SPACE, THE FINAL FRONTEARCON: THE IDENTIFICATION OF CONCURRENTLY PRESENTED EARCONS IN A SYNTHETIC SPATIALISED AUDITORY ENVIRONMENT

David K. McGookin, Stephen A. Brewster

Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow
Glasgow, Scotland, G12 8QQ
{mcgookdk, stephen}@dcs.gla.ac.uk
www.dcs.gla.ac.uk/∼{mcgookdk, stephen}

ABSTRACT

Two experiments which investigate the impact of spatialised presentation on the identification of concurrently presented earcons are described. The first experiment compared the identification of concurrently presented earcons based on the guidelines for individual earcon design and presentation of Brewster, Wright and Edwards [1] which were presented in spatially distinct locations, to the identification of non-spatially presented earcons which incorporated guidelines for concurrent presentation from McGookin and Brewster [2]. It was found that a significant increase in earcon identification occurred, as well as an increase in earcon register identification when earcons were spatially presented. The second experiment compared the identification of concurrently presented earcons based on the guidelines of Brewster, Wright and Edwards [1] which were presented in spatially distinct locations, to the identification of spatially presented earcons which incorporated guidelines for the presentation of concurrent earcons from McGookin and Brewster [2]. The incorporation of the concurrent earcon guidelines was found to significantly increase identification of the timbre attribute but did not significantly effect the overall identification of earcons.

1. INTRODUCTION

The concurrent presentation of auditory sources in an auditory display can allow both increased data presentation rates to users as well as allow for comparisons to be made between multiple data [3, 4]. However, such concurrent presentation can be problematic as auditory sources may interfere with each other, making it difficult to identify each individual source. As Norman [5] has pointed out on concurrently presented warning alarms, *"they often conflict, and the resulting cacophony is distracting enough to hamper performance"*. Such interferences are likely to be most problematic for closely related (along auditory dimensions) sounds where information is encoded within those sounds. Earcons are one such class of sound that exhibit problems when concurrently presented. Unfortunately, although a large body of work exists in psychology investigating how the human auditory system interprets multiple concurrently presented sounds [6], there is little guidance on how auditory displays can be designed to reduce such unwanted interactions between concurrently presented sounds, making it difficult for designers to exploit the advantages of concurrent presentation. Previously undertaken work [2, 7] has investigated how Auditory Scene Analysis [6] principles can be applied to the design of earcons to improve their identification. However, this work failed to take account of the impact of spatialised sound presentation, which is a key feature of ASA [6], and should therefore assist in the identification of concurrently presented earcons. There were several reasons for this, notably many current mobile devices do not support high quality spatialisation. Additionally when spatialising concurrently presented sounds, the impact of spatialisation is related to the degree of spatial separation that can be achieved between sounds. This, in many applications (where spatial location is mapped to a data parameter), may make it difficult to separate concurrently presented sounds enough. Hence the advantage of looking at the impact of spatialisation in isolation from other factors.

This paper investigates the relationship between the previously determined guidelines of concurrent earcon design and presentation (hence forth referred to as CEG) from McGookin and Brewster [2, 7], and spatialised earcon presentation. The CEG have already been shown to significantly improve earcon identification over a set of earcons which were formed from the guidelines of Brewster, Wright and Edwards [1] (BWE). The purpose of this paper is therefore to identify if presenting BWE earcons in spatial locations is superior to non-spatially presented earcons incorporating the CEG, and if so, can the CEGs further improve concurrent earcon identification is spatialised environments.

2. EARCONS AND AUDITORY SCENE ANALYSIS

Earcons are *"abstract synthetic tones that can be used in structured combinations to create auditory messages"* [8]. There are several types of earcon, but the most powerful types are formed from manipulations of various auditory attributes such as timbre, rhythm, pitch etc, according to a set of rules, or "grammar". For example a particular rhythm played with a particular musical instrument may represent a word processing file, but played with a different timbre may represent a spreadsheet file. This allows a large set of descriptive sounds to be built from relatively few learned rules ("grammar"). This also means that earcons derived from the same grammar are likely to share several components such as their timbre or rhythm, making them susceptible to interfering with each other in undesirable ways. According to auditory scene analysis they will merge when concurrently presented and make it difficult to identify the data encoded in each earcon. Whilst modifying the earcons to have increased differences between concurrently presented examples is possible, there is a limited extent to such modifications as care must be taken not to modify the earcons such that the "grammar" that relates earcons together is destroyed. This means that if two earcons are concurrently presented, which share the same timbre, but perhaps have different melodies, where melody and register are mapped onto data parameters, it is not possible to arbitrarily change the timbres to promote separation of the earcons. If arbitrary modifications were introduced, even if this improved separation of the earcons, they would destroy the mapping between the earcon and the data that it represents. Hence, whilst ASA can be used to improve concurrent earcon identification, the modifications that can be applied are limited. The CEG from McGookin and Brewster [2] have identified that incorporating a 300ms onset-to-onset gap between the starts of concurrently presented earcons and presenting each earcon with a different timbre is effective at improving the identification of concurrently presented earcons. However, this research ignores separation in space which is a major factor in how sounds are grouped by the auditory system [6].

