Subjective Evaluation of Perceived Spatial Differences in Car Audio Systems Using a Graphical Assessment Language

Natanya Ford¹, Francis Rumsey¹, and Tim Nind²

¹ Institute of Sound Recording, University of Surrey, Guildford, Surrey GU2 7XH
² Harman / Becker Automotive Systems, Bridgend, Mid Glamorgan, CF31 3SH
Correspondence should be addressed to Natanya Ford (natford@genie.co.uk)

ABSTRACT
Results from preliminary investigations studying graphical elicitation techniques suggest that a graphical assessment language, whereby listeners use their own non-verbal descriptors to depict spatial attributes of a reproduced sound, may be effective for demonstrating differences in perceived image skew and scene width.

This paper investigates the use of a graphical assessment language for evaluating subjective differences in car audio systems with respect to their distortion of stereo images from sub-optimal listening locations. The study compares the image obtained from a surround processing system and conventional two channel stereo reproduction, analysing the graphical depictions obtained using conventional statistical methods. Source material for the investigation employs both time and amplitude variation to position instruments within the reproduced stereo scene.

INTRODUCTION
Historically, subjective assessment has looked towards verbal language to describe differences, distortion artifacts and sonic qualities of audio reproduction. Contemporary research into the spatial attributes of reproduced sound continues, unsurprisingly¹, to gather verbal responses from listeners [1 - 9]. However, as previously discussed in [10, 11], the subjective assessment of spatial attributes should not be limited to verbal language, as words are not our only means of communication, especially when the information sought (spatial percepts) are not inherently verbal. Furthermore, meaning in subjective assessment is open to interpretation, a situation which is partly dependent on the communication medium used.

This study employs a visual communication medium, a Graphical Assessment Language (GAL). The GAL method requires unique graphical descriptors to be elicited from individual listeners and subsequently analysed. The use of a listener’s own language when describing their own perceptions, rather than employing a prescribed language (be it verbal or graphical), is advantageous in reducing the amount of interpretation associated with subjective assessment. For although listeners can be trained in the meaning of provided scales and language, this training can potentially bias investigation results, and does not necessarily result in a listener

¹ As verbal language is a widely used communication mediums.
being able to translate between percept and prescribed response language accurately [12].

A preliminary study of the GAL method [11] highlighted the language’s effectiveness for illustrating differences in perceived spatial attributes and introduced methods for analysing graphical data. The apparent intuitiveness encountered when auditory spatial percepts were communicated visually may stem from the association between auditory and visual spatial perception within the brain [13]. The link between the two modalities suggests that no translation is needed to convert auditory spatial perception to visual spatial perception or vice versa, potentially leading to reduced interpretation between percept and response [14].

The current investigation builds on the work of the initial study and evaluates the suitability of a GAL for assessing the image skew and ensemble width of in-car audio events.

INVESTIGATION OVERVIEW

The investigation required 12 experienced listeners to provide descriptions of ensemble width and image skew for different reproduced audio events within a stationary vehicle. The listeners were asked to depict these spatial qualities on paper response sheets using their own graphical response style. A combination of variables were manipulated for each event, these being the reproduction system used, the listening location, and the source material. The listeners’ depictions were initially subject to an examination of individual characteristics, with subsequent statistical analysis once responses had been converted into numerical format. Scatter and density plots of the listeners’ depictions were also created to highlight and obtain more qualitative data about any statistical differences.

Investigation Variables - Reproduction System

Two different CD systems were employed in the study. Firstly a conventional ‘stereo’ reproduction system, which fed six of the seven loudspeakers within the vehicle with two channels of information, and secondly a multichannel surround processing system, which up-mixed a conventional two-channel feed into seven channels.

Investigation Variables - Listening Location

For in-car entertainment systems, the effect of listening location on perception is important as the driver and front seat passenger will necessarily be sat in a sub-optimal, off-centre, location for a conventional stereo reproduction. Spatial impression is compromised due to the precedence effect whereby the listener hears the reproduction to be steered or ‘skewed’ towards the nearest door (where loudspeakers are conventionally located). Although the inclusion of a central loudspeaker can somewhat compensate for this steering, it is not entirely satisfactory [15], furthermore, rear listening locations need to be considered. For the reasons outlined above, three listening locations were investigated, namely the driving seat (front left), the rear centre location (which should be less affected by steering), and rear left location (to obtain information about image skew and ensemble width for a rear sub-optimal listening position).

