inter-noise 2004
Prague • Czech Republic • August 22-25

The 33rd International Congress and Exposition
on Noise Control Engineering

The evaluation of vehicle sound quality using an NVH simulator

M. Allman-Ward[a], R. Williams[b], G. Dunne[c], and P. Jennings[d]

[a,b]Sound & Vibration Technology, Station Lane, Millbrook, Bedfordshire. MK45 2YT. UK.
[c]W/1/009, Jaguar Engineering Centre, Abbey Road, Whitley, Coventry. CV3 4LF. UK.
[d]Warwick Manufacturing Group, University of Warwick, Coventry. CV4 7AL. UK.

[a,b][mallman;rwilliams]@sovibtech.com;[c]gdunne@jaguar.com;
[d]paul.jennings@warwick.ac.uk

Abstract [515] Rigorous evaluation of sound is a critical requirement of the NVH vehicle development process. This applies early in a programme when defining the correct sound for customer and brand perception, during the ensuing target cascade to subsystem and component level, and later when evaluating and improving the design.

Traditionally NVH evaluations have been made whilst driving a car on the road, or by auditioning fixed test conditions, such as wide open throttle acceleration, in a listening room. Both methods have significant drawbacks: On-road testing is often inconsistent and unrepeatable due to poor control over the test conditions, and in-room jury evaluations are difficult for non-experts and often biased due to evaluation of sound in isolation.

A full scale NVH Simulator has been developed to overcome these problems. It provides a realistic in-car environment where the subject can experience multiple stimuli (accurately reproduced sound and vibration as well as visual) whilst either driving or being driven.

This paper describes an initial study on six luxury vehicles that were subjectively evaluated by expert and non-expert assessors on-road, in-room and in the NVH simulator using both preference and rating methods. Objective and subjective data is presented which compares the results of these approaches and clearly demonstrates the benefits of using an NVH simulator in the vehicle development process. The findings from the study have also been used to define the role of the simulator in future academic research investigating the psychology and methodology of sound evaluation in a multi-modal environment whilst freely driving.

1 INTRODUCTION

Sound evaluation techniques, based on in-room Jury analysis, are an accepted method in the automotive industry to form a consensus of opinion when making sound quality decisions. A typical example is the evaluation of 2nd gear wide open throttle acceleration (2GWOT) of competitor cars against the target sounds for a new vehicle. However, there are issues with this approach because, in

reality, the sounds would be evaluated whilst driving a vehicle, and in the presence of other stimuli, all of which significantly influence the perception of sound.

A new approach was therefore required to enable subjective evaluation of vehicle interior noise in the right context. A driving simulator with accurate sound and vibration replay, called NoViSim, has been developed for this purpose. NoViSim allows people to assess the experience of driving the sound and vibration of a vehicle, as opposed to listening to the fixed, pre-prepared sound of a repeatable vehicle operating condition such as third gear full load acceleration.

The outcome of this development is an improvement in the effectiveness of NVH engineering, by enabling:

- Improved sound quality target setting capability
- Evaluation of component or sub-system contributions to interior sound quality
- Evaluation and selection of the best concepts for delivering the sound quality target
- Review of the measured or predicted performance during the detailed development phase of a new vehicle.
- Greatly improved speed and accuracy of subjective appraisal for NVH data.

In short, better decisions can be made earlier and throughout the vehicle programme, enabling improved efficiency and less re-engineering.

This paper describes results from an initial study, which is still on-going, to develop a robust evaluation method that allows the full benefits of NoViSim as a decision making tool in the engineering process to be fully realised.

2 THE NVH SIMULATOR

NoViSim is a fixed base vehicle simulator, meaning that it comprises a car stationed in a room with a projection screen in front of it, on which is projected a virtual environment (Figure 1).

Figure 1: The NVH Simulator

The assessor “drives” the vehicle through this virtual environment by interacting with the vehicle controls in the same manner as if they were driving the vehicle in the “real world”. Whilst performing this task of driving, the assessor also interacts with the correctly reproduced sound and

vibration of the particular vehicle that they are being asked to evaluate. A detailed description of the simulator is given in [1]. Some of the key features most relevant to this paper are described below.