The research described in this paper investigates how the identification of concurrently presented earcons is affected by presenting them in a spatial audio environment. We also investigate how the guidelines on concurrent earcon presentation from McGookin and Brewster [2] perform in a spatialised environment.

3. SPATIAL VS. NON-SPATIAL EARCONS

3.1. Motivation

Previous research [2] has shown that the identification of concurrently presented earcons based on the guidelines of Brewster, Wright and Edwards (BWE) [1] can be significantly improved by presenting each earcon with a unique timbre as well as staggering the starts of each earcon by at least 300ms. It is believed that presenting each earcon in a different spatial position by the use of a synthetic 3D audio environment will further improve the identification of concurrently presented earcons. In order to determine if this is the case, the experiment in this section attempts to determine if participants can identify earcons complying only with the guidelines of BWE [1] but with each presented in a different spatial location around a participant's head, better than earcons based on the guidelines on earcon design from BWE [1] which also incorporate the guidelines on concurrent earcon design and presentation by McGookin and Brewster (CEG) [2], but with all earcons presented in a common spatial location. This is important since, if the non-spatial presentation of concurrent earcons is superior to the spatial presentation of concurrently presented earcons complying only with the guidelines of BWE, this would mean that devices which might use concurrent earcons would not need to support the computational requirements of spatialisation to effectively present concurrent earcons. On the other hand, if spatialisation is superior to the concurrent presentation guidelines, this may remove the requirement for designers to have multiple similar timbres (one for each concurrently presented earcon), and the restrictions this imposes [7].

3.2. Methodology

Sixteen participants undertook the experiment described in this section which was of a within groups design and involved two conditions, the *spatial* condition and the *non-spatial* condition. Written consent was obtained from all participants prior to the experiment, and all participants were paid £5 on completion. Participants were randomly assigned to one of two groups to determine the order in which they undertook the conditions. Each group contained the same number of participants. The conditions are fully explained below. Each condition consisted of a training and a testing phase. The order of conditions for each group is shown in Table 1. The main hypothesis investigated was that concurrently presented earcons complying only with the guidelines of BWE in a spatial audio environment would be significantly easier to identify than earcons which also incorporated the CEG from McGookin and Brewster [2, 7] on concurrent earcon presentation, but were presented in a common spatial location. The independent variable (IV) was the earcon set and presentation method, the dependant variables (DV's) were the number of earcons and their attributes correctly identified. In addition we were interested in comparing the effort required of participants between the two conditions, modified NASA task load index (TLX) questionnaires [9] were used to evaluate this.

	First Training Session	**First Testing Session**	**Second Training Session**	**Second Testing Session**
Group 1	Spatial Condition	Spatial Condition	Non-Spatial Condition	Non-Spatial Condition
Group 2	Non-Spatial Condition	Non-Spatial Condition	Spatial Condition	Spatial Condition

Table 1: Table showing the procedure for the two groups of participant undertaking the spatial vs. non-spatial experiment.

3.2.1. *Earcons Used*

The earcons used in this experiment were of the transformational type [10] and were the same as those used in previous work investigating non-spatialised concurrent earcon identification [2]. Each earcon represented a ride in an imaginary theme park. Each earcon encoded three parameters of a theme park ride. These attributes and their possible values are described in Table 2.

The attributes from Table 2 were encoded into Earcons according to the guidelines of BWE [1], and were recorded from the output of a Roland Super JV-1080 MIDI (Musical Instrument Digital Interface) synthesiser. Although the guidelines of Brewster, Wright and Edwards note the differences between the various auditory parameters required to make earcons useful, they provide little guidance on which data parameters should be mapped to which auditory parameters. Norman [11] notes a difference between additive and substitutive dimensions, i.e. those dimensions where there is some concept of linear ordering, for example price, and substitutative dimensions, where there is only choice amongst many such as ride type. In the mapping of theme park rides to earcons, wherever possible, substitutive data attributes were mapped to substitutive auditory parameters, and additive data attributes were mapped to additive auditory parameters.

