PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

The effect of reverberated speech on working memory: Toward an optimal balance of calmness and liveliness in libraries

Kazuma SHAMOTO[1]; Hiroko TERASAWA[2]; Hiroshi ITSUMURA[3]

[1, 2, 3] University of Tsukuba, Japan

[1] Ono Sokki, Japan

ABSTRACT

In order to accommodate both young talkative children and people who want a quiet library environment, we investigated the effects of the reverberation of children's voices on the working memory function of those around them. While modern libraries often hold noisy activities, unlike traditional silent libraries, their architecture has also become more open and reverberant. In this study, we tested the working memory functions of healthy participants under four sound conditions: (1) children's conversational voices without reverberation, (2) children's voices with reverberations from a traditional library environment (short and muted reverberation), (3) children's voices with reverberations from a modern library (long and echoic reverberation), and (4) silence as a control condition. Working memory capacity was measured by Operation Span Task (OST) and compared across the four sound conditions. Twenty university students participated in the experiment. As a result, OTC (i.e., Ospan total correct, the score of OST) was found to be best in silent conditions, followed by modern library conditions, traditional library conditions, and finally non-reverberant conditions. In addition, OTC was improved in conditions where participants felt comfortable. This suggests that choosing a learning space of our own preference can result in a learning space that provides good working efficiency for each individual user.

Keywords: Sound, Working Memory Capacity, Room Acoustics, Sound Environment, Library

1. INTRODUCTION

Although quiet sound environments have been idealized as a learning space, modern libraries include active, social, and interactive learning for people of all ages. While libraries today include noisy learning activities, the architecture of libraries also became more open and reverberant, reducing traditional quiet spaces. Active learning spaces in modern libraries ideally have a comfortable murmur of visitor's speaking voices (i.e., the mixture of the conversational sounds and the reverberations) that accommodates the effective functioning of working memory. In this study, we investigated the effect of children's conversation sounds under reverberation on working memory, as well as the impression of the reverberated conversational sounds.

Traditionally, ideal learning spaces were quiet. Both classrooms and libraries were expected to be silent. Quiet spaces improves working memory (1,2), and in Denmark, for example, the recommended reverberation time in school classroom is as short as 0.6 seconds (3). However, when the type of intellectual activity changes, a noisy environment may be more suitable than a quiet environment. In contrast to the research on school classrooms, some studies have shown that a lively sound environment, like a café, is richer in creativity (4). Perhaps this is why many people enjoy working on a laptop in a café.

Today, intellectual activity and learning are expected to be more creative than mere memorization. Figure 1 shows three levels of intellectual activities: (1) information processing, (2) knowledge processing, and (3) knowledge creation. The old style of learning was primarily composed of information processing and knowledge processing, but current learning trends involve knowledge

[1] shamoto.kazuma.ru@alumni.tsukuba.ac.jp
[2] terasawa@slis.tsukuba.ac.jp

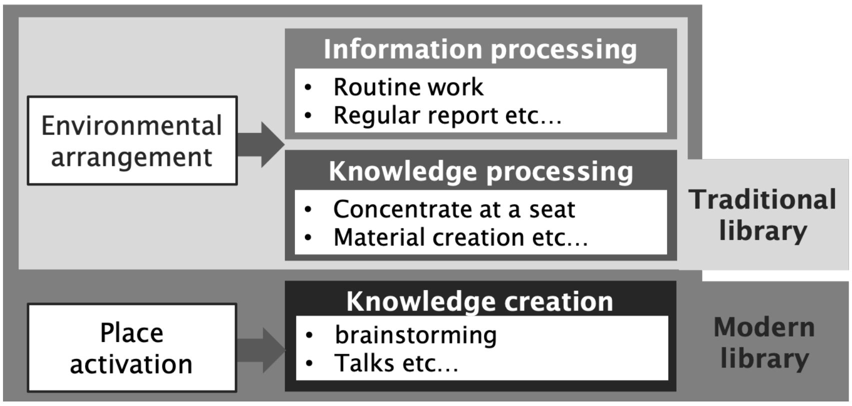

Figure 1 – The relationship between the hierarchical model of intellectual activity and architectural space (Adapted from (5))

creation. This change in learning styles calls for a change in acoustical environments and architectural styles in learning spaces. Quiet and focused learning spaces with a short reverberation time were considered to be optimal for traditional learning. However, if you want to actively exchange ideas and learn creatively, comfortable background noise may be better for encouraging oral communications.

