Audio Engineering Society

Convention Paper 6687

Presented at the 120th Convention
2006 May 20–23 Paris, France

Listener Opinions of Novel Spatial Audio Scenes

Kathryn Beresford[1], Francis Rumsey[1], and Slawomir Zielinski[1]

[1]Institute of Sound Recording, University of Surrey, Guildford, GU2 7XH, UK
k.beresford@surrey.ac.uk

ABSTRACT

Listener opinions for alternative approaches to recording multichannel classical music were investigated, particularly considering alternatives to the traditional approach. Recordings were made with pre-existing microphone arrays but alternative arrangements of musicians. These were used in a listening test to assess different attributes (timbral balance, envelopment, locatedness etc.).

From the results it was noted that naïve and trained listeners assessed the recordings in different ways. Through factor analysis, two components were identified to represent these assessments – creativity and conventionality. The naïve listeners indicated that purchasability was closely related to creativity whereas for the trained listeners, conventionality was an indicator of purchasability. A method for predicting purchasability was developed which may aid future work in the area.

1. INTRODUCTION

The multichannel audio industry is an area of continued technological development and growing consumer acceptance, particularly in the home environment with the introduction of DVD-Audio and Super Audio CD. Consumers are offered 'an emotionally engaging experience….that dwarfs the experience provided by conventional stereo' [1].

Music-only recordings in the multichannel environment have taken advantage of these developments, but despite the widespread adoption by many consumers of home theatre and multichannel systems, traditional practices and concepts are still very much in favour for commercial recording releases. The opportunity to make unusual but commercially viable classical recordings, which could be accomplished with some creative recording techniques, is often avoided.

1.1. Spatial audio for classical music

In their paper, Cohen et al [1] comment that in order for a multichannel format to become successful 'it must accommodate all kinds of music'. As this paper considers only classical music, the focus will be on this musical genre.

There are many opinions regarding the recording and reproduction of classical music. Theile [2] suggests that in its simplest form 'the reproduced sound image must correspond to the original sound image'.

One alternative suggestion by Rumsey [3] states that 'in classical music recording….the aim of high quality recording and reproduction should be to recreate as believable an illusion of "being there" as possible'.

As consumers now expect an exceedingly high standard of recorded music, Moorer [4] suggests that perhaps 'we have moved beyond realistic recordings to what we might term 'supernatural' recordings, that is, recordings so perfect that they could not have been produced in nature'.

Perhaps then, the idea that recording should be regarded as a 'different experience from natural listening' [3] is not so different from many of the recordings now available.

Theile suggests a compromise that should result in a pleasing reproduced sound; 'it should satisfy aesthetically and it should match the tonal and spatial properties of the original sound at the same time.' [2]

For cinema and home theatre sound the distribution of the soundfield is generally achieved through panpot control during the mixing stage. Sound effects, dialogue and music can be directed to any part of the loudspeaker setup to give the surround sound illusion [5].

In contrast, multichannel classical recordings usually focus on achieving distribution of the soundfield at the recording stage rather than the mixing stage. It is considered that the *3-2* stereo system is an appropriate configuration, using the surround channels to reproduce 'diffuse reverberation'. The primary signals generally expected to be presented over the rear loudspeakers are the ambient components of the recording [1]. The research conducted for this paper considers an alternative approach at the recording stage.

1.2. Multichannel reproduction

The system configuration specified in the International Telecommunications Union Recommendation BS.775-1 for 'Multichannel stereophonic sound system with and without accompanying picture' [6] indicates the loudspeaker layout and compatibility specifications, together with a variety of other suggestions for a *3-2* stereo system. The recommendation was constructed to allow compatibility with the *2-0* stereo system.

The addition of a low frequency effects channel (LFE) is included in the ITU recommendation but its use is optional for consumers and it is not used to convey the low frequency content of the main channels. The ITU recommendation states that 'the subwoofer channel is an option, at the receiver, and thus should only carry the additional enhancement information' [6].

This ensures that the consumer experience will not be inhibited if they decide not to use the LFE channel and that sound reproduction should be reasonable without the additional speaker [3]. The subwoofer is considered 'not important' by Kornacki et al. [7] for listening to classical music; its main use is for sound effects in film production. This holds relevance for the research conducted for this paper which did not use a subwoofer channel.

The *3-2* stereo loudspeaker configuration is not intended, or suited, for 'accurate 360° phantom imaging capabilities' [3] but does result in good envelopment, localisation and spatial illusions. The amplitude and time differences between the side pairs of loudspeakers are too small to create stable phantom images to the sides of the listener and the large angle between the rear pair of loudspeakers can result in a hole in the image to the rear of the listener.

Imaging is best between the 3 front loudspeakers, moderate to the rear of the listener and variable to the sides. This is a good compromise and is a considerable development and improvement over the 2-channel reproduction systems, particularly for the consumer market.

Additionally, the size of the listening area is increased when listening to music in surround in comparison to 2-channel stereo. This 'enhances the spatial listening experience' [3] and can incorporate a larger audience.

1.3. Multichannel recording

Microphone and mixing technique recommendations have not yet been specified although there are technical standards in place for the reproduction of spatial audio. The placement and positioning of microphones, ratio of direct to reverberant sound and the use of delays are of prime importance to the multichannel recording engineer [7].

A good multichannel recording should involve a delicate balance between the nature of the programme to be recorded, the technical means available and the aesthetic viewpoint of the engineer.