Ride type was mapped to timbre. Three distinct timbres were used, with a grand piano (General MIDI patch No. 000), used to

Attribute	Possible Values	Description
Ride Type	Rollercoaster, Water Ride, Static Ride	Categorises theme park rides by their properties
Ride Intensity	Low Intensity, Medium Intensity, High Intensity	How intense the ride is. Large, fast rollercoasters would be categorised as high intensity, whereas a miniature railway designed to transport customers around the park would be of low intensity.
Cost	Low Cost, Medium Cost, High Cost	How much it costs to be admitted to the ride.

Table 2: Table showing the attributes and their values encoded in the earcons used in the experiment.

represent Rollercoasters, a violin (General MIDI patch No. 040) being used to represent Water Rides, and a trumpet (General MIDI patch No. 056) being used to represent a Static Ride.

Ride Intensity was mapped to Melody, which is a combination of a rhythm and a pitch structure for that rhythm. The Melodies used for high, medium and low intensity rides are shown in Figure 1.

Figure 1: Melodies used to represent high, medium and low intensity theme park rides.

Finally the cost of the ride was mapped to the register that the melody was played in. Although the BWE guidelines [1] generally advise against register, we have used the gross differences between the registers that guidelines the recommend, as well as staggering the notes for each melody in each register to avoid the same melodies having musically harmonic intervals with each other when played in different registers. Register was mapped in such a way that the low cost rides were played in the lowest register, medium cost rides were played in the next highest cost register and high cost rides were played in the highest register. The registers used respectively were the octave of C4 for low cost, the octave of C5 for medium cost and the octave of C6 for high cost.

3.2.2. *Training Phase*

The training phase provided participants with an opportunity to learn the earcons appropriate to the condition they were to perform in the testing phase. This involved a sheet describing the grammar of the earcons used, followed by participants having ten minutes of self guided training of the earcons via a Web page interface. After this time participants were asked to identify three independently presented earcons without any form of assistance or feedback. If they were unable to do so, a further five minutes of training was provided after which the test would be retaken. If participants were still unable to correctly identify the three test earcons, the participant took no further part in the experiment. When participants had successfully completed this phase they carried onto the testing phase which is described below.

3.2.3. *Testing Phase*

The testing phase comprised of participants listening to twenty sets of four concurrently presented earcons and trying to identify the attributes of each earcon. Each set of earcons were concurrently played seven times. Variations in the presentation of earcons occurred between the two conditions of the experiment, these are described in the appropriate sections where the conditions are discussed. Participants recorded their selection in a computer based dialogue box as shown in Figure 2. Participants were given no feedback as to the correctness of their responses. The same twenty stimuli sets were used for both conditions, but were randomly presented to avoid any learning effect which may confound the results. After each testing phase, each participant completed a modified NASA TLX questionnaire [9] to measure their subjective workload.

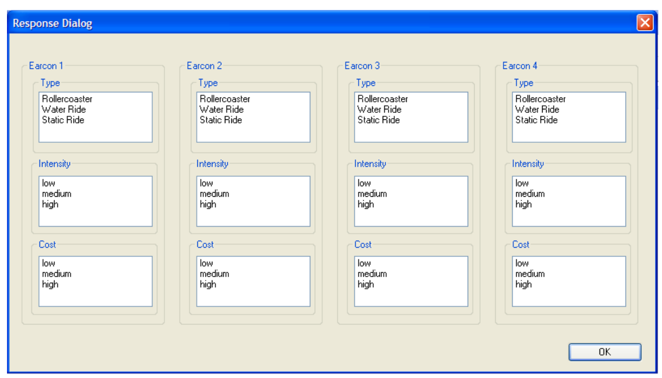

Figure 2: A screenshot of the dialogue used by participants to fill in the earcons and their attributes identified.