Many modern libraries aim to provide a comfortable space that includes all three levels of intellectual activity. Increasing the value of ‘library as a place’, and making the library a place where various intellectual activities can be performed is a new role for modern libraries (6).

As the purpose and design of libraries changed, new issues arose. In our pilot study, librarians often mentioned that they receive complaints about conversational voices in the children’s area or the learning commons. We regarded these noise complaints as an indicator that noise from knowledge creation tasks may disturb the working memory that is required for information and knowledge processing tasks.

Information and knowledge processing are constructed upon working memory. Working memory, including short-term memory, often functions in auditory format in our mind (7). External noise conflicts with the auditory information in our mind, and disturbs the functions of working memory. In this way, information and knowledge processing tasks are easily disturbed by speech sounds.

Short-term memory is more disturbed by intelligible speech sounds than by meaningless white noise (1). If we use room reverberation to transform an intelligible speech sound into ambiguous speech noise that is less intelligible, the transformed speech noise can better facilitate information and knowledge processing than the original intelligible speech sounds. If we can find an adequate degree of ambiguity for speech noise transformation to facilitate information and knowledge processing, then we can find common ground to accommodate all three levels of intellectual activity.

2. PURPOSE

In this study, we tested working memory under four sound conditions: (1) direct (children’s direct voices without reverberation), (2) traditional (children’s voices with reverberation of a traditional library), (3) modern (children’s voices with reverberation of a modern library) and (4) silence (control condition). The hypothesis underlying this experiment is that conversational speech can be transformed to a comfortable murmur by applying an appropriate reverberation; allowing working memory to function more easily. Reverberation time is closely related to speech intelligibility, background noise, and the impression of the space. Finding an appropriate reverberation time that does not spoil the function of working memory can lead to the development of guidelines for acoustic design in libraries. In this study, we targeted children's conversational sounds that were most frequently complained by library visitors.

3. PREPARATION OF STIMULI

The stimuli for this experiment were prepared according to the following procedures: (1) Measure impulse responses in traditional and modern libraries, (2) record several children’s conversational sounds, and (3) convolve the conversational sounds with impulse responses.

3.1 Impulse response measurement

Impulse responses were measured at two libraries: one traditional university library (University of Tsukuba, Kasuga Library), and another modern public library (Yachiyo City Central Library).

Figure 2 – University of Tsukuba, Kasuga Library (traditional)

Figure 3 – Yachiyo City Central Library (modern)

Kasuga Library is a traditional university library, holding many books in a quiet sound environment (Figure 2). We measured the impulse response in the quiet study area on the second floor where bookshelf is dense. The floor area is about 467 m^2 (5026.75 ft^2), the height to the ceiling is about 2.7 m (8.86 ft.), and the average reverberation time is 0.3 seconds.

Yachiyo City Central Library is a modern public library opened in 2015 in Yachiyo City, Chiba Prefecture (Figure 3). Yachiyo Library has an open design on one spacious floor. The innermost area is the quiet reading area, the middle area is the browsing area, and the outermost area is the children's corner and reference. The floor area is about 1550 m^2 (16684 ft^2), the varied ceiling height ranges from 2.9 to -9.0 m (9.5-29.5 ft.), and the average reverberation time is 1.0 seconds.

The following equipment was used for impulse response measurement: An SPL meter with recording function (Ono Sokki LA-3260), an omnidirectional speaker (solid acoustics sa-755 professional), an amplifier (YAMAHA P4050), a laptop (Macbook Pro 13-inch 2012), a soundcard (RME Babyface), and a laser distance meter (BOSCH GLM7000). The measurement was approved at the library staff meetings at each library.

Table 1 shows the distance from the sound sources, reverberation time, and definition of the impulse responses we identified. Definition (i.e., degree of speech intelligibility) is calculated as follows, where *p(t)* is sound pressure (Pa) in time.

$$Definition = \frac{\int_0^{50\text{ms}} \left(p(t)\right)^2 dt}{\int_0^{\infty} \left(p(t)\right)^2 dt} \times 100(\%) \quad (1)$$

The reverberation time for Yachiyo Library is longer, and the definition is larger for the Kasuga Library. In short, the Yachiyo Library has a long and echoic reverberation and the Kasuga Library has short and muted reverberation. We controlled the distances from the sound source to the sound receiving point to be as similar as possible, for these two impulse responses.

