} = 0.045$)

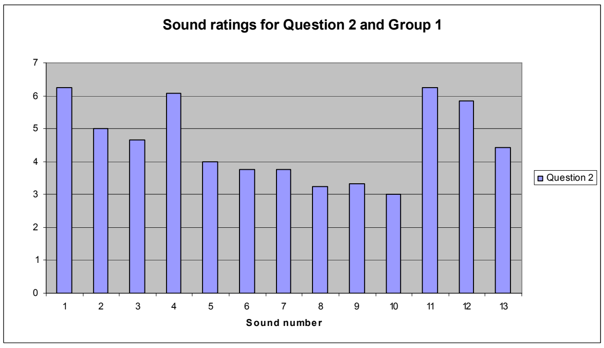

Figure 7. *Sound ratings for Question 2 and Group 1.*

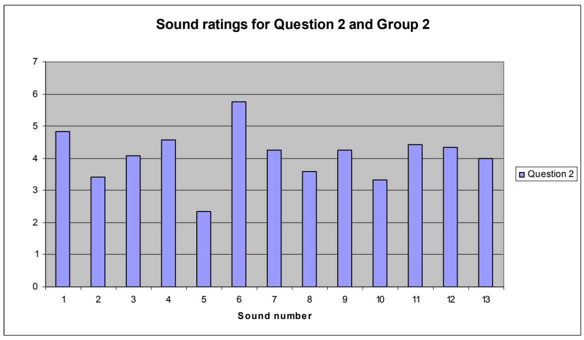

Figure 8. *Sound ratings for Question 2 and Group 2.*

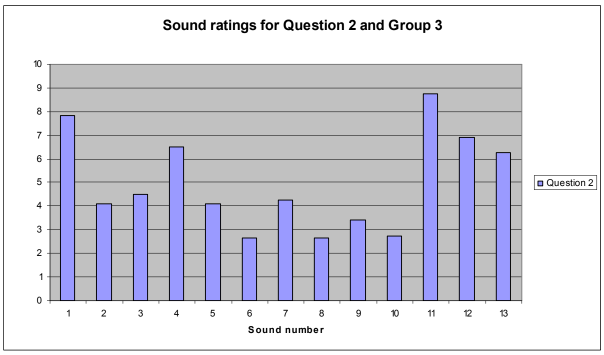

Figure 9. *Sound ratings for Question 2 and Group 3.*

These results confirm the pattern observed so far in the analysis, in that the judgments made for Question 2 were similar between Groups 1 and 3, but more acute for Group 3.

4.7.4. Ratings within sound groups

In the previous section, the ratings of each sound when paired with every other sound were analysed. The current section investigates the ratings of each sound when paired with a sound of the same sound group. There are two aims in this analysis:

- Find out whether the overall popularity of a sound is consistent with the popularity of a sound within its sound group.
- Identify the design parameters responsible for the rating differences within each sound group, if possible.

	Gr1, q1	Gr2, q1	Gr1, q2	Gr2, q2	Gr3(q2)
Sound group 1					
1	+	+	+	+	+
2	-	=	-	=	-
3	-	=	-	=	-
4	=	=	+	+	+
5	=	-	-	-	-
Sound group 2					
6	=	+	=	+	-
7	=	+	+	=	+
8	=	-	=	-	-
9	+	=	=	=	=
10	-	-	-	-	=
Sound group 3					
11	=	+	+	=	+
12	=	=	+	=	=
13	=	-	-	=	-

Table 2. *Comparison table for sound ratings within Sound groups. "+" means that the sound is, in average, preferred to sounds it has been paired with. Conversely, "-" means that the sounds it was paired with were preferred. "=" means that the preference was balanced. Statistically significant differences are indicated in bold. Greyed out "+" and "-" indicate a trend that approached significance. Gr refers to Group and q to Question.*

The most blatant result apparent in Table 2 is that Sound 1 was significantly preferred to other sounds in its group, for all subject groups and both questions. One can argue that the rhythmic pattern and tempo used were the cause of this. However, Sound 6, which is the equivalent of Sound 1 in the second sound group, was not rated so favourably. On the other hand, the least liked sounds were consistent across the sound groups: The shortest sound in each group (Sounds 5, 10 and 13) were rated in an equally negative way by all groups. In addition, in Sound groups 1 and 2, the third sound was also rated negatively. The second and third sound of the first two sound groups are the only two sounds with an irregular rhythm. This suggests that regular rhythms are preferred to irregular rhythms. However, the ratings for Sound 7 do not follow that trend.