Investigation Variables - Source Material

The influence of source material on a listener’s perception of reproduced events is well known and controversial [e.g. 16, 17, 18, 19, 20]. It was therefore necessary to consider carefully the source material used to assess listeners’ perception of ensemble width and image skew. Material was kept simple using only three instruments. These ‘instruments’, namely a cello playing a low sustained passage, a female voice repeating text in Danish, and staccato percussion, were all taken from the Archimides CD of anechoic mono recordings. A list of source material is provided in table 1. Six different two-channel stereo pieces of approximately 50 seconds in length were compiled by positioning the instruments using either time or amplitude differences between channels. By constructing dedicated source material using known instruments, it was envisaged that a degree of ecological validity would be maintained, whilst preventing listeners from relying on any predetermined knowledge of instrument location. An equal amount of reverber was added to all stimuli using a Lexicon unit.

Investigation Variables - Participants

The quality/quantity debate is common to subjective audio assessment where it is widely believed that the higher the acuity of the listener, the lower the quantity required. To improve quality, expert listeners with previous experience of listening investigations are thought to be the key to reliable results and reduced error variance, with anywhere between 12 and 15 experienced listeners being suggested [4] as an appropriate number for a descriptive study. For the purpose of this investigation, 12 listeners, all with an active interest in music and/or audio engineering, were selected. By only using participants experienced in listening, it is acknowledged that the sample is unrepresentative of all but the most critical of listening populations, and generalisation of responses to a less-expert population is unethical. However as the task is to develop a descriptive language, the importance of using experienced listeners is acknowledged. A future study is proposed to evaluate the language’s representativeness for a more diverse population.

INVESTIGATION METHOD

The first step in the investigation was to record six CD’s of source material so that six different presentation orders existed. Each CD consisted of 18 pieces of music to allow for the 12 investigation runs (three source * two panning method * two

---

2 Evans [12] suggests that an intuitive response method increases the likelihood of stability in results by reducing the amount of interpretation required between percept and response.

3 It should be noted that although a graphical response method is linked with auditory imaging, some interpretation is necessary as the listener must make known their perception in some external way (e.g. graphically depicting the auditory scene on paper).

4 Image skew refers to the amount of shift a reproduced image undergoes away from a designated reference position. For the purpose of this investigation, the reference position is taken as a line down the centre of the vehicle from front to back.

5 Ensemble width is the distance between right and left most instruments positioned within an auditory image (i.e. the outer boundaries of the image without environmental cues).

6 The seven loudspeakers were located as follows, 1: front left door, 2: centre dashboard, 3: front right door, 4: rear left door, 5: rear right door, 6: rear left boot door, 7: rear right boot door. The conventional stereo system fed all except the centre loudspeaker.

7 Mertens curves [21] were used to obtain the required time and amplitude differences for phantom image positioning. The source Merten used to identify required amplitude differences between channels was a noise signal centred at 100Hz. The curve for time differences was created for détonations.

Gabrielson et al. [22] used 14 participants none with experience of investigation. Gabrielson and Sjögren [23] suggest between 20 – 42 participants, Olive et al. [17] used 13 listeners with no previous listening experience. Gabrielson and Lindstrom [24] used 18 male participants selected according to their responses on a questionnaire, and Bech used 12 participants [25].
reproduction system) plus six allocated repeats. During each run listeners were asked to complete three response sheets, one for each of the three listening locations, resulting in 54 sheets (18 * three listening location) being completed for each of the 12 listeners.

At least a day prior to starting the investigation all listeners were presented with written instructions (appendix 1) containing full details of their task, and were shown a blank response form (Figure 1). The response form depicts a scaled representation of the auditory space (the passenger compartment of the car), and includes useful visual cues to improve translation from egocentric to external response. Space was left on the sheet so that listeners could also depict events outside of the vehicle.

