Grant Agreement No.: 687645
Research and Innovation action
Call Topic: H2020 ICT-19-2015

Object-based broadcasting – for European leadership in next generation audio experiences

D5.6: Report on Audio subjective tests and User tests

Version: v1.3

Deliverable type	R (Document, report)
Dissemination level	PU (Public)
Due date	30/04/2018
Submission date	06/07/2018
Lead editor	Nicolas Epain (b<>com)
Authors	W. Bleisteiner (BR), A. Silzle (FHG), R. Schmidt (FHG), T. Liebl (IRT), O. Warusfel (IRCAM), M. Ragot & N. Epain (b<>com)
Reviewers	N. Bogaards (elephantcandy)
Work package, Task	WP5, T5.1, T5.4
Keywords	User experience; Tests;

Abstract

Deliverable D5.6 presents the results of subjective audio quality tests and user acceptance tests performed in Tasks 5.1 and T5.4. Two different user experience tests have been conducted, which both examined how the ORPHEUS app and, more generally, object-based radio is received by naïve users. In addition, two more studies regarding the perceived quality of object-based audio content have been carried out. The first one focused on a new protocol for perceptual audio experiments involving the comparison of multiple stimuli without reference. The second one investigated the perceived quality resulting from using different schemes for object-based audio reverberation transport and rendering. This deliverable describes the aims, methodologies and results of these studies.

Document revision history

Version	Date	Description of change	List of contributor(s)
v0.1	26/03/18	Created skeleton	N. Epain
v0.2	05/04/18	Structure and draft of chapter 2	W. Bleisteiner
v0.25	10/04/18	Contribution from BR in chapter 2	W. Bleisteiner
v0.3	19/04/18	Added summary of JOSEPHS results	A. Silzle
v0.4	20/04/18	Added IRT contribution	T. Liebl
v0.5	20/04/18	Added b<>com's contribution	N. Epain, M. Ragot
v0.6	27/04/18	Consolidated chapter 2	A. Silzle, W. Bleisteiner, R. Schmidt
v0.7	17/05/18	Added IRCAM's contribution	O. Warusfel
v1.0	20/05/18	Reviewed and finalized report	N. Bogaards, N. Epain
v1.1	07/16/18	Added IRCAM test results	O. Warusfel
v1.2	13/06/18	Improved IRT section	M. Weitnauer
V1.3	06/07/18	Final editing	A. Köhler, U. Herzog

[1] http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en_US

Executive Summary

In this Deliverable, we present four studies that examined different aspects of audio perception and quality of user experience in the context of object-based audio broadcasting.

The first two studies investigated how non-expert users react to the ORPHEUS iPhone app and, more generally, to the new features available in object-based audio broadcasting when compared to standard radio. The first test took place in b<>com's facilities, while the second one took place at JOSEPHS®, a venue designed for testing new products and services located in Nuremberg. Interestingly, although the methodologies were different, relatively similar conclusions can be drawn from these two quality of user experience tests.

First and foremost, the global response from the test participants was for the most part positive. Second, the feature that was the most appreciated by the participants was to be able to set the foreground/background balance. Third, the possibility to listen to radio programmes with a binaural rendering was more appreciated by the participants than expected by the project partners. Overall, these results indicate that the general public is ready for a next-generation radio and that the work achieved in the ORPHEUS project are clearly well appreciated by the users.

The third study presented in this report focused on the Multiple Stimulus Ideal Profile methodology (MS-IPM), which is a listening test method. One of the main advantages of this technique is that it does not rely on comparing sound stimuli to a reference, contrary to the standardised Multiple Stimuli with Hidden Reference and Anchor paradigm (MUSHRA). The study presented here confirmed that MS-IPM is well suited for the object-based audio context, where there is often no reference for comparison.

Lastly, the fourth study investigates how to best represent and transport reverberation in an object-based audio context. Transmitting the reverberation as separate objects allows interesting user interactions but can be costly in terms of bitrate. This study aimed to determine the method and representation providing the best compromise between audio quality and data rate. As a prerequisite, several decorrelation methods were compared.

Table of Contents

List of Figures

List of Tables

Abbreviations

DVB	Digital Video Broadcast
EC	European Commission
HOA	Higher Order Ambisonics
UX	User Experience
QoX	Quality of Experience
MS-IPM	Multiple-Stimulus Ideal Profile Methodology

1 ORPHEUS Quality of Experience Tests at b<>com

1.1 Motivation

Studies that deal with product adoption can be divided into two fields of research. The first field of research focuses on acceptability and acceptance. Although the two terms are often used synonymously, *acceptability* refers to prospective judgments on technologies or products before use [1], whereas *acceptance* refers to judgments on and behavioural reactions to products after use [2]. The two prevailing factors in this field are the *perceived usefulness* and the *perceived ease of use.* The first factor (*perceived usefulness*) is defined as the “degree to which an individual believes that using the system will help him or her to attain gains in job performance [3]. The second factor (*perceived ease of use*) is defined as the “degree of ease associated with the use of the system” [3]. Its role as a predictor of behavioural intention has been demonstrated in several meta-analyses [4][5][6][7][8]. The second field of research deals with user experience (UX) and focuses on the factors that affect the user’s judgment on products. Perceived innovation, for example, can be related to the s*timulation* factor, which is attached to the need for challenge and novelty, and would appear to be a prerequisite for personal development [9].

Within the Orpheus project, it seems necessary to understand the judgment of future users before and after use of the currently developed app. For this purpose, different user tests were conducted, with multiple objectives in mind. The primary motivation is to gather the opinion of users, in this case representatives of the general public, about the smartphone application developed within the Orpheus project. To this aim, the different features developed during the project are presented to non-expert users in order to measure several variables: *appetite for new technology*, *perceived usefulness*, *perceived ease of use*, *perceived innovation*, *intention to use* (or *behavioural intention)*, etc. These measures may then be used to confirm or invalidate certain design choices.

Simultaneously, the objective of these tests is also to compare the user’s perceptions before use (i.e., acceptability) and after use (i.e., acceptance). To the best of our knowledge, user assessments before use have already been modelled. Paradoxically, most studies of judgments before use have only collected data after the use of the product [10][11][12]. Based on several theoretical models (expectation and confirmation models, acceptability / acceptance continuum, etc.), these user tests aim to provide a better understanding of the adoption criteria of new technology in the audio field, by apprehending the evolution of user perception and comparing the acceptability before use and acceptance after use of the Orpheus app.

1.2 Methodology

The test involved a total of 21 French participants and took place during the months of December 2017 and January 2018, in b<>com’s "Usage and Acceptability" laboratory. The experimental protocol consisted of three main steps, as illustrated in *Figure 1* below:

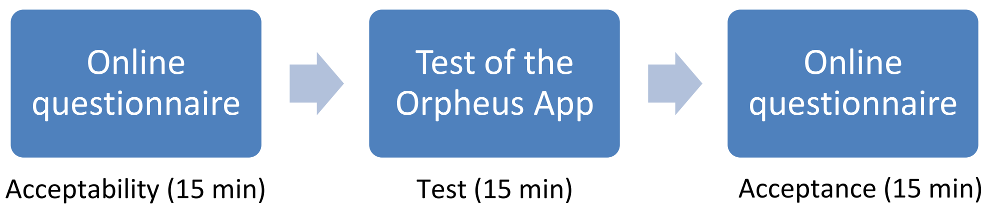


Figure 1: Experimental protocol of b<>com’s QoX study

First, an initial online questionnaire was taken by the test subjects. The objectives of this test were: a) to gather information about the users' habits in terms of radio, their appetence for new technologies, etc; and b) to evaluate their expectations with regard to the features of the ORPHEUS iOS app (acceptability). Second, the test subjects were asked to test the ORPHEUS app through specific tasks. This test was performed several days after replying to the initial questionnaire in order to reduce the fatigue of the participants and reinforce the consistency of their answers. Third, immediately after testing the ORPHEUS app, the test participants were asked to take a second questionnaire. The aim of this questionnaire was to rate the different features of the app after use (acceptance).

Note that, out of the 21 participants who took the initial questionnaire, only 16 finished the test and replied to the final questionnaire. Therefore, the results shown in the following concern these 16 subjects only. Screenshots of both the acceptability and acceptance questionnaires are provided in their original format (in French) in Appendix A at the end of this report.

1.2.1 Initial questionnaire

The initial questionnaire (acceptance questionnaire) started with a brief description of the ORPHEUS iOS app, including the following features:

1. Navigate between chapters within a programme
2. Display a text transcript of the audio programme
3. Set the "audio clarity" (foreground/background balance)
4. Interact with the audio content (change the listening perspective on the scene or move sound sources)
5. Choose the programme language
6. Set the audio rendering format (mono/stereo/binaural)
7. Adapt the length of the content

The questionnaire then consisted of questions regarding the participant's profile (interest for new technologies, etc) and the acceptability of the different app features.

1.2.2 Test of the ORPHEUS app

In order to test the ORPHEUS app, the participants were instructed to perform 6 tasks with each task focusing on a particular feature.

Chapter navigation – To test this feature the participants were asked to select the programme named "Experience Object-Based Audio" and jump from chapter to chapter.

Text transcript – The participants were instructed to select the "Art of Foley" programme and activate the text transcript feature.

Audio clarity – The participants were asked to select the "Live Football" programme and listen to the effect of changing the foreground/background balance.

Interaction – The participants were instructed to select the programme named "Passo Avanti: Mozart Gigue in 360°" and try the different versions of this content**.**

Multi-language – The participants were instructed to select the "Art of Foley" programme and switch language to English.

Audio presets and rendering – The participants were instructed to configure a profile that used binaural rendering.

Note that the original test instructions are provided in Appendix A at the end of this report.

1.2.3 Final questionnaire

Following the application test, the participants were asked questions to gather their opinions on the application and its specific features. The first eight questions were aimed at assessing the user's overall acceptance of the ORPHEUS app (How useful is this app? Is it easy to use? Etc.). The remaining questions addressed the acceptance of the specific features demonstrated through the tasks. For each of these features, the participants had to rate on a 0-10 scale how much they agreed with the following statements:

- "This feature is useful for me" (Perceived usefulness)
- "Using this feature is easy" (Perceived ease of use)
- “This feature is innovative" (Perceived stimulation)
- "I intend to continue using this feature" (Intention to use)

1.3 Results

1.3.1 Test participants

16 participants (11 men, 5 women) filled out both questionnaires, before use and after use. The average age of the participants was 28.6 years (SD = 5. 9). All declared they owned a smartphone: 6 being on iOS and 10 on Android system.

Most of the users reported that they listened to music with their smartphone, as shown in Figure 2 below.

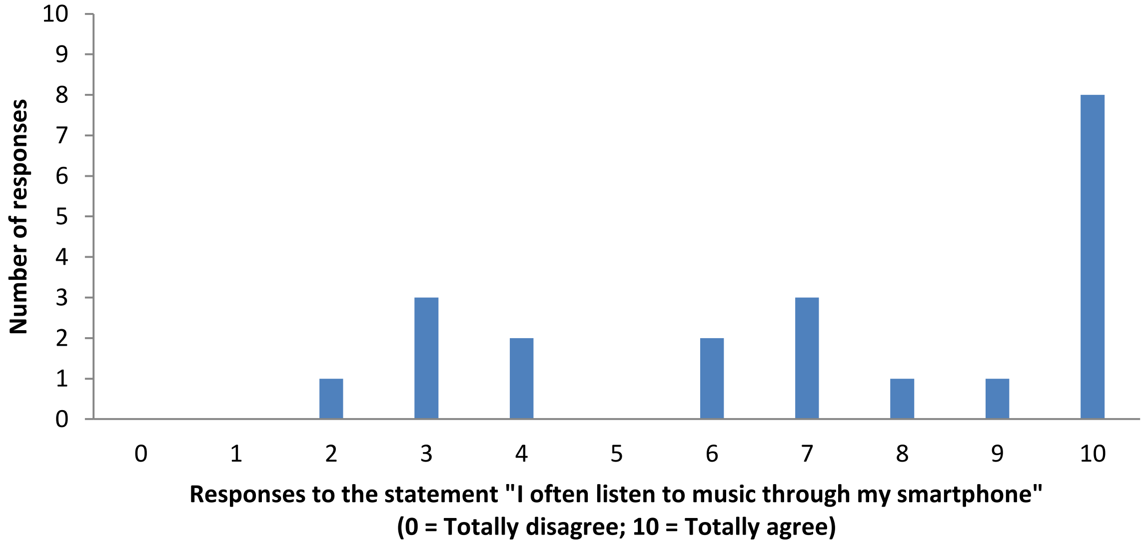


Figure 2: Answers to the statement: "I often listen to music through my smartphone”

However, most of the participants reported that they rarely listened to the radio using their smartphone. This is illustrated in *Figure 3* below.

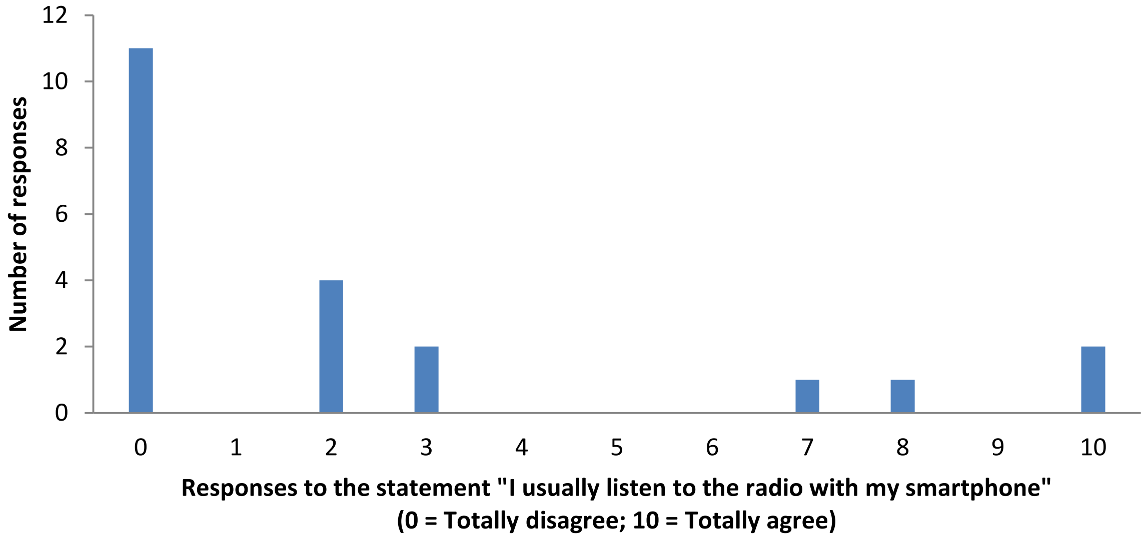


Figure 3: Answers to the statement: "I usually listen to the radio with my smartphone"

In addition, a majority of participants considered themselves as technology enthusiasts, as illustrated in *Figure 4*.

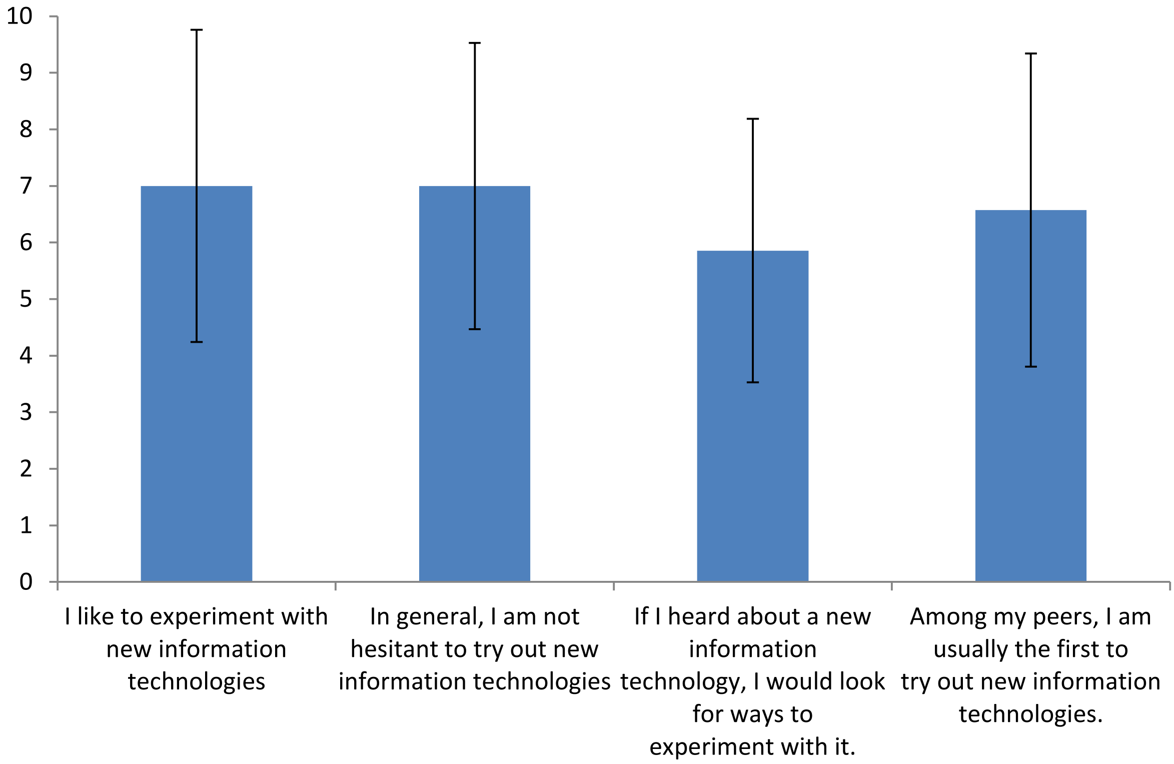


Figure 4: Answers to the "appetence for technology" questions

In summary, the participants were relatively young, frequent smartphone users and in general enthusiastic regarding new technologies. Relatively few of them were familiar with the idea of listening to the radio using their smartphone.

1.3.2 Overall acceptance/acceptability

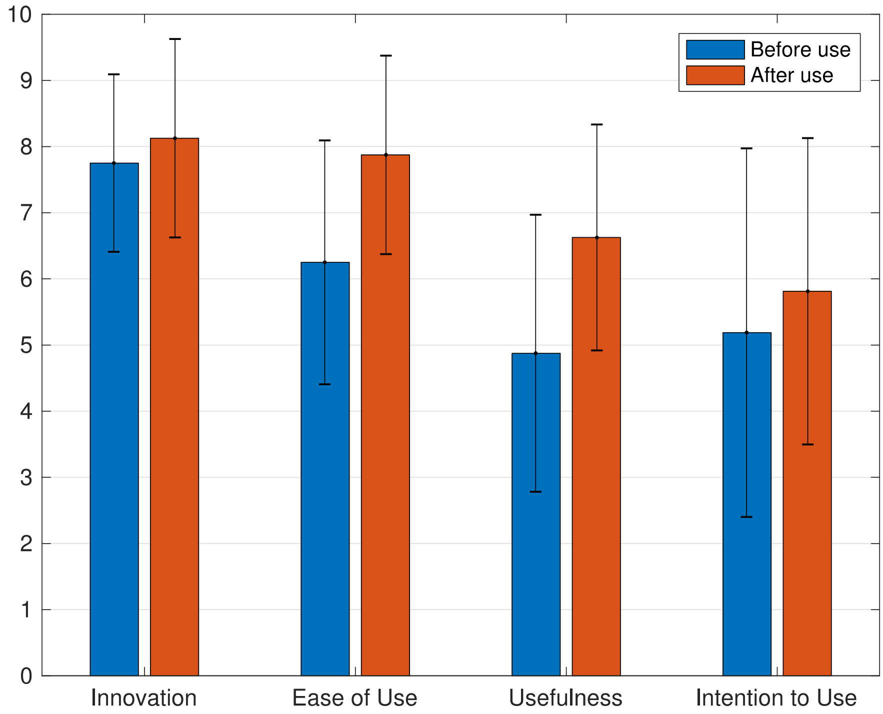


Figure 5: Overall acceptance and acceptability of the ORPHEUS app

Figure 5 presents the overall user ratings for the ORPHEUS app before and after use. Overall, the users got a positive impression of the app and the ratings increased after trying it. The increase in ratings is particularly large for the "Ease of Use" and "Usefulness" criteria, which indicates that the participants did not expect the app features to be as easy to use and useful as they found during the test. By contrast, the increase in "Intention to Use" is moderate, which is probably related to the fact that a majority of the participants do not use their smartphone to listen to radio programmes. Note that the increase in the "Innovation" rating is also small but the initial rating was already very close to the maximum score.

1.3.3 Feature acceptance/acceptability

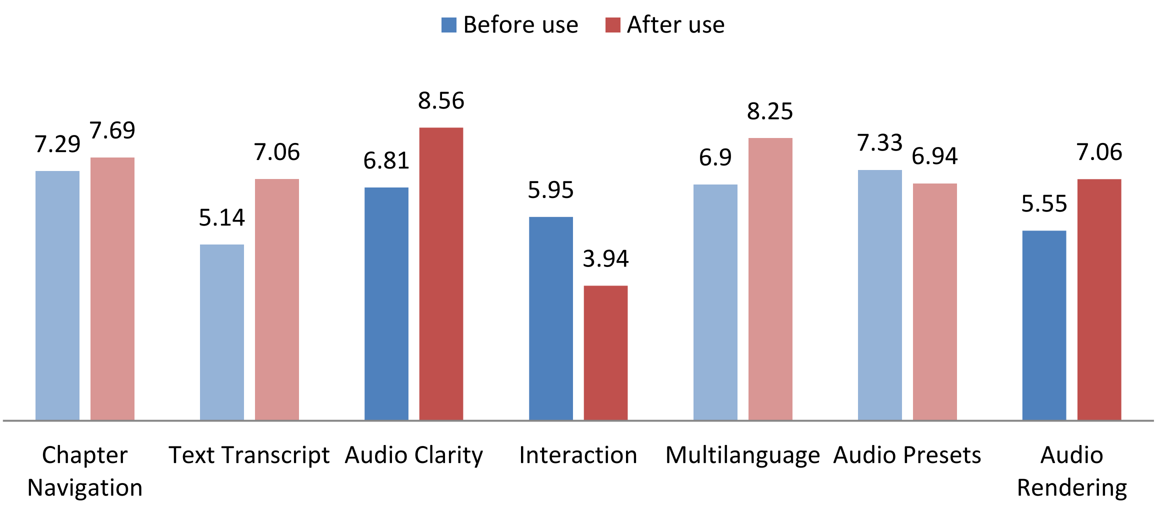


Figure 6: Perceived feature usefulness between before and after use. Note: the bars with more saturated colours indicate that the change in rating was found to be statistically significant.

Figure 6 presents the perceived usefulness before and after using the ORPHEUS app. The majority of the features were perceived as more useful after trying the app than before. The increase in the perceived usefulness is particularly significant for the “Audio Clarity” (foreground/background balance) and “Audio Rendering” (switching from stereo to binaural) features. The two features that were found to be the most useful after use were the “Audio Clarity” and “Multilanguage” features.

Lastly note that the usefulness ratings for the “Interaction” feature (ability to change the perspective in the audio scene or move sound sources) decreased significantly after use. Participants reported that they could not hear the difference between the different versions of the programme, which could have been caused by a problem in the app or in the content itself.

1.4 Conclusions

This study, conducted with 21 users, aimed to compare the perception of users before (acceptability) and after (acceptance) use. This comparison was done both on a global level (global user assessment of the app) and on a feature-specific level (user assessment of specific features within the app). The objective was to confirm or refute certain design choices, in order to better understand the strengths and weaknesses of the application, as well as the elements that can be improved. The different remarks collected from users also provide valuable contributions to improve the ergonomics of the Orpheus mobile app.

The overall acceptability of the application was relatively high, which shows that the application is perceived positively by the users. The innovative nature of the application is rated the highest by participants, followed by the ease of use. However, the perceived usefulness and intention to use, before use, are relatively low.

Overall, the users found the app easier to use and more useful after they had a chance to try it. This is a very positive result for the ORPHEUS project, which demonstrates that the app convinced the users of the advantages of Object-Based Audio radio programmes. However, the fact that ratings increased significantly after use indicate that it is important for users to experience these advantages themselves. Moreover, the participants’ intention to use the app did not increase very much after use, which could be related to the fact that most of them rarely listen to radio programmes using their smartphone.