Table 1 – Characteristics of the impulse responses we used for the stimuli

Library	Distance (m)	Distance (ft)	Reverberation Time (s)	Definition
Kasuga Library (traditional)	5.59	18.3	0.3	0.9
Yachiyo Library (modern)	4.95	16.2	0.4	0.7

3.2 Recording children's voices

We recorded children's conversational sounds in order to simulate the noise from active children in the libraries. We used a field recorder (R-05 Roland) in PCM formant, with a 44.1 kHz sampling rate. Five families with children ranging in age from two to ten years old volunteered, with prior consent to the intended research. The children's voices were recorded in the dining or living room, which has little reverberation and enabled us to record direct voice sounds. These locations may not be acoustically ideal as anechoic chamber, but among the areas at home where children can conduct natural conversations, they are the best choices. In order to record natural conversation, the parents of each family recorded the children's voices without the experimenter present. We explained the recording procedure to each of the parents, and instructed them to adhere to the following procedures: (1) Record when the children are gathered together and having fun, such as during tea time or a party, (2) place the recorder where it will pick up the children's voices clearly, such as in the centre of the table, and (3) record for about one and a half hours in total to allow for later editing.

3.3 Stimulus sound without reverberation

We edited the conversational sounds recorded in each household, and prepared the stimulus sounds before reverberation in the libraries according to the following procedures: (1) Delete sections of the recording that contain private information, that are distorted by loud sounds nearby, and any quiet sections, (2) calibrate the volume of sounds across the households. (3) divide the sounds about every 1-1.5 minutes and randomly concatenate the sounds of each household. (4) a fade-in/fade-out is edited in for eight seconds at each transition between household audio clips.

We created three stimulus sounds with a duration of 9-9.5 minutes following the above procedure. In this research, the calibration of volume is based on ITU-R BS. 1770-4 (8).We calculated the LKFS value of the sounds which represents loudness, and adjusted the amplitudes of the audio files, so that their LKFS values are at an equal level, differing less than $\pm 0.05\%$ from each other.

3.4 Stimulus sound with reverberation

We defined four reverberation conditions in this experiment: (1) direct (children's direct voices without reverberation), (2) traditional (children's voices with reverberation of Kasuga Library), (3) modern (children's voices with reverberation of Yachiyo Library) and (4) silence (control condition).

We convolved the stimulus sounds with the impulse responses from Yachiyo and Kasuga Libraries in order to achieve the desired reverberation. After reverberation, we calibrated the volume across all the stimuli equally. A total of nine types of stimulus sounds (three conversational sounds and three reverberation) were created, and used in a random order in the experiment, along with the silence control condition.

4. WORKING MEMORY EXPERIMENT

4.1 Procedure

We tested working memory using the Operation Span Task (OST (9)) under four sound conditions. The procedure of the OST is shown in Figure 4, and is designed similarly to that of previously published research (10), except as to the number of presented letters[4]. First, we presented a calculation problem, which the participant answered on the next screen. Next, we presented an alphabet and another calculation problem. After repeating this, the participant had to recall all of the presented alphabets in the order they were presented. The feedback was then displayed. One session consisted of this trial repeated for 12 times. An experiment consisted of four sessions (i.e., four different sound conditions in a random order but without a counter balance). We set about 5 minutes break between sessions. Before starting the experiment, we conducted a practice session of three OST loops for about five minutes to explain the procedure to the participants, and to acquire baseline data to individually adjust the experiment parameters. An experiment took about 1.5 hours. This experiment was approved by institutional review board at University of Tsukuba.

[4] The number of presented letters is (3, 3, 4, 4, 5, 5, 6, 6, 7, 7, 8, 8) in this study, where the original research employed (3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 7, 7, 7). The number of letters to present in the first loop is randomly selected from this list.

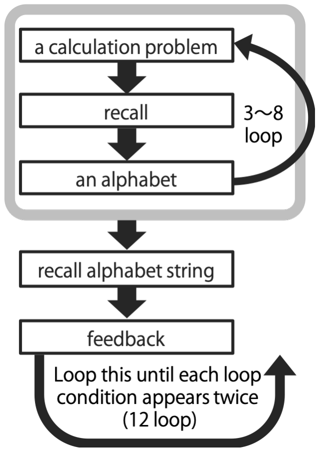

Figure 4 – Procedure of Operation Span Task

4.2 Experimental conditions

Twenty-two undergraduate and graduate students (12 males and 10 females), aged 18-25 at University of Tsukuba and possessing normal hearing abilities participated in the experiment.

