4th International Workshop on Perceptual Quality of Systems (PQS 2013)
2-4 September 2013, Vienna, Austria

Measuring Sound Experience with In-Vehicle Speaker Systems

David Wilfinger[1], *Nicole Perterer*[1], *Arno Laminger*[1], *Alexander Meschtscherjakov*[1], *Manfred Tscheligi*[1]

[1]ICT&S Center, University of Salzburg, Austria
[firstname.lastname]@sbg.ac.at

Abstract

Measuring the perceived quality of speaker systems in the car together with potential users is important to evaluate new speaker developments. Existing sound perception questionnaires suffer from being difficult to answer for human beings because audio perceptions are, for the average user, hard to express in words. In this paper, we present a questionnaire that focuses on user experiences that drivers and passengers have with in-car speaker systems. It uses seven items each representing a dimension of car sound user experience. We describe the development of the questionnaire and a first study, which compares the experiences of two speaker systems. Our questionnaire showed a high reliability and enabled us to compare the level to which the speaker systems evoked positive experiences.

Index Terms: user experience, car, speaker, evaluation, questionnaire

1. Introduction

Sound research in the vehicle has long been established. So far, the focus has mainly been on the sound perception of the car engine (e.g., [9]) and other sources of noise such as the sound that tires create on the road. An important research field is also sounds that are created by technology, such as speaker systems. Our approach to sound perception in the vehicle is based on aspects of human computer interaction (HCI) and user experience (UX) that are evoked when using sound technology. In the car, sound from an HCI perspective has mainly been used to convey information through parking systems, warnings, or future applications such as a system to give spatial information about surrounding of the car. Other research has focused on the perception of voice and their emotional effect [12].

While there are a limited number of published studies of speaker systems in the car available, other contexts have extensively dealt with the evaluation of perceived sound quality. Often used for the subjective assessment of sensory impressions are verbal elicitation methods. Most related methods use quantitative scales that are associated with verbal descriptors [13]. The perceived quality assessment of speaker systems relied heavily on perception to frequencies and the verbal expression of this perception [8]. Also, evaluations of sound systems in the car are often conducted by giving participants the possibility to assess frequency response, distortion, and similar other aspects [2]. When we started a project concerning the perception of car speaker systems, we, therefore, at first wanted to use a similar approach based on expressions and terms by [2].

We, nevertheless, soon realized in our pre-tests that for participants without knowledge in audio systems it would be too hard to rate sound quality based on the proposed scales (e.g., "how satisfied are you with the playback of high frequencies"). We could confirm that the downside of questionnaires like these is that they require a relatively high skillset and/or effort of the assessors [13]. Thus, we decided that we had to take a more humane and experience centered approach with concepts that would make it easier for participants to express their perception of a certain system. We decided to use questions that allowed participants to asses how they felt when listening to the sound rather than trying to rate a certain technical quality of the system. Answering questions like "it feels good to listen to a system" were intended to be easier to understand and thus to respond to than assessing technology centered aspects such as playback of frequencies.

Thus, we decided to apply a questionnaire, which is easy to understand, and which does not focus on aspects of sound in terms of technical attributes but in terms of user experience. Our questionnaire should allow us to compare in-car sound systems and measure which system is perceived better in terms of sound user experience by consumer assessors.

2. Car Sound UX Questionnaire

Since no user experience centered questionnaire aimed at sound perception in vehicles was available, we decided to develop our own set of items. For that purpose, four automotive HMI experts (backgrounds in computer science, design, and psychology) selected a set of user experience dimensions that they encountered as particularly relevant for this context in their prior work and a wide range of studies in the context (see [14] for a overview). In a next step, these dimensions were evaluated by the experts towards their applicability on the perception of sound. These choices were based on the experience of the experts.
In detail, each dimension was discussed with its definitions available in literature. The experts then worked on developing questions for each dimension in regard to sound experience in the vehicle. Dimensions were rejected when the experts could not find an intuitive match between the sound experience and the aspect of UX. Trust, for example, is a highly important dimension of experience in the vehicle. Regarding sound perception it was nevertheless rejected since trust towards machines involves risky behaviors, vulnerability and system integrity [15]. The experts mutually agreed that this does not intuitively relate to car speaker systems. Although these choices were subjective and are still up for discussion they represented a valuable starting point for our investigation and the questionnaire. Even if the list of relevant dimensions would not be complete, it would still provide an impression towards the influence the sound perception had on the UX.