2.1 Simulator driving conditions

The Simulator currently has two driving conditions or operating modes, namely "Free-Driving" and "Non-Interactive Driving".

"Free-Driving"

In this mode of operation, the Simulator responds entirely to the driver's inputs under the control of a performance prediction module. This means that the driver has to start the vehicle engine, put the vehicle in a gear of his choice, and operate the pedals and steering wheel as would normally be done when driving a real car.

"Non-Interactive Driving"

Because Free-driving requires the assessor to drive the vehicle, it follows that the driving actions are unlikely to be completely consistent between assessors, and the data analysis developed for perfectly repeatable tests is therefore not applicable. The simulator therefore also has the capability to replay a specific driving condition, for example 2GWOT, either from recorded sound and vibration time histories, or to reconstruct them from the "free-driving" data files. The replay of these fixed condition events is triggered by the juror, either using the touch screen or the throttle pedal. Driver inputs during the evaluation are ignored, and therefore the simulator reproduces the test condition with 100% repeatability for each assessor, and advanced data analysis techniques can be used to derive merit scores and other metrics.

2.2 Visual Scenarios

Experience was gained during the study about which features displayed on the screen added to the immersivity of the driving experience (e.g. using road markings to give the correct sensation of speed) and which were distractions (e.g. interesting shop fronts, other traffic) or very difficult to drive (e.g. sharp bends). Using this information, a visual scenario has been prepared which is well suited to the needs of interactive subjective evaluation of vehicle sounds.

3 ASSESSING SOUND IN THE NVH SIMULATOR

In order that the simulator can be a useful engineering tool and play a leading role in the NVH process, it has been necessary to develop two evaluation methods that are appropriate for an interactive multi-stimulus environment.

3.1 Free-Driving NVH Assessment

A version of "the slider bar" evaluation interface used throughout the NVH community has been incorporated into the NVH Simulator to control and record free-driving assessments.

Method

As shown in Figure 2, this interface provides a button for each car with a voting slider bar next to each one. The user selects which car to drive by touching one of the buttons, drives the simulation of that vehicle, and registers their subjective impression using the slider bar.

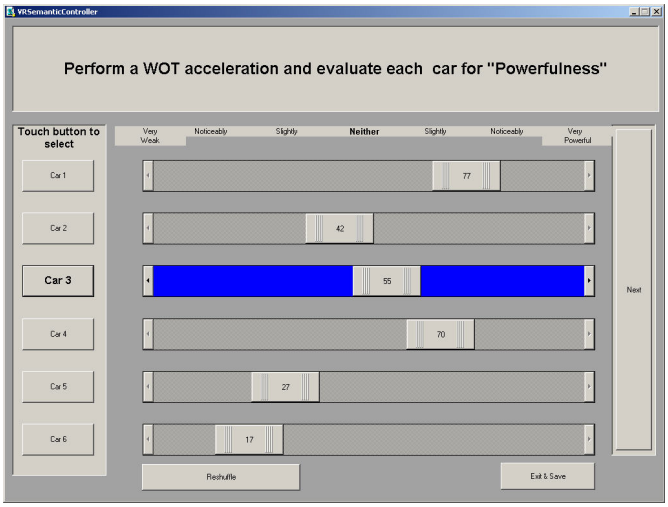

Figure 2: NVH Assessment Interface

This interface allows the test administrator to define the request displayed at the top of the screen, and the semantic words or numerical scale displayed above the bars. The 'Next' button takes the user to a new page where the car buttons and sliders are repeated, but the question and scale are different. This allows a complete simulation of an on-road 'ride and drive' to be carried out in a very efficient and controlled manner, and also enables direct back-to-back comparison of vehicles. An interesting benefit of this interface is that the main request can either be very specific, such as "Select 2nd gear, perform a full throttle acceleration from 50 km/h and assess Powerfulness" or quite free, such as "Evaluate each car for Powerfulness". Using the first form of the request, each assessor should drive each car in a very similar manner and therefore the tests are likely to be repeatable and consistent and should provide meaningful subjective data about the car sounds. Using the second form of the request, each assessor is free to drive in a way that they normally would on the road when assessing Powerfulness. Therefore, by recording how the assessors perform the test, it is possible to identify how drivers really evaluate Powerfulness (so far none out of 26 subjects have used a 2nd gear wide-open throttle test), and ensure that these 'real' driving conditions are adopted as the basis for engineering decision making regarding Powerfulness.