3.2.4. *Spatial Condition*

The spatial condition used the earcons described in Section 3.2.1, with each presented in a different spatial location on a lateral plane collocated with the listener's ears, via a head related transfer function (HRTF). Due to the lengthy procedure and specialised equipment required to create individualised HRTFs, and since many of the potential applications of this work (such as mobile devices) would make it difficult to produce individualised HRTFs, the sounds have been spatialised using a generalised HRTF. The GHRTF used was that found on the PURE Digital SonicFury PC sound

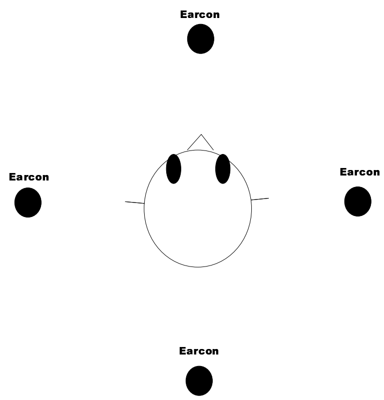

Figure 3: An illustration showing how the earcons used in the spatial condition were located relative to the listener's head.

card [1].

Each earcon was placed approx 20cm from the listener's head, one at each of the four main points of the compass. The placement of the earcons is summarised in Figure 3.

In order to overcome some of the problems of GHRTFs, a Polhemus Fastrak 6 degree of freedom tracking device [12] was used to dynamically respatialise the sounds relative to the participants as they moved and rotated their heads.

3.2.5. Non-Spatial Condition

The non-spatial condition used earcons which were based on those from Section 3.2.1, but also incorporated the concurrent earcon guidelines from McGookin and Brewster (CEG) [2]. Hence, although earcons were presented concurrently, each started 300ms after the previous one. In addition, instead of having one timbre for each ride type, each ride type had three distinct but similar (according to the guidelines of Rigas [13]) timbres which could be used interchangeably. This meant that in no set of four concurrently presented earcons would the same timbre be used for two earcons (there were no instances in the twenty sets of four earcons where all four earcons had the same ride type). The timbres used are shown in Table 3. All earcons were presented in a common spatial location, and as with the spatial condition a 1.5 sec break was incorporated between consecutive presentations of the earcons.

3.3. Results

3.3.1. Identified Earcons and Attributes

The number of correctly identified ride types, ride intensities and ride costs were collected, and from this the number of correctly identified earcons was determined. These data are summarised graphically in Figure 4.

To determine if any of the differences shown in Figure 4 were statistically significant, four within groups t-tests were carried out, one on the number of earcons correctly identified and one each on the number of correctly identified ride types, ride intensities and ride costs.

The t-test on the number of correctly identified earcons showed that participants identified significantly more earcons in the spatial condition than in the non-spatial condition (t(15) = 2.61, p = 0.020). The t-test for the number of correctly identified ride costs showed that participants identified significantly more ride costs in the spatial condition than in the non-spatial revised condition (t(15) = 2.37, p = 0.031). The t-tests on the number of correctly identified ride types (t(15) = 0.55, p = 0.591) and the number of correctly identified ride intensities (t(15) = 1.42, p = 0.176) were not significant.

Ride Type	Instrument	MIDI Patch No.
Rollercoaster	Acoustic Grand Piano	000
Rollercoaster	Electric Acoustic Piano	001
Rollercoaster	Electric Grand Piano	002
Static Ride	Tuba	058
Static Ride	French Horn	060
Static Ride	Synth Brass 1	062
Water Ride	Marimba	012
Water Ride	Shamisen	106
Water Ride	Kalimba	108

Table 3: Table showing the timbres used to represent ride types in the multi-timbre earcon set condition.

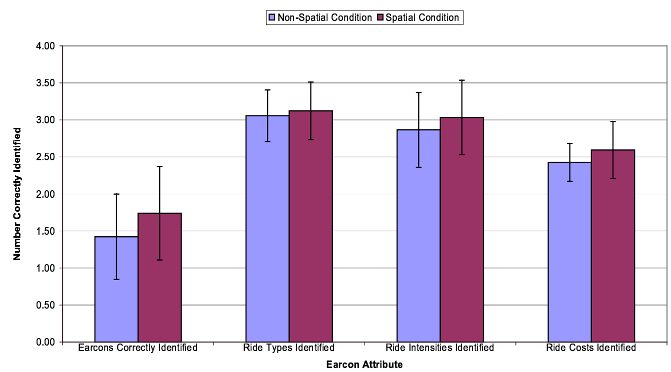

Figure 4: Graph showing the mean number of earcons, number of correctly identified ride types, ride intensities and ride costs for the non-spatial and spatial experimental conditions. Shown with standard deviations.