Rumsey [3] suggests that the challenge is to create a 'believable illusion' of a space, including its size, depth and the position of the recorded sources. The current opinion for classical music is that the 'recording is likely to adhere to the soundfield reconstruction school of thought' with the acoustic environment treated as part of the recording. Two and three channel techniques are a good basis for multichannel recordings. Most of the 2-channel techniques using spaced omnidirectional microphones are easily extended to three channels. Coincident microphone techniques can be more difficult to extend to three channel reproduction.

The majority of multichannel microphone array techniques are more complicated than similar 2-channel techniques, usually combining microphones to create a 'classic stereophonic sight' [7] with additional microphones to add depth and space to the recording.

There are two main styles of microphone array [3]:

- Singular array - all microphones in close proximity to one another. This type of array attempts to create phantom images for a 360° soundfield.

- Multiple arrays - treating the front and rear sound stages separately. One array is used for the front stage and another for ambient sounds.

1.4. Novel Spatial Audio Scenes

Various microphone techniques have been developed within these two styles but the majority use a traditional concert hall setup for recording classical ensembles.

Huber and Runstein described some mixing techniques that are currently used in professional situations for recording music in surround sound, but rarely used for commercial classical recordings [8]:

- a traditional mix in the front speakers with ambience and special effects in the rear speakers

- a move away from traditional concepts to a full 360° soundfield with instruments being placed anywhere within the mix

This paper investigates the opinions of listeners when traditional concepts were left behind. This was achieved through novel recording techniques, rather than post-production methods.

Using traditional spatial microphone techniques, but altering the artists' placement, novel spatial audio scenes were created. Listener opinions of these alternative scenes were gauged using listening tests.

2. METHOD

2.1. Recording Design

Multichannel recordings specifically tailored to the research were required for the listening test. These were made using two pre-existing singular microphone arrays (INA-5 and a Double ORTF + Centre [3, 7, 9 & 10]) and three musician arrangements. The musicians varied between a traditional concert hall setup, a partial surround set-up with the musicians located in a limited arc around the microphone array and a full surround setup with the ensembles completely surrounding the microphone array. Three different musical ensembles were used in the recordings – a chamber choir, an orchestral ensemble and a string quartet.

The recordings were carried out in Studio 1 of the Performing Arts Technology Studios (PATS) at the University of Surrey in Guildford. The audio was recorded onto digital multitrack tape via a digital mixing console and monitored in *3*-2 stereo format in Studio 3 of the PATS building. A combined microphone array was constructed on a single stand to allow both microphone arrays to be recorded simultaneously. This was elevated to a height of 1.6m for the INA-5 array and 1.9m for the ORTF array (difference in height is accounted for by the construction of the array, one set hanging from the structure and the other secured above).

All microphone gains were normalised with reference to one another to eradicate input level differences. The musical ensembles were arranged within a 2 – 5m radius (dependant on number of musicians) from the microphone array. The arrangement diagrams and photographs for the choir are shown in Figures 1 - 3.

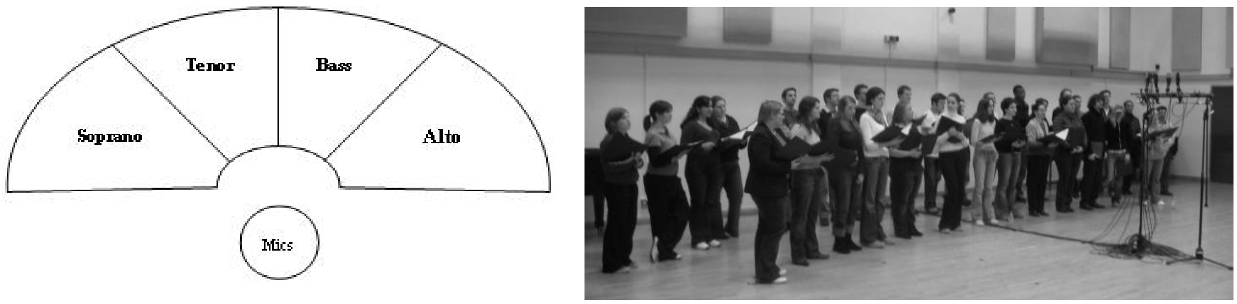

Figure 1 Choir – Traditional arrangement (A)

Figure 2 Choir - Partial arc arrangement (B)

Figure 3 Choir – Full surround arrangement (C)

The string quartet and orchestral ensemble were set out in a similar manner and music by each ensemble recorded with both arrays simultaneously.

Following the recordings a listening test was designed using representative extracts from each recording position.

2.2. Listening Test Design

The aim of the listening test was to compare certain attributes for the three musician arrangements (A, B and C) which were directly compared, using a 9-point Likert Scale to quantify the listeners' opinion. The use of a Likert Scale has several advantages. It provides a simple test procedure which can be used for all test questions; the rating scale is simple for both naïve and trained listeners and discrete answers from the test can easily be converted into quantitative data, making it ideal for analysis

The combination of programme material and microphone array, plus a repeat for a listener consistency study, resulted in 7 listening trials. In each trial, the listener was asked to assess the three musician arrangements. Eight sets of seven trials were used to answer eight questions comparing the different artist placements. This resulted in a total of 56 trials for each listener.

Four different randomised test orders were used for the tests. The question order, sample combinations (programme & array) and musician arrangements (traditional, partial arc and full surround) were randomised.

