Acoustics'08
Paris
June 29-July 4, 2008
www.acoustics08-paris.org

optimising the value of sound quality evaluations by observing assessors' driving strategies

Sebastiano Giudice[a], Paul Jennings[b], Rebecca Cain[b], Garry Dunne[c], Mark Allman-Ward[d] and Roger Williams[d]

[a]WMG, University of Warwick, Gibbet Hill Road, CV47AL Coventry, UK
[b]WMG, The University of Warwick, International Manufacturing Centre, CV4 7AL Coventry, UK
[c]Jaguar & LandRover, Jaguar Engineering Center Abbey Road, Whitley, CV3 4LF Coventry, UK
[d]Sound Evaluations Ltd, 18 Walsworth Road, Hitchin, Hertfordshire, SG49SP Hitchin, UK
s.d.giudice@warwick.ac.uk

Traditionally vehicle sound evaluations have been conducted either whilst driving a car or by auditioning fixed test conditions in a listening room. On-road testing provides the right context but the results are often inconsistent and unrepeatable. Furthermore it is not possible to evaluate prototype sounds or carry out back-to-back comparisons. In-room evaluations improve the statistical confidence, but the context of the assessments is unrepresentative.
Interactive NVH vehicle simulators enable assessments to be performed in a setting representative of real appraisals. Accurate sounds can now be generated in real-time, and assessors can adopt a driving strategy that allows their own interpretation of the attributes.
The benefits include the opportunity to understand how preferences are formed by assessors, albeit with added complexities. Different assessors may be associating their preferences on different operating conditions which have different acoustical properties. Therefore for engineers to identify the key features that influence perception they need to be able to relate the driving strategies with the subjective preferences.
This paper reports on the use of new observational methods which capture assessors' decision making strategies. It demonstrates how these help in relating subjective preferences to vehicle operating conditions and how to design a structured evaluation to reduce sources of experimental error.

1. Automotive Sound Quality

The intense competition in the automotive sector has resulted in manufacturers exploring new ways to make their vehicles more appealing. Understanding the customers' perception of all stimuli is a new frontier in product development. Factors like the sound and the feel of a new car can communicate and reinforce brand qualities, and contribute towards a positive driving experience.

This transformation in mindset has manifested itself in a change in requirements from Noise Vibration and Harshness (NVH) engineers. For many years their intention, during the vehicle development programmes, was to reduce overall noise levels. The aim however has changed. Sound Quality Engineering (SQE) processes now focus on tailoring and enhancing the vehicle's sound in order to meet or exceed the customers' expectations of the vehicle and the brand. To achieve this, manufacturers have had to adapt their traditional engineering processes to include the voice of the customer. Objective acoustical targets need to be set using the customers' subjective impressions of the sound.

Sound evaluations are used to capture the subjective impressions at the outset of the vehicle development program. The evaluations are used to compare and benchmark concept vehicle sounds against those from competitor vehicles. Customers, engineers and key decision makers listen to the interior vehicle sounds. Their preferences are captured using an array of qualitative methods such as paired comparisons, rating scales and interviews.

Traditionally the evaluations were to be conducted whilst auditioning sounds in a neutral environment such as a listening room or by driving a car. These approaches have both advantages and inherent limitations. Within the listening room the assessors are exposed to consistent and repeatable experimental conditions and therefore the data collected has greater statistical meaning. It is also possible to listen to cars back-to-back with ease. This is a desirable feature of the evaluation process as non-expert assessors are rarely able to make an objective assessment of its sound and convey their impressions in a language which has value to an engineer. On the downside for listening rooms, the assessors listen to sounds which represent fixed operating conditions, such as 2nd Gear Wide Open Throttle (2GWOT). These operating conditions are chosen because they provide a repeatable and a like-for-like means of comparison for different cars; therefore facilitating a direct comparison between competitor vehicles. The limitation however, is this driving condition is rarely used by the customer and that if error states are solved at this full load conditions then equivalent resolution at part load conditions cannot be assumed. In addition, the sound is evaluated in isolation from the other stimuli. This can lead to the assessor over-concentrating on details of the sound, more than they would in a real appraisal.