3.3.2. Workload Data

In addition to collecting data about the number of earcons and their attributes that were correctly identified, participants also completed modified NASA TLX questionnaires [9] for each condition. A summary of these data are presented in Figure 5. Overall TLX workload, calculated from a summation of each participants score for each data attribute (excluding annoyance and overall preference) was not found to be significantly different between the two conditions (t(15) = 1.60, p = 0.130). In order to determine if either annoyance or overall preference were significant, two within

[1]This card is marketed as the Turtle Beach SantaCruz in the USA

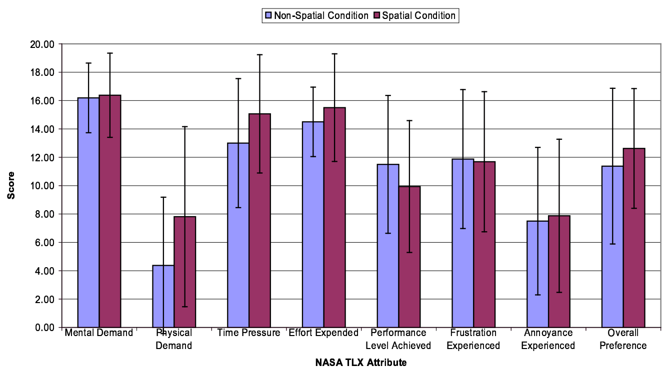

Figure 5: Graph showing the mean values for NASA TLX workload data for the non-spatial and spatial experimental conditions. Shown with standard deviations.

groups t-tests, one for each attribute, were carried out.

Participants also did not report significantly different annoyance experienced ($t(15) = 0.30$, $p = 0.765$), or express a significant overall preference for either condition ($t(15) = 0.59$, $p = 0.564$).

3.4. Discussion

The results show that having a unique spatial location for each concurrently presented earcon can significantly improve earcon identification over concurrently presented earcons which are spatially collocated. This holds even when the earcons which are spatially collocated have been modified to make them more easily identifiable than the earcons which are at spatially unique locations.

The results also show that the identification of ride cost is significantly improved in the spatial condition. This may be due to cases in the non-spatial condition where two earcons which differ only in ride cost (which is encoded as the register the earcon is played in) are concurrently presented, and the timbre and melody of the earcons then dominate in the grouping process, fusing the two sounds together and thus forming a composite pitch in a similar way to the problem of the missing fundamental [14]. As stated in Section 3.2.1, registers were chosen to avoid harmonic intervals between the earcons, however this may not have provided a strong enough grouping effect to dominate the timbre and register similarity of such earcons.

In conclusion, spatially separating the locations of concurrently presented earcons allows the listener to correctly group earcon attributes as whole earcons more easily than when all earcons are located in the same spatial location. If both spatial presentation and the concurrent earcon guidelines from McGookin and Brewster [2] were combined, would a further increase in concurrent earcons be observed?

4. SPATIAL VS. REVISED SPATIAL EARCONS

4.1. Motivation

The previous experiment has shown the importance of spatialisation when presenting concurrent earcons. The presentation of "classic" earcons, based only on the guidelines of Brewster, Wright and Edwards (BWE) [1] which were presented in a spatial audio environment, significantly outperformed non-spatially presented earcons which had been revised based on the guidelines for concurrent earcon presentation (CEG) from McGookin and Brewster [2].

This result indicates a possible divergence of guidance for designers to improve the identification of concurrently presented earcons. One possibility, is that if spatial audio presentation is available, earcons should be presented in spatially distinct locations and be designed strictly to the guidelines of BWE [1]. If spatial audio presentation is unavailable, earcons should be designed with the guidelines of BWE, as well as the concurrent earcon guidelines (CEG) from McGookin and Brewster [2]. Alternately, the identification of earcons in a spatial audio environment may be further improved by the incorporation of the guidelines of concurrent earcon identification. Since as shown by Best, vanSchaik and Carlile [15], spatial location is not a totally dominating factor in auditory scene analysis, and in applications which seek to encode some parameter of data as the position of an auditory source, sufficient spatial separation of earcons to avoid interference with each other may not be possible. It is important therefore to determine how well concurrently presented spatialised earcons which incorporate the concurrent earcon guidelines are identified.

The experiment described in this section therefore, investigates the impact on the identification of concurrently presented spatialised earcons based only on the guidelines of BWE [1], compared to concurrently presented spatialised earcons which also incorporate the CEG from McGookin and Brewster [2].