In the experiment, a GENELC 8020C loudspeaker was used to present sound, and a laptop computer (Macbook Pro 13-inch 2015) was used to display the GUI. We used Inquisit 5 to program the OST. The speaker was placed at the position in front of the participant, about 1 m (3.28 ft) away from the participant's ears and at a height of 0.95 m (3.12 ft) from the floor (the same level as the participant's ears.) We calibrated the sound pressure level of the stimulus sounds by using a SPL meter (Ono Sokki LA-3260) at these speaker and participant positions.

The experiment was conducted in a soundproof room at University of Tsukuba, with the average background noise level at 18.5 dB. The stimuli were presented with average sound pressure levels of 42 dB at the listening point. The participants seated in front of a speaker and conducted the OST while hearing the stimulus sounds.

4.3 Questionnaire

We asked the participants to answer both pre- and post-experiment questionnaires and submit to an interview. The pre-experiment questionnaire consisted of questions about the backgrounds of the participants (e.g., age, sex, study environment preference, and frequency of library use). In the post-experiment questionnaire, we asked about their impressions of the stimulus sounds and the task difficulty of the experiment. In the interview we asked which sound conditions were comfortable or uncomfortable.

4.4 Ospan Total Correct

There are two generally accepted ways to count correct memory recalls, we chose to employ the Ospan Total Correct (OTC) (10) method for this study. In order to analyse the experiment results, we used OTC score to count the number of letters that were recalled correctly. OTC represents the total number of correctly recalled alphabets in the entire task. In this study, if the participants performed perfectly, their OTC would be 66 (the total number of presented letters).

5. RESULTS

An analysis was conducted for all 20 results, excluding two participants whose accuracy rate was less than 85% under any condition. We used R for the statistical analysis. The result of the OTCs for each sound condition are shown in Figure 5. For the OTC, the direct condition had a significantly lower score than the silent condition ($p < 0.01$, corrected with Benjamini & Hochberg (BH) method (11)). Also, the difference between the direct and modern conditions, and the difference between the traditional silent conditions was not quite significant ($p < 0.09$ for both cases, also corrected with BH method). In the pre-experiment questionnaire, the participants were asked ‘Are you easily distracted by sounds while working?’ However, their resulting ANOVA showed no significant correspondence between the questionnaire responses and OTC scores.

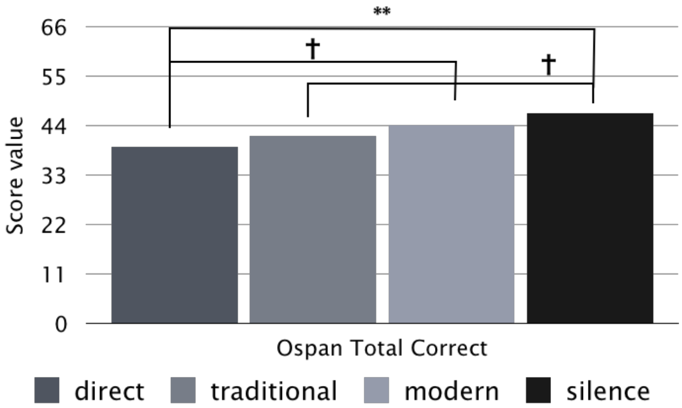

Figure 5 – OTC scores for each sound condition, where
**: p < 0.01, *: p < 0.05, †: p < 0.1 (multiple comparison corrected with BH method)