The main results of this analysis are summarized below:

- The first sound was consistently the preferred sound in its group.
- The shortest sound of each sound group was consistently the least preferred sound. This suggests that too high a density of information can be detrimental to the aesthetic judgment made by users.
- Regular rhythms are often preferred to irregular rhythms
- The differences observed were more acute for the third group of subjects.

4.7.5. Other considerations

A number of parameters entered into the design of the sounds investigated in this experiment. Considering the format of the scale of the experiment, it is difficult to make conclusive remarks regarding the impact of all the design parameters on the aesthetic judgments made by the subjects. This section reviews the various parameters that have not been addressed so far in the analysis:

- *Reverb:* This parameter was used to represent the number of unread messages in the mailbox. No evidence regarding its contribution to the judgments made were found.
- *Speech quality:* The differences found between Sound group 1 and Sound groups 2 and 3 can be partly attributed to the quality of speech sounds used. Sound groups 1 and 3 used recorded speech whereas Sound group 2 used medium quality speech synthesis. Feedback provided by subjects at the end of the experiment confirmed that they found synthesized speech unpleasant.
- *Attachment sounds:* Again, given the number of design parameters involved in the experiment, the effect of the type of attachment sounds used could not be assessed.

5. CONCLUSION AND FUTURE WORK

Auditory interface aesthetics is an issue recognized by the AD community, although tackling it has always proven a challenge. This paper has illustrated the difficulty of tackling the problem, even in a very basic case study and has uncovered some aspects of the problem by investigating the relationship between the perceived aesthetics of a simple auditory interface and tasks carried out with that interface.

From a methodological perspective, this study demonstrated that it is possible to obtain meaningful empirical results without preliminary training of the subjects. Additionally, it showed that the problematic issue of defining aesthetics could be avoided by asking subjects to compare sounds, rather than asking them to rate abstract qualities of the sounds investigated. The study also revealed that basic rhythmic elements of the design had an impact on the judgments made by subjects.

More generally and more importantly, the study has provided insight into the nature of the relationship between the functional and aesthetic properties of auditory interfaces. For instance, the level of difficulty of a task can be regarded as a factor that influences the functional/aesthetic judgment of the interface made by users: a poor functional rating for a sound, due to a difficult task, may result in a poor aesthetic rating too.

Given the importance of interaction context and tasks on the perceived aesthetics of auditory interfaces, which this paper has only started to tackle, a logical step forward would involve investigating ways to clearly articulate the relationship between sounds and tasks/context. Subsequently, it should be easier to make a distinction between the elements of a design that are linked to potential tasks and those that are independent from the tasks. In other words, an effort should be made to articulate the attributes of an auditory interface that are purely functional, those that are purely aesthetic, and those that contribute to both. This involves far more than simply declaring the mapping between information and sound and must be tackled in order to promote the design of aesthetically pleasing auditory interfaces.

This research could lead to the production of extensions to current sound design guidelines or sound design patterns.

6. REFERENCES

[1] B. N. Schenkman and F. U. Jonsson, "Aesthetics and preferences of web pages communication issues," *Behaviour and Information Technology*, vol. 19, pp. 366-377, 2000.

[2] J. Cohen, "Out to lunch: Further adventures monitoring background activities," presented at Proceedings of ICAD'94, Santa Fe, 1994.

[3] W. Gaver, "Auditory Interfaces," in *Handbook of Human-Computer Interaction*, M. G. Helander, T. K. Landauer, and P. Prabhu, Eds. Amsterdam: Elsevier Science, 1997.

[4] C. Thornton, A. Kolb, F. Gemperle, and T. Hirsch, "Audiocentric interface design: a building blocks approach," presented at Proceedings of ICAD2003, Boston, MA, USA, 2003.