4.2. Methodology

The procedure and methodology of this experiment is largely the same as that of the experiment described in Section 3. Sixteen different participants undertook the experiment. The experiment was of a within groups design, with eight participants randomly assigned to one of two groups. There were two conditions, the *"spatial(2)"* condition and the *"revised spatial"* condition. The conditions are explained below. Conditions were counterbalanced to avoid order effects. Each condition consisted of a training and testing phase. The order of the conditions for each group of participants is summarised in Table 4. Written consent was obtained prior to the experiment from all participants, each was paid £5 on completion.

The main hypothesis investigated was that concurrently presenting earcons complying with both the guidelines of BWE [1] and the non-spatial CEG from McGookin and Brewster [2] in a spatial audio environment would make identification of earcons and their attributes significantly easier than spatially presented earcons complying only with the guidelines of Brewster, Wright and Edwards [1]. The independent variable (IV) was the earcon set used, the dependant variables (DV's) were the number of earcons and their attributes correctly identified. Subjective workload was measured as before.

4.2.1. Spatial(2) Condition

The spatial(2) condition, is, as its name suggests, the same as the spatial condition described in Section 3.2. The name has been changed to avoid confusion with the results from the spatial condition.

	First Training Session	**First Testing Session**	**Second Training Session**	**Second Testing Session**
Group 1	Spatial(2) Condition	Spatial(2) Condition	Revised Spatial Condition	Revised Spatial Condition
Group 2	Revised Spatial Condition	Revised Spatial Condition	Spatial(2) Condition	Spatial(2) Condition

Table 4: Table showing the procedure for the two groups of participant undertaking the spatial vs. modified spatial experiment

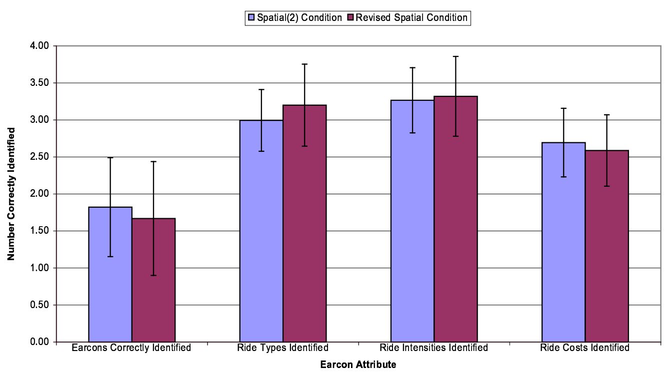

Figure 6: Graph showing the average number of earcons, number of correctly identified ride types, ride intensities and ride costs for the spatial(2) and revised spatial experimental conditions. Shown with standard deviations.

4.2.2. Revised Spatial Condition

The revised spatial condition was the same as the non-spatial condition from the previous experiment, but used the spatial placement of the earcons from the spatial condition.

Because of the 300ms onset-to-onset time difference between the start of each earcon, there would have existed a predictable pattern in the order of earcon presentation. Hence the spatial positions of the first, second, third and fourth earcons to be presented would always be the same. This may have lead to a learning effect which would have been undesirable and could confound the results. In order to overcome this issue, the spatial positions of the first, second, third and fourth earcons, were randomly alternated for each trial in the condition. However, the order of presentation was held constant during each trial.

4.3. Results

4.3.1. Identified Earcons and Attributes

As with the previous experiment, the number of correctly identified ride types, ride intensities and ride costs were collected, and from these the number of correctly identified earcons was determined. These data are summarised graphically in Figure 6.

In order to determine if any of the differences shown in Figure 6 were statistically significant, four within groups t-tests were carried out, one on the number of earcons correctly identified and one each on the number of correctly identified ride types, ride intensities and ride costs.

The t-test on the number of correctly identified ride types showed that participants identified significantly more ride types in the revised spatial condition than in the spatial(2) condition ($t(15) = 3.11$, $p = 0.007$). The t-tests for the number of earcons identified ($t(15) = 1.73$, $p = 0.104$), number of ride intensities identified ($t(15) = 0.81$, $p = 0.429$), and ride costs identified ($t(15) = 0.95$, $p = 0.356$), failed to show significance.