Figure 1: Listening room

With on-road evaluations all sensory inputs are satisfied and many operating conditions can be evaluated. Concept vehicles however cannot be evaluated unless a prototype is first built. In addition experimental control is difficult to apply as each assessor is free to drive as they wish. Consequently experimental procedures are often unrepeatable making it difficult to interpret the results. The ability to assess vehicles back-to-back, which is possible in listening rooms, is not as easy. Necessary hardware changes take too long, making it difficult for non-expert assessors to remember the acoustic characteristics of previous prototypes.

Taken together, target sounds established using these traditional evaluation techniques can lack robustness and can lead to costly hardware changes and possibly delays in the launch date.

Interactive NVH simulators have been developed to bridge the gap between these traditional evaluations. The simulator used in this study comprises of a real vehicle fixed in front of a large screen onto which a Virtual Environment (VE) is projected. The assessor drives through the VE whilst assessing many vehicles. Each vehicle can be heard at the touch of a button, making it possible to easily evaluate cars back-to-back. Binaural sound is heard via headphones and vibration is felt through the steering wheel, the floor and the seat. The sound and vibration stimuli are accurate and furthermore existing and concept vehicles can be heard in real-time in response to the assessors' driving behaviour [1].

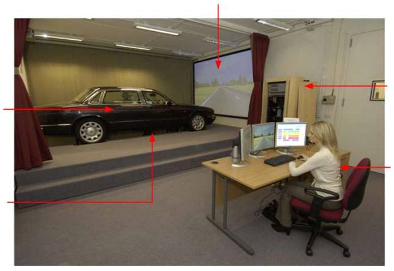

Figure 2: Exterior view of the interactive NVH simulator

Figure 3: Interior view of simulator

The assessor's responses to the vehicle sounds can be captured using a touch-screen interface located on the passenger seat. Here a variety of data capture interfaces can be displayed making it possible to explore the suitability of different methods [2].

NVH simulators provide the opportunity to evaluate concept vehicles in real-time, in a setting more representative of actual appraisal conditions, before the detailed design phases in the vehicle development programme. This reduces the possibility for expensive mistakes and can help to minimize the number of prototypes built.

There are other forms of simulator, e.g. desktop versions, and in-car systems. The desktop simulator has the same functionality as the full-vehicle simulator, although hardware set-up is similar to that of a computer game. The simulated vehicle is controlled through a video game steering wheel. This tool is a valuable engineering tool. NVH engineers can build and evaluate concept vehicles and choose which ones should be used in subjective evaluations. In-car versions of the simulator have also been recently developed. These are incorporated into real cars, with sound being heard using headphones and generated according to the operating conditions of the vehicle.

Overall interactive NVH simulators provide on-road evaluation conditions at all stages of vehicle development program and potentially allow for various degrees of experimental control to be applied. They offer the advantages of the both traditional evaluation conditions and none of their limitations.

2. Interactive Evaluations

Interactive NVH simulators have been successfully used within vehicle development programmes [3], yet there are still many opportunities to understand how to fully exploit this new technology and eventually propose best practice guidelines for their use [4]. To achieve this it is necessary to explore how the assessors perceive and interact with the simulator.

The term "interactive" encompasses the flow of information between a user and the system. When a system is interactive, an input from a user results in a response in the system. The resulting response in the system then influences the user, and so the process is an on-going cycle. There are many forms of interactivity during an evaluation in the simulator. However, the two important forms of interactivity from the evaluation perspective are the manner in which the assessors drive and how they use the data capture interface.

Previous studies concerning interactivity with the data capture interface were conducted by Baker [5] and Allman-Ward [6], in the listening room and simulator respectively and in conjunction with the paired comparison method. Using this method sounds are presented to the assessor in pairs. They are asked to pick one of the sounds relative to certain attributes. For example they may be asked to pick the most "Powerful" and / or the most "Refined" sound from the pair.

