Audio Engineering Society

Conference Paper 3

Presented at the AES 5th International Conference on Automotive Audio
2024 June 26–28, Gothenburg, Sweden

Subjective Perception based on Acoustical Parameters for In-Vehicle Virtual Sound

Minsik Choi[1], Woochul Jung[2], Sungho Kim[3], May Jorella Lazaro[4], Marcin Kalinowski[5], Jeonghwan Lee[6], Soyoun Moon[6], Myung Hwan Yun[2], and Grzegorz Sikora[5]

[1]*School of Computing, Australian National University*
[2]*Department of Industrial Engineering, Seoul National University*
[3]*Department of Systems Engineering, Republic of Korea Air Force Academy*
[4]*CX Insight Team, Samsung Electronics Co.*
[5]*Harman International, Automotive Division, Garching bei München, 85748, Germany*
[6]*Harman International Korea, Automotive Division, Seoul 06792, Korea*

Correspondence should be addressed to Minsik Choi (minsik.choi@anu.edu.au)

ABSTRACT

Automotive manufacturers are engaging with audio system providers to offer unique auditory experiences. This is achieved through the replication of acoustic environments, coinciding with the growing prominence of infotainment systems. Gaining an understanding of how acoustical parameters impact human perception is essential in the development of virtual acoustic venues within this context. Hence, this is an initial study to explore the correlation between objective measures and subjective responses in reproduced in-vehicle virtual acoustics, with the aim of augmenting the auditory experience for passengers. A jury test was conducted, and the resulting data was subjected to statistical analysis. The findings align with previous studies except for intimacy, indicating visual disparities in virtual environments. Also, while reverberance, linked to envelopment, tends to be influenced by early decay time, optimizing reverberation time within a specific range can improve the auditory experience in vehicles by enhancing naturality, which is correlated with key variables including reverberance, envelopment, and overall impression.

1 Introduction

Car manufacturers are collaborating with audio system providers to create unique auditory experiences through virtual replication of acoustic environments [1]. In the field of system interface design, where visual aspects dominate, the significance of sonic interaction cannot be overlooked, and the development of in-vehicle acoustics has become a crucial consideration with the rise of infotainment [2, 3, 4]. Music listening is a fundamental aspect of in-vehicle infotainment, and the acoustics plays a pivotal role in shaping auditory user experience with passengers [5]. Gaining an understanding of how acoustical parameters impact human perception is significant in creating virtual acoustic venues in this context. Previous studies examining the correlation between objective measurements and subjective responses in room acoustics have provided insights for the approach to in-vehicle virtual acoustics.

Barron [6] conducted a study encompassing eleven

British concert halls, which categorized subjects into two groups based on their preferences for reverberance or intimacy. The study revealed that the perception of envelopment differed between these two groups. The correlation analysis of the study identified that total sound level, early decay time (EDT), and early lateral fraction (LF) were factors influencing subjective responses from listeners. Barron et al. [7] also investigated the subjective effects of early lateral reflections through experiments, establishing a relationship between spatial impression (SI) and the early lateral energy fraction within the first 80ms after the direct sound. Additionally, Imamura et al. [8] conducted an experiment on perceived sound clarity, focusing on the arrival direction and delay time of the first reflections, and Bradley et al. [9] conducted a study on envelopment, describing six experiments that involved subjective ratings of listener envelopment (LEV) and apparent source width (ASW) in 16 concert halls.

The study of room acoustics design, particularly with respect to concert halls, and its relationship with subjective responses from listeners has yielded insights; however, the relationship between objective measures and subjective responses in the context of in-vehicle acoustics has been relatively neglected. Although insights gained from studies on concert hall acoustics have informed the design of in-vehicle acoustics, more research is needed to understand the interplay between objective measures and subjective perceptions in the automotive contexts.

Therefore, the purpose of this study is to investigate the relationship between objective acoustic parameters and subjective responses in the context of reproduced in-vehicle venues, with the goal of enhancing the auditory experience for passengers. A jury test as psychoacoustical validation was conducted by 30 musicians who possess trained abilities to detect acoustic features compared to the general population. The resulting data from the jury test were statistically analyzed to establish the validity of the findings and derive insights. The musicians completed a questionnaire based on their perception of virtual acoustics, and the results were also compared to those of previous studies conducted in actual concert halls to identify any notable differences. Any disparities in trends were investigated to identify potential causes, and some design approaches were proposed to address such differences in the in-vehicle environment.