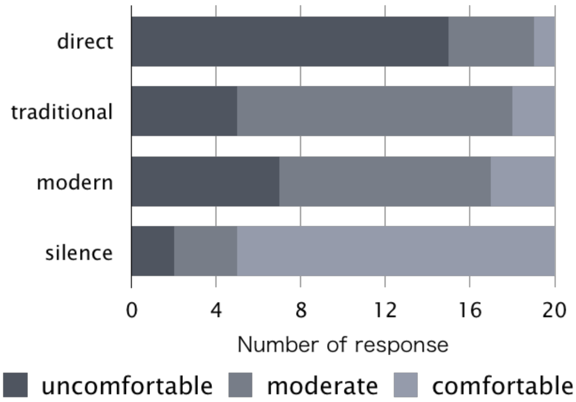

Figure 6 – Comfort judgments of sound condition

In the interview, we asked the participants which sound condition was uncomfortable and which was comfortable. Figure 6 shows the result of the uncomfortable or comfortable sound conditions as judged by the participants. Most of the participants felt that the direct condition was the most uncomfortable, and that the silent condition was the most comfortable. There were more 'comfortable' answers in the traditional and modern conditions than in the direct condition. However, there were fewer people who answered 'moderate' in the modern condition than in the traditional condition. The difference of comfort judgments between direct and silent conditions is quite evident, while the difference of comfort between traditional and modern conditions seems to be more individualized.

Figure 7 shows the Ospan Total Correct sorted by comfort judgments (rated by the participants). As a result of ANOVA, the correspondence between comfort judgments and OTC was significant ($p <$

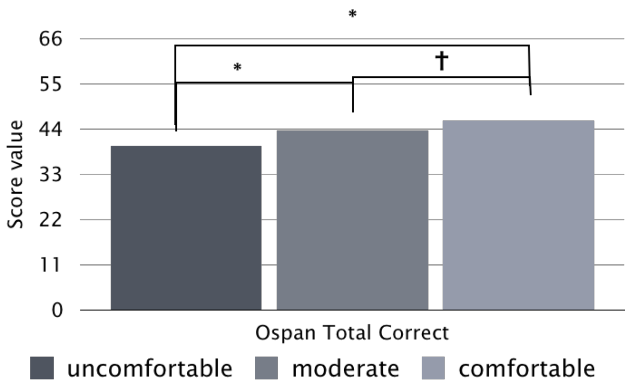

Figure 7 – OTC scores sorted by comfort judgments where
**: p < 0.01, *: p < 0.05, †: p < 0.1 (post-hoc analysis with Scheffé's method)

0.05). Post-hoc tests showed that the OTC was significantly higher in the uncomfortable condition than in moderate and comfortable conditions ($p < 0.05$, corrected with Scheffé's method (12)). Moreover, in the OTC, the score tended to be higher in the comfortable condition than in the moderate condition ($p < 0.01$, corrected with Scheffé's method).

6. DISCUSSION

6.1 Overall results

Previous studies have showed that short-term memory functions less efficiently when listening to conversational speech than in a silent environment (1, 2, 11). Our research showed a similar outcome— the difference between direct and silent conditions was significant. The modern condition tended to show better, although not significant, performance than traditional and direct conditions.

Considering the acoustic design of each library, the Kasuga Library had a small room size and was designed with many sound absorbing materials. Because the density of book volumes was high, direct sounds were absorbed by the bookshelves, and the reverberation emphasized low-frequency sounds. On the other hand, Yachiyo Library had a large room size and low bookshelf density, so the reverberation was quite broad-band. Listening to the synthesized stimuli, the conversational sounds echoed more pleasantly with the long and broad-band reverberation at Yachiyo Library, than with the short and low-frequency emphasized reverberation at Kasuga Library. We think this probably resulted in a more 'comfortable' rating for Yachiyo Library than for Kasuga Library.

These results contradict the reported disadvantages of long reverberation time in the field of class-room acoustics. For example, Klatte and Hellbrück (14) reported that environments with long reverberation times induced conversation among children. Persson, et al. (15) reported that the stress of teachers increased in an environment with long reverberation. There may be two reasons why teachers feel stressed by longer reverberation times. One is that students who are talking earn poorer grades (14), and the other is that the teacher's voice cannot be clearly transmitted in the classroom because a long reverberation makes the teacher's speech unintelligible.

We surmised that the effect of long-reverberation time differs in the context of a public library— there are no teachers or grades in a library, and people enjoy studying autonomously there. Inducing children's conversation may bring a public library a lively and positive atmosphere, which can be beneficial for knowledge creation. We also speculate that while information and knowledge processing were well considered in previous classroom acoustics research, consideration of knowledge creation was neglected. For knowledge creation, longer reverberation time is actually more desirable.

Further, public libraries are much more spacious than classrooms that accommodate diverse people in the community. For example, zoning is a common technique to spatially distribute library users based on their intended purpose of visiting, such as reading or discussion. By varying ceiling-heights and sound-absorbent materials, an ideal reverberation time can be optimally assigned for each zone within the library.