2.2.1. Listening Subjects

A mixture of naïve and trained listeners were used in the listening study, with an equal number of each listener type. All listeners had extensive experience with regards to classical music listening in the home environment.

The listeners were students and staff from the Department of Music and Sound Recording at the University of Surrey, Guildford. The students and staff from the Music and Sound Recording (Tonmeister) course were classed as trained listeners as formal technical listening training is received as part of the course. The students from the Music course were classed as naïve listeners.

2.2.2. Listening Test Environment

The listening test took place in Studio 3 of the PATS Building at the University of Surrey using the *3-2* stereo setup described in ITU-R BS.775-1 [6]. The audio was replayed from an audio editing program, using a digital mixing console as a D/A converter and level control between the computer and monitoring. This room was chosen for consistency of monitoring with the recording process.

The loudspeakers were aligned using broadband pink noise (at the level of -20dBFS). The level at the listening position was measured using an SPL meter and adjusted to 83dBC (slow weighting) for all speakers. The global playback level was adjusted to a comfortable level at the listening position and remained constant for all listeners.

2.2.3. Likert Scale and Question Design

The questions for this experiment aimed to discover listener preferences and related information regarding the placement of musicians around the surround microphone array. The following areas were covered – purchasability, listening comfort, tradition, interest, envelopment, naturalness, instrument locatedness and timbral balance. A statement was developed for each area of interest for use with a Likert Scale [11]. Two examples of the statements used are shown below, the question preceding the statement was used as a guide for the listener.

Envelopment – do you feel surrounded by the audio?

I find this recording enveloping.

Timbral Balance – is the bass, middle and treble content well balanced? Is the frequency spectrum pleasing?

The recording has a good timbral balance.

The Likert Scale was used by the listeners to show their agreement with the series of statements presented as part of the test. For each statement, 7 recording samples were presented. Three different musician arrangements (A, B and C) for each sample were compared. The listeners were instructed to show how strongly they agreed or disagreed with the presented statement for each arrangement – A, B and C. Figure 4 shows the scale used in the tests.

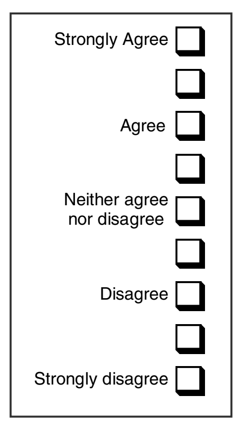

Figure 4 Likert Scale example

2.2.4. Listening Test Procedure

The test was carried out over 3 days, using 30 listeners. Each subject had time to read the instructions and was briefed about the use of the computer program. They were then able to progress through the test unaided. Subjects took between 45 minutes and 1 hour to complete the test.

Following the test, the results were entered into the statistical analysis program, SPSS for data analysis.

3. RESULTS

A brief preliminary analysis of the data suggested which specific analytical methods might yield interesting results. It was observed that differences existed between the assessment rationale of the naïve and trained listeners. Subsequently the data was divided into subsets – naïve and trained listeners. Multivariate analysis of variance (MANOVA), analysis of variance (ANOVA), correlation tests, factor analysis and regression were carried out on the data aggregated across listeners.

3.1. Summary of Preliminary Analysis

A preliminary analysis revealed that a change in Microphone Array resulted in a difference in Timbral Balance, most likely because of the different recording techniques and microphones used.

Programme Material affected Listening Comfort and Naturalness but this interaction was probably caused by bias rather than the different programme material used in the recordings. This bias may have been caused by the standard of playing on the recordings. Differences in Envelopment were noted for the different programme material which may have been due to the size and placement of the musicians within the ensembles. The choir recordings had much larger numbers of musicians than the quartet and orchestra, which affected the listeners' opinions.

The Musician Arrangements did not appear to have any effect on Listener Comfort, Instrument Locatedness or Purchasability. This was interesting as the Purchasability and Comfort were expected to alter more significantly with the Musician Arrangements. It was shown that the other five variables – Tradition, Interest, Instrument Locatedness, Naturalness and Timbral Balance – all have significant interactions with the Musician Arrangement.

The significant interaction between Array and Programme Material had the greatest magnitude of all the interactions in a MANOVA test and affected Envelopment and Purchasability. This is possibly biased by the quality of the recordings themselves for the Purchasability variable.

Musician Arrangement and Programme Material had a significant effect on Purchasability, Locatedness and Timbral Balance. These follow on from the individual effects, with Purchasability possibly being biased by the Programme Material more than the Musician Arrangement, which was shown to be non-significant in ANOVA tests. Locatedness and Timbral Balance are likely to have been affected by the Musician Arrangement more than the Programme Material.

The combination of Microphone Array and Programme Material had no significant effects on the variables.

3.2. ANOVA and MANOVA

Following on from the preliminary analysis, the data was split into two subsets – naïve and trained listeners – and ANOVA and MANOVA tests were carried out. The MANOVA data tables are shown in the Appendix.

For the naïve listeners, the MANOVA showed statistical significance, $p < 0.05$, between all interactions except Array*Arrangement. The Array*Programme interaction had the largest magnitude of these interactions (with $F(8, 246) = 11.37$, $p < 0.05$ and $\eta^2=0.27$).

The trained listeners showed similar results for the Array*Arrangement interaction but the significant interaction between Array and Programme showed a slightly greater magnitude of effect with $\eta^2=0.29$.