Results

The assessment tool described above was used in the NVH simulator to emulate a typical drive appraisal NVH assessment in the six cars. The 26 assessors who took part were asked to drive each car freely (i.e. they were not constrained to a fixed driving event such as 2GWOT) and use the sliders to record their overall perception of 'Powerfulness'. Figure 3 compares the results from the in-simulator test with an on-road drive appraisal of the same vehicles.

Vehicle	On Road Absolute Score	On-Road Ranking	In-Simulator Semantic Score	In-Simulator Ranking
Car 1	6.4	5	58	3
Car 2	7.3	1	80	1
Car 3	5.4	6	23	6
Car 4	7.3	1	62	2
Car 5	7.2	3	48	5
Car 6	7.1	4	51	4

Figure3: Comparison of on-road and simulator free drive appraisals

Conclusions

It is worth noting that an absolute numerical rating scale was used for the on-road assessments in the real cars but a semantic scale was used for the assessments of the same cars in the NVH simulator. Whilst this means it is difficult to compare the scores directly it is still possible to compare the rankings, which correlate reasonably well. There are many possible reasons why the simulator results might be different from the on-road assessment, including:

1. The driving scenario in the simulator is simpler and much more consistent than on the road.
2. Assessors can be quite sure of their judgement in the simulator as they can switch back and forth between cars at any time.
3. There is no 'badge' bias, as assessors do not know which cars they are driving.
4. NoViSim focuses on the sound and vibration stimuli – it does not reproduce whole vehicle motion.

However, the main objective of this part of the study was satisfactorily achieved, namely to provide a means of recording peoples' subjective response to the sound (and/or vibration) whilst freely driving and having the ability to switch between cars.

Work has begun on the next step, which is to develop this approach into a robust subjective evaluation method that will provide results with a high statistical confidence, and for which the psychology of evaluation in a free, multi-stimulus environment is well understood.

3.2 Fixed Driving Jury Evaluation

The Jury Evaluation methodology developed for the simulator is based on the traditional paired-comparison approach widely used in the automotive industry. The traditional method of evaluating *n* sounds consists of presenting *n(n-1)* pairs of sounds in the neutral environment of a room and, for each pair, asking the juror to indicate which sound in each pair is preferred. The assessment question can either be in terms of a measurable quantity, such as loudness, in terms of a subjective impression, such as 'sportiness', or in terms of the suitability of the sound for a particular brand image or market. The "best practice" techniques for running such experiments in a purpose built listening room and analysing their results are well documented [2] and already incorporated in the NVH processes at many vehicle manufacturers and their suppliers.

However, the adaptation of these methods to work in the complex multi-modal environment of a driveable NVH simulator was not straightforward, and an iterative methods development approach was adopted whereby feedback from a team of assessors was used to define and prioritise the improvements needed.

The study was based on measurements taken on the road in six different cars covering the full envelope of normal driving conditions. In order that the full set of stimuli could be applied it was necessary to measure vibration at the driver's sensitivity points as well as binaural sound, vehicle speed, engine speed and throttle position. This data was replayed in the NVH Simulator for the 2nd gear Wide Open Throttle driving condition using the "Non-interactive" driving mode previously described. This enabled direct comparisons to be made between "in-room" and "in-simulator" jury evaluations.

Three full evaluations of the 2nd gear Wide Open Throttle sounds were performed in this study consisting of:

In-Room Jury Evaluation

A Paired Comparison jury evaluation to assess the "powerfulness" and "refinement" of 6 cars was carried out in a Listening Room. The main purpose of this study was to provide a reference set of in-room subjective preference results for comparison with the corresponding evaluations to be performed non-interactively in the NVH Simulator. 27 jurors took part, six at a time. 25 were NVH engineers or managers, experienced in this type of evaluation, and two were naive.