6.2 The relationship between comfort and sound condition

The resulting data of this study shows that task performance was significantly higher in a comfortable sound environment than in an uncomfortable environment. This suggests that comfort and task performance are highly correlated. In other words, choosing a sound environment to perform intellectual activities based on an individual's own preference can result in the selection of a sound environment with optimal work efficiency. However, in the pre-experiment questionnaire, very few participants selected lively cafés as preferable to quiet libraries. This response was unexpected, since students often work in cafés and diners, and Kasuga Library is rarely oversubscribed. In future inquiries, we could better design our questionnaire to garner more realistic responses. In order to investigate the relationship between individual sound environment preference and working memory capacity, it may also be interesting to recruit a group of participants who prefer a lively sound environment in a future experiment.

7. CONCLUSION

In this study, we examined how children's conversational sounds in different reverberation conditions can affect working memory. We discovered that silent conditions increased working memory capacity more significantly than direct sound conditions ($p < 0.01$). The average score for working memory performance under modern library conditions seemed better than that of traditional library conditions and direct conditions, but not statistically significant. This result was analysed along with the preference ratings for the sound environments. In the comfortable sound environment, the working memory capacity increased ($p < 0.05$). This suggests that choosing a sound environment to perform intellectual activities according to one's own preference may result in selecting a sound environment with good work efficiency.

ACKNOWLEDGEMENTS

We appreciate tremendous supports from our collaborators: Dr. Hiroshi Sato (AIST), Dr. Koji Ishida (Ono Sokki), Prof. Yuzuru Hiraga and Prof. Keiichi Zempo (University of Tsukuba), Prof. Atsushi Marui (Tokyo University of the Arts), Mr. Toshihito Yagi (Yachiyo City Central Library), Mr. Hiroo Yanase (Okada Architect and Associates), and the delightful librarians from University of Tsukuba Libraries.

REFERENCES

1. Salame P, Baddeley AD. Disruption of short-term memory by unattended speech: Implications for the structure of working memory. J Mem Lang. 1982;21(2):150–64.
2. Oswald CJP, Tremblay S, Jones DM. Disruption of comprehension by the meaning of irrelevant sound. Memory. 2000;8(5):345–50.
3. Rasmussen B, Brunskog J, Hoffmeyer D. Reverberation time in classsrooms--Comparison of regulations and classification criteria in the Nordic countries. Proc Jt Balt Acoust Meet. 2012.
4. Mehta R, Zhu R, Cheema A. Is noise always bad? Exploring the effects of ambient noise on creative cognition. J Consum Res. 2012;39(4):784–99.
5. Setsuko Y, Hiroko K, Yasuhiro H. Method for Evaluating Effect of Indoor Environment on Workplace Productivity. Rep Tech Res Inst Ohbayashi-Gumi, Ltd. 2010;(74):1–6.
6. Aabø S, Audunson R. Use of library space and the library as place. Libr Inf Sci Res. 2012;34(2):138–49.
7. Szalma JL, Hancock PA. Noise effects on human performance: a meta-analytic synthesis. Psychol Bull. 2011;137 4:682–707.
8. BS I-R. BS.1770 : Algorithms to measure audio programme loudness and true-peak audio level. 2015.
9. Turner ML, Engle RW. Is working memory capacity task dependent? J Mem Lang. 1989;28(2):127–54.
10. Unsworth N, Heitz RP, Schrock JC, Engle RW. An automated version of the operation span task. Behav Res Methods. 2005;37(3):498–505.
11. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995;57(1):289–300.
12. Scheffé H. The Analysis of Variance. New York: Wiley; 1999.
13. Hygge S, Boman E, Enmarker I. The effects of road traffic noise and meaningful irrelevant speech on different memory systems. Scand J Psychol. 2003;44(1):13–21.
14. Klatte M, Hellbrück J, Seidel J, Leistner P. Effects of classroom acoustics on performance and well-being in elementary school children: A field study. Environ Behav. 2010;42(5):659–92.
15. Persson R, Kristiansen J, Lund S, Shibuya H, Møberg Nielsen P. Classroom acoustics and hearing ability as determinants for perceived social climate and intentions to stay at work. Noise Heal. 2013;15:446–53.