In terms of features, the “Audio Clarity” (foreground/background balance) was the most popular among users, followed by the “Multilingual” feature. Conversely, the “Interaction” feature (spatial changes in the audio scene) was perceived as the least useful but this seems to have been caused by a technical problem in the app or in the corresponding audio programme during the test.

2 ORPHEUS Experience Tests at JOSEPHS®

2.1 ORPHEUS' 3-Dimensional Methodology for Object-based Audio

In deliverable D5.3: Document on Methodology for Evaluation of User Experience we have collected, explained and developed in detail the various fields of interest, expertise and methodologies applied by the ORPHEUS partners that committed to "quality of user experience" investigations.

Fairly early in the course of our project we realised that in object-based broadcasting, with audio becoming 'interactive', new challenges arise to make object-based media features accessible, understandable and operable. Besides the new properties of the audio, such as immersive binaural or surround sound, the additional services and features - transcripts, additional text-based information or still pictures -, implemented by a 'presentation design', and the operability of these functionalities become integral components in the assessment of the overall media experience.

Therefore, a new challenge lies in the development of appropriate user interfaces that make human interaction to control and adjust complex technical metadata and parameters delivered alongside the audio on the different devices both appropriate and convenient. Only if this can be achieved, the user will consider object-based media technology able to provide an exciting and satisfying experience. As a consequence, domains for examining and evaluating 'quality of end user experience' that were previously evaluated separately will now have to be considered convergent and inclusive.

In D5.3, we have developed a basic approach for practical examination and evaluation of user experience within an object-based media eco-system as a holistic model, based upon the ORPHEUS project's main pillars: the user requests and use cases, the pilot architecture and the pilots themselves. This model consists of three main experience dimensions:

- Audio experience
- Usability experience
- Information experience

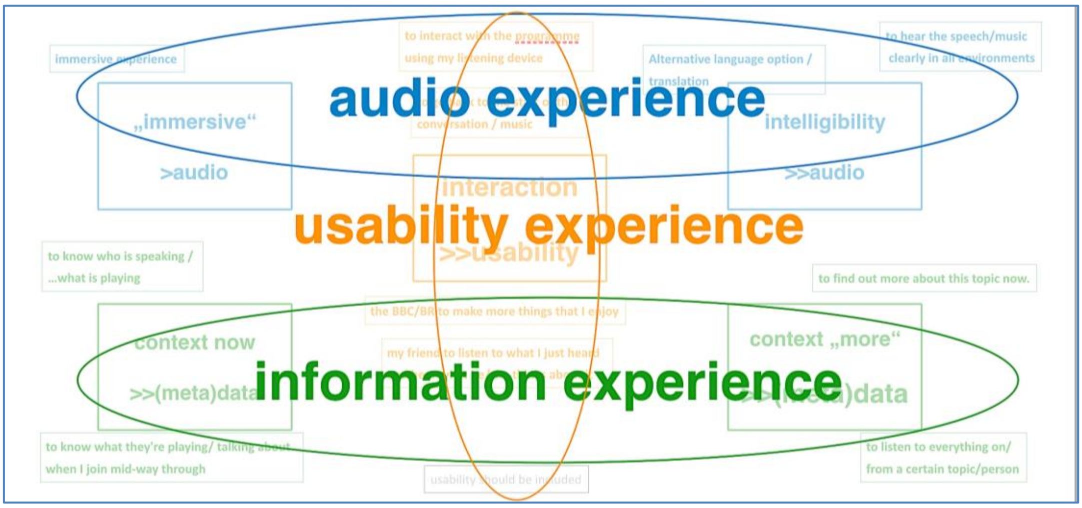


Figure 7: Experience dimensions in an object-based audio context (from D5.3)

In order to test the real-world applicability of our approach, we had to find a suitable test setting, offering low barrier access to general users in the public. The JOSEPHS® in Nuremberg and its applied concept of design-thinking and co-creation appeared to be the ideal venue for this purpose.

2.2 About JOSEPHS®: the Concept of Design-thinking and Co-creation

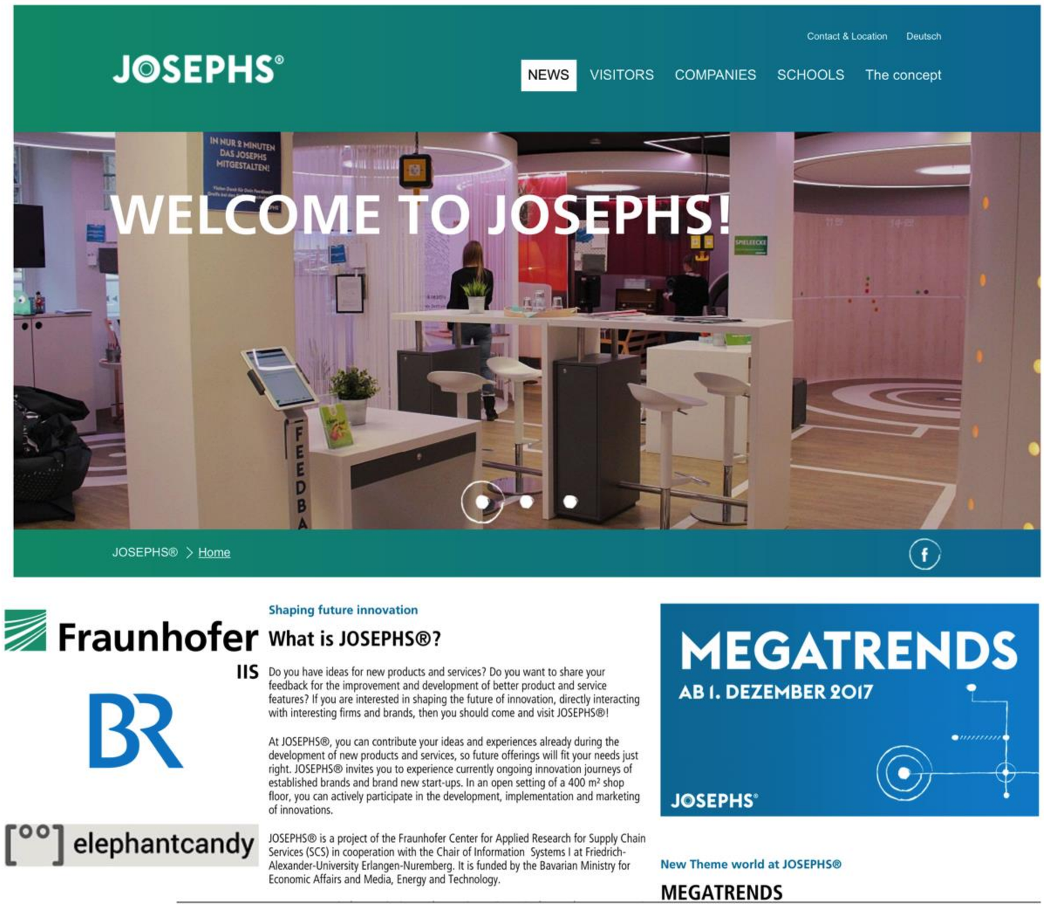


Figure 8: JOSEPHS website for Megatrends setting

The concept of design-thinking of JOSEPHS is visualized in Figure 8. It is not about representative data samples but about generating ideas. It is a physical meeting place in the centre of Nuremberg open for everybody who is interested. The design is targeted towards visitors that are interested in innovation. At JOSEPHS new ideas are created from casual conversations, without time limits, between visitor and JOSEPHS-GUIDE, stimulated by your research exhibit.

2.3 The ORPHEUS concept and setting at JOSEPHS®

The various tools and methodologies applied in JOSEPHS® offer several possibilities for the different stages of product or services development in design-thinking or co-creation processes.

In a preparatory conceptualisation workshop with the experts at JOSEPHS® we have shaped the setting for our installation to be based upon 'user stories'. There, we were able to take on our previous D5.1: *Document on user requirements*, where we had already developed such user stories from proposed use cases and created first mock-ups of the app alongside.

Hence it seemed obvious to collect feedback on the (by now) developed solutions from users, getting them into a simulation of the environment and the initially imagined usage situation.

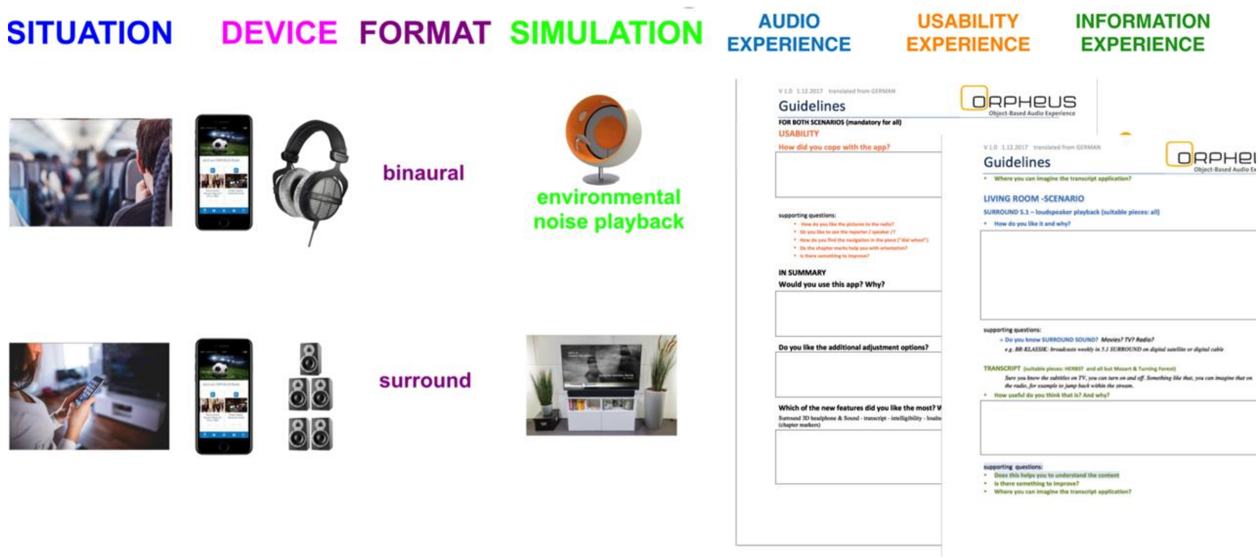


Figure 9: ORPHEUS investigation concept for JOSEPHS

We decided to focus on two usage scenarios (Figure 9) and aimed to get basic responses on our three defined dimensions – audio, information and usability experience.

- On-the go: ORPHEUS iOS app with headphones connected, offering as optimum binaural reproduction (compared to standard stereo and mono) and the foreground/background balance feature. Simulation of being on an airplane by playing typical cabin flight noise from outside via a Sonic Chair.
- At Home: iOS App with a connected AV-Receiver and a 5.1 loudspeaker setup, offering as optimum 5.1 surround sound (compared to standard stereo and monaural)

As playback device for both situations, the ORPHEUS iOS app was installed in a 'presentation mode' on 12-inch iPad Pro devices, with a German language user interface and additional instructions and tips displayed alongside the app interface (see Figure 10).

For technical reasons, it was not possible to install a complete 3D loudspeaker system (including elevated loudspeakers) for reproduction of immersive sound in the living room setup, but only a 5.1 configuration as shown in Figure 11. Studio loudspeakers were used in this scenario. The tube radio seen in the picture was for decoration only.

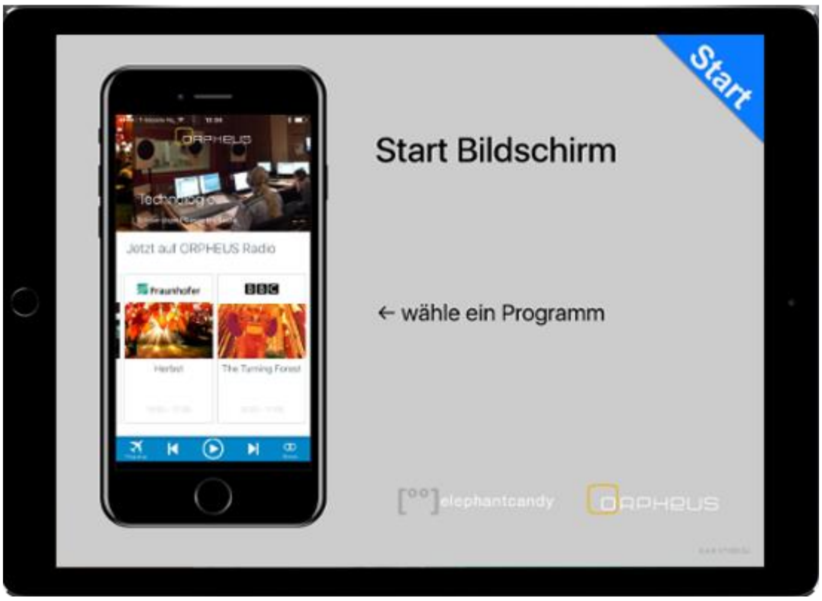


Figure 10: ORPHEUS iOS app in iPad presentation mode

As the ORPHEUS trial was in competition with four other demonstrations presented at JOSEPHS®, it was important to attract instant attention and create motivation for potential users. The ORPHEUS 'island' was located in the centre of the JOSEPHS® experience fields, occupying almost 20 m^2, with an eye-catching, inviting and cosy installation for both usage scenarios.

Figure 11: ORPHEUS installation at JOSEPHS

As content for the tests, the pieces from our ORPHEUS pilot productions were chosen, representing different types and genres in order to match the different taste and preferences of the visitors.

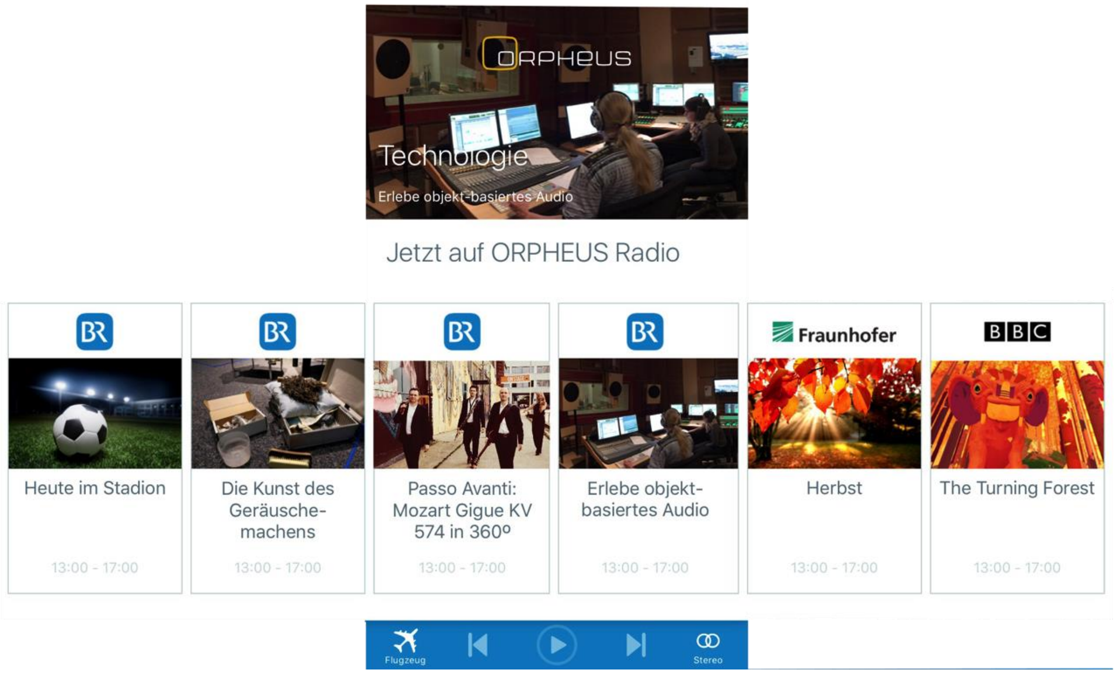


Figure 12: ORPHEUS demo content in JOSEPHS®

1. *Heute im Stadion*: a typical (live) radio report from a football match
2. *Die Kunst des Geräuschemachens* (*The art of Foley*): radio documentary
3. *Passo Avanti – Mozart Gigue*: short jazz adaption of a classical music piece
4. *Erlebe objekt-basiertes Audio (Experience object-based audio)*: informative feature
5. *Herbst (Autumn)*: *Poem with 3D* soundscape
6. *The Turning Forest*: cinematic radio drama

All pieces offered key-features of the developments in the Orpheus project:

- Audio experience[2]
  - Audio reproduction: mono/stereo/binaural/5.1 surround (1,2,3,4,5,6)
  - Foreground/background level adjustment (1,2,4,5,6)
  - Clarity/dynamic range control (1,2,3,4,5,6)
  - Multi-language (2: DE/EN/FR, 4: DE/EN, 5: DE/EN, 6: DE/EN)
  - Audio interaction (3: positioning of the instruments, 5: interactive 3D audio object)
- Information experience
  - Text Transcript (1,2,4,5)
  - Additional program info (1,2,3,4,5,6)
- Usability experience
  - dial to navigate within the currently playing content
  - points of interest /chapter markers
  - additional pictures

Entering the Orpheus area at JOSEPHS®, visitors received a short introduction to our specific subject ("audio experience of the future") and, if they were interested in participating, they were asked some basic questions about their listening habits and preferences. For this purpose, the visitors were also requested to "play" a memory game with icons of existing radio stations, podcasts and streaming apps and state which of them they know or use. This introductory task served to find out more about the subjects' familiarity with present day smartphone applications of broadcasters and other audio services providers, in order to categorise them more clearly to different target groups of users in the subsequent evaluation process.

After that, the participants were offered to go into one of the usage scenarios – "on-the-go" or "living room" – and explore the ORPHEUS iOS app and the features of object-based audio in greater detail.

The complete guidelines (questionnaire) used by the guides accompanying the participants are attached to this document in Appendix B.

2.4 The Findings from JOSEPHS®

In the three months that the experiment ran (Dec. 1 2017 - Feb 28th 2018) a total of 2766 visitors came to JOSEPHS. 1048 of them were co-creators, meaning they participated in 'islands' (installed test fields from companies). For the ORPHEUS island response data was collected from 361 persons.

The participants used either the airplane or the living room scenario. They were assigned to one of two groups: one below 40 years old and one above. A control group was defined by the ones doing both scenarios, without age differentiation, see *Table 1*.

[2] the numbers in brackets refer to the numeration of the pieces above.

	Participant	Age < 40a	Percentage
Airplane scenario	82	67	82%
Living room scenario	115	61	53%
Both scenarios	158	-	-

Table 1: Number of participants per scenario and age classification

The age distribution can be seen in *Figure 13*. The average age of the participants is 35 years. 40% of them were female and 60% male.

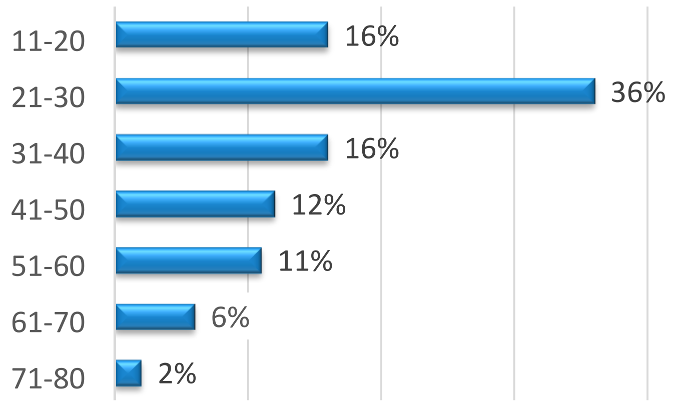


Figure 13: Age distribution of the 361 participants

The type of radio listener is presented in *Table 2*. Most of them listen radio “on the go” and “as background listener”.

“Where do you listen to the radio?”		n=243
At home	30%	
On the go	70%	
“What kind of listener are you?”		n=271
Attentive listener (HiFi-Fan)	28%	
Background listening	72%	

Table 2: Type of radio listener

As an initial task the participants should select memory cards with the logos of radio programs and streaming apps. The answers are visualized in Figure 14. 38% use apps from public service broadcasters, 54% use streaming services, 43% of them use Spotify.

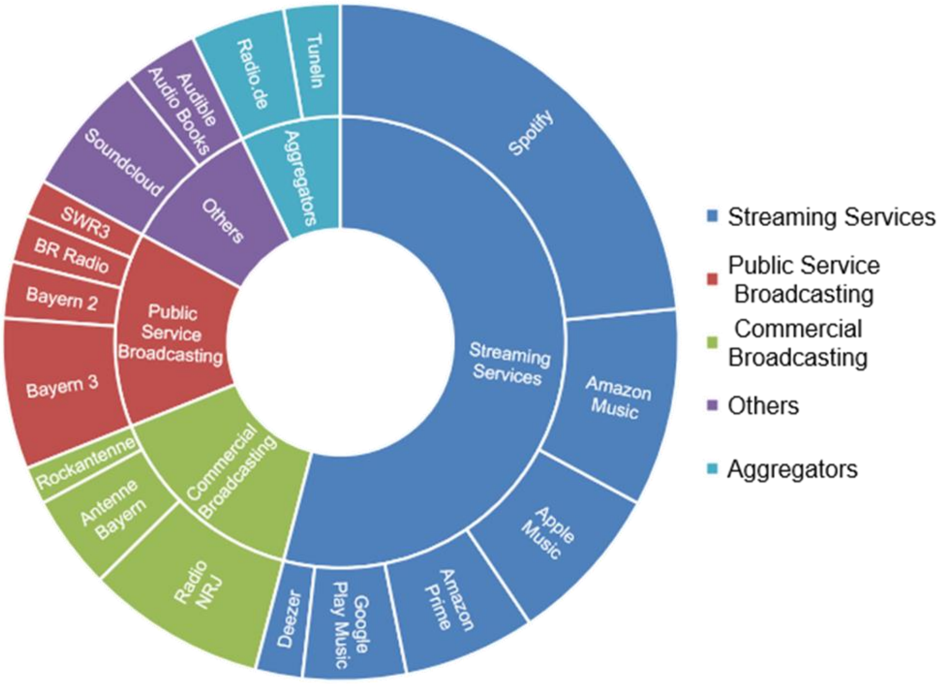


Figure 14: Which radio or streaming apps are you using, n=242 (multiple answers are allowed)

2.4.1 Audio Experience in the Airplane Scenario

In noisy environments, such as inside an airplane, spoken word is sometimes difficult to understand and quiet parts in music become less audible. The Sonic Chair (Figure 15) was used to emit a switchable airplane noise to simulate a realistic listening scenario. The audio reproduction was done over open headphones (Beyerdynamic DT990) to get full impact of the reproduced airplane noise.

Participants were asked which reproduction format they liked best. A vast majority preferred binaural reproduction compared to mono and stereo as shown in Table 3 and Figure 16.

Figure 15: Airplane scenario

	Mono	Stereo	Binaural
Which of the three sound three options do you like most?	4%	13%	83%

Table 3: Results for binaural reproduction in airplane scenario, n = 143

Figure 16 presents the co-creators perception of the binaural reproduction.

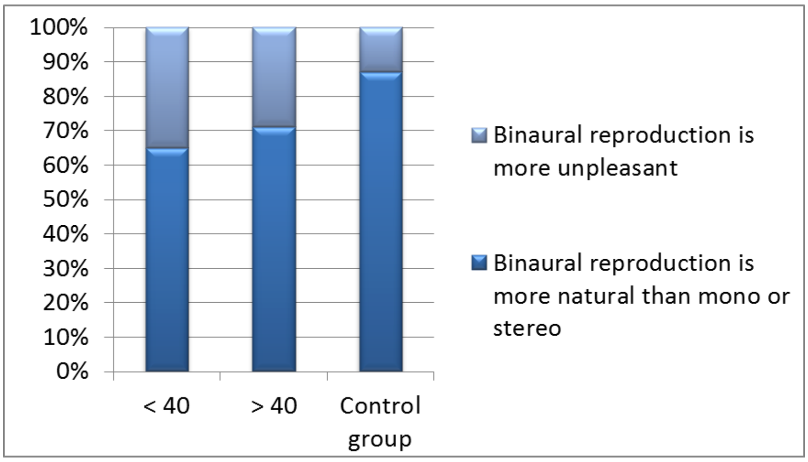


Figure 16: Results for binaural reproduction in airplane scenario per age

All questions and answers were induced in conversations between the participants and the JOSEPHS guide. Some of the retrieved comments are presented below (more can be found in the report in Appendix B).