10.21437/PQS.2013-21

In the next step we had to translate each factor into a question using descriptors that our participants would understand. Using too complicated descriptors would have rendered our efforts useless. Thus we used simplified descriptions of each dimension as statement that was rated. The following paragraphs list the selected user experience dimensions in combination with the item used in the questionnaire (translated from German) and a rationale for the selection of the question.

Presence describes the subjective experience of being in one place or environment, even when one is physically situated in another [16]. For sound this means that a system plays sound in a way that makes the listeners think they are at the place where the sound is actually recorded. The corresponding statement in the questionnaire was: *When I hear sound with this system, it feels as if I was where it was recorded (e.g., concert hall).*

Usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [11]. In the case of a sound system the interpretation of usability from our perspective is the ability of the system to play music and speech in a way that is efficient and thus effort free to understand. The opposite would be a system that requires concentration to understand the audio content (e.g., in a train or subway). The statement in the questionnaire thus was the following: *I can understand music and language through the system well.*

The dimension **aesthetics** refers to he degree to which all our senses are gratified [10]. In the case of the questionnaire, we focused on the sense of hearing. To use less complicated words, the term gratification was reduced to "feeling good". Thus the corresponding statement was: *It feels good to listen to the sound through the system.*

Pleasure refers to the deep enjoyment that users experience when interacting with a device, in case of sound systems the pleasure caused by listening to sound through the system [3]. While aesthetics focuses on the sensual perception, pleasure describes the deep inner feeling of enjoyment that eventually results in a flow experience in which a human being is completely absorbed in an activity [4]. Pleasure was rated through the following statement: *It is pleasurable to listen to sound through the system.*

The dimension of **user acceptance** can be divided into perceived usefulness and perceived ease of use, which again influence and constitute the behavioral intention to use the system [5]. We were specifically interested in this intention and thus wanted to know if participants would accept and want to use the system in their own cars if it was available to them. Thus, the related statement aimed at the intention to use: *In my car, I would use the same system to listen to sound.*

Under the dimension **self expression** the ability of the system in representing the values of the participants are subsumed. In the questionnaire we referred to the participants' taste as correspondence of values in the related statement: *The system sounds exactly how it suits my taste.*

Cars are very emotion driven objects that are seen in relation to personal and social **status**. To give participants a way to express feelings towards the expected effect the systems would have on their status we used the concept of "pride" in the following statement: *I would be proud if I had the system in my car.*

Each of the UX dimensions were rated on a 6 point Likert Scale (1:disagree to 6:agree). All descriptors leave indented room for interpretation, since concepts related to user experience are highly subjective. Aesthetics, for example, will mean something different for each listener. The goal of the questionnaire was to give that freedom in interpretation making it easier for participants to relate to each statement by filling it with their own meaning. Regardless all statements are framed in a positive sense, making a higher level of each dimension a better rating.

3. Example Study

To evaluate our questionnaire in a scenario with non-expert users we applied it in a pilot study. In this study we compared two different speaker systems built into two Volkswagen Golf 6 models. System I was a high-class speaker system that is an available upgrade when ordering the car. System II was a custom built solution aiming to reduce the amount of speakers and, thus, technical effort when installing the system in a vehicle. The goal of the comparison was to asses whether the sound experience suffered from this reduction or not.
During an open house event, we asked visitors to sit in both cars and to spend a self-chosen time in the vehicle. Before entering the car, participants were given the questionnaire and it was taken from them after they had left the vehicle. The cars were parked with the engine turned off and the doors closed. We are aware that this does not correspond with the typical usage of car speakers since the noise of the engine, the wind and tires on the road in a real driving setting strongly influence the sound perception. Nevertheless, we saw our experiment as an initial step in evaluating our questionnaire since it enabled us to gather more answers in the same timeframe due to the simplified setup that did not require users to travel in the car.
In both cars, participants listened to the same audio tracks, each of which lasts 30 seconds. The short time was chosen so that the sound quality was evaluated rather than the audio track itself. The thirty-second pieces consisted of classical, folk, rock, hip-hop music in addition to a children's song and a spoken piece of audio (news + traffic information). We chose these sound samples since they represented what we believed to be a good range of music with different sound characteristics as well as spoken language.