![Figure 1 Blank listener response form](image)

Training listeners is a contentious issue and although necessary in situations where the language used to obtain responses is unfamiliar to the participant, drawing responses may reduce or remove the need for training and limit the extent of investigator bias. Therefore, other than being presented with written instructions, listeners were not trained in any way. They were informed of which spatial attributes to depict and no restriction was placed on how they could draw the events. It was common (though not compulsory) for listeners to assess their own drawing style with the investigator at the end of the first run.

Each listener participated in the investigation individually. Before each run the listener was given three response sheets with each sheet crossed at one of the three listening location to indicate the seat order for the run. This order was randomised to avoid listeners always starting with the same seat. During the run, the CD was placed on track repeat in the mode required (either conventional stereo or surround processed) and the listener was informed to spend as long as necessary completing the response sheets. Listeners could take a break whenever tired, which was usually around run seven. If no break was requested by the listener, a break was required after run 12, to prevent listener fatigue. The average time taken by a listener to complete all 54 response sheets was approximately two hours.

**ANALYSIS OF RESPONSES**

Initially, individual graphical responses of each listener were evaluated to assess how perceived images were being depicting. Individual responses were then measured to enable statistical analysis. A qualitative evaluation of the data was made possible by plotting all listener depictions together for different combinations of variable manipulation.

**Individual Responses**

Appendix 2 contains 12 actual responses from six listeners sat (independently) in the driver’s seat, when amplitude panned material was used with both reproduction systems. Examining the depictions (figures 2 – 13) it appears that listeners were using different scales when drawing the perceived images, even though scaling information was provided on the response sheets. Furthermore a couple of depictions (figures 12 and 13) show the image to be split into several parts, with copies of the same instrument placed at different locations around the car, typically where loudspeakers were located. However there was a common link between the majority of depictions with three small circles, containing letters signifying the instrument, being drawn on the response sheet to represent the spatial qualities of the reproduction. Even though they were only asked to depict the location and width of the image, listeners provided additional information about the distance and size of relative instruments, with many listeners also asking how to depict height information (a noticeable difference between reproductions).

The analysis of individual responses raises numerous questions which ostensibly appear to have their answers in the design of the response sheet and the way in which listeners are informed about what to listen to. It could be that the current response sheet allows for an excess of information to be provided, or that definitions of ensemble width and image location were not grasped by the listener (they may even have understood the terminology but interpreted it differently) which would suggest that any explanatory language used be simplified. Alternatively the source material may have played a part in the multitude of different response styles. For example it was common for listeners to find localising the cello difficult, particularly when the cello and female voice were not separated by the percussion9. This was especially so for material where the instruments were positioned using time, differences. This difficulty in localising continuous sources was expressed informally by most listeners and could be quantified (again informally) by the increased time taken to complete runs where time differences were used.

**Analysis of Graphical Plots**

Two types of plots were created in order to analyse listener responses, with appendices 3a and 3b containing a representative sample of these plots. A single piece of source material was used for all the plots displayed in the appendices. This material was constructed with the cello 20° to the left of centre, the voice at 0° (centre) and the percussion 10° to the right of centre. Depictions for both amplitude and time panned versions of this source material have been included.

Appendix 3a (figures 14 – 25) contains scatter plots for the selected source material. Plots were created by marking the centre of each listener’s response and recording this central position onto individual plots by seat, for different reproduction system and panning method. The plots are informative for analysing amount of skew associated with each condition, and the consistency with which listeners locate responses. For example, by examining figures 14 – 17 (plots from the driver’s seat) it is apparent that listeners were reasonably consistent in their positioning of events from this front seat. Furthermore when plots for the surround system (figures 15 and 17) are compared with figures 14 and 16 for the conventional system, the image is seen to shift from the left door to the steering wheel. Sitting in the rear centre seat (figures 18 – 21) resulted in the reproduced image being depicted towards the

---

9 Examples of this effect are found in appendix 3b
middle of the car for both reproduction systems. However when listeners moved to the rear left seat (figures 22 - 25), their descriptions were more varied, listeners reported seeing the centre of the vehicle's responses for image skew and ensemble width were first converted to a statistical analysis of differences. All three listening locations are represented on each plot, the seat being determined by a number, where 1 is the driver’s seat, 2 is the rear centre and 3 the rear left position. Two plots for each instrument (figures 26, 28, 30 and 32) depict responses for the conventional stereo system, with the remaining four plots depicting the surround system. Figures 26, 27, 30 and 31 depict responses for amplitude panned material and 28, 29, 32 and 33 use the time panned material.

