TITLE

Effect of the frequency spectrum of road traffic noise on sleep: A polysomnographic study

RUNNING TITLE

Effect of road traffic noise spectrum on sleep

AUTHORS

Saana Myllyntausta[1], Jussi Virkkala[2], Paula Salo[3], Johanna Varjo[2], Laura Rekola[2], Valtteri Hongisto[4]

AFFILIATIONS

[1] Department of Public Health, University of Turku and Turku University Hospital, Turku, Finland;

[2] Finnish Institute of Occupational Health, Helsinki, Finland;

[3] Department of Psychology and Speech-Language Pathology, University of Turku, Turku, Finland;

[4] Turku University of Applied Sciences, Turku, Finland;

CORRESPONDING AUTHOR

Valtteri Hongisto

mobile phone: +358 40 5851 888

email: valtteri.hongisto@turkuamk.fi

Ordinary mail address: Turun ammattikorkeakoulu, Lemminkäisenkatu 14-18 B, FI-20520 Turku

DATE

Manuscript submitted 28th June 2019. Revised manuscript submitted 16th December 2019. Second revised manuscript submitted 2nd March 2020.

ABSTRACT

Spectrum of sound affects noise annoyance. Spectral differences of road traffic noise (RTN) transmitted indoors are usual because of spectrally different sound insulation of façades. The purpose was to compare the effect of RTN spectrum on sleep. Twenty-one volunteers slept three nights in a sleep laboratory in three sound conditions: low-frequency (LF) RTN, high-frequency (HF) RTN, and quiet (control). The A-weighted equivalent levels were 37 dB, 37 dB, and 17 dB $L_{\mathrm{Aeq,8h}}$, respectively. The nocturnal time profiles of LF and HF were equal. Sleep was measured with polysomnography and questionnaires. HF and LF did not differ from each other in respect to their effects on both objective and subjective sleep quality. The duration of deep sleep was shorter, satisfaction with sleep lower, and subjective sleep latency higher in HF and LF than in quiet. Contrary to subjective ratings given right after the slept night, HF was rated as the most disturbing condition for sleep after the whole experiment (retrospective rating). The finding suggests the sound insulation spectrum of the façade construction might play a role regarding the effects of RTN. More research is needed about the effects of spectrum on sleep because the field is very little investigated.

Keywords: road traffic noise; sound spectrum; sound insulation; sleep quality;

I. INTRODUCTION

In Europe, noise is acknowledged as the second most significant environmental pollution after air pollution. Its direct effects on both daytime activities and sleep have been summarized in important reviews.[1,2,3] Excessive nocturnal noise has been associated with, for example, increases in heart rate, arousals, sleep stage changes, motility, and cortical awakenings, as well as self-reported sleep disturbances, and increased use of sleep medication.[1,3] Road traffic is the most common source of environmental noise in residential areas. However, consensus about the most appropriate objective acoustic variable predicting sleep effects of road traffic noise (RTN) is lacking.[4]

The main acoustic properties of RTN that have been of interest in relation to sleep are equivalent sound levels, such as L_{AeqT} (equivalent A-weighted sound pressure level [SPL], during period T),[5] maximum SPLs such as L_{AFmax} (maximum level of A-weighted SPL using Fast time weighting),[6] number of noise events per night,[7] and the intermittency of noise.[8,9] The reason for the lack of consensus is probably the complex interaction between noise characteristics, individual sensitivities to noise, and the context of the explored environment, whether it is a laboratory or home.[10]

To date, very few sleep studies have considered the effects of the frequency content (later referred to as "spectrum") of noise on sleep. For example, Persson Waye et al.[11] and Öhrström and Skånberg[12] investigated the effects of both spectrally and contextually different sounds on various sleep outcomes, but did not examine the effects of the spectrum *per se*. The former study involved two sound conditions: artificial tonal ventilation noise (40 dB L_{Aeq}, tone at 50 Hz) and a façade filtered RTN (35 dB L_{Aeq}).[11] The spectra were reported within 20 −1000 Hz. The latter study involved three sound conditions: RTN filtered through an open window (39 dB L_{Aeq}), ventilation noise (40 dB L_{Aeq}), and their superposition (43 dB L_{Aeq}).[12] The spectra were reported within 20 – 8000 Hz. Because the sound conditions under comparison had different origins, different temporal patterns, different spectra, and different equivalent A-weighted SPLs, the findings cannot be attributed solely to the spectrum. Smith et al.[13] presented the first study where sleep effects of spectrum were reported. They investigated the effect of ground-borne noise from railway tunnels on sleep quality in laboratory using polysomnography. Each experimental night involved 32 noise events having the same A-weighted SPL. The spectra was reported within 31.5−1000 Hz. Half of the events, the low frequency events, had a higher SPL below 100 Hz than the rest of the events, the high frequency events. Above 100 Hz, the situation was the opposite. The experiment was conducted using three levels, 18, 20, and 22 dB L_{Aeq}. They found that high frequency events led to greater elevations of heart rate and increased arousal probability than low frequency events. Because RTN exposure is much more prevalent than railway noise exposure, it is very important to investigate whether an effect of spectrum on sleep quality could be found also for RTN. However, the spectral differences ought to be realistic.