[5] R. Peterson and T. D. Wilson, "Measuring customer satisfaction: fact and artefact," *Journal of the Academy of Marketing Science*, vol. Winter, 1992.

[6] G. Lindgaard and C. Dudek, "What is this evasive beast we call user satisfaction?," *Interacting with Computers*, vol. 15, pp. 429-452, 2003.

[7] "9241-1 ISO Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability,"., 1998.

[8] N. Tractinsky, A. S. Katz, and D. Ikar, "What is beautiful is usable," *Interacting with Computers*, vol. 13, pp. 127-145, 2000.

[9] G. Cockton, "From doing to being: bringing emotion into interaction," *Interacting with Computers*, vol. 14, pp. 89-92, 2002.

[10] J. Klein, Y. Moon, and R. W. Picard, "This computer responds to user frustration: Theory, design and results," *Interacting with Computers*, vol. 14, pp. 119-140, 2002.

[11] J. Kim and J. Moon, "Designing towards emotional usability in customer interfaces," *Interacting with Computers*, vol. 10, pp. 1-29, 1998.

[12] R. W. Picard and J. Klein, "Computers that recognise and respond to user emotion: theoretical and practical implications," *Interacting with Computers*, vol. 14, pp. 141-169, 2002.

[13] J. Scheirer, R. Fernandez, J. Klein, and R. W. Picard, "Frustrating the user on purpose: a step toward building and affective computer," *Interacting with Computers*, vol. 14, pp. 93-118, 2002.

[14] F. E. Sparshot, *The structure of Aesthetics*: Routedge and Kegan, 1963.

[15] K. Karvonen, "The beauty of simplicity," presented at Proceedings of CUU'00, Arlington VA, USA, 2000.

[16] "Ecommerce Trust Study": A joint research project by Cheskin Research and Studio Archetype/Sapient, 1999.

[17] J. Khaslavsky and N. Shedroff, "Understanding the seductive experience," *Communications of the ACM*, vol. 42, pp. 45-49, 1999.

[18] N. Tractinsky, "Aesthetics and apparent usability: Empirically assessing cultural and methodological issues," presented at Proceedings of CHI'97, Atlanta GA, USA, 1997.

[19] G. Lemaitre, P. Susini, S. Winsberg, and S. McAdams, "Perceptively based design of new car horn sounds," presented at Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, 2003.

[20] F. A. Russo, M. E. Lantz, G. W. English, and L. L. Cuddy, "Increasing effectiveness of train horns without increasing intensity," presented at Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, 2003.

[21] M. McLuhan, *The Gutenburg Galaxy: The Making of Typographic Man*. Toronto: University of Toronto Press, 1962.

[22] W. Ong, *Orality and Literacy: The Techonologizing of the World*. London and New York: Methuen, 1982.

[23] E. Somers, "Abstract sound objects to expand the vocabulary of sound design for visual theatrical media," presented at Proceedings of ICAD'2000, Atlanta, GA, USA, 2000.

[24] W. Gaver and G. Mandler, "Play it again, Sam: on liking music," *Cognition and Emotion*, vol. 1, pp. 259-282, 1987.

[25] S. Hart and L. Staveland, "Development of NASA TLX (Task Load indeX): Result of empirical and theoretical research," *Human Mental Workload*, pp. 139-183, 1988.

[26] S. Brewster, P. C. Wright, and A. D. N. Edwards, "An evaluation of earcons for the use in auditory computer-human interaction," presented at Proceedings of InterCHI'93, Amsterdam, 1993.

[27] G. Leplâtre, "The Design and Evaluation of non-Speech Sounds to Support Navigation in Restricted Display Devices," PhD thesis, *Department of Computing Science*. Glasgow, UK: University of Glasgow, 2002.

[28] G. Leplâtre and S. Brewster, "Designing non-speech sounds to support navigation in mobile phone menus," presented at Proceedings of ICAD'2000, 2000.

[29] S. E. Hudson and I. Smith, "Electronic mail previews using non-speech sounds," presented at Proceedings of CHI'96, Vancouver, Canada, 1996.