Figures 26 and 27 compare reproduction systems using amplitude panned percussion. The surround system plot of figure 27 indicates that the majority of listeners placed the percussion at the centre of the vehicle, whereas in figure 26, the conventional system, clusters of responses occurred at front left for seats 1 and 3 (driver’s and rear left seat) and front right for the central listening location, suggesting that listeners were able to locate the percussion most accurately when sat in the central listening location, a finding validated by the statistical analysis. Differences between reproduction systems were more pronounced for time panned percussion. The gap at the front of the conventional system plot (figure 28) increased, with responses for the central listening location clearly defined on the right of centre and towards the middle of the vehicle. For the combination of conventional system and driver’s seat (seat 1) images were perceived to be close to the windscreen, with depictions from seat 3 (rear left) creeping into the rear doors. The shape of figure 29, responses for the surround processed system, is very different. The image is central and no responses occur within the left door. This is again illustrative of the lack of skew apparent with the surround system.

The cello was less precisely located than the percussion in all but one condition, illustrated in figure 30, when amplitude panning was assessed alongside the conventional system. It is clear from this figure that the image was perceived to at front left, and very much in the left doors for all seats. Figure 31 plots the same instrument and panning condition for the surround system, and although there is still a left bias to the image, the depiction is larger and less clustered, with responses for the two rear seats occurring throughout the left side of the vehicle.

When the panning method was altered and the cello located using time differences (figures 32 and 33), the responses became more spread out, with the least focussed depictions occurring for the surround system (figure 33). Here the listeners perceived the cello to be all around them. It was anticipated that the combination of time panned cello and surround system would result in a less precisely localisable image, due to the surround system creating its own sense of transparency. A measurement of 0mm indicated a response centred on the middle of the car (an optimum location, where the image was not degraded by steering), with negative values highlighting a left bias to the image, and positive values denoting a right of centre image. The larger the value in either direction, the greater the skew. Due to the source material being unsymmetrical, with an instrument positioned further to the left than right, there was automatically a left bias to all images, which is noted in the statistical analysis.

A between subjects analysis was completed on the data produced by the investigation, with each independent variable (seat, system, panning method and source material) being used as a grouping variable.

An initial exploration of the data set found responses for image skew and ensemble width to be non-normally distributed, with the Kolmogorov – Smirnov test of normality significant to p > 0.001. Further examination of the data confirmed ensemble width to have a significantly flatter distribution than normal (z kurtosis = -2.124) and a right bias (z skew = -3.085). Due to the non-normal data distribution, simple non-parametric tests were used to analyse the statistical significance of any differences caused by the manipulation of the independent variables. It was found that with the exception of source material, all variable manipulation caused significant differences to be perceived for both width and skew.

All figures referred to in the following section can be found in Appendix 4, where ensemble width is displayed as a percentage width of the vehicle on the response sheet (total width being 131mm), and image skew is displayed similarly as a percentage value. Zero image skew, indicates that the image was at the centre of the vehicle and is displayed as 0 percent, with 100 percent referring to maximum skew (a value of 65.5mm away from the centre).

### Manipulating Individual Variables of Seat and System

<table>
<thead>
<tr>
<th>WIDTH</th>
<th>SKEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi_Square</td>
<td>19.695</td>
</tr>
<tr>
<td>Df</td>
<td>2</td>
</tr>
<tr>
<td>Asymp. Sig.</td>
<td>.000</td>
</tr>
</tbody>
</table>

### Table 2 Kruskal – Wallis test of difference for width and skew by listening location

When perceived width was analysed by seat using the Kruskal – Wallis test for differences, a significant result was found (p < 0.01, table 2), suggesting that listeners were perceiving width to vary according to where they were listening from. As predicted in the plots of appendix 3a, significant differences were also found for perceived skew by seat (table 2, p<0.01).

