

Comparing individual perception of timbre and reverberance

Markus VON BERG^{1,2}; Lukas PRINZ³; Jochen STEFFENS^{1,2}

¹ Institute of Sound and Vibration Engineering (ISAVE), Hochschule Düsseldorf, Germany

² Audio Communication Group, Technische Universität Berlin, Germany

³ Faculty of Electrical Engineering and Information Technology, Hochschule Düsseldorf, Germany

ABSTRACT

Room reverberation alters the spatial impression and timbre of a sound by modulating its spectral and temporal characteristics. Thus, we argue that, on a perceptual level, reverberation basically breaks down into interaural differences and spectro-temporal cues and that the separation of a perceived timbre into a sound source and a surrounding room is a purely cognitive process. To investigate the connection between the perception of reverberation cues and timbre analysis, the sensitivity for changes in reverberation was compared to timbre perception abilities. The Timbre Perception Test was used to measure the perception of the temporal envelope, spectral centroid, and spectral flux of artificial sounds. Sensitivity for changes in reverberation time was tested with a discrimination task using speech and noise with speech-alike spectral and temporal envelopes as source signals. Musical and acoustical expertise was assessed through the Goldsmiths Musical Sophistication Index and self-reports on experience with and knowledge of acoustics. There was a considerable correlation between timbre and reverberance perception ability, but timbre perception and academic experience predicted only 41% of the variance in reverberance perception. Still, perception abilities related to similar acoustical phenomena seem to be better indicators of listening skills than self-reports on acoustical or musical expertise.

Keywords: Timbre, Reverberation, Sensitivity

1. INTRODUCTION

Timbre is probably one of the most defining features of a sound, next to loudness and pitch. Although the term timbre is somewhat vague and comprises several auditory attributes, one attribute that is usually excluded is the perceived reverberation (McAdams & Giordano, 2015). The term reverberation primarily refers to the physical superposition of a sound with room reflections that create a perceptual impression of room acoustics that comprises several factors such as the so-called reverberance, but also loudness amplification, brilliance, etc. (Weinzierl et al., 2018). Looking at the acoustical cues on which the perception of both timbre and reverberance rely, the distinction between the two realms cannot be clearly made. Previous research has identified the spectral centroid, temporal amplitude envelope, and spectral changes over time as essential constituents of timbre (McAdams & Giordano, 2015). Room reverberation, however, alters all of these properties by superposing reflections that frequency-dependently amplify both the amplitude and sustain of the sound.

A major difference between reverberance and timbre is that the perception of reverberance is not only based on spectral and temporal features of a sound, but also on its spatial representation. However, strictly speaking, the only measurable acoustical cues unique to reverberation (as opposed to timbre characteristics) are interaural differences. Non-interaural localization cues, as determined by the head-related-transfer function (HRTF) are spectral changes that the auditory system associates with certain incident directions and would otherwise probably be interpreted as timbre changes. This means that any separation of reverberance from timbre based on non-interaural cues – for example when reverberance is recognized in monophonic recordings – would be a purely cognitive process that relies on previously acquired mental representations of both the sound source's timbre and reverberance.

Thus, we hypothesized that the ability to detect changes in the acoustical characteristics relevant for either of the two realms, would be at least to some degree determined by the same sensitivity for

¹markusmartin.vonberg@hs-duesseldorf.de

spectro-temporal cues. Furthermore, we assumed that it would be more difficult to attribute the perceived acoustical changes to either the room's reverberance or the source's timbre in cases of unfamiliar sound sources with unknown timbre properties compared to familiar sources?.

2. METHODS

The individual perception abilities for timbre and reverberance were compared in a laboratory study with 30 participants (14 female, 16 male) between 19 and 60 years old (M : 33.1 years, SD : 11.7 years). Because distinguishing reverberance from source timbre relies on the comparison of a sound with mental representations of both realms, the impacts on familiarity with the source on the one hand and general acoustical and musical expertise, on the other hand, were also tested. The experiment took place in an acoustically optimized room, and the stimuli were presented via electrostatic headphones.