Some of the positive comments:

„I‘m blown away by the binaural sound, really knocked out.“ (f, 23)
„The sound is especially impressive for football broadcasts.“ (m, 48)
„Once you‘ve heard binaural, you don‘t want to hear anything else.“ (m, 37)

Some of the negative comments:

„Instead of binaural, I would prefer to hear louder ambient sounds.“ (f, 26)
„The binaural sound gives me a headache.“ (m, 54)
„The sound is unpleasant. I feel like I‘m in a helmet.“ (m, 38)

Summary and interpretation: The binaural reproduction was vastly preferred. This result is more positive than from many other formal binaural listening tests. Still, a small percentage of users have different sound reproduction expectations. This is not a problem, because these listeners can just select stereo or mono reproduction instead of binaural reproduction.

With a slider in the app the level balance between the fore- and the background (f/b) part of the ORPHEUS content could be adjusted. This is a specific feature of object-based audio and is implemented in MPEG-H. The results are presented in Figure 17.

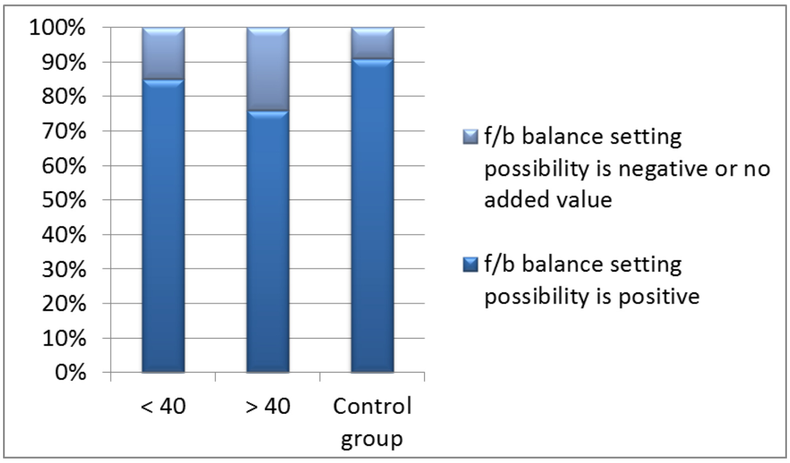


Figure 17: Results for fore-/background balance in airplane scenario

Some of the positive comments:

„The individual sound experience is most important to me." (f, 20)
„Finally you can hear the football supporter songs without commentary." (m, 48)
„Background noise no longer impairs listening pleasure." (m, 38)

Some of the negative comments:

„In everyday life I would not use the settings." (f, 23)
„The settings are not yet perfectly worked out." (m, 53)
„ I can hardly notice any differences to other sound experiences." (m, 66)

To make this foreground/background balance even easier to adjust a simple switch was offered in the user interface. It was labelled "Clarity, make quiet parts louder". Results are shown in Figure 18.

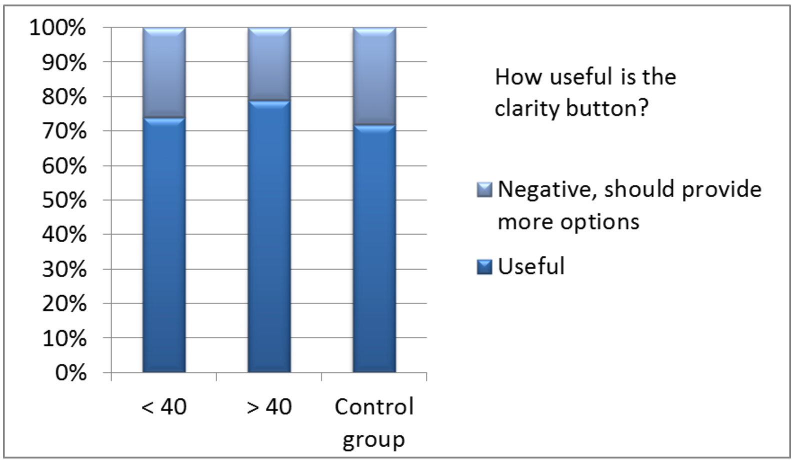


Figure 18: Results for clarity button in airplane scenario

Some of the positive comments:

„With this setting I can understand what I usually would barely hear." (m, 28)
„This setting helps you to understand the speaker better. I think that's great." (m, 48)
„I think this setting is very good because you can get more of the atmosphere." (m, 54)

Some of the negative comments:

„I'd rather have just one setting that is perfect for the purpose." (m, 23)
„I don't want to hear every slight noise." (m, 45)
„I want to hear everything as it should be. I don't need this setting." (m, 28)

Summary and interpretation: The new foreground/background feature was highly accepted. The adjustable slider was preferred to a simple switch. As an additional feature, the original intended balance setting defined by the sound engineer, tonmeister or producer could be used as default and/or marked on the scale.

2.4.2 Living Room Scenario

2.4.3 Audio Experience in the Living Room Scenario

The first question was: How do you like the 5.1 surround reproduction and why. Results are presented in Figure 19.

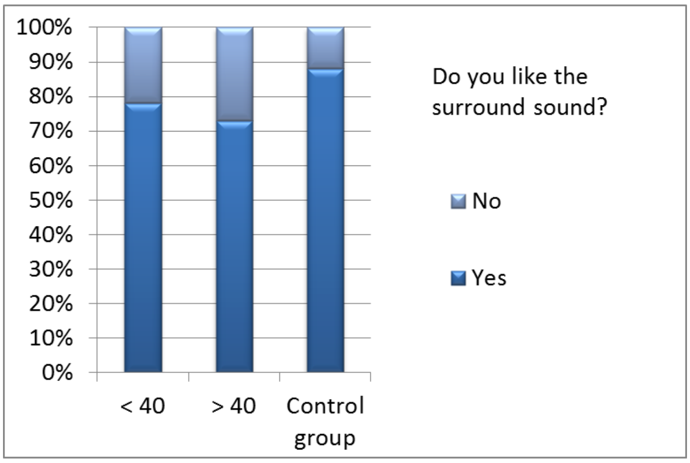


Figure 19: Results for surround sound in living room scenario

Some of the positive comments:

„Very cool. You can really feel it.“ (f, 22)
„I could imagine that in a car.“ (f, 54)
„I find the feeling of the music around myself very pleasing.“ (f, 61)

Some of the negative comments:

„I miss the bass.“ (m, 23)
„The good old stereo sound is sufficient for me.“ (m, 45)
„I don't like the sound because I prefer listening to music through headphones.“ (m, 24)

Summary and interpretation: A large majority of listeners preferred 5.1 surround sound. From the comments given it seems that many people had never really experienced surround sound reproduction in this context, even though it has been on the market for more than 20 years. Other tests have shown even the advantage of 3D reproduction over 2D reproduction[3], which is possible with the new NGA codecs.

2.4.4 Information Experience

The "Live text" transcript presents the spoken language as readable text. Backward and forward navigation can be facilitated by scrolling up and down the text feed, (or with transport wheel), as shown in Figure 20.

[3] Silzle, A., S. George, and T. Bachmann. Experimental Setups for 3D Audio Listening Tests. International Conference on Spatial Audio (ICSA). 2011. Detmold, Germany.

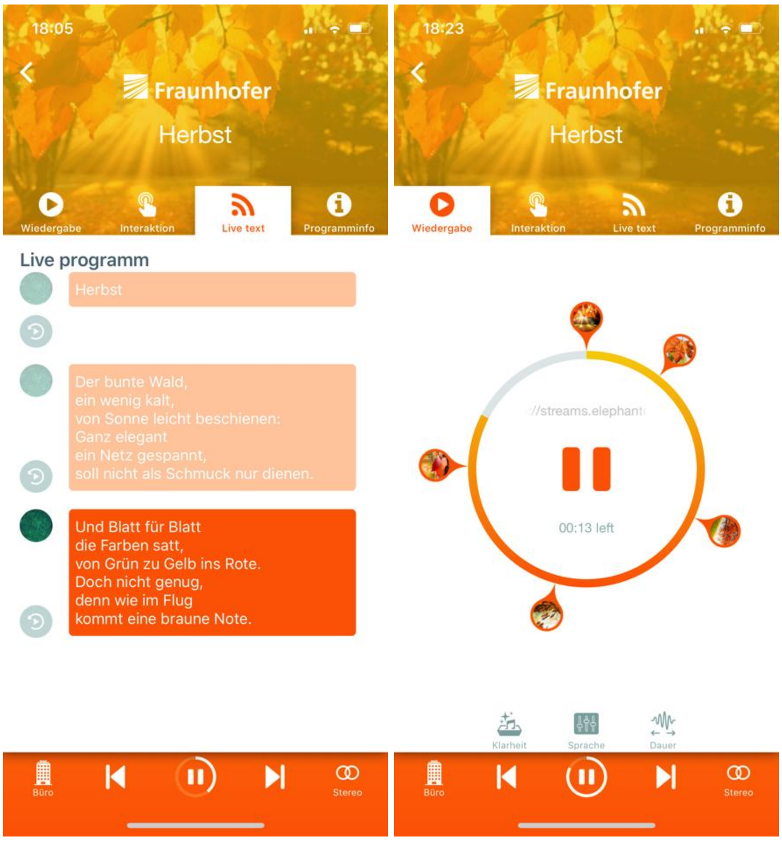


Figure 20: Transport wheel and transcript of the spoken word

The participants were asked how useful text transcript is and why. Results are shown in Figure 21.

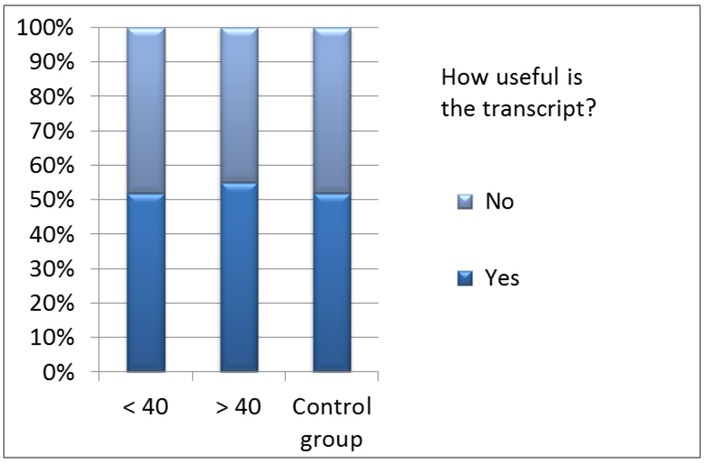


Figure 21: Results for information experience *regarding transcript*

Some of the positive comments

„Useful when I can‘t turn on my smartphone‘s sound.“ (m, 11)
„I can‘t hear so well anymore. The transcript would make listening to the radio easier for me.“ (m, 60)
„When my boyfriend interrupts, I can just read it all.“ (f, 31)

Some of the negative comments:

„I listen to podcasts on the go and don‘t want to read along." (f, 20)
„If I want to read, I'm reading a book, not a transcript." (f, 58)
„When I read, I read. When I listen, I listen." (m, 24)

Summary and interpretation: About half of the participants found the text transcript feature useful. Several special cases were mentioned, where this feature is useful: going back and reading it again, people with hearing difficulties, learning a language, translation. Others did not see the benefit of the feature.

2.4.5 Usability experience

Concerning the usability, one of the questions was: "How do you find the dial for navigating within the piece?". Answers are reported in Figure 22.

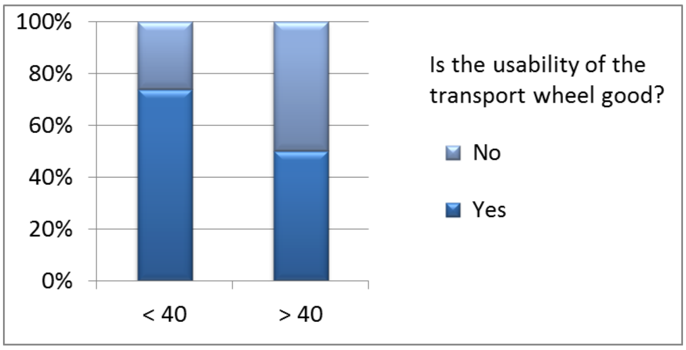


Figure 22: Results for usability of transport wheel

Some of the positive comments:

„I find the design and the operation of the dial attractive." (f, 21)
„The dial and apps are eye-catchers and easy to use." (m, 54)

Some of the negative comments:

„If you've missed something, you can rewind intuitively." (m, 28)
„I could only use the transcript with some guidance." (m, 56)

Summary and interpretation: Regarding the usability questions, co-creators responded mainly on the navigation dial. This feature was mainly appreciated by the younger co-creators. Only half of the older ones were able to use it intuitively. Co-creators also spontaneously added some more features on a wish list: e.g. search function, a user interface for people with bigger fingers, music or program adaptive presets, and user-defined presets.

Do you like the new user features? Yes	88%
Do you like the new user features? No	12%

Table 4: Results for usability in general, n = 196.

Would you use the app? Yes	58%
Would you use the app? No	42%

Table 5: Results for usability, usage of the app, n = 116.

Table 4 and Table 5 suggest that most people like the new features of the app, provided by the object-based approach. However only 58% expressed they would use such an app.

Some of the positive comments:

- binaural sound
- access to interesting radio programmes
- appealing, innovative design
- individuality, because you can create your own profile
- intuitive and easy to use
- use: yes, if free of charge and data protection is guaranteed

Some of the negative comments:

- too many functions
- no added value because I am more a casual listener
- the app is not useful because it does not provide presets, you have to set everything yourself
- cumbersome handling
- not much different from existing apps
- use of the app: only if it is integrated in Spotify, otherwise not useful

The most liked functions are listed in Figure 23.

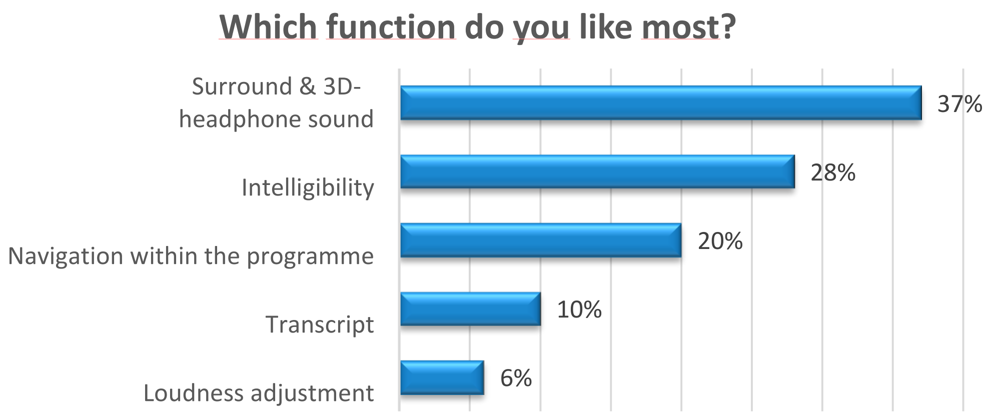


Figure 23: Which function do you like most? n = 194

2.4.6 Summary of the JOSEPHS Results

The open-question investigation approach has delivered results that would probably not been obtained in a more pre-determined way, with fixed questions. The setting in a public accessible venue and a run-time of three months resulted in a high number of participants (361), covering all ages and types of users. The audio quality evaluation is less strict than in a standard listening test but may represent the judgment of a “normal” user more closely.

Summary list:

- The majority of the co-creators listen to music or spoken word programs casually or on the go.
- Streaming services, such as Spotify, are the most popular audio apps among the co-creators.
- Most of the co-creators rated binaural sound as better and more natural than mono or stereo.
- Co-creators rate the foreground/background balance predominantly positive.

- 75% of co-creators find the setting “turn up/down the volume of specific sounds, e.g. the announcer" useful.
- The majority of co-creators liked the surround sound over speakers as played in the living room setup. However, some co-creators find this sound too intense.
- Only half of the co-creators would use the transcript.
- The transcript is deemed useful for people with hearing impairments, for language learning and translation.
- Co-creators up to the age of 40 found the transcript and dial more intuitive and easier to use than older users.
- Some co-creators would like an easier navigation and search function as well as more presets within the app. All in all, controls and font should be bigger.
- The majority of co-creators liked the new app functions. Surround and 3D-headphone sound are the most popular.
- Just over half of the co-creators would use the app.

The general acceptance of the new features and functions provided by object-based audio is very high. Much potential is recognized, but usability can still be improved for some of the users. A very positive surprise is that the sound quality with surround sound or binaural reproduction impressed the listeners most. Until now, this aspect of “reality-like listening” was apparently underestimated in audio production and commercial products. The importance of speech intelligibility for Radio (and TV) is a long known and often discussed issue. But to achieve improvements now with object-based audio seems to have never been easier.

3 Subjective evaluations conducted at IRT

This section describes two evaluations which were performed at IRT within the scope of ORPHEUS. The first subsection describes a comparison of soundbars and the second subsection illustrates an evaluation of renderers over different loudspeaker setups.

Subjective evaluations of audio content are often conducted using a method that compares stimuli with an explicit and/or hidden reference, such as ITU-R BS.1534 [13] or ITU-R BS.1116 [14]. For certain applications, however, a methodology without a reference is required if e.g. different reproduction systems (loudspeaker setup vs. headphones) shall be evaluated and compared. For the object-based audio context, such a methodology is also considered to be useful for e.g. comparing different renderers. For multiple comparisons of audio stimuli without explicit reference, the ITU-R – which is the relevant standards body for broadcasters – has currently no published Recommendation for an appropriate methodology, but one was proposed recently which is called MS-IPM.

The multiple stimulus ideal profile method (MS-IPM), a new method to assess advanced sound systems without an explicit reference, was introduced in 2016 by Zacharov, Pike et al [15]. Being involved in the standardization of such methodologies, IRT conducted two experiments to gather experience and knowledge with MS-IPM. The first was an evaluation of the audio quality of soundbars compared to an ordinary TV setup and a 5.1 speaker system. The purpose of the second was to assess the characteristics and performance of different object-based audio renderers on different loudspeaker layout configurations.

3.1 Evaluation method MS-IPM

The MS-IPM is designed to evaluate various systems without an explicit reference and provides measures of overall subjective quality, as well as characterising the nature of the systems by using attributes.

The MS-IPM uses the multiple stimulus presentation approach to compare the sound systems under test similar to the MUSHRA [13] approach. The assessors are asked to provide their overall impression of the systems on a 100-point basic audio quality scale. A multiple stimulus comparison is also used for the rating of the attributes. Relevant attributes to describe the differences between the systems are selected by experts prior the test from established lexica. Additionally, the method seeks to establish how well the sound systems under evaluation compare to an envisaged ideal. For this purpose, the assessors are asked to rate the ideal level of each attribute, a hypothetical ideal system based on their wishes and experience. Depending on the nature of the systems under test and the attribute ratings, the ideal point may vary from the ratings of the systems. It should not be assumed to yield the same results as the preferred system.

The test should be conducted in the following manner:

1. Assessor instruction
2. Basic audio quality familiarization
3. Basic audio quality rating
4. Ideal point and attribute familiarization
5. Ideal point and attribute rating

For the first step, the assessors are provided with written and verbal instructions about the test in general and a detailed description of the task. In the second step, the assessors have time to listen to the test samples and familiarize themselves with the systems and the software for the basic audio quality (BAQ) rating (Figure 24, left). The rating of the BAQ in step three is conducted for all systems under test in a multiple stimulus comparison. Each trial comprised one test sample. The order of the samples and the systems should be randomized for each assessor.

After the BAQ rating the assessors have time to familiarise themselves with the attributes and the ideal point rating (Figure 24, right). A detailed description of the attribute under test should be included in the test software. For each attribute the order of the samples and the systems should be randomized.

The combination of the basic audio quality and attribute rating allows an in-depth analysis and interpretation of the quality of the systems under test.

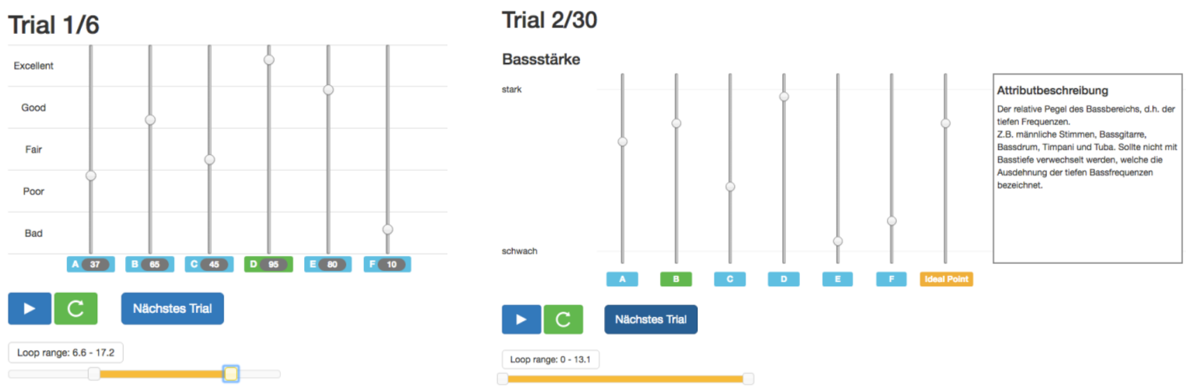


Figure 24: Basic audio quality test (left) and attribute test (right) graphical user interfaces

3.2 Soundbar evaluation

3.2.1 Experimental setup

The purpose of the evaluation was to find out whether soundbars may be used as a good alternative for a 5.1 speaker system in a living room environment and whether they can improve the audio reproduction quality compared to a common TV set significantly.

Eight soundbars were selected for the test and compared to a common TV set with integrated speakers and a high quality 5.1 speaker system to evaluate their audio quality. The differences of the playback devices were studied in two separated tests for stereo and 5.1 content. Typical German TV content from different genres (sport, documentation, TV-show, movie and music) was used to test the systems.

To select relevant attributes for the test, four expert assessors familiarized themselves with the sound systems and the test samples in the listening room. They reviewed available attributes from a lexicon [16] and discussed the selection. They agreed on five attributes which were considered to best describe the differences and characteristics of the systems under test. The selected attributes were envelopment (for 5.1), width (for stereo), tinny, natural, detailed and bass strength.

The test was conducted in the order described in 3.1. In the attribute rating, the assessors rated all samples for one attribute before continuing with the next.

3.2.2 Test results

A number of analyses were performed on the collected data. The Shapiro-Wilk test showed that the data was normally distributed. The applied ANOVA showed significant influence of the systems and no significant influence of the assessors.

Figure 25 and Figure 26 illustrate the average BAQ scores for 5.1 and stereo content averaged over all six samples and 24 assessors.

In order to obtain a more detailed view on the data, the attribute and ideal point data was studied. The ideal point ratings for each attribute were averaged over all systems and assessors. This creates an ideal profile which illustrates an envisaged ideal system provided by the assessors. The ratings for 5.1 content for all attributes and each system averaged over the 24 assessors and six samples are presented in combined spider plots in Figure 27. This data collection explains the performance of the systems better and in more detail. For example, the TV set is found to lack not only transparency characteristics, but there is also a lack of envelopment. Moreover, the system appears to be very tinny with nearly no bass strength. The 5.1 system and SB 2 come closest to the ideal profile. This separates them from the rest of the soundbars for most of the attributes, whilst only for bass strength more of the soundbars seem to reach the ideal point. Figure 28 and Figure 29 show detailed results of the attribute ratings for envelopment and bass strength.