2 Jury Test

2.1 Participant

Participants who possessed musical expertise were recruited. Prior investigations in the field of music cognition have suggested that musicians possess enhanced auditory perception abilities compared to non-musicians [10, 11]. This advantageous trait not only aids in the accurate validation of acoustic stimuli but also access to their preferences as users. Zhang [12] defined musicians as individuals holding at least an undergraduate degree, having undergone a minimum of six years of professional music training, and spending a minimum of one hour per week to musical practice. For the purpose of our experiment, we selected participants who satisfied all three criteria above. Ethical approval was obtained from ANON Institutional Review Board.

Furthermore, we devised a screening test including ten questions to assess each participant's logical perception of sound details. Subjects who correctly answered a minimum of eight questions and have normal hearing abilities in both the left and right ears were considered eligible. The screening test utilized musical excerpts from Walter Rabl's Quartet for Piano, Violin, Clarinet, and Cello, Op. 1, and Bach's Invention No. 13 in B minor, Op. 784. The hearing assessment exploited a tester to evaluate participants' proficiency in discriminating between left and right sounds, as well as pitch perception at frequencies of 0.5k, 1k, 2k, and 4k Hz, while maintaining a constant loudness level.

A total of 32 musicians meeting the established criteria were initially recruited for this study, with 30 of them successfully participating in the experiment after the screening test (M = 6, F = 24). As for the age of the participants, 16 people were the most between 25 and 30 years old, 12 were 25 years old and under, and 2 were the least between 30 and 35 years old. In terms of the participants' academic backgrounds, their majors spanned across a diverse range of disciplines. 3 participants specialized in piano, 3 in western classical vocal, 2 in composition, 5 in music theory, 10 in strings, 5 in winds, and 2 in Korean traditional music. The number of years dedicated to professional music training varied, with 10 individuals having studied for more than 15 years but less than 20 years, 9 for more than 10 years but less than 15 years, 7 for more than 20 years, and 4 for more than 6 years but less than 10 years. All participants attended concerts at least once a month and listen to music through the in-vehicle system at least once a week. Participants adhered to pre-experiment guidelines, refraining from smoking, consuming alcohol, or caffeine, and engaging in strenuous exercise. No signs of fatigue were observed among the participants as well.

2.2 Virtual Environment

A test environment was established in the experimental setup to create virtual acoustic venues, as outlined in the research by von Tuerckheim and Münch [13]. Six venues, including a reference, were prepared using Virtual Venues software, developed by Harman. The reference venue represents the original sound from the car audio system devoid of any virtual effects, while remaining five venues consist of varied acoustic parameter values within the reverberation core A/B, enabling the presentation of virtual venues with distinct acoustic environment perceptions described in Figure 1. To conduct the jury test, a mid-size sedan (Genesis G70) was selected as the designated test car. The car seats were leather, and the passenger volume was 95.9 cu ft. Four microphones and twenty-three speakers were installed within the car, following the speaker and microphone configurations proposed in the aforementioned study [13]. The speakers were driven by an external class D amplifier.

Once the hardware installation for the speaker, microphone, and amplifier was completed in the test car, the acoustic tuning process was carried out by experienced acoustic engineers with over a decade of expertise in automotive acoustics. Tuning parameters such as gain, equalization, and delay for each channel were controlled by Virtual Venue software. The objective of this process was to achieve a flat frequency response within the reference venue. Moreover, a symmetric tuning approach was implemented for both the driver and passenger seats to ensure consistent sound performance across both seating positions.

A set of four acoustic parameters was primarily selected to form the whole parameter pool, based on discussions conducted with industry experts possessing over a decade of experience in room acoustics. These parameters included RT via T30 method, EDT, Definition (D50), and LF, which have been commonly utilized in previous studies to examine their correlation with clarity and reverberance. Additionally, C80 and Ts were incorporated into the parameter pool. To