Timbre perception was assessed by means of the Timbre Perception Test (TPT) by Lee and Müllensiefen (2020). The test measures the ability to adjust the spectral centroid, temporal envelope, and temporal flux to match a given target sound. Participants were presented a target sound as well as a second sound for which they could change one of the three mentioned acoustical characteristics with a slider. The task was to use the slider to adjust the second sound to sound exactly like the target.

To test the individual reverberance perception, a discrimination task for changes in reverberation time that has been previously used by von Berg et al. (2021) was adapted to follow the same paradigm as the TPT. Again, two sounds and a slider were presented. Both sounds had the same source signal, but one had a static (target) reverberation time and the other one's reverberation time could be adjusted by moving the slider. This was realized by rendering 26 binaural room impulse responses with increasing reverberation times in RAVEN (Schröder & Vorländer, 2011). At each slider position, a different room response was applied to the source signal so that the reverberation time increased along the slider.

This does not necessarily mean that there was an equal increase in the perceived reverberance, since reverberation time is not an ideal predictor for reverberance (Lee et al., 2010). However, also the more favored early decay time (Soulodre & Bradley, 1995) and strength, clarity and IACC according to ISO 3382-1 (2009) changed monotonously in the simulated room responses. Thus, it was assumed that all perceptual qualities that were affected by the manipulation of reverberation time also changed monotonously along the slider range, even though there was now previous perceptual evaluation of the room responses.

Six test items were created by selecting six target reverberation times with randomized ranges of reverberation covered by the slider. Two source signals were used for each of the six items to test whether familiarity with the sound source would make it easier to distinguish the added reverberation, leading to a total of 12 test items. The familiar source signal was an anechoic recording of an English-speaking female voice, taken from the "Music for Archimedes" recordings by Hansen and Munch (1991). For the unfamiliar source signal, modulated noise was utilized to obtain a similar frequency spectrum and temporal envelope as the voice recording, and thus a comparable excitement of the virtual rooms by both signals. After the slider task, participants were asked to rate how much timbre and reverberance were changed by the slider on a five-point scale from "not at all" to "very much".

Musical expertise and activities were measured by the three subscales "active engagement", "perceptual abilities" and "musical training" of the Gold-MSI (Müllensiefen et al., 2014), a comprehensive self-report inventory that assesses *musical sophistication*, expressed in various musical activities ranging from everyday listening behavior to formal training on an instrument. The three subscales comprised a total of 25 items. To inquire previous experience with acoustics, participants were asked to rate their involvement with acoustics as part of their academic curriculum, musical performances, professional or semi-professional audio recording and clinical or audiological activities. All questions on previous experience were answered on a five-point scale from "never" to "very often".

3. RESULTS

The TPT results were calculated by applying a factor analysis to the bin-scored responses of the single items to retrieve three subscores for each of the parameters spectral centroid, temporal envelope and spectral flux, as well as a general timbre perception score. All four scores were transformed to a score between 0 and 100 (see Lee & Müllensiefen, 2020 for a detailed explanation). A similar

procedure was applied to the results of the reverberance perception test. The responses (i.e., the adjusted slider positions) were converted into reverberation time differences between the selected sound and the target (with a value of zero representing an exact match). The distances were transformed into bins of approximately even size, except for the first bin containing all zero responses (exact matches) that occurred more often than any other response value. Bin scoring yields the possibility to group the responses based on a perceptual scale derived from the given response behavior.

A linear mixed-effects model was calculated to test if the source signal had an impact on the response behavior. Interestingly, the items' mean response error was slightly higher when voice was used as a signal, although the difference was not significant ($p = .157$). Furthermore, participants consistently rated the reverberance to change more along the slider ($M: 4.04, SD: 1.05$ on a five-point scale) than timbre ($M: 2.06, SD: 1.02$).