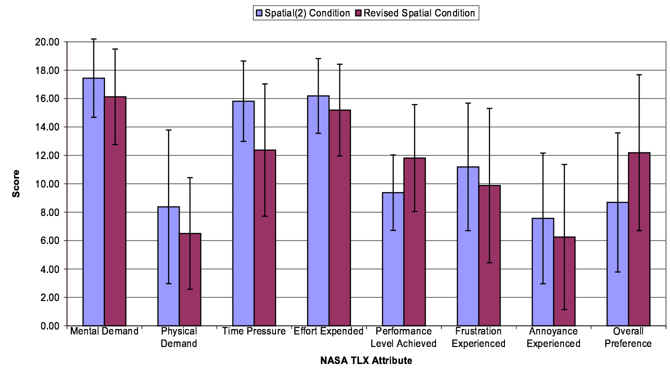

Figure 7: Graph showing the mean values for NASA TLX workload data for the spatial(2) and revised spatial experimental conditions. Shown with standard deviations.

4.3.2. Workload Data

In addition to collecting data about the number of earcons and their attributes that were correctly identified, participants also completed modified NASA TLX questionnaires [9] for each condition. A graphical summary of these data are presented in Figure 7. Overall TLX workload, calculated from a summation of each participants score for each data attribute (excluding annoyance and overall preference) was found to be significant ($t(15) = 3.32$, $p = 0.005$). In order to determine which of the differences were significant, eight within groups t-tests were carried out, one for each NASA TLX attribute.

Participants reported that performance level achieved ($t(15) = 3.36$, $p = 0.004$) was significantly higher in the revised spatial condition than in the spatial(2) condition. Participants reported significantly lower physical demand ($t(15) = 2.75$, $p = 0.015$) in the revised spatial condition than in the spatial(2) condition. Participants also reported significantly lower time pressure ($t(15) = 2.52$, $p = 0.024$) in the revised spatial condition than in the spatial(2) condition. The t-tests for mental demand ($t(15) = 1.50$, $p = 0.154$), effort expended ($t(15) = 1.25$, $p = 0.231$), frustration experienced ($t(15) = 1.18$, $p = 0.258$) and annoyance experienced ($t(15) = 1.52$, $p=0.150$), failed to show significance at the 5% level. Participants did not express a significant overall preference for either condition ($t(15) = 1.63$, $p = 0.123$).

4.4. Discussion

The results show that the identification of earcons in a spatialised audio environment can be improved by the application of the concurrent earcon guidelines from McGookin and Brewster [2]. The identification of ride type, encoded by timbre, was significantly improved in the revised spatial condition. In addition, the subjective physical demand of participants was significantly lower in the revised spatial condition, indicating that less head movement (using the headtracking device) was required for participants to perceptually separate the earcons. This may be useful in reducing fatigue if a display incorporating spatial earcons and headtracking is to be used for a prolonged period. Participants also reported significantly lower time pressure and higher perceived performance in the revised spatial condition, this may make any interface which uses such a technique a more enjoyable experience for users.

Although the number of ride types that were correctly identified was significantly greater in the revised spatial condition than in the spatial(2) condition, the actual number of earcons identified was not significantly different in either condition. In many ways this is predictable, given the previous work on non-spatial earcon modifications that the revised spatial condition incorporated [2], where modifications to the design and presentation of earcons tended to increase the number of earcon parameters that were successfully identified, rather than the total number of earcons identified. The significant reduction in physical demand for the revised spatial condition however, indicates that applying the concurrent earcon guidelines makes the task easier by requiring less physical movement of the participant's head in order to adequately perceptually separate the earcons. As discussed in Section 4.1, spatialisation is not a totally dominating parameter in ASA, hence having more cues available to separate different earcons from each other is advantageous in reducing the reliance solely on spatial location to separate earcons.

5. OVERALL DISCUSSION AND CONCLUSIONS

The two experiments described in this paper have significantly contributed to our knowledge of participants ability to identify earcons when they are concurrently presented. The first experiment described in this chapter investigated the effect of concurrent earcon presentation when those earcons were designed solely using the guidelines of Brewster, Wright and Edwards [1], but were presented in spatially distinct locations (the spatial condition), to the non-spatial presentation of concurrent earcons which were based on the guidelines of Brewster, Wright and Edwards [1], but also included the CEG from McGookin and Brewster [2], which have already been shown to significantly improve the identification of non-spatial, concurrently presented earcons compliant only with the guidelines of BWE (the non-spatial condition). Participants were found to have identified significantly more earcons in the spatial condition than in the non-spatial condition. These results showed that the separation of earcons in space is an effective way to concurrently present them. Indeed it is a superior method than if those earcons were non-spatially presented but included presentation and design modification which have already been shown to improve earcon identification in concurrent situations.