The spectrum of sound affects noise annoyance.[14,15,16] Nilsson[14] studied the effect of RTN spectrum on annoyance in the overall sound level range 47−77 dB L_{Aeq}. He suggested that the relative increment of low frequency contribution in RTN increases the noise annoyance although the A-weighted SPL remains the same. However, the investigated high sound levels were higher than typical levels of environmental noise indoors (i.e. <35 dB L_{Aeq} and <45 dB L_{AFmax}). Their findings cannot be transformed to levels usually found indoors (under 45 dB L_{Aeq}) because the sensitivity to hear different frequencies of sound is different at low levels than at high levels according ISO 226 standard.[17] For example, a level change of 5 dB at 20 Hz produces a change in perceived loudness similar to a level change of 10 dB at 1000 Hz. Indeed, high-frequency wide-band sound was found to be more annoying than low frequency wide-band sound in an experiment where the annoyance of spectrally different wide-band sounds where compared at a constant level of 42 dB $L_{\mathrm{Aeq.}}$[15] As the spectrum of sound affects annoyance, it is justified to expect that the spectrum may also affect sleep. However, no previous research exists regarding the effect of RTN spectrum on sleep.

The spectrum of RTN indoors depends strongly on the sound insulation performance of the façade,[16,18] but also on the traffic speed, type of vehicles, and road surface. The proportion of high frequency noise (i.e. tire and aerodynamic noise) increases with increasing traffic speed and with decreasing proportion of heavy vehicles,[16] as well as with dense asphalt instead of porous asphalt.[19] The indoor SPL depends also on the reverberation time of the room and the measurement position in the room especially below 200 Hz.[18,20]

Regulations for indoor SPL are usually given using $L_{\mathrm{Aeq,T}}$. In many countries, it is mandatory to dimension the sound insulation of the façade of residential dwellings to achieve the regulated indoor SPL. The requirement for the façade's sound insulation depends on the A-weighted SPL outdoors ($L_{\mathrm{Aeq,out}}$), the regulated value indoors ($L_{\mathrm{Aeq,in}}$), the façade area, and the floor area of the room under inspection.[21] Requirements for the façade sound insulation are nearly always expressed using single-number quantities (SNQ) of sound insulation, such as those defined in ISO 717-1 standard.[22] The SNQs are calculated from the frequency-dependent sound insulation values by specific rules. However, the same value for a SNQ can be achieved with completely different sound insulation spectra. For example, **Fig. 1a** depicts the sound reduction indices (the physical quantity describing airborne sound insulation) of two façades, HF and LF, and corresponding SNQ values according to ISO 717-1.[22] The façade LF transmits more low frequency sound and less high frequency sound indoors than façade HF. However, both façades have the same value for the SNQ $R_{\mathrm{w}}+C_{\mathrm{tr}}$ (weighted sound reduction index against RTN) although the sound reduction indices are extremely different at different frequencies. Because sound spectrum affects noise annoyance,[14,15,16] it is highly relevant to investigate whether sleep quality is differently affected by RTN if the RTN is transmitted indoors through two nominally equivalent (equal values of $R_{\mathrm{w}}+C_{\mathrm{tr}}$) but spectrally different façades.

A psychoacoustic experiment[16] suggested that $R_\mathrm{w}+C_\mathrm{tr}$ did not explain the annoyance caused by RTN transmitted through the façade constructions in the best possible way. Instead, another single-number quantity of ISO 717-1 standard,[22] $R_\mathrm{w}+C_{50\text{-}3150}$, explained the annoyance better. It is of high relevance to confirm their suggestion in a sleep study.