When listeners’ depictions for perceived ensemble width and skew were analysed with respect to the manipulation of reproduction system, differences were once again significant (p<0.05, table 3).

<table>
<thead>
<tr>
<th>WIDTH</th>
<th>SKEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>20233.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>43669.000</td>
</tr>
<tr>
<td>Z</td>
<td>-2.385</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.017</td>
</tr>
</tbody>
</table>

### Table 3 – results of non-parametric assessment of differences for width and skew by system

The above tables of significance have a limited attraction, restricted as they are by the amount of information they can provide.
provide, with only the consequence of manipulating a single variable being investigated. Were the data normally distributed, an ANOVA would be able to provide information about the significance of any interaction between independent variables, thus clarifying how the systems and listening locations combined to produce the significant results obtained. With non-parametric analysis, interactions between independent variables may still be investigated using clustered error bar charts.

**Interaction of Seat and System**
When width values for seat and system are plotted together on a single graph, interactions between the variables can be described. An interaction between rear central seat and the conventional reproduction system produced the widest image (figure 34), however in both other seats, the surround system was perceived as wider than the conventional system. This was especially so for the driver’s seat, where a noticeable difference in width was apparent between the two systems. The confidence intervals associated with this difference (figure 35) suggests it to be significant, with the difference for the interaction of rear left seat and system being potentially significant. However the confidence intervals overlap slightly and a parametric multiple comparison test would be required to confirm this significance.

Regardless of system, listeners perceived the ensemble to be wider from the central rear seat than from any other listening location, a finding very much in keeping with previous findings [11], and the assumption that image degradation is less for a central listening location. The confidence intervals for system and seat interaction could also lead to the hypothesis that perceived ensemble width was more greatly affected by listening location when the conventional reproduction system was used, as the surround system performed similarly regardless of listening location.

The amount of skew perceived away from the central reference location, and illustrated in appendix 3a, can also be described in terms of the interaction between listening location and system. Figures 36 and 37 display this interaction, and demonstrate that the skew for the conventional reproduction system was more left biased than the surround system for both the driver’s seat and rear left listening location. Both systems coped well with the central rear location, producing only a small skew to the left (characteristic of the source material’s left bias). Two channel stereo theory predicts an increase in skew for a conventional system coupled with off centre listening location, however the surround system produced more stable images for off centre listening locations, this stability being measured as a reduced pull away from the centre. The error bars in figure 37 confirm this difference between systems to be significant. No significant difference can be assumed between rear centre responses by system (they both produced images with a small percentage of skew, indicating their closeness to centre) and there were no significant differences between the surround system’s results for rear centre and rear left listening location.

**Manipulating Panning Method, System and Seat.**
Statistically significant results were found for ensemble width and image skew when amplitude differences were employed to locate the instruments within the source material opposed to when a time based panning law was applied. (table 4)

<table>
<thead>
<tr>
<th>Method</th>
<th>WIDTH</th>
<th>SKEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>18930.000</td>
<td>20385.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>42566.000</td>
<td>43821.000</td>
</tr>
<tr>
<td>Z</td>
<td>-3.390</td>
<td>-2.269</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.001</td>
<td>.023</td>
</tr>
</tbody>
</table>

*Table 4 – results of non-parametric assessment of difference in width and skew by panning method*

At first, reasons for the significance were unintuitive as many different variables were involved in obtaining the result and little difference is visible in the plots of appendix 3a. However differences became more intelligible when interactions between panning method and listening location, plus panning method and reproduction system, were investigated and the plots of appendix 3b were cross-referenced.

Plotting ensemble width for panning method and listening location on the same graph (figure 38) illustrated a difference between methods for the driver’s seat which could be interpreted as significant due to the separation between confidence intervals in figure 39. Listeners perceived amplitude panned material as significantly wider than time panned material from this sub-optimal listening location, however when they moved to the rear seats, no such difference existed between panning methods. For time panned source material, the ensemble was notably wider from the rear centre seat than the driver’s seat for the same material.