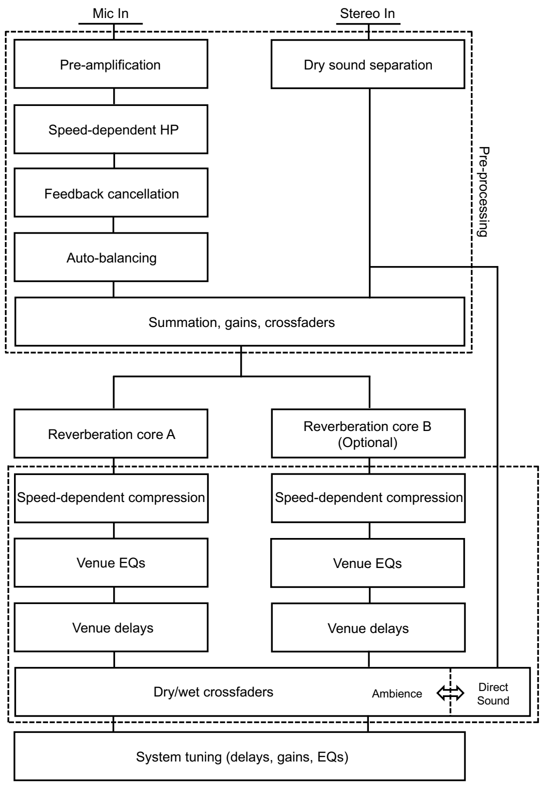

Fig. 1: This figure shows the design architecture for creating the virtual venues outlined in [13].

capture a comprehensive analysis, measurements were performed across three frequency ranges: low (bass) frequencies (125 Hz, 250 Hz), mid-frequencies (500 Hz, 1000 Hz, 2000 Hz), and the entire frequency spectrum (125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz), drawing upon the groundwork established in Barron's earlier research [6]. The objective acoustic parameters for each venue measured in the car are presented in Table 1.

The measurement of acoustic parameters in the test car was conducted using an ambisonic microphone and Marshall Day Acoustics IRIS, following the guidelines by ISO3382-1:2009 [14], described in Figure 2. The microphone was positioned at ear level in the driver's seat for accurate recordings. Analysis of the measurement results for the reference venue revealed the lowest values for RT using the T30 method, EDT, and the highest values for C80 and D50 among the whole virtual venues. Based on previous research by Barron [6], the reference venue will exhibit the highest clarity scores and lowest reverberance and envelopment scores in the jury test. Conversely, the room 5 venue, which is most tuned, demonstrated the highest values for RT and EDT, along with the lowest values for C80 and D50. Consequently, the room 5 venue will receive the lowest clarity scores and highest reverberance scores compared to other venues. Furthermore, the measurement results for Room 4 indicated the highest value for LF, leading to the highest envelopment score.

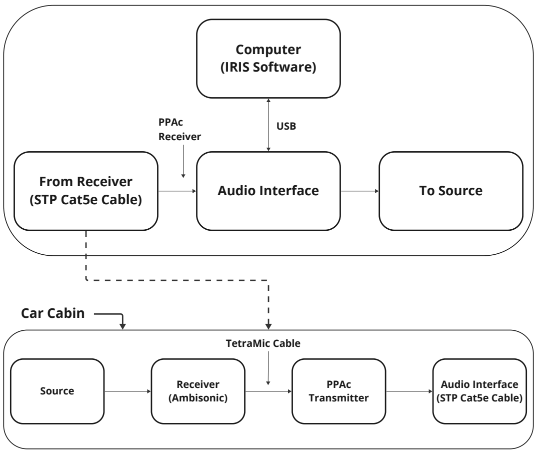

Fig. 2: This figure illustrates the complete process of measuring acoustic parameters in the test car as described in [15].

2.3 Sound Stimuli

Two classical music pieces in WAV format, with a sample rate of 44.1 kHz and a bit depth of 16, were used as sound stimuli for the jury test. The primary piece chosen for the main experiment is the overture from Mikhail Glinka's opera, Ruslan and Lyudmila, Op. 5. The musical excerpts utilized in the experiment were sourced from a recorded album within the controlled environment of an anechoic chamber [19]. The initial 1 minute and 35 seconds of the introduction were extracted for the stimuli. The overture captivates listeners with its striking opening, characterized by the vibrant combination of brass and timpani instruments. It further unfolds with a prominent first theme performed by the high strings, followed by a contrasting second theme presented by the low strings. Noteworthy woodwind solos suitably bridge the two themes, providing a unique opportunity for participants to discern the acoustic properties of different registers and timbres. Moreover, to facilitate the practice session, the opening 30 seconds of the overture from Wolfgang Amadeus Mozart's opera The Marriage of Figaro, Op. 492, were used. This orchestral piece was also recorded within an anechoic chamber [19], enabling participants to explore both the lowest and highest ranges of reverberation and set a standard for evaluation within the experiment environment. The SPL of the peak of each stimulus was set to 80dB(A) at the driver's position.