Figure 25: Average" Basic Audio Quality" of all ratings for 5.1 content

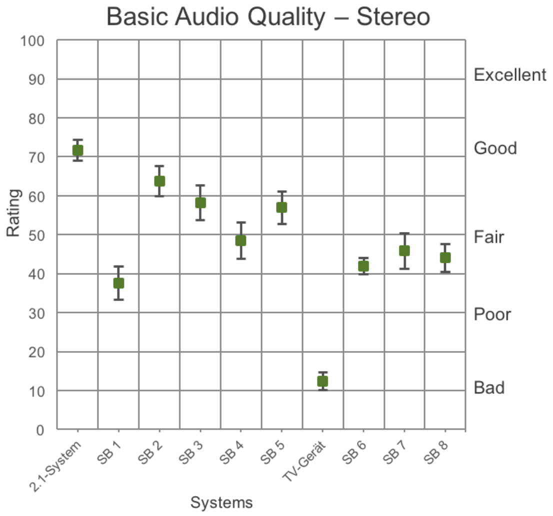


Figure 26: Average "Basic Audio Quality" of all ratings for Stereo content

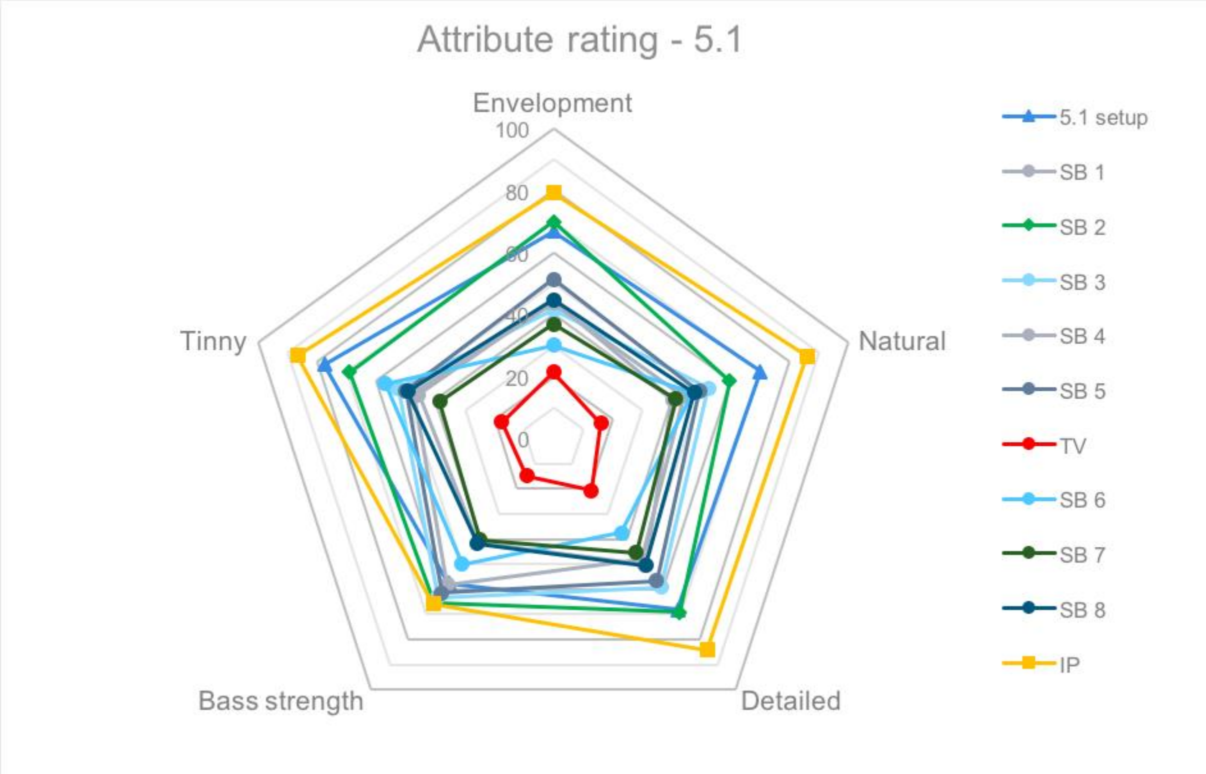


Figure 27: Combined spider plots of the attribute rating per system with 5.1 content, averaged over all 24 assessors and all samples

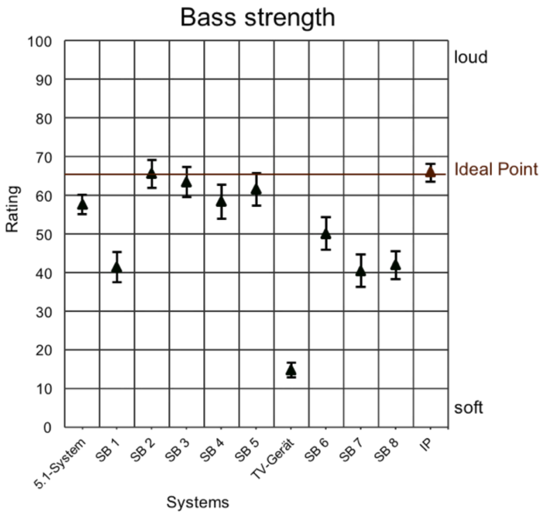


Figure 28: Attribute and Ideal Point ratings for "Bass strength" with 5.1 content, average over all 24 assessors and all samples

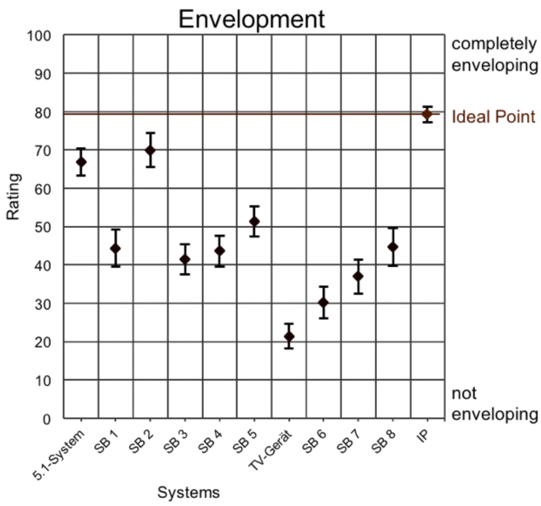


Figure 29: Attribute and Ideal Point ratings for "Envelopment" with 5.1 content, average over all 24 assessors and all samples

3.2.3 Conclusion

The test results show that soundbars can improve the audio quality of an ordinary TV set in a living room environment significantly for typical TV content. For both 5.1 and stereo content, the basic audio quality as well as the attribute ratings show significant better results for most of the soundbars. Some of the soundbars could even be an alternative for a high level 5.1 speaker system in this environment. Over all, the soundbars showed slightly better ratings for stereo content compared to 5.1 samples.

The MS-IPM proved to be the right choice for the given experiments. The attribute and ideal point ratings provided a better understanding of the quality of the systems under test and the assessor's expectations in the context of the test. The absence of a reference was a challenge for the assessors but gave better insight in the relations of the systems among themselves and to the assessor's expectations.

3.3 Renderer evaluation

3.3.1 Experimental setup

The focus of this evaluation was to assess the similarities, characteristics and performance of different Next Generation Audio renderers in a range of broadcast and standard listening rooms, using different loudspeaker layout configurations and in absence of an explicit reference. The renderers under test were a commercially available product and an internal development, both foreseen to be used in the production of NGA and especially object-based programs. The test was conducted within the scope of an EBU working group by multiple institutions (BBC, France TV, NHK, Force Technologies and IRT) in their listening rooms.

The experiment was designed to compare seven systems with six original broadcast audio items. The systems comprised of a combination of channel layouts (0+2+0, 0+5+0, 4+7+0, 9+10+3), in accordance to ITU-R BS.2051-1, two different renderers and one down-mix. The identical double-blind test design was performed in five different laboratories, comprising of either ITU-R BS.1116-3 compliant listening rooms or broadcast listening labs. In total 58 assessors participated in the study across the five laboratories.

When an explicit reference is not available, as in this context, the only suitable ITU recommendation is ITU‑R BS.1284 using paired or multiple comparison methods but the goal of this test was to go beyond the overall quality. For this reason, the MS-IPM was chosen: the overall quality still remains the primary indicator of performance but the analysis of the attribute ratings and that of the ideal profile provide insight into the most perceptually pertinent characteristics (attributes) and strengths or weaknesses of the technology under test.

One additional reason to conduct this evaluation was also to collect critical feedback from the assessors in order to improve the methodology in the standardization process.

The same test was set up in each laboratory with instructions, attributes and the test user interface translated into the local language. Six programme items were selected to represent a broad range of broadcast content including sports, radio dramas, classical, and electronic music. The evaluations were performed using the following attributes, selected specifically for this experiment (with associated descriptions and scales) and response variables:

- Basic audio quality
- Envelopment
- Scene depth
- Localisation accuracy
- Tone color
- Clarity

Additionally, for each attribute the assessors were asked to envisage the ideal characteristic they might desire and provide a rating of this ideal level for each attribute.

The test was conducted in the order described above in 3.1. Contrary to the soundbar evaluation, here the assessors rated all attributes for one sample before continuing with the next.

3.3.2 Test results

As the primary research question of the study was to investigate the similarity of renderers for a range of listening conditions (different listening rooms, loudspeaker types and equalisation strategies), we studied the aggregated performance across all laboratories.

A strict post-screening was performed for basic audio quality rating using the method provided in Report ITU-R BS.2300 and the best 35 assessors passed the post screening were included for the subsequent analysis. The univariate analysis was applied to study the research question for basic audio quality and each attribute individually. Additionally, a combined analysis was conducted using a multivariate analysis (PCA).

A first analysis of the data showed that the majority of the assessors were reliable on all attributes. The data is normally distributed and shows differences between the systems. Whilst this section describes the results of the evaluation in a brief and illustrative way, a full report of the evaluation is currently being drafted by the EBU working group and will be published in near future.

Figure 30 illustrates the average BAQ scores for all systems, averaged over all labs, the 35 post screened assessors and all samples. Overall there’s a significant difference between the systems, but a larger difference between the different layouts than the systems. Between the different samples there were only small differences.

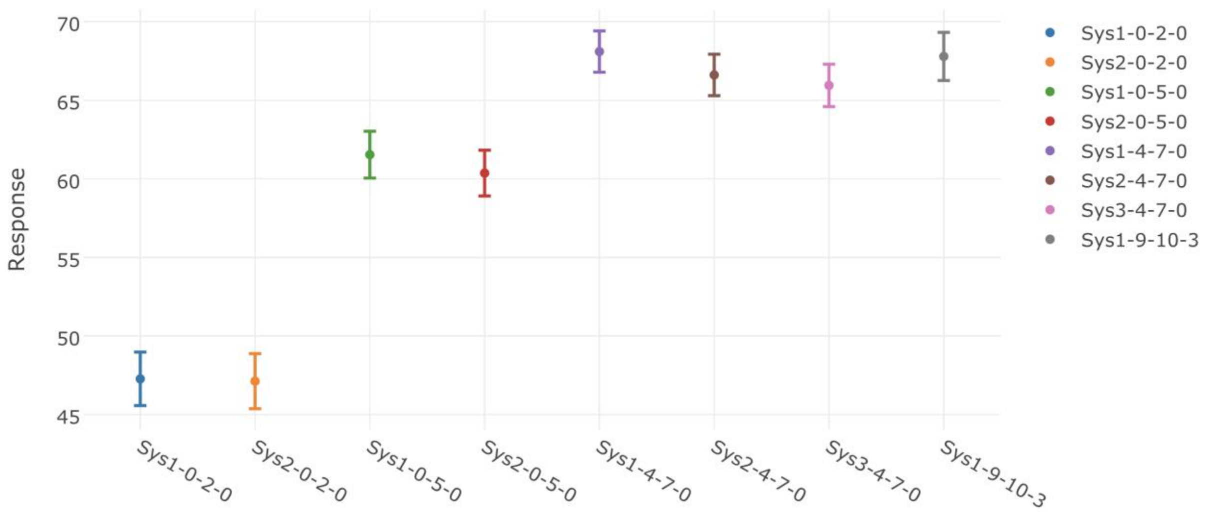


Figure 30: Overall “Basic audio quality” scores for systems including the 95% confidence intervals, averaged over all labs, the 35 post screened assessors and all samples (Source: FORCE Technology)

For a more detailed view on the data, the attribute and ideal point data was studied. The ideal profile which illustrates an envisaged ideal system provided by the assessors and the raw attribute ratings for all attributes and each system averaged over the 35 post screened assessors and six samples are presented in combined spider plots in Figure 31. This data collection explains the performance of the systems in more detail. Figure 32 and Figure 33 show detailed results of the attribute ratings for envelopment and scene depth.

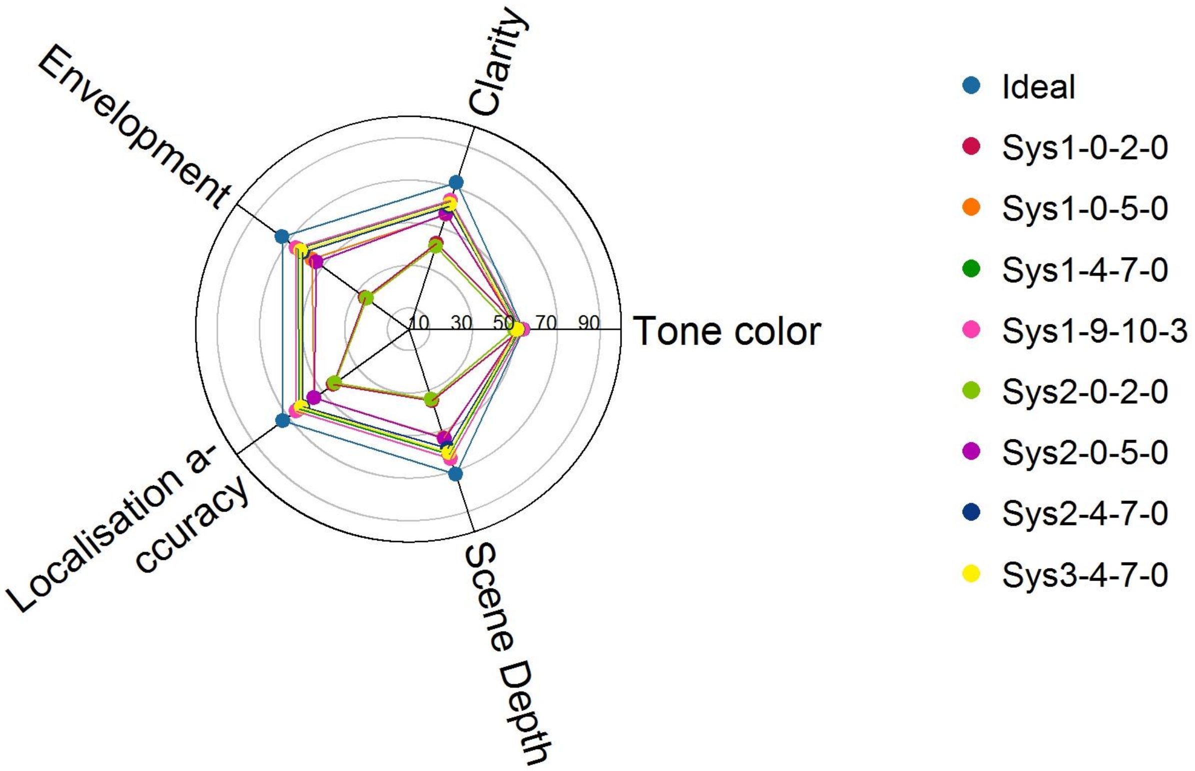


Figure 31: Combined spider plots of the attribute rating per system, averaged over all labs, all assessor and all samples (Source: FORCE Technology)

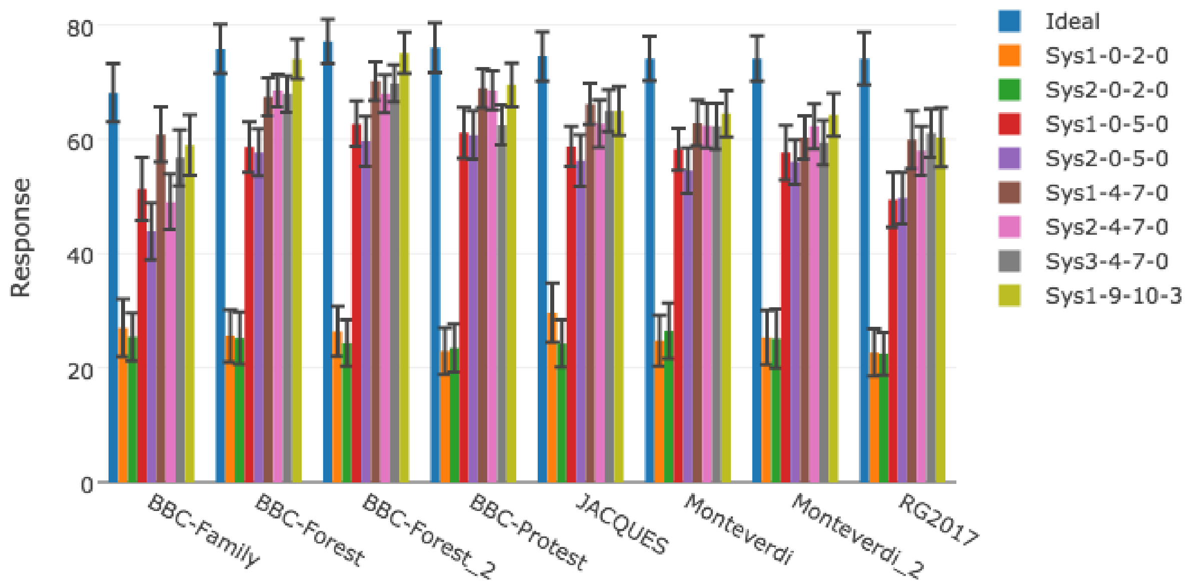


*Figure 32: "Envelopment" scores for systems*samples, averaged over all labs all assessors (Source: FORCE Technology)*

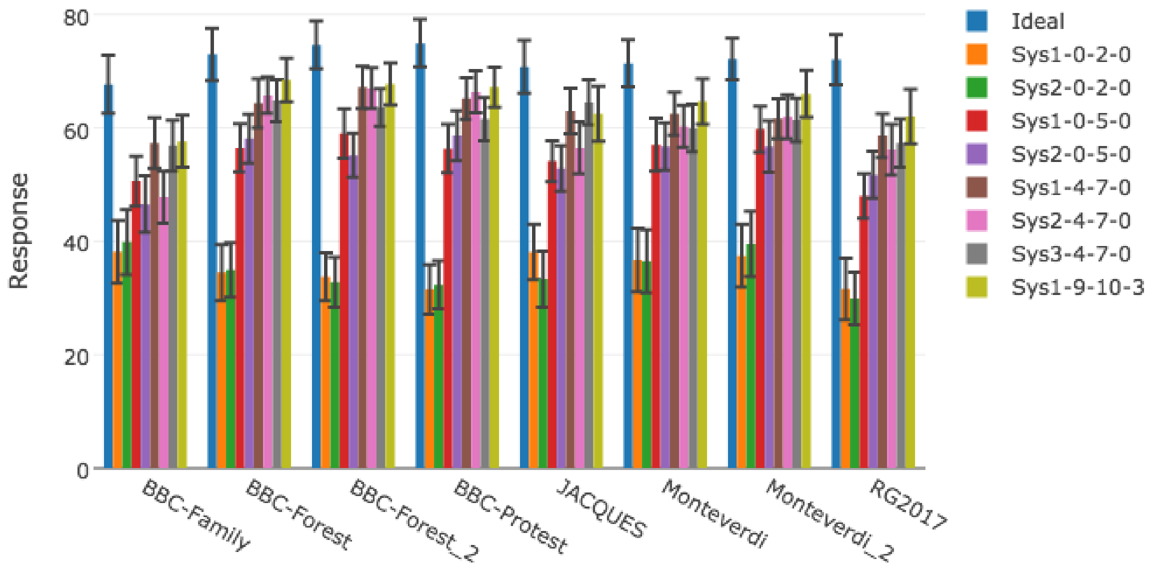


*Figure 33: "Scene depth" scores for systems*samples, averaged over all labs all assessors (Source: FORCE Technology)*

When averaging across all assessors, laboratories and programme items, statistically significant differences are shown between the seven systems. For basic audio quality, scene depth, envelopment, localisation accuracy, and clarity the systems differences are primarily between the channel layout. With the exception of tone colour, all attributes were clearly and reliably discriminating system differences. For any given channel layout or attribute, no statistically significant differences are found between renderers.

The data was analyzed further by applying a principle components analysis (PCA) to the attribute data and projecting the ideal profile data into the PCA space. The multivariate analysis, taking all attributes into the analysis, is illustrated in Figure 34 and Figure 35. This analysis provides an overview of dominating perceptual characteristics and differences in the dataset. The first dimension dominates 55% of the variance of the data and relates to spatial attributes. The second dimension shows 20% of the variance linked to tone colour. The 3rd dimension, explains a further 11% of the variance, but has not been discussed in detail in this paper. The 2-dimensional PCA is suitable to explain 75% of the variance, as illustrated in the biplot. The 95% confidence ellipses allow us to evaluate the statistical similarity of systems under study. For dimensions 1 and 2, for any given channel layout or attribute, no statistically significant differences are found between renderers. Similarities can be found between the systems with overlapping 95% confidence ellipsis, e.g. Sys1-0-2-0 and Sys2-0-2-0.

The ideal profile can be studied from the raw data (Figure 31) or the PCA analysis (Figure 34) and illustrates an envisaged ideal system provided by the assessors in context of the systems under evaluation. The ideal profile is not absolute, but indicates the degree of similarity between the ideal and the systems under test. Figure 34 shows that all of the systems are, from a statistical standpoint, significantly different from the ideal rating for dimensions 1 and 2.

Further, interpretation of the system performance per programme item, averaged over all assessors and all labs, provides insight into system differences for each programme item. In a few specific cases with certain programme items, attributes, and speaker configurations significant differences can be identified between renderers.

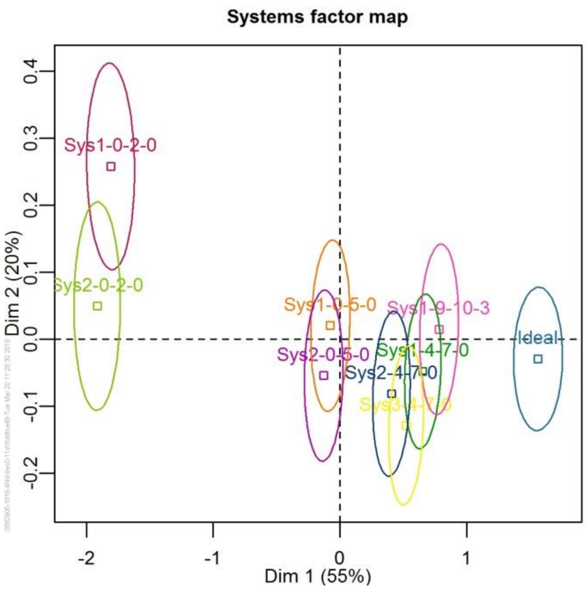


Figure 34: Principal component analysis (PCA), system factor map for dimensions 1 and 2 (Source: FORCE Technology)

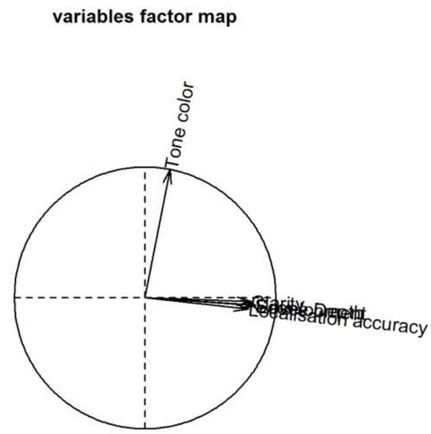


Figure 35: Principal component analysis (PCA), variables factor map for dimensions 1 and 2 (Source: FORCE Technology)

3.3.3 Conclusion

The results of the performed subjective evaluation confirmed the main research hypothesis, i.e. the similarity of the renderers.

The data shows statistically significant differences between some of the systems under test. There are large and statistically significant differences between loudspeaker layouts but no statistically significant differences between the renderers under test. Also, all loudspeaker system and renderer combinations were found to be significantly different from the ideal profile.

Even if the test can be considered as a difficult one due to the similarity of the systems, the statistical analysis showed discriminating results for all attributes but the tone color. The data also shows that some attributes were more relevant than others. Therefore, a better attribute selection could have improved the test results.

A full report of the evaluation will be published by the EBU in near future. The received feedback was very helpful for the further refinement of the methodology within the standardization process.