For the ANOVA tests, some variables had a greater magnitude of partial eta squared for naïve listeners and some for trained listeners. These demonstrate which variables are of importance to the two groups of listeners.

For the Microphone Arrays, the naïve listeners gave a greater magnitude of effect for the significant result for Listener Envelopment but the trained listeners gave larger magnitudes of effect for Purchasability and Timbral Balance. Trained listeners showed a significant interaction between the Array and Naturalness whereas the naïve listeners showed a nonsignificant interaction.

The interaction between Array and Programme for the trained listeners showed significance across all variables. The naïve listeners only showed significance for five of the eight variables, with nonsignificant interactions for Listening Comfort, Tradition and Interest.

Of great interest is the interaction between Programme Material and Musician Arrangement, which for naïve listeners showed significance for Purchasability but did not for trained listeners. The trained listeners did however show significance for the Timbral Balance variable whilst the naïve listeners showed a nonsignificant interaction. This implies that trained listeners are more attentive to the timbral quality of the recordings for Musician Arrangement and Programme Material. In contrast, naïve listeners concentrate on these factors where purchasability is concerned.

These findings indicate that naïve and trained listeners assess recordings in a different manner and that this might affect which recordings would be considered purchasable.

3.3. Correlations

The correlations between the individual variables for all the listeners were particularly informative. A high correlation between Purchasability and Timbral Balance can be seen. (A perfect correlation is indicated by a Pearson Correlation = 1, no correlation is indicated by a Pearson Correlation = 0) When the correlations for the two subsets of listeners were examined, it was seen that the correlation was much greater for the trained listeners (Pearson Correlation = 0.832) than for the naïve listeners (Pearson Correlation = 0.675).

For trained listeners, the correlation between Timbral Balance and Purchasability had the greatest magnitude of correlation of all the variables (Pearson Correlation = 0.832). This indicated that the timbral quality of the recording was the most important aspect for consideration the trained listeners when purchasing a recording. In addition, Naturalness (Pearson Coefficient = 0.756) and Listening Comfort (Pearson Coefficient = 0.675) were considered important by trained listeners – shown by their correlation with Purchasability.

For the naïve listeners, Timbral Balance gave the highest correlation with Purchasability (Pearson Correlation = 0.675) but this was not as high as for the trained listeners. The naïve listeners also considered Listener Envelopment of importance when purchasing a recording.

Other highly correlated variables included: Timbral Balance and Comfort with Tradition for naïve listeners; Naturalness with both Comfort and Tradition for trained listeners.

From these correlations several points can be concluded. Trained listeners show a high correlation between Naturalness, Listening Comfort and Timbral Balance with Purchasability. This shows that trained listeners place emphasis on comfort and naturalness when choosing recordings to purchase, with greatest emphasis on the timbral quality. Naïve listeners place more emphasis on envelopment as well as timbral quality but overall the correlations between variables for naïve listeners are less than for trained listeners.

3.4. Factor analysis

The correlation test was followed by factor analysis. The same three sets of data were explored (all listeners, naïve listeners and trained listeners). The results for the naïve and trained listeners are reported.

For naïve listeners, two components are able to account for 67% of variances in the data and 72% for trained listeners. (Using only two components allows the variables' relation to these components to be simply plotted in a 2 dimensional plot. Each component can be labelled to relate its significance to the data.)

A factor analysis plot in rotated space can show how each variable relates to the two components. The labels 'Conventionality' and 'Creativity' were chosen for the two components. These were selected by observing the variables which were correlated with the components – Tradition and Naturalness were considered to be related to conventionality; Interest and Envelopment were representative of creativity. The graph in Figure 5 shows the plot for naïve listeners.

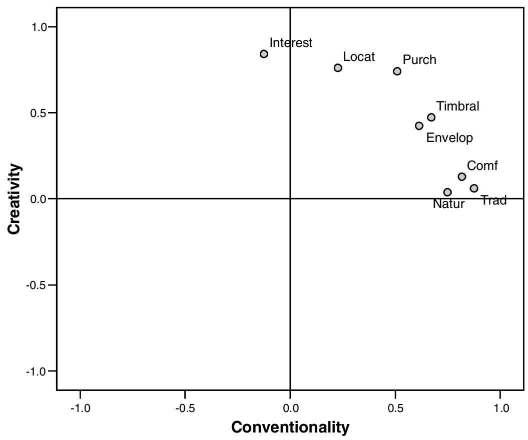

Figure 5 Component plot for naïve listeners

The naïve listener plot shows that three variables are closely related to the 'Conventionality' of the recording – Naturalness, Tradition and Listening Comfort. Of these, Naturalness is the closest in relation. The 'Creativity' component is quite closely related to the Interest variable and Instrument Locatedness is more closely related to creativity than conventionality. For naïve listeners, Purchasability appears to be more dependent upon creativity than conventionality.

Figure 6 shows the same components, this time for the trained listeners.