Initial Simulator Jury Evaluation

The first Jury Evaluation in the NVH Simulator attempted to reproduce the in-room evaluation methodology (i.e. no juror interaction with the sound or the order in which sounds were played) but in a more realistic driver's environment. Voting was carried out using a keypad, exactly as for the in-room evaluation. 21 of the same jurors took part in the evaluation. After completing the evaluation, each juror was asked to describe the experience, to identify any concerns and to make suggestions for improving the methodology and/or the simulator. The major findings were then prioritised and used to define improvements to the hardware and software of the simulator.

Improved Simulator Jury Evaluation

A significant number of modifications were incorporated into the simulator and jury test to take into account the feedback from the initial study. These included:

1. Allowing the juror to be able to choose freely either vehicle in a pair for replay and only to vote when ready
2. The introduction of a touch screen, located in the simulator, which allowed the juror to receive clear instructions (Figure 4) and easily register a vote

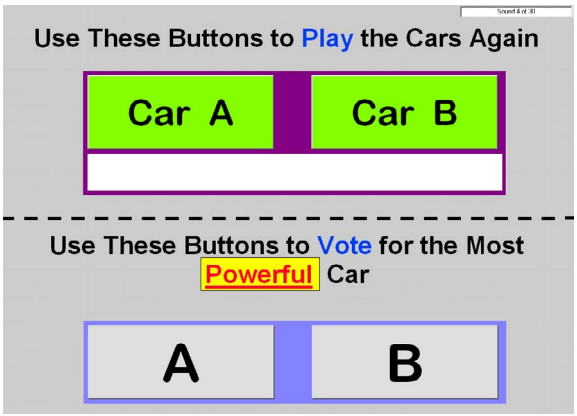

Figure 4: Jury Evaluation Interface

Feedback from assessors has indicated that the ease with which they could control and record responses during an evaluation more than compensated for any adverse effect that the presence of the screen has on the feeling of immersivity.

3. Changes to the visual and driving scenarios to minimise distractions whilst maintaining immersivity
4. Fine tuning of the flow of the experiment. A neutral splash screen is used after each event to separate the events, and smooth the transition from the end of the first 2GWOT (driving at high speed) to the beginning of the next one (at low speed).

22 jurors took part in the evaluation, including 18 who had carried out the initial evaluation.

Results and Conclusions

The voting data from the in-room and improved simulator evaluation methods was assessed objectively using a number of different techniques. Some examples are given for the question "which car sounds the most powerful?" Figure 5 presents the Kendall Coefficient of Consistency (based on cyclic triads) for the three experiments [3]. Since the most common causes of poor consistency are that the juror either lost interest during the test or found it very difficult, this table indicates that the free-play simulator method is a significant improvement.

	In-Room	Initial Simulator (Fixed-Play)	Improved Simulator (Free-Play)
Percentage of jurors better than 80% Consistent	51%	71%	95%

Figure 5: Consistency Results

This result, together with other significance tests, confirmed that in spite of the complex, multi-stimulus environment that exists in the NVH simulator, its results were at least as reliable as the in-room data. Figure 6a is a plot of the mean and standard deviation of the Total Preference Scores for all jurors for the Powerfulness question and 6b gives the corresponding normalised Overall Merit Scores.