3.4 Feedback

MS-IPM is a relatively new method and has not been used very often for the evaluation of audio content. One of the reasons to conduct this test was also to get especially critical feedback from the assessors in order to improve the methodology specification. The summarized feedback is as follows:

- The concept of the test method, measuring the relationship between total audio quality and individual attributes at the same time, was considered very interesting by the assessors.
- Several assessors found it tedious to listen to the same sample for several trials, e.g. for multiple attributes in the renderer evaluation. To avoid this, the presentation of samples and attributes should be randomized for each assessor.
- Some assessors commented that the original samples were sometimes too long and the

scenes too complex in nature. It would be best to use short samples that ensure consistency across the content.

- Some assessors commented that the training and familiarization was very useful. Written and verbal instructions for the experiment, with clear and well understood attribute names and definitions were very important and helped a lot to understand the task. Sufficient time for the training and familiarization of all the test stimuli and attributes prior to the main test should be planned for each assessor.
- Some assessors commented on the difficulty of using certain attributes for certain samples. Other assessors commented on the similarity of some attributes. It should be ensured that pertinent attributes are selected that allow assessors to discriminate the systems and samples under evaluation well.
- The attribute rating was considered to be very helpful to rate a system more thoroughly than just by means of the basic audio quality. Therefore, the attributes are a key step to getting good results and should be selected with care.
- Some assessors found it challenging to estimate the ideal and requested for guidance on the usage of the ideal rating scale. The meaning and usage of the ideal should be well described to the assessors. Sufficient time for explanation and training regarding the usage of the ideal rating should be included. It is helpful if the assessors are familiar with the field of application of systems under test, such that their expectations are based on experience.
- For reliable ratings, only expert assessors with listening experience of the technology under evaluation should be employed in such tests.

This feedback will be taken into account during the specification process of the methodology within the scope of ITU-R and EBU.

3.5 Conclusions

The MS-IPM methodology was considered by both assessors and test organisers as a very valuable tool for the subjective assessment of the sound quality of advanced sound systems or even sound systems in general, when an explicit reference is not available or appropriate. The method can be considered when current ITU-R methodologies cannot be applied or do not provide sufficient depth of analysis.

The selection of relevant attributes is a key step to getting meaningful results. A lot of consideration should be taken in this step. Furthermore, the familiarisation of the assessors with the stimuli and test interface is very important, especially if they are not yet familiar with this kind of test.

IRT will continue to use this methodology for further subjective evaluations in future.

4 Subjective evaluations conducted at IRCAM

4.1 General motivations

Within an object-based audio scheme, the most straightforward way to convey reverberation and room effects is to encode them as a specific multichannel bed [17][18], which is independent from objects conveying direct sound tracks, such as that originating from microphone trees or individual spot microphones. In the ADM format, for instance, such a reverberation bed is typically represented as an object of type “audioObject” comprising several channels, for which desired directions are evenly distributed in 2D or 3D and specified in the associated “AudioPackFormat”. It can also be represented as a scene-based object of type “HOA”.

The expected advantages of such a structure is to guarantee a more faithful reproduction of the desired sound scene, independent of the rendering setup. Each direct sound track may be processed with a very accurate rendering of its specific direction (e.g. direction specific and personalised HRTF in the case of binaural reproduction on headphones) whereas the rendering of the reverberation channels does not require to reproduce directions accurately provided that the overall sensation of the late room effect is diffuse. Thus, rendering algorithms can be adapted and optimised not only according to the rendering setup (2D or 3D loudspeakers distribution, binaural reproduction on headphones), but also according to the nature of the sound object attributes (e.g. diffuseness).

Another advantage is to allow for some content interaction and personalisation at the end-user side. For instance, the direct/reverberant balance can be tuned in order to improve intelligibility, or the direction and apparent distance of the different sources can be modified. In the Orpheus deliverable 3.2, several reverberation implementation schemes have been discussed and have been shown to allow for different degrees of interaction.

However, a critical aspect of this approach is to decide how many channels should be used to encode the late reverberation. Nowadays, professional rendering setups may easily comprise ten to twenty or even more loudspeakers. In order to convey a reverberation bed that can be optimally decoded on such systems it would however be prohibitive to convey as many reverberation channels. A compromise has to be found between the complexity of the object-based scene, i.e. the number of transmitted channels, and the desired level of interaction/personalisation. A common idea is then to transmit a limited number of reverberation channels and ask the rendering devices to create as many mutually uncorrelated copies of those reverberation channels in order to feed each loudspeaker; or to create enough mutually uncorrelated copies and distribute them spatially in a way that creates the desired sensation of diffuse reverberation.

The aim of the test described below is to compare various strategies for rendering a reverberation field from a limited number of reverberated channels.

4.2 Test on Decorrelation filters

There is a significant literature on decorrelation algorithms. Most of them have been originally introduced as a refinement of panning techniques over multiple loudspeakers, in order to blur the apparent direction of the phantom source or to control its apparent spatial extent. Another main application is related to perceptual coding and upmixing techniques [25][26] where ambience sounds extracted from a low-order channel-based stream (e.g. stereo) needs to be distributed on a larger number of loudspeakers in order to provide a convincing immersive sensation, even for off-centred listeners.

Within the ADM format specification, some attributes may require decorrelation processing at the rendering side. Besides their position in space, Objects may also be described by their *Width* and *Height* corresponding to their horizontal and vertical extent, respectively. Although there is no specific attribute for reverberation signals, the attribute *diffuse* can be associated to such audio tracks in order to indicate that they may require decorrelation processing if the number of loudspeakers is larger than the number of available reverberation channels.

Different decorrelation algorithms have been proposed according to the targeted application (e.g.

control of the apparent source width or synthesis of a diffuse sound field). The general approach is to modify the phase coherence between the original signal and its replica, while preserving their frequency magnitude spectrum as much as possible. If the preservation of the frequency spectrum is a mandatory criterion in any case, the alteration of the time structure may be more or less critical according to the specific application. Taking the analogy of room reverberation, a straightforward approach to create decorrelation is to convolve the original signal by a short noise sequence, possibly modulated by a decaying envelope. This tends to blur the time transients of the processed signal, which may be critical if the aim is only to control the apparent source width, but could still be acceptable when being applied to reverberation channels. Such considerations are taken into account when selecting the different decorrelation algorithms in the following study as it only considers the context of reverberation processing.

The main categories of decorrelation algorithms are:

- Using a combination of comb filters [19][20]. This technique, which is among the pioneering approaches, is however known to create colouration artefacts and will not be considered in this study.
- Processing the signal by a filterbank and applying different delays to each frequency band [22][23][28][29]. This technique will not be considered here as it is mainly used for the control of the apparent source width over two channels and may be difficult to generalise over complex loudspeaker layouts.
- Designing an allpass filter in the frequency domain, with a unit magnitude across all frequency bins and a uniformly distributed random phase [21]. The filter is implemented in the time domain after inverse Fourier transform and truncation. This technique is integrated in the study as it is often referenced in the literature. Moreover, it is proposed in the ADM baseline renderer.
- Convolving the reverberation channels (or the diffused composed of an omnidirectional RIR) with a short, exponentially-decaying Gaussian noise burst [25]. Refinement of the technique uses different decays in different frequency bands [26]. This approach is included in the present study as it has been especially introduced for the decorrelation of reverberated signals.
- Filtering the original signal with a cascade of biquad allpass filters, described each by a randomized angle and radius of the pole-zero pairs [24]. This approach is included as it has been shown to provide better behaviour than the allpass filter design using random phase distribution. Moreover it offers efficient control parameters (number of allpass cells, radius range of the pole-zero pairs).

4.2.1 Protocol

The test follows the MUSHRA protocol [13]. For each sound excerpt, the participant is asked to rate the quality of the different methods under examination and compare them to the reference method. The participant is free to switch at any time between the different methods and the reference by pressing the labelled buttons (Figure 36). The reference method is labelled R, and the tested methods are randomly relabelled at each run, in order to avoid any routine in the ranking. One of the tested methods is a copy of the reference method in order to check the participant’s ability to discriminate the methods. One of the stimuli is generated with a low-quality algorithm and exhibits clearly audible artefacts in order to help the participant using the full rating range.

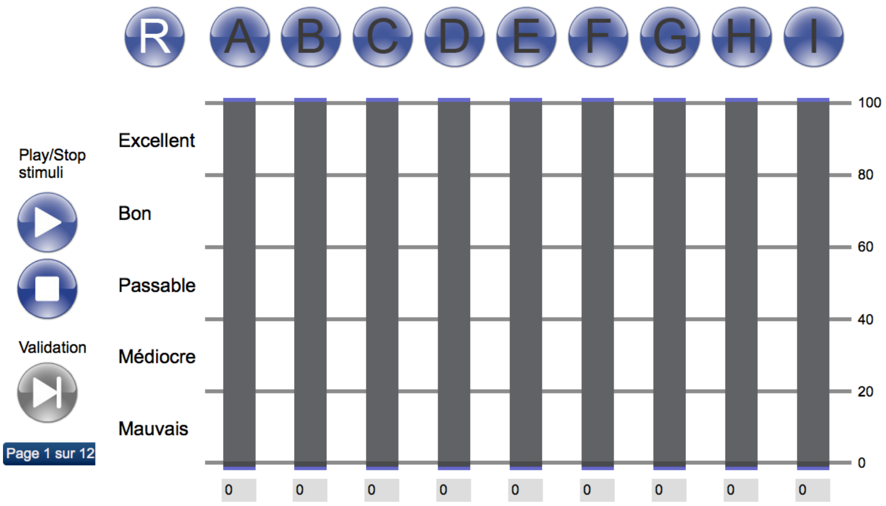


Figure 36: User Interface of the MUSHRA test

4.2.2 Equipment

All stimuli and decorrelation methods are played back on the same quasi-hemispheric loudspeaker setup. The loudspeaker coordinates (azimuth, elevation and distance) are given in *Table 6*. For practical reasons, the distance of the loudspeakers to centre (listening point) vary from 1.60 m to 1.95 m. All loudspeakers were equalised for level and propagation delay differences with respect to the listening point (i.e. realigned on a virtual hemisphere).

Ring 1	Elevation 0°	**Azimuth (°)**	-158	-117	-84	-45	0	47	83	117	157
		Distance (m)	1.69	1.83	1.59	1.63	1.61	1.65	1.61	1.86	1.7
Ring 2	Elevation 31°	**Azimuth (°)**	-20	-60	-102	-141	180	138	100	63	20
		Distance (m)	1.62	1.65	1.65	1.67	1.65	1.65	1.68	1.64	1.62
Ring 3	Elevation 59°	**Azimuth (°)**			-135	-71	0	71	134		
		Distance (m)			1.91	1.92	1.91	1.94	1.93		
Ring 4	Elevation 90°	**Azimuth (°)**					0				
		Distance (m)					1.95				

Table 6: Coordinates of the loudspeakers of the quasi-hemispheric layout

Figure 37: Layout of the loudspeakers

4.2.3 Selection of tested methods and stimuli

4.2.3.1 Tested methods

The reference method applies reverberation processing to the different sound excerpts. The reverberation processor uses IRCAM's real-time spatial audio processing library Spat~. The reverberation module is based on a Feedback Delay Network (FDN) using 32 internal feedback channels. Twenty-four of these decorrelated channels are directly fed to the loudspeakers of the rendering system.

All the other methods only use a limited number (one to four) of reverberation channels, then create decorrelated copies of these channel(s) and distribute them on the rendering setup.

- Methods #1 & #2 are based on 24 allpass filters designed in the frequency domain with unit magnitude and random phase across the frequency bins. The length of the filters is 512 samples (i.e. 10ms). However, as the magnitude is only defined at the discrete frequencies, it does not guarantee the magnitude behaviour in between frequency bins. Method #2 attempts to reduce the strong zeros in between the FFT bins. Twenty-four filters are generated independently and their outputs are connected to the twenty-four loudspeakers respectively.

  Note that Method #1 is equivalent to that proposed for the ADM baseline renderer.

- Methods #3 & #4 are based on a zero-mean white Gaussian noise process, which is modulated by a decaying envelope $w(t)$:

  $$w\ (t) = \frac{e^{\frac{t_{max} - t}{t_{max}}} - 1}{e - 1}$$

  where t_{max} is the length of the noise burst.

  Twenty-four filters are generated independently and their outputs are connected to the twenty-four loudspeakers respectively.

  - Method #3: $t_{max} = 43ms$, i.e. 2048 samples
  - Method #4: $t_{max} = 171ms$, i.e. 8192 samples

- Methods #5 and #6 are based on a chain of biquad allpass filters. In each processing cell, the pole and zero angles are chosen randomly. The radius of the poles is also randomly chosen within a specified range (the radius of the associated zeros is calculated accordingly). The two methods differ in terms of number of cells and radius range. By construction, these recursive filters are IIR. Truncation at a dynamic range of 90dB results in the filter lengths given below. Twenty-four filters are generated independently and their outputs are connected to the twenty-four loudspeakers respectively. Method #5 will play the role of a low anchor.

    - Method #5: 128 cells, radius min = 0.1, radius max = 0.9, length ~ 1024 samples
    - Method #6: 1024 cells, radius min = 0.8, radius max = 0.9, length ~ 2048 samples

- Methods #7 & #8 investigate the spatial distribution of the decorrelated channels when their number is lower than the number of loudspeakers and when being applied to a limited number of original reverberation channels.

    - In method #7, two original reverberation channels are used to feed a 1st-order Ambisonics stream. The first reverberation channel is sent to a decorrelation filter and feeds the Y_{00} component. The second reverberation channel is decorrelated three times in order to feed the three components Y_{1-1}, Y_{10} and Y_{11}, respectively. This 1st order Ambisonics stream is then decoded on the twenty-four loudspeakers. Note that this decorrelation method uses the same first four filters as in method #6.
    - In method #8, four original reverberation channels are used to feed the sixteen components of a 3rd-order Ambisonics stream, which is then decoded on the twenty-four loudspeakers. The first channel feeds the Y_{00} component through a decorrelation filter. The second channel feeds the three 1st-order components through three decorrelation filters. The same procedure is then repeated for higher orders.

Method	Description	Length (samples)	Spatial scheme
0	FDN (REF)	N.A.	[24 ----->24]
1	Allpass filters (random phase – unit magnitude	512 samples	[1-> 24->24]
2	Allpass filters (random phase – unit magnitude – avoid zeros)	512 samples	[1-> 24->24]
3	Noise burst	512 samples	[1-> 24->24]
4	Noise burst	2048 samples	[1-> 24->24]
5	Allpass biquads (preset #1)	~1024 samples	[1-> 24->24]
6	Allpass biquads (preset #2)	~2048 samples	[1-> 24->24]
7	Allpass biquads (preset #2)	~2048 samples	[2 -> 4 -> 24]
8	Allpass biquads (preset #2)	~2048 samples	[4-> 16-> 24]

Table 7: Reverberation methods that were tested.

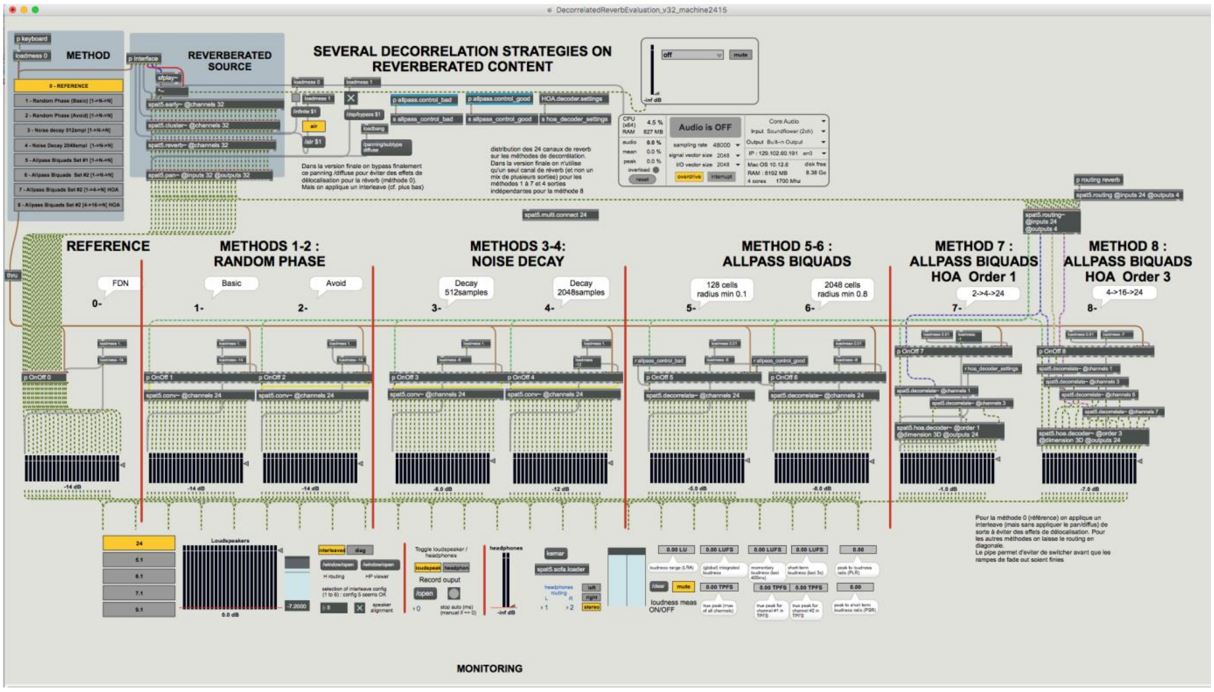


Figure 38: Max Patch implementation of the different decorrelation methods used in the test.

4.2.3.2 Objective characterisation

In order to anticipate the interpretation of the perceptual studies, several objective evaluations have been conducted to reveal the perceptual artefacts that are expected, whether they are related to time, frequency or space. This characterisation also served for the design of the experiment, i.e. to select the different tested conditions. Indeed, as it is impossible to vary every control parameter for every method, the perceptual test only helps to verify if the objective evaluation methods can serve as good criteria to assess the perceptual quality of the decorrelation methods.

The different characterisations may be first conducted on the decorrelated channels that are generated by the different methods and feed the loudspeakers. Another possibility is to analyse binaural signals generated using a virtual speaker approach. To this end, the 24 channels loudspeaker signals are convoluted with KEMAR HRTFs corresponding to the loudspeaker directions.

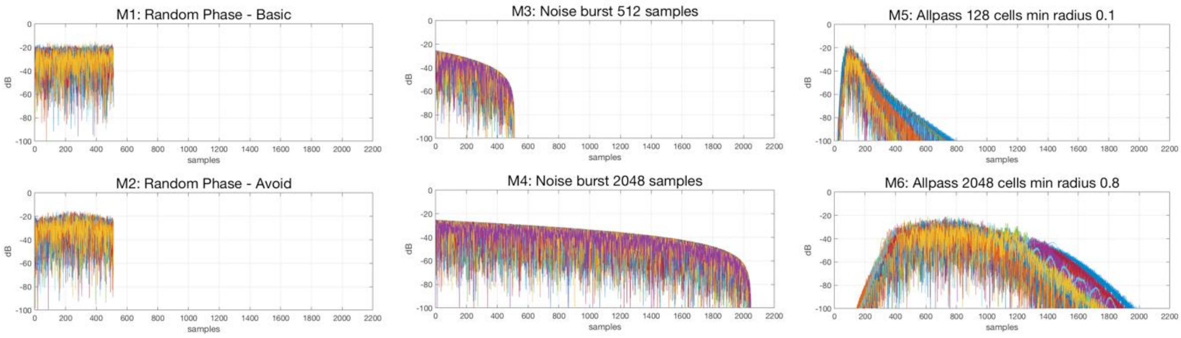


Figure 39: Superimposition of the twenty-four impulse responses of the decorrelator filters for each decorrelation method. Left: Random phase method. Mid: Noise bursts. Right: Allpass biquad filters.

Time domain: The impulse responses of the different filters are presented in Figure 39. For each method, the twenty-four responses of the filters are superimposed. The main observation is related to the behaviour of the allpass filters built with a series of biquads (methods 5 & 6). Their impulse response is infinite and shows a long onset, which increases with the number of cells. We can expect a lower efficiency compared to the more time-compact methods 1 and 2.

Spectral dimension: In order to reveal possible colouration effects, the skewness of the spectral distribution of the signal is characterized by its Kurtosis and compared to that of an ideal Gaussian noise or to that of the original reverberation channels.

In order to evaluate the performance of the different methods, the reverberator is tuned with an infinite reverberation time and is excited with a single pulse. The generated outputs should be close to a Gaussian noise. They can be directly distributed to the virtual loudspeakers in order to represent the 'Reference' situation, or they can be first sent through the different decorrelation processors. For each method, as well as for the reference, the Kurtosis is measured for the Fourier transform of every loudspeaker signal as well as for the binaural signals (i.e. the left and the right ear signals) that have been derived with the virtual loudspeakers approach (see above). Figure 40 depicts the Kurtosis of the real part (^ symbol), the imaginary part (v symbol), and the magnitude (* symbol). For the twenty-four channel signals only the minimum and maximum values across the different loudspeakers are displayed (cyan and green symbols, respectively). For binaural signal the Kurtosis is displayed for both the left and the right-ear signals (blue and red symbols, respectively).

It can be observed that the reference method (outputs of the FDN reverberator) fulfils the expected behaviour, i.e. Kurtosis is close to that of white Gaussian noise. This behaviour is observed for the real and the imaginary parts (ideally Gaussian distributions), as well as for the magnitude (ideally Rayleigh distribution). This is not only true for the reverberated signals but also for the left and the right ear signals.

The selected methods show a small augmentation of the Kurtosis values of the loudspeaker feed signals, except for the Noise bursts. In contrast, most of them show a significant increase of the Kurtosis when being assessed at both ears of the virtual listener. This is especially the case for the methods based on allpass filters irrespective of whether they are built in the frequency domain (methods 1 & 2) or with a series of allpass biquad cells (methods 5 and 6). There is no clear evidence about the threshold of the Kurtosis for which a perceptually significant colouration is expected. However, according to this index, method 5 with a limited number of biquad cells may play the role of a low anchor. A comparison of methods 6, 7 and 8 shows that applying the same decorrelation filters to a small number of reverberation channels (two and four for methods 7 and 8, respectively), instead of applying them to a single channel, the criterion improves significantly. The Kurtosis is significantly reduced and is even very close to the reference for method 8.

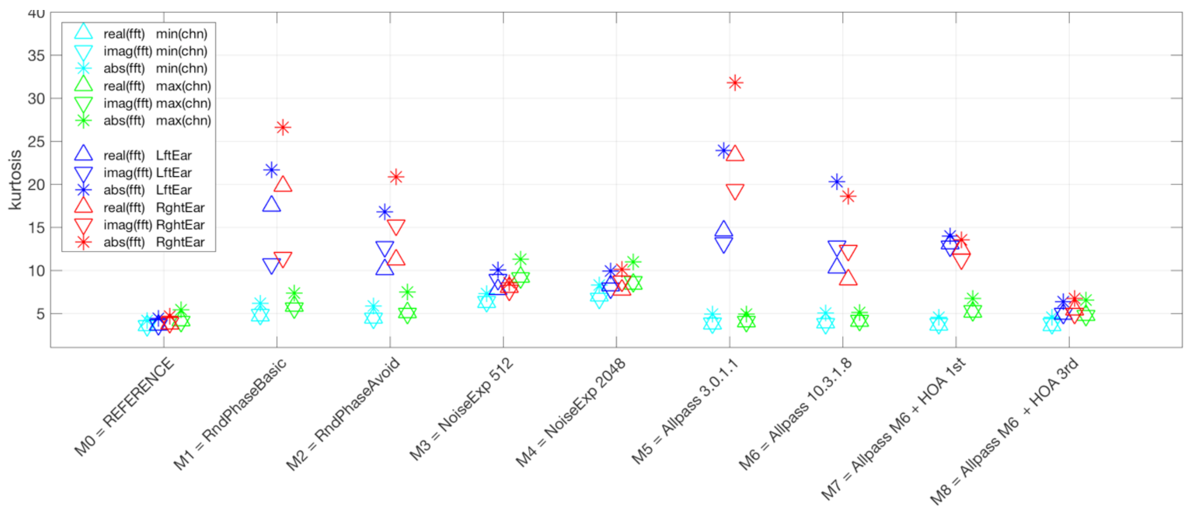


Figure 40: Analysis of the spectral colouration obtained with the reference and the different decorrelation methods. The Kurtosis is calculated for the real (^) and imaginary (v) parts as well as for the magnitude (*) of the signals. For each method, the minimum (cyan) and maximum (green) values are calculated across the different output channels as well as for the left (blue) and right (red) ears of the listener after virtualisation of the loudspeaker setup.