When ensemble width was plotted for panning method by reproduction system (figures 42 and 43) it was clear that the conventional system was affected by the panning method employed in the source material, with a notably wider image being produced when amplitude panned material was used. No such differences exist for the surround system, which performed equally for both methods of instrument positioning.

Figures 40 and 41 illustrate the interaction between panning method and listening location for perceived image skew. When analysed by panning method, significant differences between the listening locations can be suggested by the independence of the confidence intervals in figure 41. For both methods, the driver’s seat produced images with a greater amount of skew than the rear left which was in turn perceived as more left biased than the rear centre. The only difference between panning methods occurred from the driver’s seat, where time panned material was skewed further to the left than amplitude panned material auditioned from the same seat.

The most noticeable difference obtained from figures 44 and 45 (image skew assessed by panning method and system) occurs between systems. For both time and amplitude panned material, the conventional reproduction system produced images which were significantly more skewed than the surround processed system. No appreciable difference exists for skew when panning method was manipulated.

**CONCLUSION and FURTHER WORK**
This investigation employed a visual communication medium, a Graphical Assessment Language (GAL) to obtain pictorial representations of image skew and ensemble width from individual listeners. Depictions were analysed statistically and qualitatively to establish the suitability of the method for assessing the spatial qualities of in-car audio events when a combination of different variables were manipulated.

Analysis of listener depictions found that, for width:
- Listeners perceived the ensemble to be widest from the central listening location
- An interaction between rear central seat and the conventional reproduction system produced the widest image
- The surround system resulted in images being perceived as wider that the conventional system from the driver’s seat and the rear left listening location (i.e. the two sub-optimum listening locations)
- The width of conventional reproductions was more greatly affected by listening location than that of surround processed material, (width for surround system was a similar value for all three listening locations)
Amplitude panned material was perceived as significantly wider than time panned material from the driver’s seat. Envelope width was greater for time panned material when auditioned from the rear left seat than from the driver’s seat. Ensemble width for the surround system was similar for both panning methods.

And for skew:

- Differences in skew for reproduction system by seat were not significant for the rear centre seat, with both systems producing responses located close to the middle.
- Images for the conventional system were significantly more skewed than the surround system for both the driver’s seat and rear left listening location, suggesting that the surround system was more stable for off-centre listening.
- No significant difference was obtained for image skew when responses for rear centre and rear left seats were compared for the surround system.
- For the combination of conventional system and driver’s seat images were perceived close to the front left, with depictions from the rear left seats creeping into the rear doors.
- For the combination of surround system and driver’s seat images were located predominantly the steering wheel.
- The least focused depictions occurred for time panned cello and the surround system with listeners perceiving the cello to be all around them. (This effect was anticipated due to the way in which the surround system creates its surround channel feeds.
- There was little other obvious difference in skew when panning method was manipulated.
- A greater amount of skew was visible for the driver’s seat than the rear left seat, which was more left biased than the rear centre for both panning methods.
- From the driver’s seat, time panned material was skewed further to the left than amplitude panned material.

When listeners’ responses were analysed initially they appeared to be different in style and scale, suggesting that training in the use of the graphical language may have been beneficial. However these individual elicited depictions have enabled a full evaluation of differences in perceived ensemble width and image skew to be undertaken, producing significant results which are very much in line with conventional methods. These findings indicate that a Graphical Assessment Language is an effective means of evaluating these spatial characteristics of reproduced sound.

As training listeners is a time consuming and contentious issue, it is proposed that the GAL be improved further (minimising the extraneous variances apparent in responses from this investigation) by the careful creation of the response sheet and the simplification of verbal language used to provide instruction to listeners. The choice of source material is also implicated in the multitude of different response scales with listeners finding the cello very hard to localise, and depictions for this instrument being more varied than those for the percussion. As listeners did appear to take longer when responding to stimuli that were more difficult to localise, it would be of interest to repeat the procedure and record the time listeners take to produce their responses, to obtain further information about this correlation.