2.4 Measurements

The jury test employed subjective measures for acoustics evaluation: clarity, reverberance, envelopment, intimacy, naturality, and overall impression. Most of these measures were derived from previous research that primarily concentrated on assessing the acoustic attributes of concert halls, as outlined by Barron [6]. Clarity refers to the ability to perceive sonic detail. Reverberance refers to the degree of perceived reverberation in a temporal sense, and envelopment refers to the spatial aspect of the perceived sound, the degree to which one feels surrounded by the sound. Intimacy refers to the extent of identification with the stimuli, whether one feels acoustically involved or detached from it. Overall impression refers to one's overall impression of the acoustics in the virtual venues for the given stimuli; however, naturality was specifically introduced for this study based on the outcomes of discussions conducted with experts. Naturality serves as an indicator of how closely the virtual acoustics in the vehicle replicate the natural acoustics experienced in concert halls. This measure was included due to the underlying principle of maintaining reverberation in concert halls, as well as the fact that this study originated from a larger investigation into concert hall acoustics evaluation.

The questionnaire utilized in this study encompassed two main components elaborated in Appendix A. The first part is a pre-experiment survey that sought to gather participants' basic information. The second part focused on the actual evaluation and consisted of 6 multiple-choice items presented on a 7-point Likert scale (ranging from 1—strongly disagree to 7—strongly agree) for each venue. The questions related to five subjective measures, excluding overall impression, were specifically designed to discern objective sound features from the standpoint of sound

Table 1: This table displays the acoustical parameters of the generated virtual venues, including reverberation time (T30) in seconds, early decay time (EDT) in seconds, center time (Ts) in milliseconds, clarity of music (C80) in decibels with an 80 ms integration time, clarity of speech (D50) in decibels with a 50 ms integration time, and early lateral energy fraction (LF), as defined in [16, 17, 18].

	Frequency Range	Frequency	Reference	Room 1	Room 2	Room 3	Room 4	Room 5
T30 (s)	Tb	125, 250 Hz	0.34	0.51	1.19	2.25	1.56	3.42
	Tm	500, 1000, 2000 Hz	0.07	0.52	1.15	2.21	1.45	2.59
	Tt	125, 250, 500, 1000, 2000 Hz	0.17	0.52	1.16	2.23	1.50	3.00
EDT (s)	Db	125, 250 Hz	0.16	0.36	0.50	0.62	1.05	1.66
	Dm	500, 1000, 2000 Hz	0.08	0.51	1.30	1.61	1.62	4.06
	Dt	125, 250, 500, 1000, 2000 Hz	0.11	0.45	0.98	1.21	1.39	3.10
Ts (ms)	Sb	125, 250 Hz	19.41	30.16	45.89	49.30	78.20	111.39
	Sm	500, 1000, 2000 Hz	5.49	16.43	39.98	40.66	88.59	139.81
	St	125, 250, 500, 1000, 2000 Hz	11.06	21.92	42.34	44.11	84.43	128.44
C80 (dB)		125, 250, 500, 1000, 2000 Hz	36.89	12.96	8.00	7.22	3.03	3.15
D50		125, 250, 500, 1000, 2000 Hz	0.99	0.88	0.74	0.81	0.51	0.64
LF		125, 250, 500, 1000 Hz	0.20	0.23	0.22	0.20	0.24	0.20

validation, while overall impression sought to capture participants' personal preferences. Validation entails identifying distinct aspects of sound characteristics that are recognizable to individuals, rather than focusing on subjective taste.

2.5 Procedure

Prior to the experiment within a stationary vehicle, participants were provided with comprehensive information regarding the research objectives, test procedures, questionnaire format and composition. Continuously, participants underwent a screening test, including a hearing assessment. Any queries or concerns regarding their involvement in the experiment were addressed by providing further clarifications upon boarding the vehicle. Once participants had acquired a foundational understanding of the experiment, two participants, accompanied by at least one moderator, entered the vehicle. Both participants occupied the front seats, while the moderators are positioned in the back seat. To ensure a seamless process, the sound stimuli for the experiment were externally controlled. Meanwhile, the moderator focused on internal situations, maintaining communication with the external moderator.