Spatial dimension*:* two indexes have been used to assess the expected spatial diffuseness of the sound field that is reconstructed from the playback of the decorrelation filter outputs. The first index is based on the multichannel cross-correlation coefficients (MCCC). The MCCC is derived from the determinant of the spatial cross-correlation matrix [31]. It is typically used in microphone array signal processing for estimating the time difference of arrival (TDOA) in noisy environments. Here, the MCCC is estimated from the spatial cross-correlation matrix of the loudspeaker feed signals. For the reference method M0 these signals are directly the outputs of the FDN reverberator tuned with an infinite reverberation time and triggered by a single impulse. For the different decorrelation methods, these signals are the outputs of the decorrelation filters. The second index is based on the observation of the Interaural cross-correlation (IACC) estimated at the ears of the virtual listener.

The MCCC and IACC functions are displayed on the left and right sides of Figure 41, respectively. The behaviour of the different methods is also clear. The MCCC of the reference method presents the lowest maximum (< 0.1) and its shape rapidly vanishes with increasing time lag between the different FDN output channels. The same behaviour is also noticed for the IACC (max = 0.13).

Both random phase methods (Methods 1 & 2) provide comparable MCCC functions with a maximum of 0.6. For both methods, the IACC function presents a maximum slightly above 0.2 and does not tend to decrease over the entire delay range (mean value of 0.1). We may expect a lack of spaciousness.

For the noise burst methods (Methods 3 & 4) although the MCCC show a significant maximum value (0.8 and 0.3 respectively), the resulting IACC functions estimated at the ears of the virtual listener seem to be slightly more favourable than for methods 1 & 2. They are centred on zero and present lower maximum values (0.16 and 0.13 for the 512 samples and 2048 samples noise bursts, respectively). However their maximum occurs for an interaural delay unequal to zero, which may lead to some residual delocalization effects as the listening situation do not present any direct sound.

The allpass biquads method (Methods 5 & 6) shows a strong dependency on the number of biquad cells. With 2048 cells (Method 6) it provides results close to Methods 1 & 2. In contrast, when the number of cells is limited to 128, the maximum of the MCCC is close to 1.0 (0.97) and the IACC function shows very high values with a maximum close to 0.4. It confirms that method 5 will play the role of low anchor.

Once again, the comparison between methods 6, 7 and 8 is interesting. In the case of methods 7 and 8, the total number of available decorrelated channels is lower than the number of loudspeakers, which explains why the MCCC function presents a maximum of 1 (the twenty four output channels are linear combinations of the four or sixteen decorrelated channels, respectively). However as they are exploiting a limited number of original reverberation channels (two and four for methods 7 and 8, respectively) instead of a single one, the resulting IACC function presents a better behaviour. For method 8, the IACC function is now close to that of the reference method.

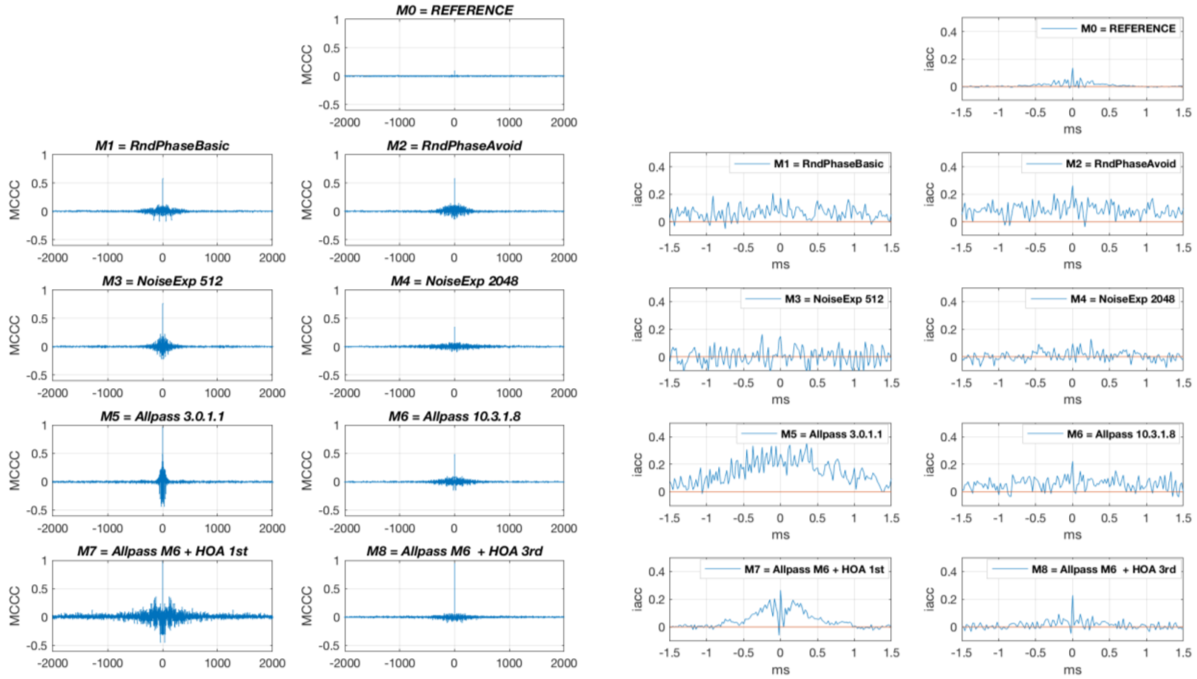


Figure 41: Estimation of the inter-channel cross-correlation for the reference M0 and the different methods M1 to M8. Left: Multichannel cross-correlation (MCCC) estimated across the twenty-four output channels as a function of time lag (from -2000 to +2000 samples). Right: Interaural cross correlation (iacc) estimated at the listener position and displayed from -1.5 to 1.5 msec.

4.2.3.3 Stimuli

The stimuli are short sound excerpts convolved with the different decorrelation methods. Three different reverberation conditions are tested with a reverberation time (RT) of 0.8, 1.6 and 3.2 s respectively. Each reverberation condition is tested with two different sound excerpts (*Table 8*). The first one is an impulse and the second is a more realistic excerpt chosen in line with the reverberation condition. A speech excerpt is chosen for the short RT, a saxophone example for the mid RT and a marimba example for the long RT.

RT (seconds)	Click	Speech	Saxophone	Marimba
0.8	S1	S2		
1.6	S3		S4	
3.2	S5			S6

Table 8: Tested stimuli according to the reverberation time condition

4.2.4 Results

Twenty listeners participated to the test (17 males and 3 females). Most of them are working in the field of acoustics, audio signal processing or music but only six of them could be qualified as audio experts (sound engineers). Following the ITU-R BS.1534 recommendations, one participant was discarded from the analysis, as his results suggest that the task was not fully understood. Although this participant scored the hidden reference higher than all the other methods in the majority of the conditions (> 80%), the score of the hidden reference was however often below 90 (90% of the conditions) and even below 80 (25% conditions). Thus, the analysis is conducted with the nineteen remaining participants.

Figure 42 depicts the distributions (median and interquartiles) of the data collected for each method across all stimuli conditions and averaged over the two repetitions. From this picture, the overall ranking of the different decorrelation methods appears already clearly.

The right-tailed Wilcoxon test was conducted on the scores of the different decorrelation methods averaged across the audio stimuli. The Wilcoxon test has been chosen since these averaged scores are not all normally distributed. This is especially the case for the hidden reference (Kolmogorov-Smirnov test rejected $p < 0.0001$). By construction, its scores show a pronounced negative skewness (< -2.0) since they are limited to 100. To a smaller extent, this is also the case for the method M8 (skewness < -1.0).

The right-tailed Wilcoxon test shows that the score of the hidden reference is significantly higher than any other method ($p < 0.0001$). The score of Method 8 is higher than any others ($p < 0.0001$), except the reference and Method 4 is higher than the rest of the others ($p < 0.0001$). All methods exhibit higher scores than Method 5 ($p < 0.0001$).

These first results confirm that the reference was discriminated from all other methods and that Method 5 played the role of a low anchor. As expected by the objective study, Method 8 is the closest to the reference, followed by Method 4.

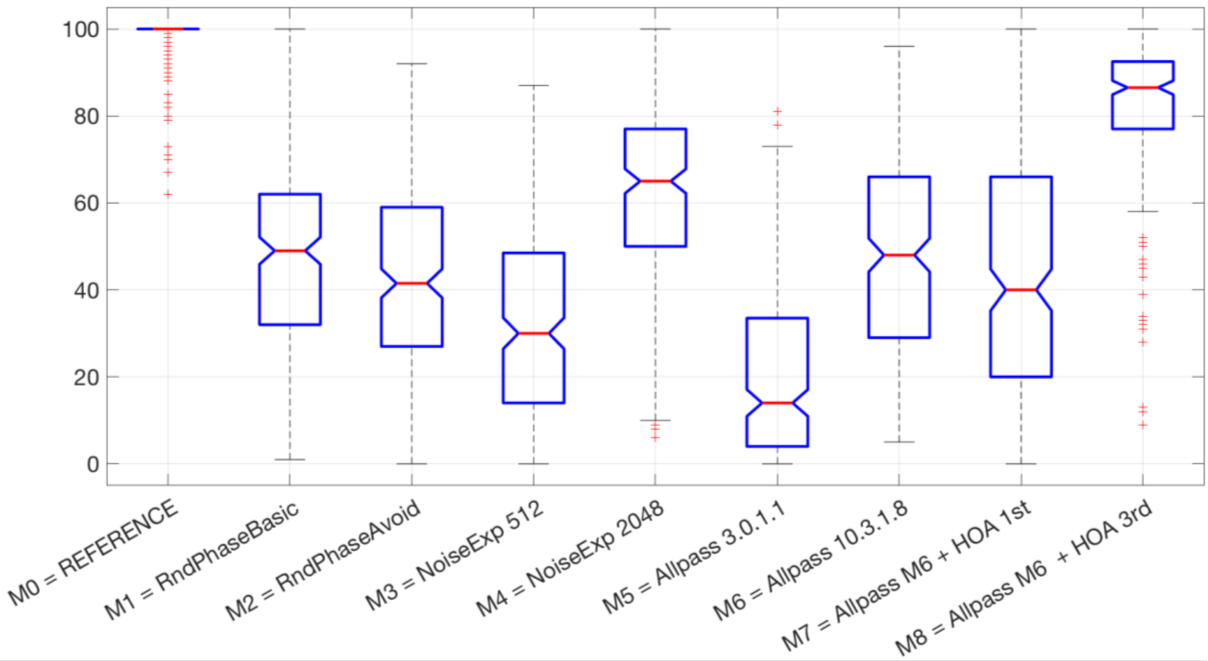


Figure 42: Boxplot of the scores collected among participants and grouped according to the methods.

A repeated analysis of variance (ANOVA) was conducted across conditions to investigate the main effects and two-way interactions of the factors METHODS (M1 to M8) and STIMULI (S1 to S6). The hidden reference was excluded from the analysis, as its scores did not follow a normal distribution ($p < 0.05$) and were highly discriminated from all the other methods. For all the other conditions, the Kolmogorov-Smirnov test did not reject the normal hypothesis ($p > 0.25$). The factor METHODS was shown significant ($F(7,126) = 72.6$, $p < 0.00001$) as well as the factor STIMULI ($F(5,90) = 8.7$, $p < 0.0001$). There was a significant interaction effect between factors METHODS and STIMULI ($F(35,630) = 3.6$, $p < 0.00001$).

A post-hoc test (Bonferroni) conducted on the factor METHODS shows that the scores of the methods can be parcelled out into five homogeneous groups ($\alpha = 0.5$). All methods are discriminated from each other ($p < 0.05$) except for the group formed by methods M1, M2, M6 and M7 (*Figure 43*).

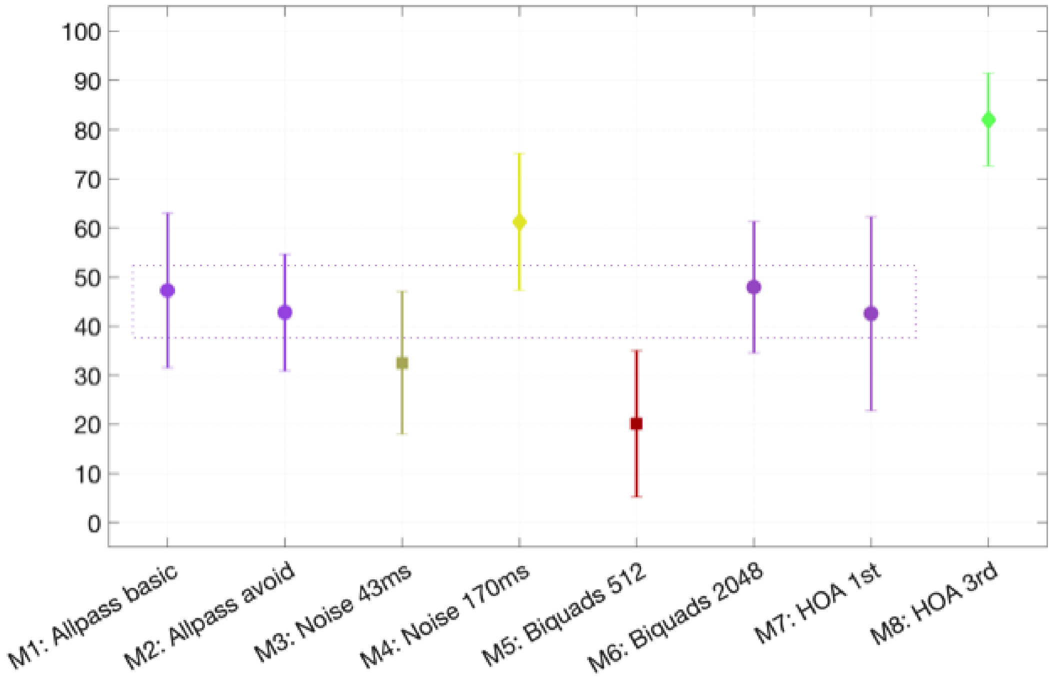


Figure 43: Mean scores and confidence intervals (0.95) of the different methods (except REF) computed across participants and stimuli. Colours represent homogeneous groups (Bonferroni test)

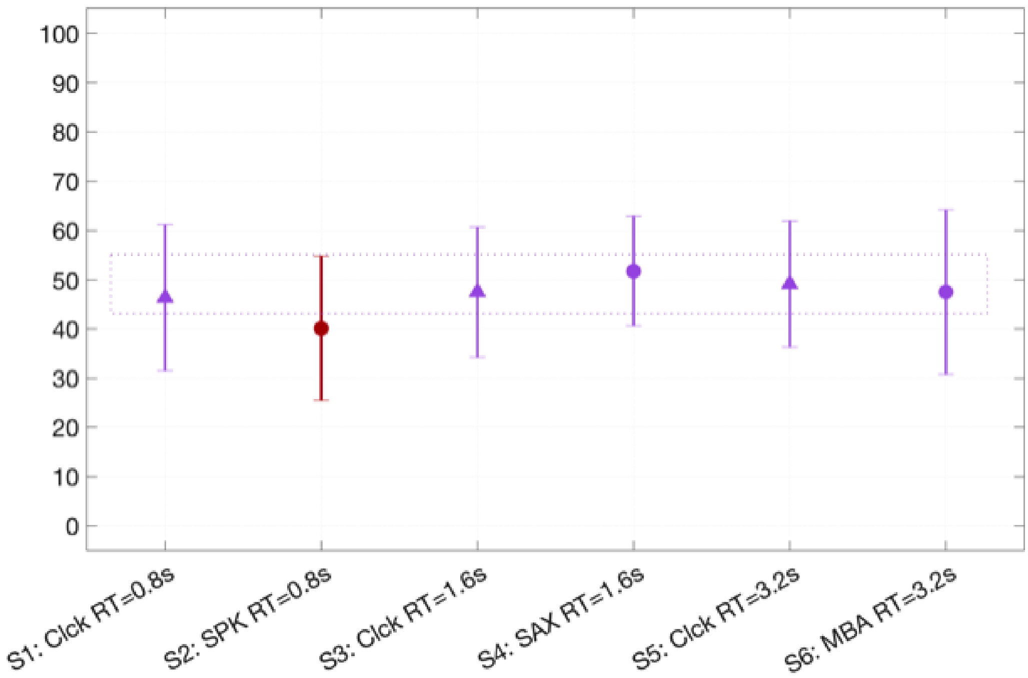


Figure 44: Mean scores and confidence intervals (0.95) of the different stimuli computed across participants and methods. Colors represent homogeneous groups (Bonferroni test).

A post-hoc test (Bonferroni, with $\alpha = 0.5$) conducted on the factor STIMULI shows that only the speech stimulus with short RT (S2) is significantly different from all the others and present and exhibits the lowest mean score (*Figure 44*). This tendency seems to be linked to the timbre of this male voice excerpt, which tone colour was even emphasized by most of the decorrelation methods. This effect can also be observed on *Figure 45*, which depicts the means and confidence intervals calculated across all participants for each test condition (method and stimulus). The score of the speech stimulus (S2) is significantly lower than stimuli S4 and S5 for method M8. The other noticeable effect of the interaction between method and stimulus is shown for method 7 where the score of the stimulus S4 (Saxophone with mid RT) is significantly higher than for the other stimuli.

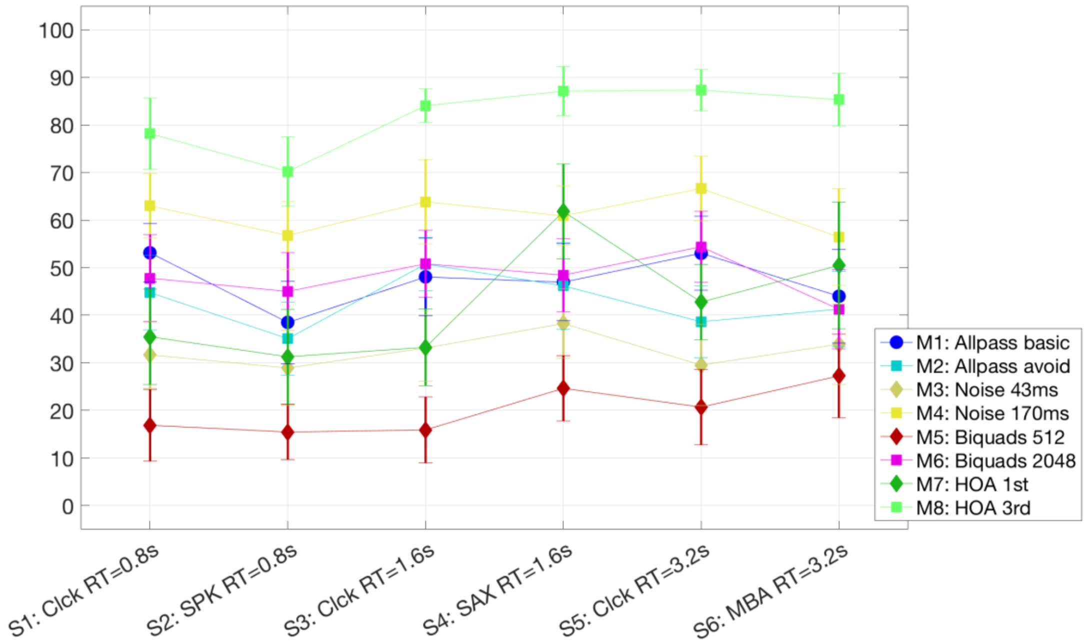


Figure 45: Mean scores and confidence intervals (0.95) collected among participants for each test condition (9 methods x 6 stimuli)

4.2.5 Discussion

The observations show the difficulty to obtain a fully convincing diffuse field from decorrelation techniques, i.e. without exhibiting spatial or timbre artefacts. Methods aiming at reconstructing a diffuse field from a single reverberation channel (M1 to M6) just obtain fair results. It must be noticed however that the test was critical as the sound scene were all restricted to the late reverberation without any direct sound or first reflection sections that could mask some artefacts.

Not surprisingly, the length of the filters improves the performance, as can be seen when comparing methods M4 and M3 or methods M6 and M5. Although the test did not exhibit any effect of the reverberation time, increasing the length of the filters may present a risk when applied to short reverberations.

No significant difference could be noticed for the two allpass methods based on random phase design. The method M2 that attempts to reduce the zeros in between the FFT bins does not seem to modify significantly the performances of the method.

The impact of the stimulus was significant but no general tendency could be drawn from the test. The only exception was for the male speech excerpt, which obtained lower scores. The MUSHRA test does not allow examining whether or not the reference was also judged less natural for this excerpt, since, by construction, participants were enjoined to rate the hidden reference with the highest score.

The significantly highest score was obtained for the method M8 that keeps four original reverberation channels and generates a 3rd-order HOA stream with sixteen decorrelated channels. Although still discriminable from the reference, it was rated excellent on average (score > 80). In a significant number of times it was even confused with the reference (rated equally or above the hidden reference in 14% of 228 tested conditions). This observation could have some implication for the ADM format. In its current state the flag "diffuse" is a sub-element of an AudioChannel. It could be interesting to define this property at a higher level (AudioPack), i.e. where the diffuse property is considered globally for the whole AudioPack (reverberation channel bed, reverberation HOA bed) and not individually for each Audiochannel. Note that the choice of using a 3rd-order HOA stream to convey the output of the sixteen decorrelation filters is probably not crucial in the relatively good performance of the method. It was just a convenient way to distribute a sixteen-channel diffuse stream to a larger number of loudspeakers.

The perceptual ranking derived from the test confirms the relevance of the objective characterisation described in section 4.2.3.2. Although the test does not provide evidence on the respective role of the colouration and the spatial properties of the filters, it shows that both aspects are important. The method M4, which provides interaural correlation performances very close to that of the reference, was only judged "good" (score > 60), probably because its behaviour in the spectral domain was not satisfying. Only method M8, which could reach simultaneously good objective performances on the spectral and the spatial indices, was judged excellent. The ability of these indices to estimate the performance can be used to tune the parameters of the different methods under estimation or to optimise them.

5 Conclusions

In this deliverable we have reported about the different user tests and perceptual experiments that occurred over the course of the Orpheus project. These tests focused on different aspects of the project.

First, two separate studies investigated the quality of user experience offered by the Orpheus iOS app, which implements various features made possible by object-based audio and allows to experience original content produced during this project. Overall, a majority of users had no difficulty using the app and found the demonstrated features useful, which is a very positive result for this project. These studies also provided hints on how to improve the Orpheus iOS app and, more generally, how to approach object-based audio with non-expert users.

The final two studies focused on more general matters related to object-based audio. The third study examined the MS-IPM listening test methodology, which can for instance be used to compare different audio rendering techniques for standardisation purposes. The results of this study suggest that MS-IPM provides a reliable method for comparing audio stimuli in the absence of a reference.

The fourth study investigated the issue of transmitting and rendering reverberation in an object-based audio context. More specifically it compared various strategies used for rendering a diffuse reverberation field from a limited number of transmitted reverberation channels. Different objective criteria were proposed to quantify and improve the perceptual performances of the decorrelation methods. The study also suggests some modification of the ADM format in order to guide the rendering of the "diffuse" flag attribute.

References

[1] Schade, J., and Schlag, B. “Acceptability of urban transport pricing strategies”. Transportation Research Part F: Traffic Psychology and Behaviour, 6(1), 45–61, 2003

[2] Schuitema, G., Steg, L., and Forward, S. “Explaining differences in acceptability before and acceptance after the implementation of a congestion charge in Stockholm”. Transportation Research Part A: Policy and Practice, 44(2), 99–109, 2010.

[3] Venkatesh, V., Morris, M., Davis, G., and Davis, F. D. “User Acceptance of Information Technology: Toward a Unified View”. Management Information Systems Quarterly, 27(3), 2003.

[4] Dwivedi, Y. K., Rana, N. P., Chen, H., and Williams, M. D. “A Meta-analysis of the Unified Theory of Acceptance and Use of Technology (UTAUT)”. In M. Nüttgens, A. Gadatsch, K. Kautz, I. Schirmer, & N. Blinn (Eds.), Governance and Sustainability in Information Systems. Managing the Transfer and Diffusion of IT (pp. 155–170). Springer Berlin Heidelberg, 2011.