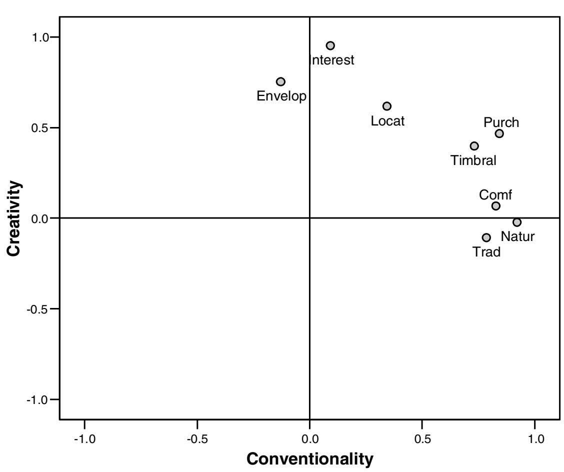

Figure 6 Component plot for trained listeners

Once again, conventionality is closely related to Tradition, Naturalness and Listening Comfort. Timbral Balance is more closely related to conventionality than creativity for trained listeners. For naïve listeners, Timbral Balance is more related to creativity.

The creativity component is dependent upon Listener Envelopment and Interest. The envelopment variable is of note, because naïve listeners find this more related to conventionality than creativity – the opposite opinion to the trained listeners.

For trained listeners, purchasability is more dependent upon conventionality than creativity – once again the opposite of naïve listeners. This indicates that naïve and trained listeners assess the recordings in a very different manner and that purchasability is based upon these assessments.

3.5. Regression

The next step in the analysis was to look at the variables in relation to Purchasability. This enabled an equation to be formed predicting Purchasability from a selection of the other variables. This was achieved by regression of the variables using Purchasability as the dependent variable.

For the naïve listeners, purchasability can be predicted using the other seven variables and result in a Pearson Correlation of 0.887. This is quite high and therefore quite an accurate predictor of Purchasability.

The Purchasability Predictor can be plotted against the measured Purchasability to give a correlation plot between the two factors. This is shown in Figure 7.

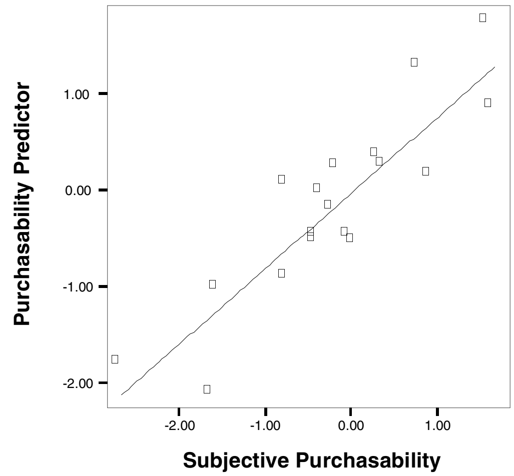

Figure 7 Subjective and predicted purchasability correlation plot for naïve listeners

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.531	.294		-1.810	.100
	Comf	.123	.294	.090	.418	.685
	Trad	-.628	.470	-.362	-1.336	.211
	Interest	.476	.388	.238	1.225	.249
	Envelop	.566	.212	.577	2.665	.024
	Locat	-.038	.312	-.025	-.122	.905
	Natur	.128	.302	.079	.424	.680
	Timbral	.717	.334	.545	2.149	.057

a. Dependent Variable: Purch

Table 1 Regression statistics for naïve listeners

A simplistic equation can be formed from the significant variables (Sig. <0.05) from the regression analysis.

The grey cells in Table 1 indicate the constant and the envelopment variable which is significant at $p < 0.05$ level. For the naïve listeners, the simplistic equation will therefore be

$$\text{Purchasability} = 0.556 \times \text{Envelopment} - 0.531 \qquad (1)$$

The data from the trained listeners is treated in the same way. The correlation between the Purchasability Predictor (calculated from the seven other variables) and the subjective Purchasability is much greater than for the naïve listeners. The Pearson Correlation is 0.968 for the trained listeners. The correlation plot demonstrates this graphically in Figure 8.

For the trained listeners, the number of variables which are statistically significant (Sig. <0.05) is greater than for the naïve listeners, with three variables instead of just one. Table 2 shows these statistics and enables the simplistic Purchasability Predictor equation to be worked out.

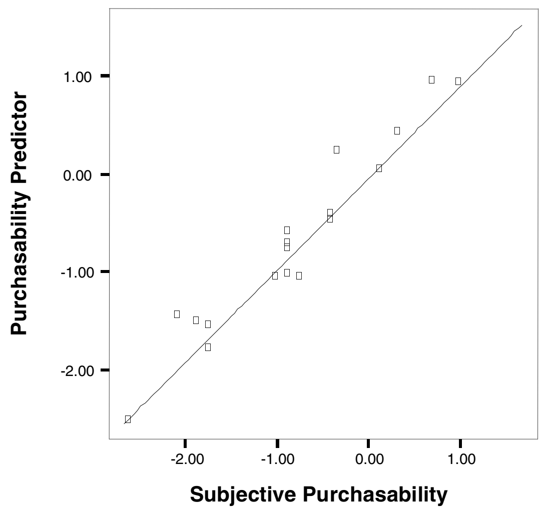

Figure 8 Subjective and predicted purchasability correlation plot for trained listeners

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.573	.165		-3.476	.006
	Comf	-.006	.140	-.006	-.044	.966
	Trad	-.129	.160	-.112	-.802	.441
	Interest	-.415	.361	-.212	-1.150	.277
	Envelop	.255	.129	.257	1.975	.076
	Locat	.742	.245	.394	3.027	.013
	Natur	.534	.184	.469	2.905	.016
	Timbral	.576	.128	.560	4.504	.001

a. Dependent Variable: Purch

Table 2 Regression statistics for trained listeners

$$\text{Purchasability} = 0.742 \times \text{Locatedness} + 0.534 \times \text{Naturalness} + 0.576 \times \text{Timbral} - 0.573 \qquad (2)$$

The three variables in the grey cells are Instrument Locatedness, Naturalness and Timbral Balance; these are used to form the Predictor equation which is shown in (2).