Initially, a practice session was conducted to familiarize the participants with the acoustic environments, using sounds from reference and room 5 venues as a baseline. Following the practice session, participants evaluated the six virtual acoustic venues twice, with the order randomized for the main experiment. If desired, participants had the option to replay the sounds during each evaluation. A mandatory five-minute break with ventilation was taken between the two trials. The first and second trials were identical with the only distinction being the seat position alternating between driver and passenger. In order to reduce external noise interference, the experiment took place within a serene laboratory environment. The doors and windows were closed, and only the vehicle's sound systems were activated, without the engine running.

2.6 Data Analysis

The quantitative data gathered from the experiment were subjected to analysis employing mean differences and linear correlations. To address the research questions pertaining to in-vehicle acoustic validation and the interrelation among different variables, One-Way Analysis of Variance (ANOVA) was used to recognize mean differences, while correlation analysis (Pearson) was used to examine linear associations between the measures. For the one-way ANOVAs, the null hypothesis states that there are no significant differences among the group means for each of the six subjective measures across the six venues. For the correlation analysis, the null hypothesis is that there is no significant correlation between any pair of the subjective measures. The data collected from the experiment underwent Z-score normalization. Z-score normalization considers the mean and standard deviation of the overall scores given by each participant, allowing for the identification of outliers and the determination of true individual score values. Normality assumptions were also verified by examining skewness, kurtosis, and adhering to the central limit theorem [20], and appropriate models were selected based on the results of homoscedasticity test (Levene's test). Additionally, to enhance the validity of the experimental data, a comparison between the driver's and passenger's seats was conducted, and no significant differences were observed across all dependent variables based on the seat position; however, the data exclusively from the driver's seat were only analyzed to ensure a higher level of result validity.

3 Result

3.1 Descriptive Analysis

Both reference (0.85) and room 1 (0.80) exhibited particularly higher mean clarity scores compared to the other venues, and higher mean reverberance scores were observed for room 4 (1.12) and room 5 (1.02) shown in Table 2. Room 4 displayed a lower RT value (1.50) compared to room 3 (2.23), while exhibiting a higher EDT value (1.39) than room 3 (1.21). In contrast to the unpredictable variations in RT across different venues, a consistent upward trend in EDT signifies a discernible pattern, and a modest (due to room 5) tendency between reverberance and EDT can be established, taking into account the concurrent increments in both variables. The mean envelopment scores for room 4 (0.83) and room 5 (0.65) yielded similar results to those of reverberance as well. Meanwhile, room 1 recorded the highest mean intimacy score (0.39) and room 4 exhibited a relatively higher mean intimacy score (0.35) surpassing the other intermediate venues. Establishing a clear relationship between intimacy and the acoustic parameters proves challenging, warranting further investigation through ANOVA to ascertain statistically significant differences between the mean intimacy scores of venues.

The mean naturality scores for room 2 (0.53), room 3 (0.26), and room 4 (0.40) were relatively higher compared to room 1 (-0.40) and room 5 (-0.26). The mean score for the reference venue (-1.45) was the lowest value. The venues with higher naturality scores, room 2, room 3, and room 4, recorded RTs ranging between 1.16 seconds and 2.23 seconds. This observation suggests a bounded correlation between naturality scores and RTs, highlighting the influence of reverberation on the perception of naturality. In addition, the mean score for overall impression shows an increasing trend from the reference venue to room 2, followed by a decrease in scores for room 3, room 4, and room 5. Particularly, room 2 records the highest mean score for overall impression (0.62). It is worth noting that a previous study has established a correlation between overall impression and RT, and the overall impression in the car also demonstrates a similar trend, but only within a specific range of RT (less than 1.16 seconds). Across all measures, including naturality, the obtained results accorded closely with our initial expectations, except for intimacy, which causes further discussion.

3.2 Analysis of Variance

One-way ANOVA was conducted to confirm the descriptive variations among the six venues for the six measures detailed in Table 3. Welch's test was employed for intimacy and naturality based on each homogeneity of variance. In accordance with the findings in the descriptive analysis, significant differences were observed for all the subjective measures, except for intimacy, at a significance level of 0.01. Post-hoc tests (Turkey, Games-Howell) further confirmed that the reference venue obtained the highest score for clarity, room 4 exhibited the highest score for reverberance and envelopment, and room 2 achieved the highest score for naturality. These outcomes align with the anticipated trends identified in the earlier analysis. In addition, the outcome for overall impression indicates the presence of significant differences among the venues. Further examination through post-hoc test revealed that room 2 obtained the highest score for overall impression, as mentioned in the descriptive analysis section.