3.1. Questionnaire Quality Criteria

In total, 114 visitors (84m, 30f) participated in the study by filling out questionnaires after sitting in both cars. Participants who did not complete both questionnaires were excluded from the analysis. The mean age of the subjects was 37 (SD= 37.1), the youngest person was 9, the oldest 75 years. In a first step of analysis, we conducted a factor analysis (using SPSS) on the seven items to find out whether the dimensions were measuring one or several constructs. As the study had a within setup, meaning that the participants filled in the questionnaire after each of the two conditions, we conducted a principal component analysis (PCA) with orthogonal (varimax) rotation; the decision

criterion was an Eigenwert > 1 on both subsets (condition I and condition II). The Kaiser-Meyer_Olkin measure verified sampling adequacy for the analysis (KMO = .90 in condition) and .92 in condition II), meaning "superb" values [7]. All KMO values for individual items were > .85, which is above the acceptable limit of .5 [7]. Bartlett's test of sphericity χ^2 (21) = 681.30, p = < .001 (condition I) and χ^2 (21) = 929.29, p = < .001 (condition II) indicates that the correlations between the items were sufficiently large for a PCA. The results confirm a single-factor construct, explaining 72.8 % of the total variance in the first condition and 80.9 % in the second, both indicating a satisfying construct. The PCA thus revealed one common underlying factor for all items, i.e. the items measure the same phenomenon.

Furthermore, we were interested in the internal consistency of our questionnaire. In order to assess the reliability, we chose the Cronbach's α coefficient, which is "… a measure that is loosely equivalent to splitting data in two in every possible way and computing the correlation coefficient for each split" in order to overcome the problem of arbitrary split-half procedures [7]. Cronbach's α was computed for the 7 items again separately for each subset, resulting in a value of .94 for condition I, and .96, for condition II. These values refer to a very high reliability of the questionnaire, indicating a good internal consistency. Additionally, inter-item correlations show that the items are selective, i.e. they are not redundant (correlations between .41 and .82 in condition I and between .63 and .93 in condition II).

3.2. Results of the Example Study

After assuring the reliability of the questionnaire, we computed the experience ratings to compare systems I and II. Combining all perceived UX related items, System I was rated lower than System II on average. Regarding the various perceived UX items our results show that the mean of all UX dimensions were rated higher with System II than with System I (see table 1). The relevant items are illustrated in Figure I measured with a six point Likert scale 1 meaning, "strongly disagree" and 6 "strongly agree").

	System 1	System 2
Combined	M = 4.36; SD = 1.43	M = 4.79, SD = 1.42
Presence	M = 4.04, SD = 1.39	M = 4.50, SD = 1.49
Usability	M = 5.01, SD = 1.14	M = 5.04, SD = 1.19
Aesthetics	M = 4.61, SD = 1.30	M = 4.88, SD = 1.38
Pleasure	M = 4.50, SD = 1.27	M = 4.88, SD = 1.37
User Acceptance	M = 4.11, SD = 1.55	M = 4.64, SD = 1.65
Self Expression	M = 4.16, SD = 1.56	M = 4.64, SD = 1.49
Status	M = 4.06, SD = 1.63	M = 4.46, SD = 1.67

Table 1: Means of the relevant user experience items

When computing a paired T-test over all 7 user experience related items, we found that these differences were not statistically significant (t(112) = -1,913 p < .058). To test the single UX aspects, we conducted t-tests which unveiled significant differences in the factors Presence (t(111) = -2.15 p < .034), Usability (t(109) = -2.47 p < .015), and Self Expression (t(111) = -2.14 p < .034), regarding System I and System II. This means that participants indicated that they could understand music and language trough System II better than through System I. Additionally, System II supports the feeling of presence better than System I and sounds more the way that suits the participants' taste.

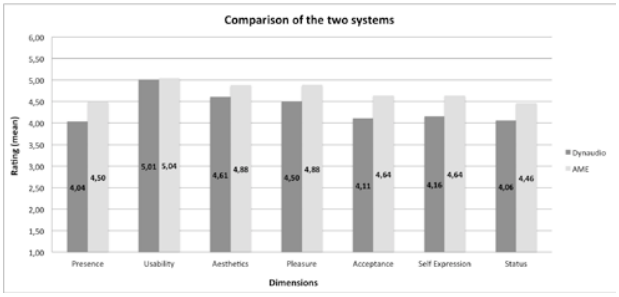

Figure 2: Means of the relevant user experience items

In these cases, System II was rated significantly better than to System I. We did not find cases in which System II, which is the system that uses less speakers and is thus easier to install, was inferior to System I. The questionnaire, therefore, succeeded in showing that the experience with System II did not suffer from the reduction of speakers.

Overall, we can conclude that the questionnaire was partly successful in measuring differences in the sound experience as perceived by our participants. Despite a decent sized sample, we could not find a statistically significant difference in the overall user experience with both systems. We assume that both systems provided a very high sound quality and that therefore the effect of the sound system on differences in perception was limited. Nevertheless we achieved our study goal in showing that System II could achieve a similar (an partly better) sound experience with less technical effort compared to System I.