Conventional paired comparison evaluations are referred to as "fixed-play". This means the assessor can only listen to each sound in the pair once, before making a decision. [5] Compared fixed-play with "free-play". With this mode, the assessor was allowed to listen to each sound as many times as they wished before making a choice. A comparison of the two modes was conducted. In both the listening room and simulator, higher levels of consistency were achieved using the free-play mode. Furthermore interactivity leads the assessor to being more confident, as they can assess each vehicle as many times as they wish before making a decision. Hence interactivity in this context can be a benefit to the evaluation process.

The second main form of interactivity concerns the capability to interact with the vehicle. For this two levels exist, "non-interactive" and "fully interactive". With non-interactive mode, assessors listen to the sound of fixed driving conditions such as 2GWOTs. They cannot interact with the steering wheel and the foot pedals. They still feel vibration and the VE gives the impression of movement. With this mode experimental conditions available through the listening room can be applied, although there is the added benefit of conducting the evaluation in the right context. From the evaluation perspective the non-interactive mode is ideal. The evaluation conditions are constant for each assessor and can be easily repeated. The drawback however, is that this mode does appropriate levels of context; as 2GWOTs are not suitable representations of the vehicle sound.

Fully interactive evaluation allows the assessors to choose the driving conditions which satisfy their own interpretation of the attributes being assessed. For example, the assessment of the "Refinement" attribute may be related to the reduction of internal noise at cruising conditions, whereas "Powerfulness" is a dynamic character related to the engine sound, better assessed when performing Wide Open Throttle (WOT) maneuvers.

The assessor's freedom to choose the driving conditions they believe, represent the attribute can introduce an extra source of variability. This occurs as the assessors make decisions based on the sounds they hear, and the sounds they hear are based on how they drive. An issue related to this is that the sound heard at different driving conditions is generated from different mechanical components. It is therefore necessary to capture how the assessors drove during an evaluation as well as capturing the manner in which they score the vehicles. Without such information the subjective opinions captured will only relate to an overall impression of the car. This does not make it possible to relate the subjective opinions to key acoustic features that influenced perception. Without this missing link, the manipulation of engineering components can be conducted in a misguided fashion.

Through previous work delivered a driver observation module was built into the simulator. This captured data related to how the simulator was driven through the virtual environment and how they used the data capture interfaces.

2.1. Driver Behaviour

Large amounts of useful data can be generated from recording observational data from fully interactive evaluations. Analysis methods have been developed so the relationship between assessment strategy and subjective preference can be established. Figure 4a shows typical data collected from observational studies. It illustrates an assessor's changes in speed during the course of the evaluation. The blue line represents the assessment strategy employed when evaluating Powerfulness. The red line demonstrates the assessment strategy for Refinement. The colour bands at the top of the graph indicate which car is being driven at that particular time. The faint vertical lines indicate the assessor's interactions with the data capture interface. Typical actions would include, selecting or rating cars.

Displaying the data in this form, highlights that behaviour tends to be repeated following interactions with the data capture interface, and that assessment strategy can change depending on the attribute being assessed. Following a review of this data it was possible to make this assumptions:

1. Had the assessor made a conscious effort to consistently evaluate the vehicles then aspects of the driver behaviour made between interactions with the data capture interface should be very similar. Identifying these aspects of driver behaviour would illustrate how the assessor made decisions.

A method has been developed, referred to as "Overlapping", which helps pin-point the assessment strategy most used to assess the vehicles [7]. The method is applied by considering the assessment of each attribute. Using the assessors interactions with the data capture interface the speed time profile for each attribute is then divided into a number of individual "events". Each event represents an action which resulted from a decision made by the assessor (figure 4b). An algorithm was developed which shifts each of the individual events on top of each other so that there is maximum overlap between them. The overlapped events are then displayed onto a 3D frequency distribution chart (figure 4c). Here the colour intensity gives an indication of how much of the assessor's time was spent at that driving condition

Once the methods were available to understand what assessors do in the evaluation it was possible to compare the influence of interactivity on the evaluation process and demonstrate how driver behaviour can be linked to assessors' opinions.