Prior to the factor analysis, all items with a mean sampling accuracy below the recommended threshold of .60 were excluded by iteratively removing the item with the lowest MSA until all items exhibited an MSA above or equal to .60. The final item selection comprised 8 items – four with noise and four with voice as a source signal – with a KMO of .71. Internal consistency was measured in terms of Cronbach's α , which was at .81, and McDonald's ω , which was at .82. A factor analysis with one factor was calculated representing the individual reverberance perception ability. The suitability of a single factor was confirmed by the Kaiser criterion and the Scree test.

Individual factor scores were correlated with the general TPT score and the individual measures of musical expertise and previous experience (see Table 1). Significant correlations of reverberance perception ability with the TPT score ($r = .61, p > .001$), previous experience with acoustics as part of the academic curriculum ($r = .42, p = .019$) and the “musical training” subscale of the Gold-MSI ($r = .36, p = .049$) were observed.

Finally, a stepwise regression model predicting reverberance perception ability was computed. Here, a model with the TPT score and previous academic experience as predictors yielded the best fit in terms of the sample-corrected Akaike information criterion (AICc). The model revealed a significant, positive effect of timbre perception ability on reverberance perception ability, while the prediction effect of academic experience was not statistically significant ($p = .061$). The adjusted R^2 was at .41.

Table 1 – Correlation matrix of the perception abilities and reported acoustical and musical expertise

	RPA	TPT	EA	EM	ER	EC	PA	AE	MT
Reverberance perception ability	1.00								
TPT-score	0.61*	1.00							
Experience: academic	0.43*	0.26	1.00						
Experience: musical	0.31	0.17	0.77*	1.00					
Experience: recording	0.18	0.06	0.84*	0.74*	1.00				
Experience: clinical	-0.02	-0.30	0.56*	0.52*	0.67*	1.00			
Perceptual abilities	0.33	0.35	0.48*	0.53*	0.47*	0.13	1.00		
Active engagement	-0.05	0.00	0.44*	0.59*	0.66*	0.46*	0.43*	1.00	
Musical training	0.36*	0.50*	0.74*	0.75*	0.61*	0.18	0.72*	0.36	1.00

4. DISCUSSION

In our study, we compared the perception abilities of two elementary auditory features, timbre and reverberance, with musical activities, musical training and experience with acoustics. The results revealed a considerable correlation between both perception abilities. More concretely, a regression model predicting reverberance perception ability by timbre perception ability and academic experience with room acoustics showed a moderate amount of explained variance. The results therefore support the hypothesis that the perception of timbre and reverberance partially relies on the same sensitivity for spectro-temporal cues. Yet, there is still a lot of unshared variance between the

two perception abilities, although results suggest that timbre perception ability is a more suitable predictor of reverberance perception ability than the applied self-reports on acoustical and musical expertise.

One explanation for the unshared variance might be the participants' internal conceptualization of how the slider changed the sounds in the different listening tests. In the reverberance perception test, the change of the slider corresponded to a natural, comprehensible change of an acoustical feature participants knew from their daily lives. The rating of reverberance as the dominant change between stimuli confirms that most participants clearly understood what acoustical feature they were adjusting. This might have biased stimulus comparison, because participants focused on the auditory features they associated with reverberance, possibly neglecting other features like, e. g., the stimuli's loudness, which also increased with reverberation time. By contrast, in the TPT, where more abstract adjustments were applied to artificial sounds, participants probably only had a diffuse idea of how the slider changed timbre. Thus, they would have stronger relied on the mere audibility of differences between the adjusted sound and the target. Conclusively, the TPT was supposedly more difficult and the assessed timbre perception ability depended more on the actual sensitivity for spectro-temporal changes than the measured reverberance perception ability that was also affected depended on individual conceptualization of reverberance.

This could also explain why the prediction of reverberance perception ability slightly improved by including academic experience in the model, while there was no significant correlation of academic experience with the TPT scores. Previous findings of von Berg et al. (2021) showed a similar prediction effect of knowledge on room acoustics on abilities estimated from a previous version of the reverberance perception test. Because participants were aware that they were adjusting the room reverberation, a better understanding of how reverberation alters sound was apparently beneficial. It should be noted, that in this sample, academic experience with acoustics and musical training were highly correlated, so that participants might have also benefitted from practical experience from playing an instrument in different acoustic surroundings.