[5] Khechine, H., Lakhal, S., and Ndjambou, P. “A meta-analysis of the UTAUT model: Eleven years later”. Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences de l’Administration, 33(2), 138–152, 2016.

[6] King, W. R., and He, J. “A meta-analysis of the technology acceptance model”. Information & Management, 43(6), 740–755, 2006.

[7] Ma, Q., and Liu, L. “The Technology Acceptance Model: A Meta-Analysis of Empirical Findings”. Journal of Organizational and End User Computing, 16(1), 59–72, 2004

[8] Schepers, J., & Wetzels, M. “A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects”. Information & Management, 44(1), 90–103, 2007.

[9] Hassenzahl, M. “The Thing and I: Understanding the Relationship Between User and Product”. In M. A. Blythe, K. Overbeeke, A. F. Monk, and P. C. Wright (Eds.), Funology (Vol. 3, pp. 31–42). Dordrecht: Kluwer Academic Publishers, 2005.

[10] Karahanna, E., Straub, D. W., and Chervany, N. L. “Information Technology Adoption Across Time: A Cross-sectional Comparison of Pre-adoption and Post-adoption Beliefs”. MIS Q., 23(2), 183–213, 1999.

[11] Šumak, B., and Šorgo, A. “The acceptance and use of interactive whiteboards among teachers: Differences in UTAUT determinants between pre- and post-adopters”. Computers in Human Behavior, 64, 602–620, 2016.

[12] Yang, S., Lu, Y., Gupta, S., Cao, Y., and Zhang, R. “Mobile payment services adoption across time: An empirical study of the effects of behavioral beliefs, social influences, and personal traits”. Computers in Human Behavior, 28(1), 129–142, 2012.

[13] ITU Radiocommunication Sector Recommendation BS.1534-3 - Method for the subjective assessment of intermediate quality level of audio systems (2015).

[14] ITU Radiocommunication Sector Recommendation ITU-R BS.1116-3 - Methods for the subjective assessment of small impairments in audio systems (2015).

[15] Zacharov, N., Pike, C., Melchior, F., and T. Worch. “Next Generation Audio System Assessment using the Multiple Stimulus Ideal Profile Method”. Proceedings of QoMEX 2016, Lisbon, Portugal, 2016

[16] Pedersen, T. H., and Zacharov, N. “The Development of a Sound Wheel for Reproduced Sound.” In 138th Convention of the Audio Engineering Society, 2015

[17] Coleman, P., Franck, A., Jackson, P., Hughes, R., Remaggi, L., and Melchior, F., “On Object-Based Audio with Reverberation”, *in Proc. 60th AES Int. Conf.*, 2016.

[18] Coleman, P., Franck, A., Jackson, P., Hughes, R., Remaggi, L., and Melchior, F., “Object-Based Reverberation for Spatial Audio”, J. Audio Eng. Soc., vol. 65(1/2), pp.66-77, 2017.

[19] M. Gardner, “Image fusion, broadening, and displacement in sound location,” *J. Acoustical Society of America*, vol. 46 (339), pp. 339–349, 1969.

[20] M. Schroeder, “Natural sounding artificial reverberation,” *J. Audio Eng. Soc.*, vol.10(3), pp. 219–223, 1962.

[21] G. Kendall “The Decorrelation of Audio Signals and its Impact on Spatial Imagery,” *Computer Music Journal*, vol. 19, pp. 71–87, 1995.

[22] M. Bouéri and C. Kyriakakis “Audio Signal Decorrelation Based on a Critical Band Approach,” *in Proc. of the 117th Convention of the Audio Eng. Soc*., Oct. 2004.

[23] R. Penniman, “A general-purpose decorrelation algorithm with transient fidelity,” *in Proceedings of the 137th Convention of the Audio Eng. Society*, 2014.

[24] E. Kermit-Canfield and J. Abel, Signal Decorrelation using perceptually informed all pass filters, *in Proceedings of the 19th Int. Conf. on Digital Audio Effects (DAFx)*, Sept. 2016

[25] J. Merimaa and V. Pulkki “Spatial Impulse Response Rendering I: Analysis and Synthesis,” *J. Audio Eng. Soc.*, vol. 53, pp. 1115–1127, Dec. 2005.

[26] V. Pulkki “Spatial Sound Reproduction with Directional Audio Coding,” *J. Audio Eng. Soc*., vol. 55, pp. 503–516, June 2007.

[27] D. Romblom, P. Depalle, C. Guastavino, and R. King “Diffuse Field Modeling Using Physically-Inspired Decorrelation Filters and B-Format Microphones: Part I Algorithm,” *J. Audio Eng. Soc.*, vol. 64, pp. 177–193, April 2016.

[28] F. Zotter, M. Frank, G. Marentakis, and A. Sontacchi, « Phantom source widening with deterministic frequency dependent time delays, » *in Proc. of the 14th Int. Conf. on Digital Audio Effects (DAFx*), 2011.

[29] G. Potard, I. Burnett, « Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays », 2004

[30] F. Baumgarte and C. Faller, « Binaural cue coding part i: Psychoacoustic fundamentals and design principles, » *IEEE Trans. on Speech and Audio Processing*, vol. 11(6), pp. 509-519, 2003.

[31] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 24, no. 4, pp. 320-327, Aug. 1976.

Appendix A Questionnaires for b<>com's QoX study

In this appendix we provide screenshots of the questionnaires presented to the subjects who participated to the b<>com's Quality of Experience test. These questionnaires are provided in French language.

A.1 Screenshots of the acceptability questionnaire

A.1.1 Description of the project

Le projet Orpheus

ORPHEUS est un projet de recherche européen. Son objectif est d'inventer une nouvelle manière de produire et de diffuser des contenus audio permettant de s'adapter aux différentes plateformes et conditions d'écoute, tout en essayant d'offrir à l'utilisateur une expérience plus immersive et interactive.

Le questionnaire suivant a pour but de recueillir certaines informations d'ordre général. Cela vous prendra entre 5 et 10 minutes pour le compléter.

ORPHEUS Object-Based Audio Experience

SUIVANT

Page 1 sur 17

A.1.2 Generation of a participant ID

Le projet Orpheus

*Obligatoire

Votre identifiant

Afin de respecter la confidentialité et l'anonymat des données, tout en conservant la possibilité de mettre en lien les réponses que vous apporterez aux différents questionnaires, nous vous proposons de créer un identifiant unique.

Celui-ci sera à rappeler aux différents questionnaires auxquels vous répondrez au sein de cette étude.

Pour créer cet identifiant, nous vous proposons de noter:
- les deux derniers chiffres de votre année de naissance
- suivis des 3 premières lettres de votre ville de naissance
- conclus par votre numéro de rue.

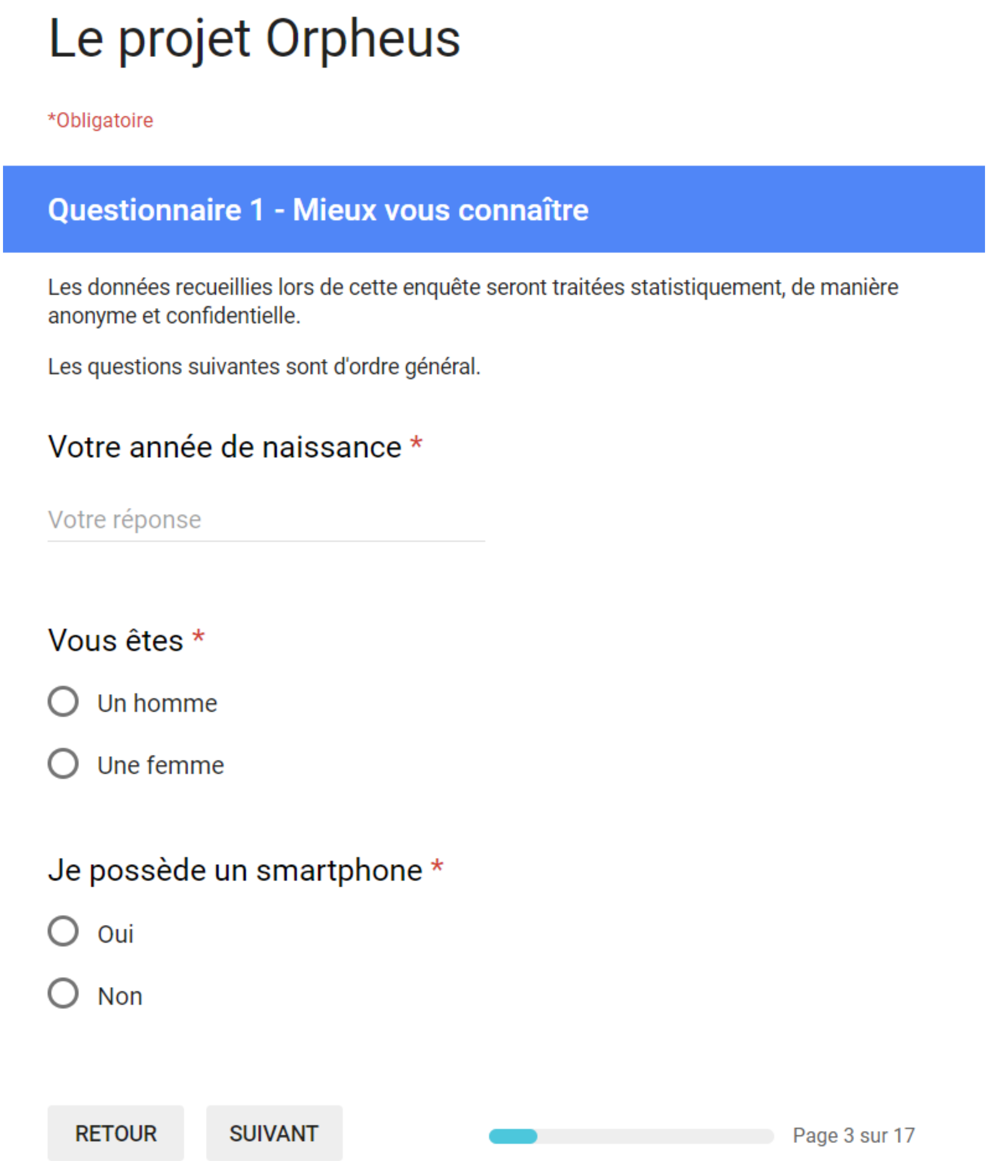
Exemple d'identifiant: 87REN2 (pour une personne née en 1987, à Rennes, et habitant le 2, rue Victor Hugo)

Votre identifiant *

Votre réponse

RETOUR SUIVANT Page 2 sur 17

A.1.3 General information questions



Le projet Orpheus

*Obligatoire

Questionnaire 1 - Mieux vous connaître

Les données recueillies lors de cette enquête seront traitées statistiquement, de manière anonyme et confidentielle.

Les questions suivantes sont d'ordre général.

Votre année de naissance *

Votre réponse

Vous êtes *

- Un homme
- Une femme

Je possède un smartphone *

- Oui
- Non

RETOUR SUIVANT Page 3 sur 17

A.1.4 Questions regarding smartphone use

Le projet Orpheus

*Obligatoire

Questionnaire 1 - Mieux vous connaître

Les questions suivantes portent sur votre smartphone

Quelles applications utilisez-vous sur votre smartphone pour écouter de la musique ? (Plusieurs réponses possibles)

- ☐ Deezer
- ☐ Spotify
- ☐ Soundcloud
- ☐ Google play music
- ☐ Apple music
- ☐ Bandcamp
- ☐ Youtube
- ☐ Autre : ____

Mon smartphone fonctionne sous *

- ○ Android (Samsung, Sony, etc.)
- ○ iOS (iPhone)
- ○ Autre : ____

RETOUR SUIVANT Page 4 sur 17

A.1.5 Questions regarding radio and music listening habits

Le projet Orpheus

*Obligatoire

Questionnaire 1 - Mieux vous connaître

Les questions qui suivent concernent vos habitudes en termes de musique et de radio.

Veuillez répondre le plus spontanément et le plus sincèrement possible aux questions suivantes. Il n'y a pas de bonnes ou mauvaises réponses.

Si vous écoutez la radio sur votre smartphone, quelles applications utilisez-vous ?

Votre réponse

J'écoute très souvent des podcasts d'émissions de radio via mon smartphone. *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

J'ai l'habitude d'écouter très fréquemment la radio à l'aide de mon smartphone *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

J'écoute très souvent de la musique via mon smartphone. *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 5 sur 17

A.1.6 Questions about the participants' appetence for new technologies

Le projet Orpheus

*Obligatoire

Questionnaire 1 - Mieux vous connaître

Les questions suivantes concernent votre rapport aux nouvelles technologies

En général, je n'hésite pas à essayer de nouvelles technologies. *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Si j'entends parler d'une nouvelle technologie, j'essaie de l'expérimenter rapidement. *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Dans mon entourage proche, je suis habituellement le (la) premier(ière) à explorer les nouvelles technologies. *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

J'aime découvrir et tester de nouvelles technologies. *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 6 sur 17

A.1.7 Description of the ORPHEUS iOS app

Le projet Orpheus

Description de l'application

Dans le cadre du projet Orpheus, une nouvelle application radio pour mobile est développée.

En plus des fonctionnalités habituelles d'une application mobile radio, celle-ci permettra :

- de naviguer au sein des différents éléments d'un programme, d'une émission, séparés en chapitres ;

- d'accéder à une retranscription textuelle, en temps réel ou a posteriori, de tous les échanges verbaux des interlocuteurs radiophoniques ;

- de modifier la clarté du rendu sonore (amplification de parties sonores peu audibles, ajustement du niveau sonore des éléments présents au premier plan et/ou à l'arrière-plan ; exemple: augmenter / diminuer le bruit du public, etc.) ;

- d'interagir directement avec le contenu audio (Exemple: choisir à quel endroit virtuel nous voulons nous placer durant un concert afin d'en modifier la perspective d'écoute) ;

- de modifier la langue du contenu écouté (français, anglais, etc.) ;

- de choisir le rendu audio (mono, stéréo, écoute des sons en 3D, etc.) ;

- d'adapter le contenu des podcasts/émissions en fonction de la durée d'écoute désirée (exemple: choisir d'écouter les 10 meilleures minutes d'une émission durant initialement 1H30).

RETOUR SUIVANT Page 7 sur 17

A.1.8 Questions about the app's overall acceptability

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Votre avis global sur l'application

L' application mobile vient de vous être décrite. Nous aimerions recueillir votre avis global à ce propos.

Veuillez répondre le plus spontanément et le plus sincèrement possible aux questions suivantes, en vous basant sur la description de l'application. Il n'y a pas de bonnes ou mauvaises réponses.

L'utilisation de cette technologie semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

La probabilité que j'utilise cette application est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette application pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

L' application Orpheus me semble novatrice *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 8 sur 17

A.1.9 Questions regarding intention to use

Le projet Orpheus

*Obligatoire

Évaluation de l'application

Si vous deviez évaluer globalement cette technologie, quelle note lui donneriez-vous ? *

Sélectionner

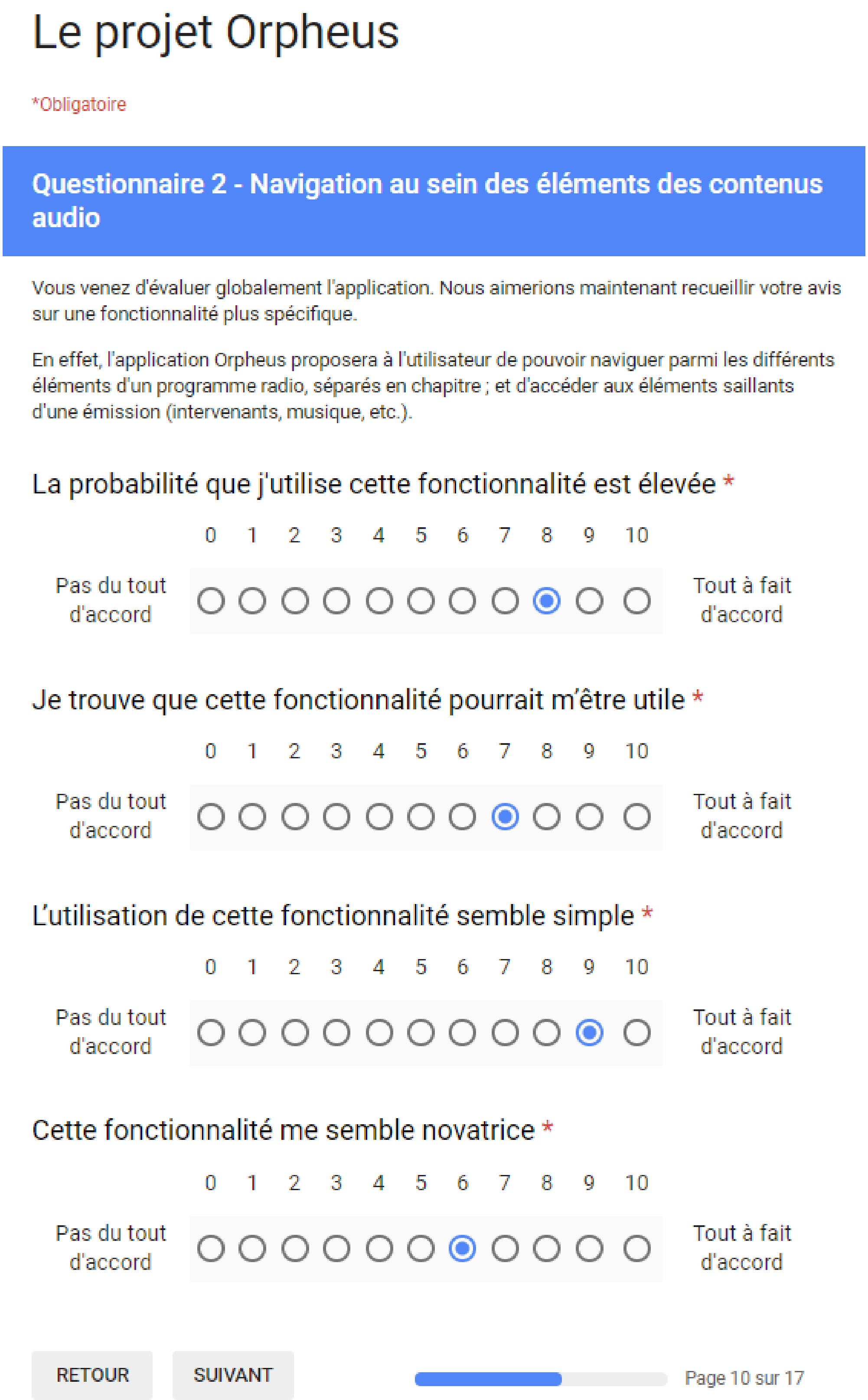
Je souhaiterais tester cette application *

- Oui
- Non
- Je ne sais pas

RETOUR SUIVANT

Page 9 sur 17

A.1.10 Questions regarding the acceptability of the chapter navigation feature



Le projet Orpheus

*Obligatoire

Questionnaire 2 - Navigation au sein des éléments des contenus audio

Vous venez d'évaluer globalement l'application. Nous aimerions maintenant recueillir votre avis sur une fonctionnalité plus spécifique.

En effet, l'application Orpheus proposera à l'utilisateur de pouvoir naviguer parmi les différents éléments d'un programme radio, séparés en chapitre ; et d'accéder aux éléments saillants d'une émission (intervenants, musique, etc.).

La probabilité que j'utilise cette fonctionnalité est élevée *

0 1 2 3 4 5 6 7 8 9 10

Pas du tout d'accord — Tout à fait d'accord

Je trouve que cette fonctionnalité pourrait m'être utile *

0 1 2 3 4 5 6 7 8 9 10

Pas du tout d'accord — Tout à fait d'accord

L'utilisation de cette fonctionnalité semble simple *

0 1 2 3 4 5 6 7 8 9 10

Pas du tout d'accord — Tout à fait d'accord

Cette fonctionnalité me semble novatrice *

0 1 2 3 4 5 6 7 8 9 10

Pas du tout d'accord — Tout à fait d'accord

RETOUR SUIVANT

Page 10 sur 17

A.1.11 Questions regarding the acceptability of the textual transcript feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Retranscription textuelle automatique du contenu audio

L' application Orpheus proposera également à l'utilisateur d'accéder à une retranscription textuelle, en temps réel ou a posteriori, de tous les échanges verbaux des interlocuteurs radiophoniques.

L' utilisation de la retranscription semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

La probabilité que j'utilise cette fonctionnalité est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité me semble novatrice *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 11 sur 17

A.1.12 Questions regarding the acceptability of the "audio clarity" feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Modification de la clarté du rendu sonore

De même, l'application Orpheus permettra à l'utilisateur de modifier lui-même la clarté du rendu sonore. Vous pourrez ainsi amplifier les parties sonores peu audibles si cela est nécessaire, ajuster le niveau sonore des éléments présents au premier plan et/ou à l'arrière-plan (exemple: augmenter / diminuer le bruit des supporters lors d'une rencontre sportive, augmenter les dialogues lors d'une pièce de théâtre, etc.).

L' utilisation de cette fonctionnalité semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que modifier la clarté du rendu sonore pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

La probabilité que j'utilise cette fonctionnalité est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

La fonctionnalité de modification de la clarté du rendu sonore me semble novatrice *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 12 sur 17

A.1.13 Questions regarding the acceptability of the "interaction" feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Interaction(s) avec le contenu audio

L' application Orpheus permettra aussi à l'utilisateur d'interagir directement avec le contenu audio. Ainsi, lors d'un concert retransmis à la radio, par exemple, vous pourrez choisir où vous positionner virtuellement, afin de changer de perspective d'écoute.

La probabilité que j'utilise cette fonctionnalité est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

L' utilisation de cette fonctionnalité semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve qu'interagir de cette manière avec le contenu audio pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité me semble novatrice *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 13 sur 17

A.1.14 Questions regarding the acceptability of the multilanguage feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Modification de la langue

L' application Orpheus prévoit de laisser à l'utilisateur la possibilité de choisir et modifier la langue (français, anglais, allemand, etc.) des contenus audio écoutés (émissions de radio, rencontres sportives, pièces de théâtre, etc.).

La probabilité que je modifie la langue choisie d'un programme radio est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

L' utilisation de cette fonctionnalité semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Pouvoir modifier la langue des programmes radio me semble novateur *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 14 sur 17

A.1.15 Questions regarding the acceptability of the audio preset feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Définition de pré-réglages audio

En parallèle, l'application proposera à l'utilisateur de définir des pré-réglages audio, activés automatiquement en fonction des conditions d'écoute de l'utilisateur: types d'activité (bureau, train, sport, etc.), types de lieux (maison, bureau, etc.), types de connexion internet, de moyens d'écoute (écouteurs, casques, enceintes, etc.).

Je trouve que cette fonctionnalité pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

L' utilisation de cette fonctionnalité semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

La probabilité que j'utilise cette fonctionnalité est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité, est, à mes yeux, innovantes *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 15 sur 17

A.1.16 Questions regarding the acceptability of the “audio rendering” feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Choix du rendu audio

Pour chaque contenu, l'application proposera à l'utilisateur plusieurs formats audio qu'il pourra modifier en fonction du contexte. Ces formats sont de plusieurs types : mono, stéréo ou encore binaural (format audio qui donne l'impression d'entendre les sons en 3D).

L' utilisation de cette fonctionnalité semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

La probabilité que j'utilise cette fonctionnalité est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 16 sur 17

A.1.17 Questions regarding the acceptability the variable-length content feature

Le projet Orpheus

*Obligatoire

Questionnaire 2 - Adaptation de la durée du contenu

Nous aimerions recueillir votre avis sur une dernière fonctionnalité.

En effet, l'application proposera d'adapter le contenu des podcasts/émissions en fonction de la durée d'écoute désirée.
Par exemple, si vous ne disposez que de 10 minutes, vous pourrez choisir d'écouter les 10 meilleures minutes d'une émission radio d'une heure et demie.

La probabilité que j'utilise cette fonctionnalité est élevée *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité pourrait m'être utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

L' utilisation de cette fonctionnalité semble simple *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité me semble novatrice *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR ENVOYER Page 17 sur 17

A.2 Screenshots of the user task descriptions

A.2.1 Reminder of the procedure for creating an identifier

Test et évaluation de l'application Orpheus

Vous allez pouvoir tester une nouvelle application, créée dans le cadre du projet Orpheus.