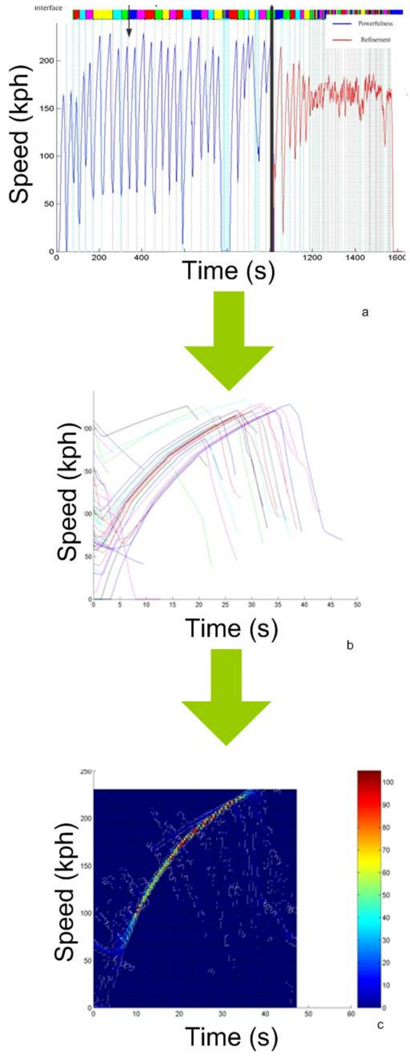

Figure 4: Overlapping method

2.2. Influence of Interactivity

To understand the implications of using fully interactive evaluations over non-interactive ones, a comparison of the outcomes of the two evaluation modes was conducted. 36 assessors were asked to rate 6 cars according to how powerful and how refined they believed each one to be. All the assessors experienced both evaluations modes. In addition a questionnaire was conducted at the end of the fully interactive evaluation. This was used to find out if the assessors had knowingly chosen to change their assessment strategy according to the attributes assessed.

The mean scores for each car were compared for each level of interactivity and the relationship between each mode was tested using Pearson's correlation. For Powerfulness the relationship was significant at the 0.01 level. For Refinement the relationship was not significant.

The responses from the questionnaire were analysed using Content Analysis. This is a method used to quantify the value of themes in transcripts. For Powerfulness the assessors claimed to concentrate on the assessment during acceleration and at full throttle. For Refinement they attempted steady speed assessments.

To fully understand how the stimuli influenced subjective opinion it was necessary to observe the assessors' driver behaviour. Each of the assessors' driving strategies was visualised using the Overlapping method. The assessors were placed into three groups according to how they drove. The first group contained those assessors who drove the vehicles at Wide Open Throttle conditions (WOT). The second group contained those assessors who drove at constant speed, and the final group contained those assessors whose assessment strategy was not sufficiently consistent for the Overlapping algorithm to identify any clear patterns. The percentage of people in each category is shown in Table 1 and 2.

Powerful Sample size 34	Constant speed	WOT	No Pattern
% Participants	9%	68%	24%
Mean Speed	35 mph	N/A	N/A
Mean Starting Speed	N/A	15 mph	N/A
Mean Final Speed	N/A	88 mph	N/A

Table 1: Assessment strategy for powerfulness

Refined Sample size 36	Constant speed	WOT	No Pattern
% Participants	19%	53%	28%
Mean Speed	58 mph	N/A	N/A
Mean Starting Speed	N/A	11 mph	N/A
Mean Final Speed	N/A	90 mph	N/A

Table 2: Assessment strategy for refinement

It is evident that the driver behavior results confirmed the findings from the questionnaire responses. For powerfulness 68% of the assessors performed WOT maneuvers and only 9% performed a constant speed assessment. For refinement 53% of assessors performed a WOT assessment, whereas 19% assessed the cars at constant speed. This confirms the relationship between the two modes of interactivity may be related to shape of the speed-time profile. In fact, if those assessors who chose to evaluate refinement at constant speed are not considered in the comparison between the two evaluations modes, for this attribute, then the relationship between the two modes is also significant at the 0.01 level.