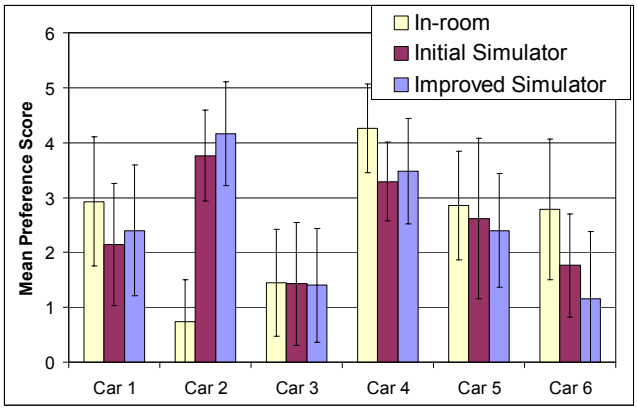

Figure 6a: Preference Scores

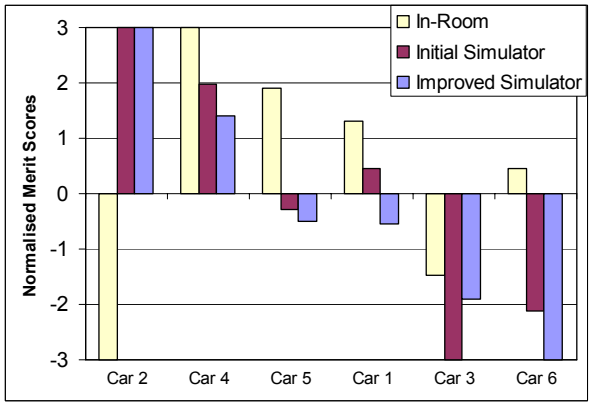

Figure 6b: Merit Scores

It is clear from both of these plots that there are differences between in-room and in-simulator preferences. There is an obvious anomaly with the results for Car 2 in-room data, which is being investigated. Evidence from these initial studies however suggests that the simulator results, especially the preference scores, are in good agreement with the on-road findings. Extensive

academic studies are now under way to confirm this and to improve the robustness of the evaluation process.

For the purposes of this development study, subjective feedback from the jurors was as important as the objective comparisons of the methods. The most significant comments were:

- The inclusion of Free-play made for a much more relaxed test. Jurors no longer felt under pressure and also felt much more satisfied with their decision making process.
- The holistic nature of the environment inside the NVH simulator made for a much more natural and believable context. Vibration, visuals and operating instruments all added to the ability to evaluate the sound correctly (in particular, the sensation of rate of change of speed was important for Powerfulness and vibration had an influence on Refinement).
- In comparison with the on-road tests, the ability to switch instantly from one car to another highlighted differences that were not apparent during ride and drive events.
- One very important result was that nearly all the jurors who participated indicated that they would be willing to take part in regular evaluations in the simulator. After one or two evaluations in a room, it becomes hard to attract jurors back, partly due to the difficulty of the test and partly due to its perceived value (since it seems so unrelated to the real driving experience).

4 SUMMARY

The rationale behind the development of the NVH simulator was to ensure that it had sufficient functionality to be a fundamental engineering tool in the NVH process of the future whilst, at the same time offering a proven capability for the rigorous evaluation of vehicle sound and vibration.

Several hundred evaluations have now been performed in the NVH Simulator involving NVH experts, vehicle programme managers and typical customers. The resulting feedback has been universally positive, and has been used to guide the development within the project as well as shaping future plans for the engineering application of the Simulator.

The subjective and objective results from this preliminary study indicate that NoViSim offers a significant improvement over current non-interactive evaluation techniques used in the NVH process. Future academic research [4, 5] will be aimed at improving the subjective evaluation methods used in the NVH Simulator. Particular emphasis will be placed on the psychology of multi-modal perception and the development of robust statistical analysis methods.

REFERENCES

[1] M. Allman-Ward, J. Venor, R. Williams, M. Cockrill, H. Distler, A. Crewe, and T. Heinz, "The Interactive NVH Simulator as a Practical Engineering Tool", in *2003 Noise and Vibration Conference Proceedings,* SAE, 2003, paper n. 2003-01-1505.

[2] N. Otto, S. Amman, C. Eaton and S. Lake, "Guidelines for Jury Evaluations of Automotive Sounds", *Proceedings of the 1999 Noise and Vibration Conference*, paper n. 1999-01-1802.

[3] MTS Jury Evaluation Software Manual, version 2.0. MTS Software Systems, pp. 123-124.

[4] J. Fry, P. Jennings, R. Williams and G. Dunne, "Understanding how customers make their decisions on product sound quality", *Proceedings of the 33rd International Congress and Exposition on Noise Control Engineering*, paper n. 170.

[5] S. Baker, P. Jennings, G. Dunne, R. Williams, "Improving the Effectiveness of Paired Comparison Tests for Automotive Sound Quality", *Proceedings of the 11th International Congress on Sound and Vibration*, 2004.