The second experiment answers the obvious question to come out of the results of the first: Can the use of earcons which incorporate the non-spatial concurrent earcon guidelines from McGookin and Brewster [2] be used in a spatial audio environment to further improve the identification of concurrently presented earcons? This is important as in some applications it may not be possible to sufficiently separate earcons in space to allow them to be uniquely identified, especially if sound source location is mapped to some data attribute (e.g. cartographic data). Best, vanSchaik and Carlile [15] have shown that interference may still occur with up to 60^o separation between concurrently presented audio sources. The experiment therefore compared identification of earcons based only on the guidelines of Brewster, Wright and Edwards [1], which were spatially presented (the spatial(2) condition), to earcons which were also spatially presented but incorporated the concurrent guidelines for earcon presentation of McGookin and Brewster [2] (revised spatial condition). The results showed that participants identified significantly more ride types in the revised spatial condition, and that participants subjective physical demand and time pressure were significantly lower in the revised spatial condition than the spatial(2) condition. The lower physical demand indicates that less physical movement of participants' heads was required when more cues for the separation of sounds existed, as was the case in the revised spatial condition.

From these experiments, the following guidelines, which extend those previously developed by McGookin and Brewster [2, 7] have been identified. These guidelines can be used for future designers of interfaces which use concurrent earcon presentation.

- The use of spatialised presentation with headtracking significantly improves the identification of concurrently presented earcons. Therefore spatialised presentation should be employed whenever practically possible in cases where earcons will be concurrently presented.
- Whilst spatial audio significantly increases the number of concurrent earcons that can be identified, it does not always guarantee that earcons will be adequately separated, hence the maximum amount of angular (in azimuth) separation between concurrently presented earcons should be used. The guidelines of McGookin and Brewster [2] should also be incorporated when using spatial presentation to improve earcon identification.

The experiments which have been described in this paper have shown how spatial location can be used to improve the identification of concurrently presented earcons. This should allow future designers of auditory displays to more effectively use concurrent presentation of earcons in their applications.

6. ACKNOWLEDGEMENTS

This work was supported by EPSRC studentship 00305222.

7. REFERENCES

[1] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "Experimentally derived guidelines for the creation of earcons," in *Proccedings of BCS-HCI '95*, Huddersfield, UK, 1995, vol. 2, pp. 155–159, Springer.

[2] D. K. McGookin and S. A. Brewster, "An investigation into the identification of concurrently presented earcons," in *ICAD 2003*, Boston, MA, 2003, pp. 42–46, ICAD.

[3] N. Sawhney and C. Schmandt, "Nomadic radio: Speech audio interaction for contextual messaging in nomadic environments," *ACM Transactions on CHI*, vol. 7, no. 3, pp. 353–383, 2000.

[4] L. Brown, S. Brewster, R. Ramloll, W. Yu, and B. Riedel, "Browsing modes for exploring sonified line graphs," in *BCS-HCI 2002*, London, UK, 2002, vol. 2, pp. 6–9, BCS.

[5] D. A. Norman, *The Design of Everyday Things*, vol. 1, MIT Press, Boston, Massachusets, 1 edition, 1998.

[6] A. S. Bregman, *Auditory Scene Analysis*, MIT Press, Cambridge, Massachusetts, 1994.

[7] D. K. McGookin and S. A. Brewster, "Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition," *to appear in ACM Transactions on Applied Perception*.

[8] S. A. Brewster, *Providing a structured method for integrating non-speech audio into human-computer interfaces*, Ph.D. thesis, Department of Computer Science, University of York, 1994.

[9] S. Hart and L. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Human Mental Workload*, P Hancock and N Meshkati, Eds., vol. 1, pp. 139–183. North Holland, Amsterdam, 1988.

[10] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.

[11] D. A. Norman, "Cognitive artifacts," in *Designing Interaction: Psychology at the Human-Computer Interface*, John M Carroll, Ed., Cambridge Series on Human-Computer Interaction, pp. 17–38. Cambridge University Press, Cambridge, 1991.

[12] Polhemus Inc, "http://www.polhemus.com," 2001.

[13] D. I. Rigas, *Guidelines for Auditory Interface Design: An Empirical Investigation*, Ph.D. thesis, Loughborough University, 1996.

[14] B. C. J. Moore, *An Introduction to the psychology of hearing*, Academic Press, London, 4th edition, 1997.

[15] V. Best, A. van Schaik, and S. Carlile, "Two-point discrimination in auditory displays," in *ICAD 2003*, Boston, MA, 2003, vol. 1, pp. 17–20, ICAD.