3.3 Correlation Analysis

As anticipated from the descriptive analysis, clarity shows strong negative correlations with reverberance

Table 2: This table shows a summary of descriptive statistics for the six subjective measures across the six virtual venues; impression represents overall impression.

		Clarity	Reverberance	Envelopment	Intimacy	Naturality	Impression
Reference	Mean	0.846	-1.608	-1.179	0.131	-1.448	-0.581
	SD	0.941	0.543	0.670	1.084	0.837	0.998
Room 1	Mean	0.795	-0.683	-0.358	0.391	-0.402	0.444
	SD	0.510	0.698	0.841	0.730	0.924	0.836
Room 2	Mean	0.289	0.391	0.308	0.157	0.532	0.616
	SD	0.740	0.471	0.619	0.642	0.615	0.690
Room 3	Mean	0.109	0.670	0.645	0.008	0.264	0.144
	SD	0.716	0.621	0.598	0.900	0.794	0.866
Room 4	Mean	-0.376	1.122	0.827	0.348	0.399	-0.085
	SD	0.815	0.446	0.733	0.961	0.847	0.888
Room 5	Mean	-0.454	1.024	0.650	-0.141	-0.257	-0.705
	SD	0.884	0.692	0.779	0.989	1.103	0.958

Table 3: This table shows the results of one-way ANOVAs for the six subjective measures; impression represents overall impression.

	SS	df	MS	F	P
Clarity	46.424	5	9.285	15.254	.000
Reverberance	174.697	5	34.939	101.481	.000
Envelopment	90.091	5	18.018	35.082	.000
Intimacy	6.053	5	1.211	1.503	.191
Naturality	80.926	5	16.185	21.593	.000
Impression	43.062	5	8.612	11.166	.000

(-0.93) and envelopment (-0.89) shown in Table 4. Reverberance demonstrates a strong correlation with envelopment (0.99), which is higher than the correlation value (0.74) observed in the previous study by Barron [6] for individuals preferring reverberance in real venues. This suggests that the correlation between reverberance and envelopment in the in-vehicle virtual environment can be stronger than in physical venues. Furthermore, naturality shows strong relationships with reverberance (0.84) and envelopment (0.87). Due to the strong correlations between reverberance, envelopment, and naturality, these factors are interconnected and associated with RT via naturality. Maintaining an appropriate RT range is expected to enhance naturality, along with controlling reverberance and envelopment; however, extreme RT values, as observed in reference (0.17 sec) and room 5 (3.00 sec), result in lower naturality values. Naturality also reveals the strongest relationship with overall impression (0.59), although the p-value exceeds the significant level (0.22).

4 Discussion

EDT tends to exhibit a stronger correlation with reverberance when compared to RT, reflecting the concurrent increments in both reverberance scores and EDT values between the venues. Similarly, the mean envelopment scores demonstrated consistent results with the reverberance scores, as anticipated based on the measured acoustic parameters outlined in Section 2.2. Correlation analysis further revealed a correlation coefficient of 0.993 between reverberance and envelopment, indicating a strong relationship between the two variables. In the meantime, the RT values for room 2, room 3, and room 4, which received higher naturality scores, ranged from 1.16 to 2.23 seconds, suggesting a bounded correlation between naturality and reverberation. Corre-

Table 4: This table shows the results of correlation analysis between pairs of the six subjective measures; impression represents overall impression.

	Clarity	Reverberance	Envelopment	Intimacy	Naturality	Impression
Clarity	1	-.929**	-.886*	.387	-.594	.293
		.007	.019	.448	.214	.573
Reverberance	-.929**	1	.993**	-.270	.836*	.063
	.007		.000	.604	.038	.906
Envelopment	-.886*	.993**	1	-.222	.873*	.147
	.019	.000		.673	.023	.781
Intimacy	.387	-.270	-.222	1	.078	.545
	.448	.604	.673		.884	.264
Naturality	-.594	.836*	.873*	.078	1	.587
	.214	.038	.023	.884		.221
Impression	.293	.063	.147	.545	.587	1
	.573	.906	.781	.264	.221	

lation analysis unveiled strong relationships between naturality and reverberance as well as naturality and envelopment.