As only a small group of experienced listeners participated in the investigation, it is not yet known if the similarity in style observed in some of the depictions for this group would be replicated by a wider population. Therefore a future study is proposed to evaluate the language’s representativeness for a more diverse population.

Additionally it would be informative to follow up the observation that listeners were prepared to provide height information had they had an alternative response form which would have enabled them to do so.

ACKNOWLEDGEMENTS

The authors would like to thank Russell Mason for his assistance in producing the source material, Viv Railton and Rosemary Kriescher for their assistance during the study, and all the listeners for their good humour and willing participation.

REFERENCES

15 Nind, T (2001) Multimedia in Cars : The Application of Logic 7 Surround Processing to Provide Surround Sound in Cars from 2
Channel and Encoded 5.1 Sources. Audio Engineering Society 110th Convention Preprint
APPENDIX 1

Information For Listeners
Please read the following instructions carefully, making sure you fully understand what is required during the investigation. Should you have any questions or queries, please ask!

In a moment you will be asked to listen critically for spatial qualities (specified below) of car audio reproductions and respond by DRAWING what you have heard on the response sheets provided. You will be asked to move between three seats within a car and provide responses from each seat.

It is important to note that there are NO CORRECT ANSWERS, instead the investigation is looking for appropriate visual depictions of what YOU have perceived. Therefore, assume what you have heard to be correct and draw this as best as possible.

What you will hear & what qualities to depict
You will be played pieces of music specially recorded for the purpose of this investigation. Each piece consists of three instruments (voice, cello and percussion) positioned within an ensemble.

You are asked to provide the following spatial information on your response sheets:

Location of instruments within space
Draw the three individual instruments on the response sheet at the location where YOU perceive them to be.

Width of ensemble within space
Whilst placing the three instruments on the response sheet, think about the width of the ensemble (this is the total width of the three instruments) and draw the ensemble with this width.

APPENDIX 2

Individual graphical responses for surround and conventional reproduction systems, using amplitude panned source material.
APPENDIX 3a

Scatter Plots of Listeners Depictions

Figure 14  Plot from driver’s seat for amplitude panned source and conventional system

Figure 15  Plot from driver’s seat for amplitude panned source and surround system

Figure 18  Plot from centre seat for amplitude panned source and conventional system

Figure 19  Plot from centre seat for amplitude panned source and surround system

Figure 16  Plot from driver’s seat for time panned source and conventional system

Figure 17  Plot from driver’s seat for time panned source and surround system

Figure 20  Plot from centre seat for time panned source and conventional system

Figure 21  Plot from centre seat for time panned source and surround system
Appendix 3b Plots by Individual Instrument
Numbers on plots refer to listening location. 1 = driver’s seat, 2 = centre rear seat, 3 = rear left seat.

Figure 22 Plot from rear left seat for amplitude panned source and conventional system
Figure 23 Plot from rear left seat for amplitude panned source and surround system
Figure 24 Plot from rear left seat for time panned source and conventional system
Figure 25 Plot from rear left seat for time panned source and surround system
Figure 26 Amplitude panned percussion, conventional system
Figure 27 Amplitude panned percussion, surround system
APPENDIX 4 Graphs to Accompany Statistical Analysis

**Figure 34** – ensemble width for interaction between seat and reproduction system

**Figure 35** – 95% confidence intervals for ensemble width, for interaction between system and listening location

**Figure 36** – image skew for interaction between listening location and system

**Figure 37** – 95% confidence intervals for image skew, for interaction between system and listening location

**Figure 38** – ensemble width for interaction between panning method and listening location

**Figure 39** – 95% confidence intervals for ensemble width, for interaction between panning method and listening location
Figure 40 – skew for interaction between panning method and listening location

Figure 41 – 95% confidence intervals for image skew, for interaction between panning method and listening location

Figure 42 – ensemble width for interaction between panning method and system

Figure 43 – 95% confidence intervals for ensemble width, for interaction between panning method and system

Figure 44 – image skew for interaction between panning method and system

Figure 45 – 95% confidence intervals for image skew, for interaction between panning method and system