Furthermore by splitting the scores for refinement according to how the assessor drove, it is then possible to observe how different driving conditions were scored as illustrated in Figure 1. The vertical axis illustrates the mean scores associated with the sound by those assessors who drove at constant speed. The horizontal axis gives the mean scores by those assessors who assessed the vehicles at WOT conditions.

The results portrayed within this figure are interesting because car C is in-fact a simulated concept vehicle. This vehicle is based on an existing vehicle, known as the "donor". The donor would have been modified and manipulated to simulate engineering changes. This was conducted with the knowledge of the key acoustic features that influence the perception of the customer. However much of the sound quality engineering effort for this concept sound was focused on its characteristics during WOT conditions. Its characteristics at constant speed had not been considered for listening evaluations. Evidently the modifications made have paid off as the vehicle performs better than its competition at WOT conditions. However car B has out performed all of its competition at a constant speed assessment.

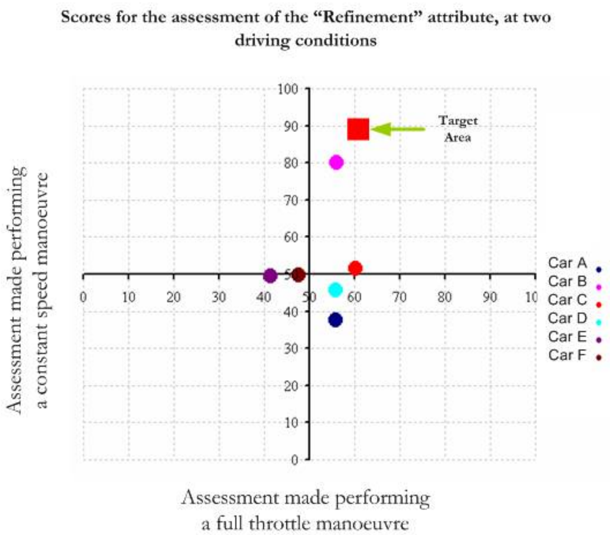

Figure 5: target setting approach

This illustrates that there is a need to change from processes which focus on one operating condition so that they cater for all conditions delivering whole vehicle sound quality to the customer. With the use of the simulator and the observational methods discussed above, it is now possible to link driver behaviour with subjective impressions, allowing for all driving conditions to be assessed in sound evaluations in the initial stages of the vehicle development program.

3. Next steps: Structured Evaluations

A limitation of conducting fully interactive evaluations is illustrated in tables 1 and 2; it can be seen that between 23% and 28% of assessors do not assess each vehicle in a consistent manner. In order to avoid this, it is possible to apply constraints to the VE so that each assessor performs the same assessment strategy. For example, to help the assessors evaluate the sound at constant speed whilst allowing them to drive freely, it is possible to ask them to follow virtual traffic, which can have a pre-determined speed. On the other hand, for the evaluation of higher load conditions, assessors can be asked to perform overtaking manoeuvres. By controlling the speed of both the car in front and that of oncoming traffic, it is possible to create suitable overtaking conditions.

To achieve an optimized process, more must be learned about the ideal driving conditions to evaluate. This can be done by using the findings from the observation studies reported within this paper as guidelines. However it is more appropriate to observe assessment strategies used in the real world and to implement those within the simulator. This approach is the subject of ongoing research. New studies which observe assessor behaviour in real cars and on real roads are being conducted. Back-to-back assessments cannot be performed in real cars, and rating scales are not effective for non-experts. Instead, assessors are being asked to comment on aspects of the sound which they like and dislike, whilst driving.