To sum up, for naïve listeners, envelopment is used and results in a fairly high correlation between the Subjective and Predicted Purchasability. For trained listeners, the correlation between the predicted and actual scores is much higher and the regression equation is formed from three variables.

4. DISCUSSION

The results analysis for this research has shown some interesting results which suggest further work, albeit with some major adjustments to the experimental procedure.

A few limitations were discovered during the course of the investigation. Although these did not directly affect the results, their improvement would aid further testing.

The length of the listening test was a slight problem which was mentioned by several test subjects. It took most listeners between 45 minutes and 1 hour to complete the test and many commented that this was perhaps a little long to hold their attention. Although it was not possible to shorten the test without reducing the amount of information gained as a result, any future work or testing in this area may require considerable revisions of the test. With the Purchasability Predictor it would be possible to reduce the number of variables under test, thus reducing the number of trials needed in each test. A reduction in variables would be possible in a more focussed investigation. Further work in the area could focus on a more specific part of this investigation, perhaps looking at only one type of programme material. The construction of an improved listening test would be possible with further developments and experimentation.

The programme content of the recordings was a distraction for many of the candidates due to the performance standard rather than the musical content. It was very difficult to find enough experienced and competent players to make the recordings which resulted in some 'less than professional' standard playing on the recordings – perhaps influencing some listeners' opinions. Unfortunately, some level of bias may have been introduced to the test due to this issue but this was unavoidable. For further work, more experienced musicians might help to make the recordings more suitable and reduce the potential bias of the listeners.

Issues with image stability, as mentioned in the Introduction, did not cause any serious problems for the investigation. This was mainly due to the design of the listening test which was tailored to gain the listener opinions of the recordings rather than a more specific investigation of spatial characteristics. As this was an introductory experiment in the area, this technique seemed more appropriate and further work in the area could look more specifically at the spatial imaging and characteristics of certain microphone array recording techniques for different musician arrangements. Future work may also look into techniques using greater numbers of loudspeakers to create a more stable surround soundfield.

The choice of positions for the musicians was purely experimental for the Partial Arc and Full Surround arrangements, but the placements were based upon musical experience. It was attempted to use the microphone techniques to the artists' advantage, focusing on specific instruments and creating a sense of familiarity for the listener.

The intermediate step of the Partial Arc arrangement was there to provide a stepping-stone between the Traditional and Full Surround techniques. This simply extended the outer edges of the ensemble around the microphone array. The musicians within the Partial arrangement were essentially kept in the same relative positions within the ensemble to keep a familiar element to the musician positioning for the listener whilst still altering the envelopment characteristics.

The Full Surround array completely changed the positions of the musicians, but did try to place some instruments within the scope of familiarity, for example, placing the violins of the orchestra and quartet to either side of the listener. Some sources were placed centrally to create a hard central source to act as an anchor, i.e. the sopranos in the choir.

Different musician placement techniques could be attempted, perhaps positioning the musicians directly relative to the loudspeaker positions in the 5-channel system rather than in a 360° soundfield. This may improve the spatial imaging and have an effect upon the Envelopment and Instrument Locatedness variables.

The variety of analytical tests indicates particular areas which may be of interest to investigate further. Of most interest is the Purchasability Predictor, which allows an indication of Purchasability to be calculated from a small number of variables and gives a high correlation to the subjective Purchasability. This would enable an investigator to conduct a shorter test with fewer variables and calculate the Purchasability from the results.

5. CONCLUSION

A new method for recording classical music in multichannel surround sound was developed, using existing microphone techniques but altering the musician arrangements away from the traditional concert hall setup. Recordings were made for a variety of programme material and used for a subjective listening test. The aim of the research was to discover listener opinions and determine if unusual recording methods are commercially viable for classical music.

The listening test asked for listener opinions in response to a variety of questions concerning attributes of the recordings. These included Listening Comfort, Instrument Locatedness and Tradition as well as five other variables.

The data collected from the experiment enabled a wide variety of analytical tests to be carried out, assessing the listener opinions of the recordings and indicating connections and correlations between the variables. A system for predicting Purchasability from a selection of question variables was developed.

Although the 'purchasability' attribute cannot be confirmed to reflect consumer behaviour, the results clearly indicate that naïve listeners are prepared to purchase classical music recordings which use novel spatial scenes which envelop the listener. Experienced listeners would be interested in more traditional recordings in which timbral fidelity, naturalness and locatedness are optimised.

In additional work, recording methods and listening test procedures could be refined and the Purchasability Predictor could be used to reduce the number of test variables. This may help to establish further whether unusual recording techniques in the commercial classical industry may have a niche within the consumer market.

6. ACKNOWLEDGEMENTS

The authors would like to express their thanks to Professor Sebastian Forbes and members of the University of Surrey Orchestra and Chamber Choir for participating in the recordings.

7. REFERENCES

[1] Cohen, Elizabeth and Eargle, John, 'Audio in a 5.1 Environment', presented at the AES 99th Convention, New York, USA, 1995 September, Preprint 4071.