Significant mean differences were observed among the subjective measures, except for intimacy. The post-hoc test revealed that the reference venue attained the highest score for clarity, room 4 exhibited the highest score for reverberance and envelopment, and room 2 obtained the highest score for naturality. While the descriptive analysis indicated that room 1 received the highest intimacy score, the ANOVA results revealed no significant differences between the other venues. Previous studies have highlighted the correlation between intimacy and factors such as source-received distance [21] and sound level [6]; however, in the in-vehicle environment, it becomes challenging for listeners to accurately evaluate such factors. Moreover, the visual disparities between real and virtual venues could potentially impact the results for intimacy. This implicates that intimacy is the most intricate subjective measure to control for the impact level of reproduced venues within a car setting.

Clarity exhibits a strong negative correlation with both reverberance and envelopment, and reverberance demonstrates a solid correlation with envelopment, surpassing the correlation value observed in the previous study conducted by Barron [6]. This indicates that the relationship between reverberance and envelopment in the context of virtual venues within a car setting could be even stronger compared to real venues. The correlation analysis results for reverberance and envelopment imply that these two subjective measures can be mutually controlled, especially in the context of in-vehicle virtual venues. In addition, the newly introduced measure, naturality, exhibits a strong relationship with both reverberance and envelopment. Given the strong correlation between reverberance and envelopment, they can be considered a combined factor influenced by RT through their relationship with naturality. Consequently, optimizing RT within an appropriate range contributes to increased naturality for in-vehicle virtual venues illustrated in Figure 3. Naturality also exhibits a strong relationship with overall impression, although the significance level exceeds. The correlation analysis might be subject to limitations due to the relatively small venue sample size. Future studies should incorporate a larger number of virtual venues in the jury test. This will enable a more reasonable examination of the relationship between naturality and overall impression, yielding insights for the design direction.

Overall, the results of this study accord with the findings of the previous study conducted by Barron [6], apart from intimacy. It is worth noting that the visual disparities between real and virtual venues could potentially impact the outcomes of the correlation analysis

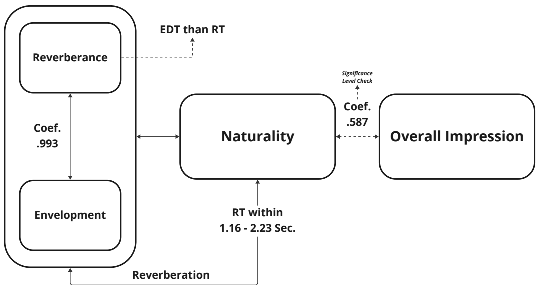

Fig. 3: This figure illustrates the dynamics among three key measures—reverberance, envelopment, and naturality—that potentially influence overall impression through bounded RT.

for intimacy, and that RT emerges as the most influential objective measure, displaying an interconnected relationship with reverberance, envelopment, and naturality. This concludes that by appropriately adjusting RT values, the levels of the subjective measures, namely reverberance, envelopment, and naturality can be efficiently improved in the context of in-vehicle virtual acoustics, and these improvements, in turn, contribute to an enhancement in overall impression.

5 Conclusion

This initial study aimed to examine the relationship between objective parameters and subjective responses regarding in-vehicle virtual sound, employing a combination of jury test and statistical analysis. To assess the subjective auditory experience, a set of key measures in the preceding study were selected in this investigation. Additionally, considering the software-based processing to replicate the acoustic environments, experts proposed the inclusion of naturality as an additional subjective measure, recognizing its potential influence on the valid assessment. Overall, the observed trends across different in-vehicle virtual venues not only align with the findings of the previous studies, but also demonstrate its distinctive features. The subjective factors of in-vehicle virtual acoustics, such as reverberance, envelopment, and naturality, should be acknowledged as decisive considerations, while following a design methodology akin to real venue acoustics to enhance the in-vehicle user experience. Future studies could explore contextual factors such as emotions and vehicle types, incorporating diverse participant profiles and musical genres to mitigate potential biases.

References

[1] Toole, F., "The measurement and calibration of sound reproducing systems," *Journal of the Audio Engineering Society*, 63(7/8), pp. 512–541, 2015.

[2] Marshall, D. C., Lee, J. D., and Austria, P. A., "Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness," *Human factors*, 49(1), pp. 145–157, 2007.

[3] Nees, M. A. and Walker, B. N., "Auditory displays for in-vehicle technologies," *Reviews of human factors and ergonomics*, 7(1), pp. 58–99, 2011.

[4] Tardieu, J., Misdariis, N., Langlois, S., Gaillard, P., and Lemercier, C., "Sonification of in-vehicle interface reduces gaze movements under dual-task condition," *Applied ergonomics*, 50, pp. 41–49, 2015.