Afin de respecter la confidentialité et l'anonymat des données recueillies lors de ce test, tout en conservant la possibilité de mettre en lien les réponses que vous apporterez aux différents questionnaires, nous vous avions proposé de créer un identifiant unique, lors du premier questionnaire.

Pour rappel, votre identifiant, déjà créé, débute:
- par les deux derniers chiffres de votre année de naissance
- suivis des 3 premières lettres de votre ville de naissance
- conclus par votre numéro de rue.

Exemple d'identifiant: 87REN2 (pour une personne née en 1987, à Rennes, et habitant le 2, rue Victor Hugo)

*Obligatoire

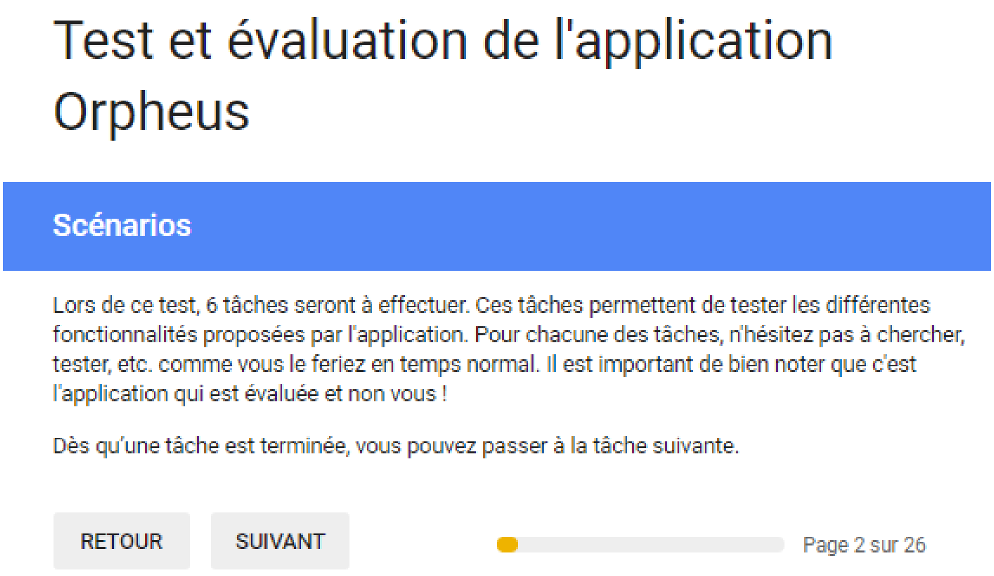
Votre identifiant *

Votre réponse

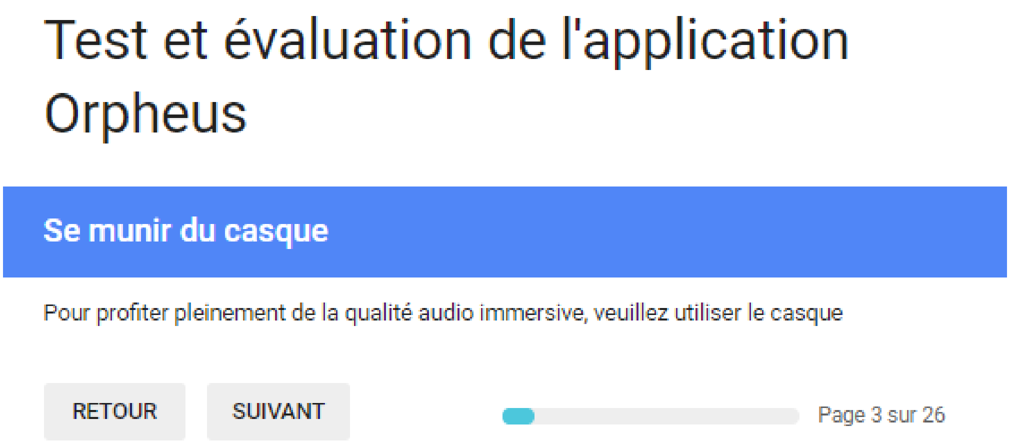
SUIVANT

Page 1 sur 26

A.2.2 Introduction



A.2.3 Instructions related to headphones



A.2.4 Description of the task related to the chapter navigation feature

Test et évaluation de l'application Orpheus

Tâche 1/ Naviguer parmi les éléments des contenus audio

L'application Orpheus propose à l'utilisateur de pouvoir naviguer parmi les différents éléments d'un programme radio, séparés en chapitre ; et d'accéder aux éléments saillants d'une émission (intervenants, musique, etc.).

Pour tester cette fonctionnalité, sélectionnez le programme intitulé "Expérience Audio orientée objet".

Après avoir manipulé à plusieurs reprises cette fonctionnalité, cliquez sur "suivant" (via ce formulaire) afin d'accéder à la tâche suivante.

A.2.5 Description of the task related to the textual transcript feature

Test et évaluation de l'application Orpheus

Tâche 2/ Accéder à la retranscription textuelle

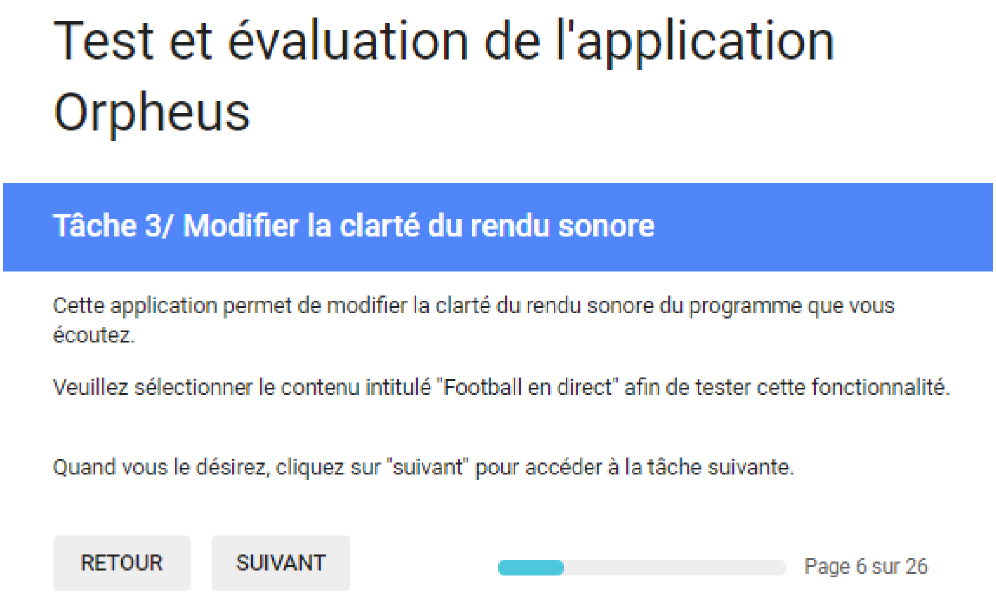
Via l'application, vous pouvez accéder en direct au contenu écrit d'un programme radio.

Pour tester cette fonctionnalité, sélectionnez le programme intitulé "l'art du bruitage" puis lisez/écoutez quelques échanges radiophoniques.

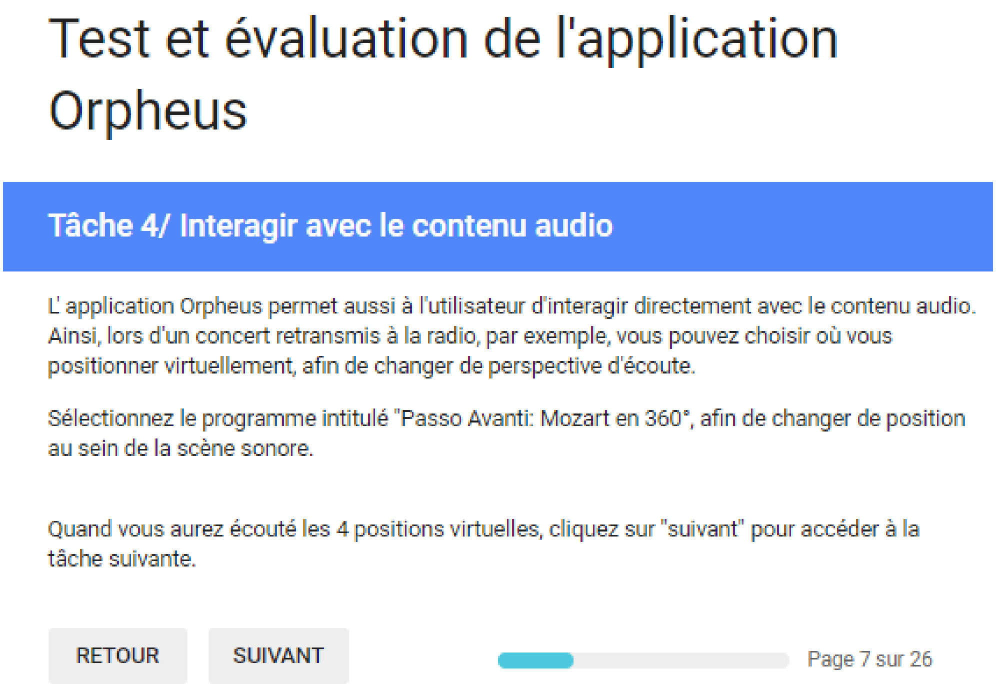
Quand vous le voulez,, cliquez ensuite sur "suivant" (via ce formulaire) pour accéder à la tâche suivante.

RETOUR SUIVANT Page 5 sur 26

A.2.6 Description of the task related to the clarity feature



A.2.7 Description of the task related to the “audio interaction” feature



A.2.8 Description of the task related to audio preset feature

Test et évaluation de l'application Orpheus

Tâche 6/ Définir les pré-réglages audio

L' application propose à l'utilisateur de définir des pré-réglages audio, activés automatiquement en fonction des conditions d'écoute de l'utilisateur: types d'activité (bureau, train, sport, etc.), types de lieux (maison, bureau, etc.), types de connexion internet, de moyens d'écoute (écouteurs, casques, enceintes, etc.).

Pour tester cette fonctionnalité, nous vous proposons d'éditer un profil avec les caractéristiques suivantes:
Nom : bcom
Connexion: Wifi de mon travail
Sortie: Casque
Activité: Stationnaire
Format audio: Binaural

Quand vous pensez avoir bien pris connaissance des pré-réglages possibles, veuillez cliquer sur "suivant".

RETOUR SUIVANT Page 9 sur 26

A.2.9 Description of the task related to the multilanguage feature

Test et évaluation de l'application Orpheus

Tâche 5/ Modifier la langue

L' application Orpheus prévoit une option multilingue. Ainsi, l'utilisateur a la possibilité de choisir et modifier la langue (français, anglais, allemand, etc.) des contenus audio écoutés (émissions de radio, rencontres sportives, pièces de théâtre, etc.).

Pour tester cela, sélectionnez le programme intitulé "l'art du bruitage" afin de modifier la langue d'écoute par l'anglais.

Cliquez sur "suivant" pour accéder à la tâche suivante.

RETOUR SUIVANT Page 8 sur 26

A.3 Screenshots of the acceptance questionnaire

A.3.1 Overall acceptance questions

Test et évaluation de l'application Orpheus

*Obligatoire

Votre avis général sur l'application

Vous venez de tester l'application. Merci pour cela. Nous aimerions maintenant recueillir votre avis global.

Veuillez répondre le plus spontanément et le plus sincèrement possible aux questions suivantes. Il n'y a pas de bonnes ou de mauvaises réponses.

J'aime l'apparence visuelle de cette technologie *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

J'estime que cette technologie est utile pour moi *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Se servir de cette application est facile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Les fonctionnalités proposées par cette technologie sont innovantes *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 10 sur 26

A.3.2 Instructions on how to leave comments

Test et évaluation de l'application Orpheus

Votre avis général sur l'application

Vous pouvez ajouter des commentaires libres sur votre ressenti global par rapport à l'application.

Commentaires

Votre réponse

RETOUR SUIVANT

Page 11 sur 26

A.3.3 Questions regarding intention to use

Test et évaluation de l'application Orpheus

*Obligatoire

Évaluation de l'application

Si vous en avez la possibilité, pensez-vous continuer à utiliser cette application ? *

- ○ Oui
- ○ Non

Si vous deviez évaluer globalement cette application, quelle note lui donneriez-vous ? *

Sélectionner ▾

J'ai l'intention de continuer à utiliser cette application *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 12 sur 26

A.3.4 Questions regarding the acceptance of the chapter navigation feature

Test et évaluation de l'application Orpheus

*Obligatoire

Navigation au sein des éléments des contenus audio

Vous venez d'évaluer globalement l'application. Nous aimerions maintenant recueillir votre avis sur une fonctionnalité plus spécifique, que vous avez pu tester.

En effet, l'application Orpheus propose à l'utilisateur de pouvoir naviguer parmi les différents éléments d'un programme radio, séparés en chapitre ; et d'accéder aux éléments saillants d'une émission (intervenants, musique, etc.).

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Se servir de cette fonctionnalité est facile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité est utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 13 sur 26

A.3.5 Questions regarding the acceptance of the textual transcript feature

Test et évaluation de l'application Orpheus

*Obligatoire

Retranscription textuelle automatique du contenu audio

Comme vous avez pu le voir, l'application Orpheus propose également à l'utilisateur d'accéder à une retranscription textuelle, en temps réel ou a posteriori, de tous les échanges verbaux des interlocuteurs radiophoniques.

Utiliser cette fonctionnalité est facile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est utile pour moi *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve cette fonctionnalité innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 15 sur 26

A.3.6 Questions regarding the acceptance of the sound clarity feature

Test et évaluation de l'application Orpheus

*Obligatoire

Modification de la clarté du rendu sonore

L' application Orpheus permet également à l'utilisateur de modifier lui-même la clarté du rendu sonore. Vous pouvez ainsi amplifier les parties sonores peu audibles si cela est nécessaire, ajuster le niveau sonore des éléments présents au premier plan et/ou à l'arrière-plan (exemple: augmenter / diminuer le bruit des supporters lors d'une rencontre sportive, augmenter les dialogues lors d'une pièce de théâtre, etc.).

Cette fonctionnalité est facile à utiliser *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité est utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT Page 17 sur 26

A.3.7 Questions regarding the acceptance of the "interaction" feature

Test et évaluation de l'application Orpheus

*Obligatoire

Interaction avec le contenu audio

L'application Orpheus permet aussi à l'utilisateur d'interagir directement avec le contenu audio. Ainsi, lors d'un concert retransmis à la radio, par exemple, vous pouvez choisir où vous positionner virtuellement, afin de changer de perspective d'écoute.

Cette fonctionnalité est utile pour moi *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est facile à utiliser *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité est innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 19 sur 26

A.3.8 Questions regarding the acceptance of the multilanguage feature

Test et évaluation de l'application Orpheus

*Obligatoire

Modification de la langue

L'application Orpheus prévoit de laisser à l'utilisateur la possibilité de choisir et modifier la langue (français, anglais, allemand, etc.) des contenus audio écoutés (émissions de radio, rencontres sportives, pièces de théâtre, etc.).

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est facile à utiliser *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Je trouve que cette fonctionnalité m'est utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 21 sur 26

A.3.9 Questions regarding the acceptance of the audio presets feature

Test et évaluation de l'application Orpheus

*Obligatoire

Définition de pré-réglages audio

L' application propose à l'utilisateur de définir des pré-réglages audio, activés automatiquement en fonction des conditions d'écoute de l'utilisateur: types d'activité (bureau, train, sport, etc.), types de lieux (maison, bureau, etc.), types de connexion internet, de moyens d'écoute (écouteurs, casques, enceintes, etc.).

Se servir de cette fonctionnalité est facile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord												Tout à fait d'accord

Je trouve que cette fonctionnalité est utile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord												Tout à fait d'accord

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord												Tout à fait d'accord

Cette fonctionnalité est innovante *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord												Tout à fait d'accord

RETOUR SUIVANT

Page 23 sur 26

A.3.10 Questions regarding the acceptance of the "audio rendering" feature

Test et évaluation de l'application Orpheus

*Obligatoire

Choix du rendu audio

Pour chaque contenu, l'application propose plusieurs formats audio que vous pouvez modifier en fonction du contexte (bureau, train, sport, etc.). Ces formats sont de plusieurs types : mono, stéréo ou encore binaural (format audio qui donne l'impression d'entendre les sons en 3D).

Si cela est possible, j'ai l'intention de continuer à utiliser cette fonctionnalité *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Se servir de cette fonctionnalité est facile *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est utile pour moi *

	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

Cette fonctionnalité est innovante *

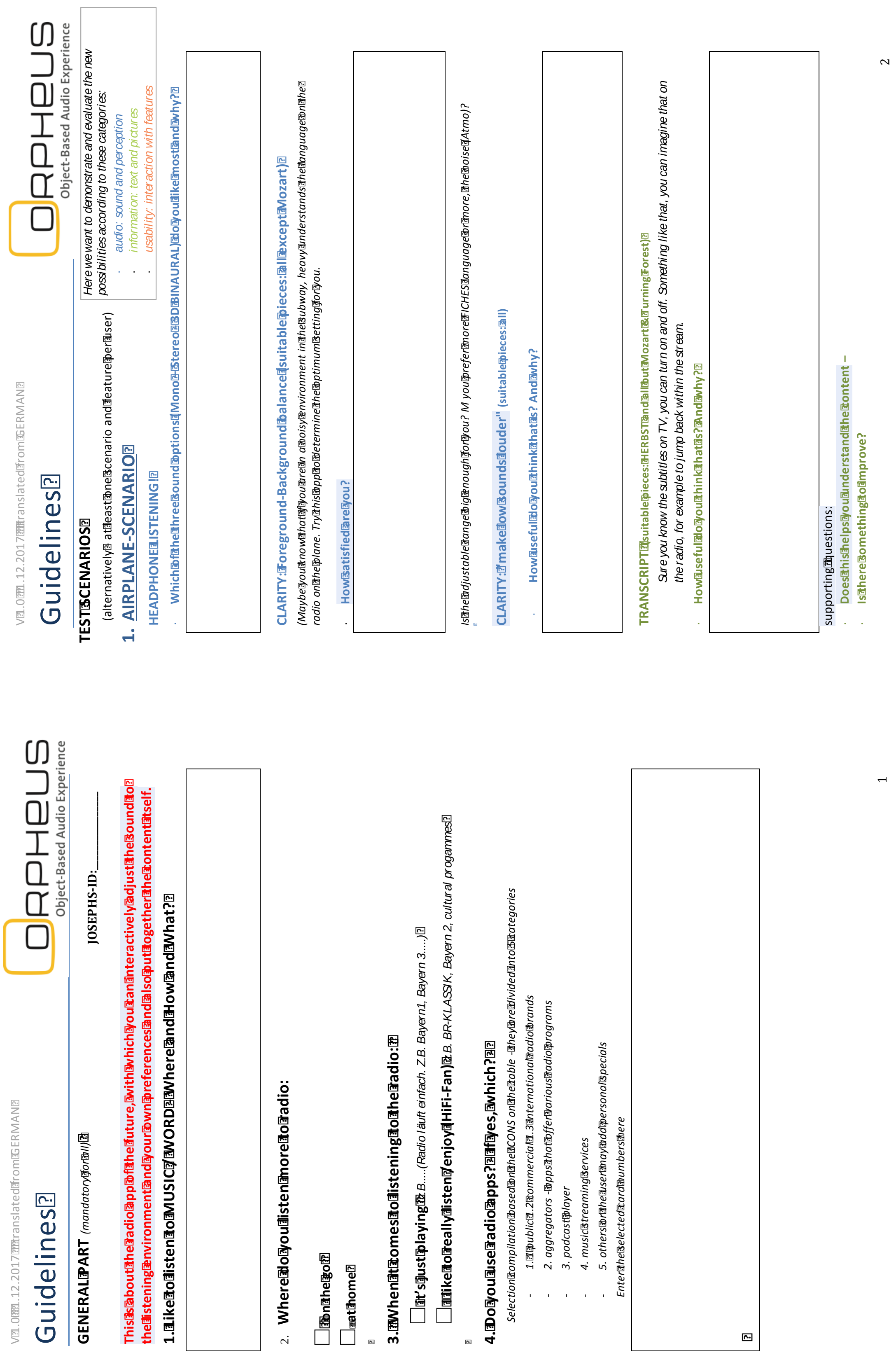
	0	1	2	3	4	5	6	7	8	9	10	
Pas du tout d'accord	○	○	○	○	○	○	○	○	○	○	○	Tout à fait d'accord

RETOUR SUIVANT

Page 25 sur 26

Appendix B Guidelines developed for JOSEPHS QoX tests

These guidelines - here translated from German – were used by the accompanying guides at JOSEPHS®. Mind: the colours refers to the respective dimension or category of investigation.



V 1.0 1.12.2017 translated from GERMAN

Guidelines

GENERAL PART *(mandatory for all)*

JOSEPHS-ID:____________

This is about the radio app of the future, with which you can interactively adjust the sound to the listening environment and your own preferences and also put together the content itself.

1. Like to listen to MUSIC / WORD Where and How and What?

2. Where do you listen more to radio:

- on the go
- at home

3. When it comes to listening to the radio:

- it's just playing *B.....(Radio läuft einfach. Z.B. Bayern1, Bayern 3.....)*
- I like to really listen / enjoy (HiFi-Fan) *z.B. BR-KLASSIK, Bayern 2, cultural progammes*

4. Do you use radio apps? If yes, which?

Selection compilation based on the ICONS on the table - they are divided into 5 categories

- *1. public 1.2 commercial 1.3 international radio brands*
- *2. aggregators - apps that offer various radio programs*
- *3. podcast player*
- *4. music streaming services*
- *5. others or the user may add personal specials*

Enter the selected card numbers here

1

V 1.0 1.12.2017 translated from GERMAN

Guidelines

TEST SCENARIOS

(alternatively at least one scenario and feature per user)

Here we want to demonstrate and evaluate the new possibilities according to these categories:

- *audio: sound and perception*
- *information: text and pictures*
- *usability: interaction with features*

1. AIRPLANE-SCENARIO

HEADPHONE LISTENING!

- Which of the three sound options (Mono - Stereo 3D BINAURAL) do you like most and why?

CLARITY: Foreground-Background balance (suitable pieces: all except Mozart)

(Maybe you know that if you are in a noisy environment in the subway, heavy understands the language on the radio on the plane. Try this app to determine the optimum setting for you.

- How satisfied are you?

Is the adjustable range big enough for you? M you prefer more FICHES language or more, the noise (Atmo)?

CLARITY: "make low sounds louder" (suitable pieces: all)

- How useful do you think that is? And why?

TRANSCRIPT (suitable pieces: HERBST and all but Mozart & Turning Forest)

Sure you know the subtitles on TV, you can turn on and off. Something like that, you can imagine that on the radio, for example to jump back within the stream.

- How useful do you think that is? And why?

supporting questions:

- Does this helps you understand the content –
- Is there something to improve?

2

V 1.0 1.12.2017 translated from GERMAN

Guidelines

- Where you can imagine the transcript application?

LIVING ROOM SCENARIO

SURROUND 5.1 - loudspeaker playback (suitable pieces: all)

- How do you like it and why?

supporting questions:

> Do you know SURROUND SOUND? *Movies? TV? Radio?*

e.g. BR-KLASSIK: broadcasts weekly in 5.1 SURROUND on digital satellite or digital cable

TRANSCRIPT (suitable pieces: HERBST and all but Mozart & Turning Forest)

Sure you know the subtitles on TV, you can turn on and off. Something like that, you can imagine that on the radio, for example to jump back within the stream.

- How useful do you think that is? And why?

supporting questions:

- Does this helps you to understand the content
- Is there something to improve?
- Where you can imagine the transcript application?

3

V 1.0 1.12.2017 translated from GERMAN

Guidelines

FOR BOTH SCENARIOS (mandatory for all)

USABILITY

How did you cope with the app?

supporting questions:

- How do you like the pictures to the radio?
- Do you like to see the reporter / speaker /?
- How do you find the navigation in the piece ("dial wheel")
- Do the chapter marks help you with orientation?
- Is there something to improve?

IN SUMMARY

Would you use this app? Why?

Do you like the additional adjustment options?

Which of the new features did you like the most? Why?

Surround 3D headphone & Sound - transcript - intelligibility - loudness adjustment - Navigation in the program (chapter markers)

Thank you!

4

[end of document]