[2] Theile, Günther, 'On the Naturalness of Two Channel Stereo Sound', J. Audio Eng. Soc., Vol. 39, No. 10, p761-767 (1991 October)

[3] Rumsey, Francis, 2001 'Spatial Audio' (Oxford, Focal Press)

[4] Moorer, James A, 'Audio in the New Millennium', J. Audio Eng. Soc., Vol. 48, No. 5, p490-498 (2000 May)

[5] Borwick, John, 1990 'Microphones: Technology and Technique' (London, Focal Press)

[6] International Telecommunications Union, Recommendation ITU-R BS.775-1 'Multichannel stereophonic sound system with or without accompanying picture', 1994 (ITU)

[7] Kornacki, Artur, Bozena, Kostek, Odya, Piotr and Czyzewski, Andrzej, 'Problems Related to Surround Sound Production', Presented at the AES 110th Convention, Amsterdam, The Netherlands, 2001March, Preprint 5374

[8] Huber, David Miles and Runstein, R E, 2001 'Modern Recording Techniques' (Oxford, Focal Press)

[9] Mason, Russell and Rumsey, Francis, 'An Investigation of Microphone Techniques for Ambient Sound in Surround Sound Systems', Presented at the AES 106th Convention, Munich, Germany, 1999 May, Preprint 4912

[10] Theile, Günther, 'Multichannel Natural Recording based on Psychoacoustic Principles', Presented at the AES 108th Convention, Paris, France, 2000 February, Preprint 5156

[11] Usability Company, The, 2005 'Likert Scale (The Usability Company)' http://www.theusabilitycompany.com/resources/glossary/likert-scale.html, 2005 February

8. APPENDIX

Between-Subjects Factors

		Value Label	N
Microphone Array	1	Array 1	270
	2	Array 2	267
Programme	1	Orchestra	177
	2	Quartet	180
	3	Choir	180
Musician Arrangement	1	Traditional	179
	2	Midway	179
	3	Full Surround	179

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.319	29.999[a]	8.000	512.000	.000	.319
	Wilks' Lambda	.681	29.999[a]	8.000	512.000	.000	.319
	Hotelling's Trace	.469	29.999[a]	8.000	512.000	.000	.319
	Roy's Largest Root	.469	29.999[a]	8.000	512.000	.000	.319
Array	Pillai's Trace	.198	15.842[a]	8.000	512.000	.000	.198
	Wilks' Lambda	.802	15.842[a]	8.000	512.000	.000	.198
	Hotelling's Trace	.248	15.842[a]	8.000	512.000	.000	.198
	Roy's Largest Root	.248	15.842[a]	8.000	512.000	.000	.198
Prog	Pillai's Trace	.193	6.839	16.000	1026.000	.000	.096
	Wilks' Lambda	.815	6.884[a]	16.000	1024.000	.000	.097
	Hotelling's Trace	.217	6.929	16.000	1022.000	.000	.098
	Roy's Largest Root	.153	9.838[b]	8.000	513.000	.000	.133
Arrang	Pillai's Trace	.115	3.902	16.000	1026.000	.000	.057
	Wilks' Lambda	.887	3.964[a]	16.000	1024.000	.000	.058
	Hotelling's Trace	.126	4.026	16.000	1022.000	.000	.059
	Roy's Largest Root	.111	7.128[b]	8.000	513.000	.000	.100
Array * Prog	Pillai's Trace	.297	11.203	16.000	1026.000	.000	.149
	Wilks' Lambda	.712	11.860[a]	16.000	1024.000	.000	.156
	Hotelling's Trace	.392	12.521	16.000	1022.000	.000	.164
	Roy's Largest Root	.356	22.812[b]	8.000	513.000	.000	.262
Array * Arrang	Pillai's Trace	.026	.854	16.000	1026.000	.624	.013
	Wilks' Lambda	.974	.854[a]	16.000	1024.000	.624	.013
	Hotelling's Trace	.027	.854	16.000	1022.000	.624	.013
	Roy's Largest Root	.021	1.325[b]	8.000	513.000	.228	.020
Prog * Arrang	Pillai's Trace	.132	2.197	32.000	2060.000	.000	.033
	Wilks' Lambda	.873	2.211	32.000	1889.758	.000	.033
	Hotelling's Trace	.139	2.222	32.000	2042.000	.000	.034
	Roy's Largest Root	.082	5.263[b]	8.000	515.000	.000	.076
Array * Prog * Arrang	Pillai's Trace	.096	1.575	32.000	2060.000	.022	.024
	Wilks' Lambda	.907	1.575	32.000	1889.758	.022	.024
	Hotelling's Trace	.099	1.575	32.000	2042.000	.022	.024
	Roy's Largest Root	.045	2.881[b]	8.000	515.000	.004	.043

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c. Design: Intercept+Array+Prog+Arrang+Array * Prog+Array * Arrang+Prog * Arrang+Array * Prog * Arrang