[5] Kim, M., Seong, E., Jwa, Y., Lee, J., and Kim, S., "A cascaded multimodal natural user interface to reduce driver distraction," *IEEE Access*, 8, pp. 112969–112984, 2020.

[6] Barron, M., "Subjective study of British symphony concert halls," *Acta Acustica united with Acustica*, 66(1), pp. 1–14, 1988.

[7] Barron, M. and Marshall, A. H., "Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure," *Journal of sound and Vibration*, 77(2), pp. 211–232, 1981.

[8] Imamura, H., Marui, A., Kamekawa, T., and Nakahara, M., "Influence of directional differences of first reflections in small spaces on perceived clarity," in *Audio Engineering Society Convention 136*, Audio Engineering Society, 2014.

[9] Bradley, J., Reich, R., and Norcross, S., "On the combined effects of early-and late-arriving sound on spatial impression in concert halls," *The Journal of the Acoustical Society of America*, 108(2), pp. 651–661, 2000.

[10] Koelsch, S., *Brain and music*, John Wiley & Sons, 2012.

[11] Hodges, D. and Thaut, M., *The Oxford handbook of music and the brain*, chapter 5, Oxford University Press, 2019.

[12] Zhang, J. D., Susino, M., McPherson, G. E., and Schubert, E., "The definition of a musician in music psychology: A literature review and the six-year rule," *Psychology of Music*, 48(3), pp. 389–409, 2020.

[13] von Tuerckheim, F. and Münch, T., "Automated Sound Optimization of Car Audio Systems Using Binaural Measurements and Parametric IIR Filters," in *Audio Engineering Society Convention 137*, Audio Engineering Society, 2014.

[14] ISO, E., "3382-1: 2009," *Measurement of room acoustic parameters–Part*, 1, 2009.

[15] IRIS, "Hardware Overview," `https://docs.iris.co.nz/v2/hardware/overview/`, 2024, accessed: 2024-06-03.

[16] International Organization for Standardization, "Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces," ISO Standard, 2009, iSO Standard ISO 3382-1 Ed. 1.

[17] International Organization for Standardization, "Acoustics – Measurement of room acoustic parameters – Part 2: Reverberation time in ordinary rooms," ISO Standard, 2008, iSO Standard ISO 3382-2 Ed. 1.

[18] IRIS, "Acoustic Parameters," `https://docs.iris.co.nz/v2/other/acousticparams/`, 2024, accessed: 2024-06-03.

[19] Glinka, M. and Mozart, W. A., "Anechoic Orchestral Music Recording," CD, Compilation, Stereo, 1988, label: Denon – PG-6006. Series: Pure Gold Collection. Country: Japan. Genre: Classical.

[20] Kline, T. J., *Psychological testing: A practical approach to design and evaluation*, Sage publications, 2005.

[21] Hawkes, R. and Douglas, H., "Subjective acoustic experience in concert auditoria," *Acta Acustica united with Acustica*, 24(5), pp. 235–250, 1971.

A Appendix: Questionnaire

A.1 Instruction

First, understand the definitions of six subjective measures and whether each evaluation tends to be objective or subjective. Then, select the most suitable figure based on your judgement. Objective evaluation means assessing sound attributes, typically as high or low, without regard to personal tastes. Conversely, subjective evaluation centers on personal preferences, such as likes or dislikes. The Likert scale, represented as low/high or negative/positive, consists of a seven-point response format.

The measures utilized were explained here; we employed the same definitions as those outlined in Section 2.4 for the six subjective measures. The definitions were consistently provided in Section A.3 for the convenience of the participants.

A.2 Participant Information

(1) Name

(2) Seat type: driver seat or passenger seat

(3) Subject number

Section A.2 was provided once at the onset of the experiment. We gathered more demographic information during the recruitment phase, which included criteria indicative of musical background (e.g., time spent studying music, etc.). See the participant criteria details in Section 2.1.

A.3 Evaluation

(1) How would you rate the clarity of the sound provided? (objective, likert: low/high)

(2) How would you rate the reverberance of the sound provided? (objective, likert: low/high)

(3) How would you rate the envelopment of the sound provided? (objective, likert: low/high)

(4) How would you rate the intimacy of the sound provided? (objective, likert: low/high)

(5) How would you rate the naturality of the sound provided? (objective, likert: low/high)

(6) How would you rate the overall impression of the sound provided? (subjective, likert: negative/positive)