4. Conclusions

In this paper, we describe the development and application of a simple sound questionnaire for a quick assessment of in-car speaker systems. The goal was to make it applicable at events like fairs, our approach thus aimed at using questions that are easy to understand for participants with no experience in sound assessment. Thus our questionnaire focused on aspects of user experience relevant for sound and the car. For that purpose we introduced seven user experience dimensions and addressed each dimension with a question. Our questionnaire was reliable and allowed us to unveil some differences in the perception of a speaker system.

We are aware that our questionnaire is a very simplified approach that is not able to uncover detailed qualities of a speaker system such as the rating of sound at certain frequencies. Together with a simple evaluation it allows a quick evaluation of user opinion in an easy understandable manner. We nevertheless believe that we can make the questionnaire even easier to understand for users. Thus we plan to iterate the questions using attributes that the users choose themselves [1] based on their own and not our interpretation of each user experience dimension. Additionally we will use the questionnaire in a study with sound systems that are more diverse in quality to assess the usefulness of the questionnaire.

Finally we want to highlight the usefulness of our questionnaire for large events, which are common in the automotive field (e.g., motor shows). Within a short timeframe a good estimate on the users perception can be gathered. And in the end this short term experience with a sound system is what may make users decide for one or the other system.

5. Acknowledgements

The authors thank Verena Fuchsberger for her support in the statistical analysis. The financial support by the Federal Ministry of Economy, Family and Youth, the National Foundation for Research, Technology and Development and AUDIO MOBIL Elektronik GmbH is gratefully acknowledged (Christian Doppler Laboratory for “Contextual Interfaces").

6. References

[1] Altinsoy, M.E. and Jekosch, U. (2012). “The Semantic Space of Vehicle Sounds: Developing a Semantic Differential with Regard to Customer Perception,“ Journal of the Audio Engineering Society, Vol. 60, No. 1/2, pp. 13-20, Jan.-Feb. 2012

[2] Bahne, A. “Perceived Sound Quality of Small Original and Optimized Loudspeaker Systems.” Journal of the Audio Engineering Society 60.1/2. 29-37. 2012

[3] Blythe, M. and Hassenzahl, M. “The semantics of fun: differentiating enjoyable experiences”. Funology, Kluwer Academic Publishers, Norwell, MA, USA 91-100. 2005.

[4] Csikszentmihalyi, M. and Csikszentmihalyi, I. “Flow: The Psychology of Optimal Experience.” Harper and Row, New York. 1990

[5] Davis, F. D. “Perceived usefulness, perceived ease of use, and user acceptance of information technology.” MIS Quarterly 13, 3, 319–340. 1989.

[6] Farina, A. and Ugolotti, E. "Subjective comparison of different car audio systems by the auralization technique." Audio Engineering Society Convention 103. 1997.

[7] Field, A. “Discovering Statistics Using SPSS”. Sage, 2009.

[8] Gabrielsson, A. and Lindström, B. "Perceived Sound Quality of High-Fidelity Loudspeakers". Audio Eng. Sac. 33, 33-53. 1985

[9] Gonzalez, A., Ferrer, M., De Diego, M., Pinero, G., and Garcia-Bonito, J. J. “Sound quality of low-frequency and car engine noises after active noise control”. Journal of Sound and Vibration, 265(3), 663-679. 2003

[10] Hekkert, P. Design aesthetics: Principles of pleasure in product design. Psychology Science 48, 2, 157–172. 2006.

[11] ISO 9241-11. International Organization for Standardization.

[12] Li, H., Seaman, S. and Young, R. "Effect of emotional speech tone on driving from lab to road: fMRI and ERP studies." Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 2010.

[13] Lorho, G. “Perceived quality evaluation: an application to sound reproduction over headphones.” 2010.

[14] Meschtscherjakov, A., Wilfinger, D., Gridling, N., Neureiter, K., and Tscheligi, M. “Capture the car!: qualitative in-situ methods to grasp the automotive context.“ In Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 105-112). ACM. 2012

[15] Wang, Y. D., Emurian, H. H. “An overview of online trust: Concepts, elements, and implications” Computers in Human Behavior, Volume 21, Issue 1, Jan. 2005.

[16] Witmer, B. G. and Singer, M. J. Measuring presence in virtual environments: A presence questionnaire. Presence 7, 3, 225–240. 1998.