Table 3 MANOVA tables for all listeners

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.275	11.640[a]	8.000	245.000	.000	.275
	Wilks' Lambda	.725	11.640[a]	8.000	245.000	.000	.275
	Hotelling's Trace	.380	11.640[a]	8.000	245.000	.000	.275
	Roy's Largest Root	.380	11.640[a]	8.000	245.000	.000	.275
Array	Pillai's Trace	.190	7.192[a]	8.000	245.000	.000	.190
	Wilks' Lambda	.810	7.192[a]	8.000	245.000	.000	.190
	Hotelling's Trace	.235	7.192[a]	8.000	245.000	.000	.190
	Roy's Largest Root	.235	7.192[a]	8.000	245.000	.000	.190
Prog	Pillai's Trace	.265	4.697	16.000	492.000	.000	.133
	Wilks' Lambda	.749	4.754[a]	16.000	490.000	.000	.134
	Hotelling's Trace	.315	4.810	16.000	488.000	.000	.136
	Roy's Largest Root	.233	7.177[b]	8.000	246.000	.000	.189
Arrang	Pillai's Trace	.161	2.683	16.000	492.000	.000	.080
	Wilks' Lambda	.844	2.701[a]	16.000	490.000	.000	.081
	Hotelling's Trace	.178	2.719	16.000	488.000	.000	.082
	Roy's Largest Root	.135	4.143[b]	8.000	246.000	.000	.119
Array * Prog	Pillai's Trace	.300	5.423	16.000	492.000	.000	.150
	Wilks' Lambda	.708	5.765[a]	16.000	490.000	.000	.158
	Hotelling's Trace	.400	6.107	16.000	488.000	.000	.167
	Roy's Largest Root	.370	11.365[b]	8.000	246.000	.000	.270
Array * Arrang	Pillai's Trace	.061	.974	16.000	492.000	.485	.031
	Wilks' Lambda	.939	.973[a]	16.000	490.000	.485	.031
	Hotelling's Trace	.064	.973	16.000	488.000	.486	.031
	Roy's Largest Root	.047	1.448[b]	8.000	246.000	.177	.045
Prog * Arrang	Pillai's Trace	.164	1.326	32.000	992.000	.108	.041
	Wilks' Lambda	.845	1.325	32.000	905.111	.108	.041
	Hotelling's Trace	.174	1.324	32.000	974.000	.109	.042
	Roy's Largest Root	.089	2.769[b]	8.000	248.000	.006	.082
Array * Prog * Arrang	Pillai's Trace	.156	1.255	32.000	992.000	.157	.039
	Wilks' Lambda	.853	1.250	32.000	905.111	.162	.039
	Hotelling's Trace	.163	1.243	32.000	974.000	.168	.039
	Roy's Largest Root	.068	2.099[b]	8.000	248.000	.036	.063

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c. Design: Intercept+Array+Prog+Arrang+Array * Prog+Array * Arrang+Prog * Arrang+Array * Prog * Arrang

Table 4 MANOVA tables for Naïve listeners

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.422	22.125[a]	8.000	242.000	.000	.422
	Wilks' Lambda	.578	22.125[a]	8.000	242.000	.000	.422
	Hotelling's Trace	.731	22.125[a]	8.000	242.000	.000	.422
	Roy's Largest Root	.731	22.125[a]	8.000	242.000	.000	.422
Array	Pillai's Trace	.257	10.475[a]	8.000	242.000	.000	.257
	Wilks' Lambda	.743	10.475[a]	8.000	242.000	.000	.257
	Hotelling's Trace	.346	10.475[a]	8.000	242.000	.000	.257
	Roy's Largest Root	.346	10.475[a]	8.000	242.000	.000	.257
Prog	Pillai's Trace	.172	2.857	16.000	486.000	.000	.086
	Wilks' Lambda	.835	2.852[a]	16.000	484.000	.000	.086
	Hotelling's Trace	.189	2.848	16.000	482.000	.000	.086
	Roy's Largest Root	.117	3.568[b]	8.000	243.000	.001	.105
Arrang	Pillai's Trace	.115	1.847	16.000	486.000	.023	.057
	Wilks' Lambda	.886	1.880[a]	16.000	484.000	.020	.059
	Hotelling's Trace	.127	1.914	16.000	482.000	.017	.060
	Roy's Largest Root	.117	3.565[b]	8.000	243.000	.001	.105
Array * Prog	Pillai's Trace	.393	7.419	16.000	486.000	.000	.196
	Wilks' Lambda	.637	7.662[a]	16.000	484.000	.000	.202
	Hotelling's Trace	.525	7.904	16.000	482.000	.000	.208
	Roy's Largest Root	.414	12.561[b]	8.000	243.000	.000	.293
Array * Arrang	Pillai's Trace	.032	.499	16.000	486.000	.948	.016
	Wilks' Lambda	.968	.497[a]	16.000	484.000	.949	.016
	Hotelling's Trace	.033	.496	16.000	482.000	.949	.016
	Roy's Largest Root	.023	.709[b]	8.000	243.000	.684	.023
Prog * Arrang	Pillai's Trace	.179	1.435	32.000	980.000	.057	.045
	Wilks' Lambda	.831	1.439	32.000	894.047	.056	.045
	Hotelling's Trace	.192	1.442	32.000	962.000	.054	.046
	Roy's Largest Root	.104	3.200[b]	8.000	245.000	.002	.095
Array * Prog * Arrang	Pillai's Trace	.186	1.497	32.000	980.000	.038	.047
	Wilks' Lambda	.823	1.515	32.000	894.047	.034	.048
	Hotelling's Trace	.204	1.531	32.000	962.000	.031	.048
	Roy's Largest Root	.126	3.853[b]	8.000	245.000	.000	.112

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c. Design: Intercept+Array+Prog+Arrang+Array * Prog+Array * Arrang+Prog * Arrang+Array * Prog * Arrang

Table 5 MANOVA tables for trained listeners