In a similar manner to the simulator based evaluations, many forms of data are being collected. Impressions of the sound are captured through audio recordings of assessors' comments and video capture of their facial expressions and road conditions. Data from the vehicle, such as vehicle and engine speed, and the route the assessor chose are being captured and at the same time the sound of the car is being recorded. All data are time-stamped allowing the reactions of the assessor to be related to driving conditions and the sound of the vehicle.

Preliminary results have highlighted the need for advanced analysis capabilities. Assessors' subjective impressions can take many forms, such as subtle facial expressions and head movement and hand gestures. The terminology used in their commentary depends on the type of assessor. For example NVH experts and key decision makers describe sound with technical language, although this will vary depending on their job function and their relationship with the vehicle. Moreover they are more sensitive to minor engineering changes and error states than customers. Non-experts, on the other hand lack the same level of articulation. Often they are more descriptive and relate the impressions to other experiences, using onomatopoeias as a means of expression.

To date, the driving conditions performed on road show similarities with those conducted in the simulator. The main difference is that assessors are not able to conduct repetitions of the same driving manoeuvres so frequently, which was a distinct aspect of the evaluation in the simulator. This is caused by real-world constraints such as the presence of other traffic or the type of road.

In summary, this research will ensure that sound quality evaluations conducted in a simulated environment will have the same outcome as if the evaluation was conducted in the real world, through the development of further guideline improvements on the appropriate use of interactive NVH simulators

4. Conclusions

The observational techniques presented within this paper have allowed the capture of assessors' evaluation strategies using an interactive NVH simulator. This allows subjective impressions of the vehicle to be more closely related to the operating conditions assessed during the evaluations. By linking this approach to new knowledge of on-road appraisal strategies, new insight will be provided to NVH engineers. Using a simulator, they will be able to identify the key acoustic features that influence drivers' perceptions of a vehicle, and crucially will be able to make better decisions regarding the design of the components that contribute to the sound. In doing this, the delivered vehicle's sound will be more likely to exceed the customers' expectations.

Acknowledgements

The authors acknowledge the UK Engineering and Physical Sciences Research Council for supporting the research through the Warwick Innovative Manufacturing Research Centre and the Engineering Doctorate Programme.

Reference

[1] M. Allman-Ward, J. Venor, R. Williams, "The interactive NVH simulator as a practical engineering tool" *Proceedings of SAE (2003) paper number 2003-01-1505*

[2] S. Giudice, P. Jennings, J. Fry, R. Williams, G. Dunne, "Development and Evaluation of a new Sound Quality evaluation approach for interactive NVH simulators" *Proceedings of Internoise (2006)*

[3] R. Williams, F. Henderson, M. Allman-Ward, G. Dunne, M.Cockrill and S. Adlard, "Using and Interactive NVH Simulator for target setting and concept evaluation in a new vehicle programme" *Proceedings of SAE 2005 paper number 05NVC-156*

[4] P. Jennings, S. Giudice, J. Fry, R. Williams, M. Allman-Ward and G. Dunne. "Developing Best Practice for Sound Quality Evaluations using an Interactive NVH Simulator" *Japanese Society of Automotive Engineering. (2006).*

[5] S. Baker, P. Jennings, et al. "Improving the effectiveness of paired comparison tested for automotive sound quality". *17th International Congress on Sound and Vibration, St. Petersburg, Sound and vibration (2004).*

[6] M. Allman-Ward, R. Williams, G. Dunne and P. Jennings, "The evaluation of vehicle sound quality using an NVH simulator" Proceedings of the 33rd *International Congress and Exposition on Noise Control Engineering 2004.*

[7] S. Giudice, P. Jennings, R. Williams, G. Dunne, "Using and interactive NVH simulator to understand driver behaviour during sound evaluations" *Proceedings of SAE (2007) paper number 07NVC-344.*