Another reason for the observed differences in timbre and reverberation perception abilities might be the processing of and sensitivity for interaural differences. Previous research has identified individual differences and training effects for interaural sensitivities among different listeners (Spencer et al., 2016), which could have created additional variance in reverberance perception, but expectably would not affect timbre perception.

It was also hypothesized, that discrimination reverberation times would be more difficult with unfamiliar sound sources. In fact, the test seemed to be slightly easier, when noise was used as an unfamiliar source signal, although the difference was not statistically significant. Even though the noise was modulated to have the same overall spectrum and temporal envelope as the speech signal, the noise simultaneously excited the room responses at frequency bands that appeared one after another in the voice signal, for example when a vowel followed a consonant. Thus, in the short-time domain, the noise created a more broadband excitation of the room responses than the voice, possibly making the room responses easier to discriminate. Using artificial source signals that also keep the temporal order of different spectral components of speech (containing vowel- and consonant-alike sequences) might give more insight into the importance of source familiarity on reverberance perception.

Despite the unshared variance in the perception abilities for timbre and reverberance, the experiment showed that such a test performance can be better predicted from another listening ability than from self-reports on a-priori-defined expertise criteria. If the perceptual abilities of a test sample are of interest, it might be advisable to determine these based on other listening tests that measure abilities related to similar acoustical phenomena rather than relying on self-reports.

References

1. Hansen, V., & Munch, G. (1991). Making Recordings for Simulation Tests in the Archimedes Project. *The Journal of the Acoustical Society of America*, 39(10).
2. ISO 3382-1 (2009). *Measurement of Room Acoustic Parameters Part 1: Performance Rooms*. International Organization for Standardization. Geneva, Switzerland.

3. Lee, D., Densil, C., & Martens, W. L. (2010). Equal reverberance matching of running musical stimuli having various reverberation times and SPLs. In *20th International Congress on Acoustics*, Sydney, Australia.
4. Lee, H., & Müllensiefen, D. (2020). The Timbre Perception Test (TPT): A new interactive musical assessment tool to measure timbre perception ability. *Attention, Perception & Psychophysics*, *82*(7), 3658–3675. <https://doi.org/10.3758/s13414-020-02058-3>
5. McAdams, S., & Giordano, B. L. (2015). The Perception of Musical Timbre. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (2nd ed.). Oxford University Press.
6. Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PloS One*, *9*(2), e89642. <https://doi.org/10.1371/journal.pone.0089642>
7. Schröder, D., & Vorländer, M. (2011). Raven: A Real-Time Framework for the Auralization of Interactive Virtual Environments. In *Proceedings of Forum Acusticum 2011: 27 June - 01 July, Aalborg, Denmark*. Spanish Acoustical Society.
8. Soulodre, G. A., & Bradley, J. S. (1995). Subjective evaluation of new room acoustic measures. *The Journal of the Acoustical Society of America*, *98*(1), 294–301.
9. Spencer, N. J., Hawley, M. L., & Colburn, H. S. (2016). Relating interaural difference sensitivities for several parameters measured in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *140*(3), 1783. <https://doi.org/10.1121/1.4962444>
10. von Berg, M., Steffens, J., Weinzierl, S., & Müllensiefen, D. (2021). Assessing room acoustic listening expertise. *The Journal of the Acoustical Society of America*, *150*(4), 2539–2548. <https://doi.org/10.1121/10.0006574>
11. Weinzierl, S., Lepa, S., & Ackermann, D. (2018). A measuring instrument for the auditory perception of rooms: The Room Acoustical Quality Inventory (RAQI). *The Journal of the Acoustical Society of America*, *144*(3), 1245. <https://doi.org/10.1121/1.5051453>