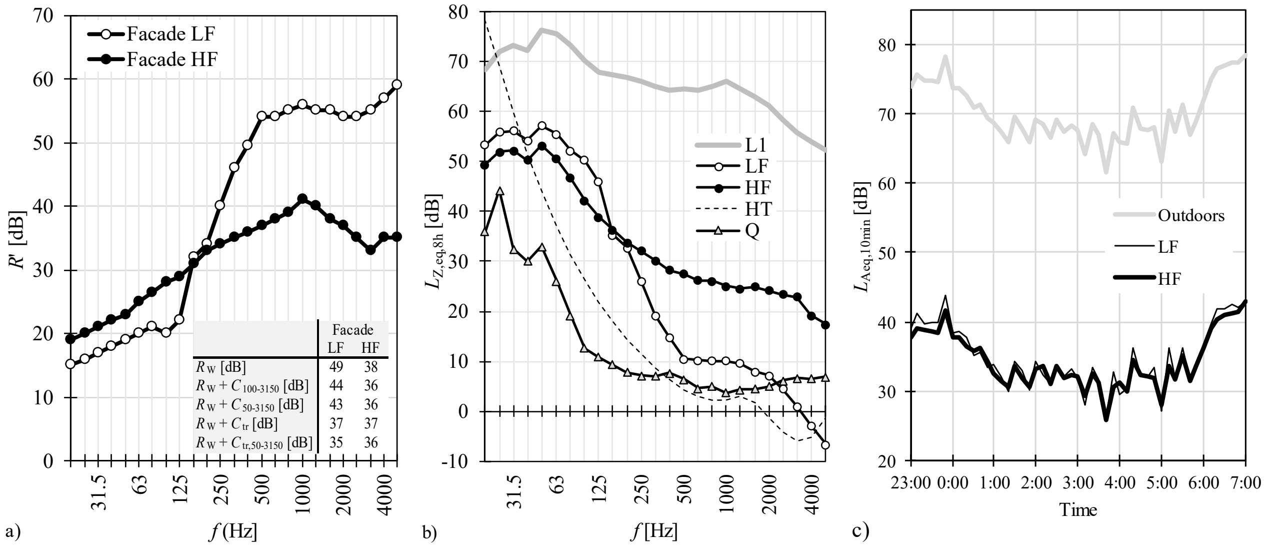

FIG. 1. a) The sound reduction index, R', as a function of frequency, *f*, for two façade constructions investigated in our study. The most important single-number quantities of ISO 717-1[20] are shown in the embedded table. b) The equivalent unweighted SPL, $L_{Zeq,8h}$, of the *conditions* HF, LF, and Q as a function of frequency during the 8-hour nocturnal exposure. Hearing threshold, HT, according to ISO 226,[17] and the outdoor level (L1) are shown for reference. c) The nocturnal time profile of RTN before filtering (outdoors) and after filtering for experimental *conditions* LF and HF. The profiles are presented using 10-min equivalent A-weighted SPL, $L_{Aeq,10min}$.

The primary purpose of our study was to compare the effects of two spectrally different RTN conditions on sleep among healthy adults with no sleep problems. The condition LF had more emphasis on low frequencies (20−125 Hz) while condition HF had more emphasis on middle and high frequencies (250–2000 Hz). The A-weighted equivalent SPL of both conditions was 37 dB L_{Aeq}. These two spectra correspond to two alternative conditions that occur indoors for two nominally equivalent façade constructions (equal values of R_w+C_{tr}) having completely different sound reduction indices (see **Fig. 1a**). We were primarily interested in the effects on overall quality of sleep of a full night of exposure to these RTN conditions. The secondary purpose was to investigate the effect of RTN at 37 dB L_{Aeq} on sleep by comparing the LF and HF conditions to the quiet condition (17 dB L_{Aeq}) where all sound was absent (control condition). Both objective and subjective measures of sleep quality were examined.

2 MATERIALS AND METHODS

A. General Outline

The study sample comprised 21 subjects who slept for four consecutive weekday nights in a sleep laboratory. The first night was an acclimatization night in quiet. The participants were exposed to three experimental sound conditions the three following nights. The independent variable of the laboratory experiment was the *condition*. The experiment was conducted using repeated measures design, i.e. each participant was exposed to each *condition*, one after the other. The exposure was arranged in six pseudo-randomized orders to minimize false findings due to order effects (**Table S1**[23]). The effects of the *conditions* on sleep quality were measured with polysomnography and questionnaires (dependent variables).

B. Participants

Eighteen to thirty-year old participants were recruited from student organizations in southern Finland. The invitation letter informed that the sleep study dealt with the effects of RTN. The inclusion criteria were stated in the recruitment letter and included normal hearing, a regular sleep-wake rhythm (from 22–00 to 06–08), and the volunteer regarding themselves as a normal sleeper (e.g. does not have a diagnosis of sleep apnea). Participants were required to live in an apartment building close to or nearby a busy road, as we wanted to avoid volunteers who are usually not exposed to any neighbor or environmental noise. It was also required that the participants did not do shift or night work, or take naps regularly. These inclusion criteria were selected to control for confounding factors that might affect participants' sleep or exposure to noise. A physician was consulted regarding any medications the volunteers reported using and whether using them could interfere with the experiment. The study sample consisted of 21 healthy volunteers (90% women), aged 20–30 years. The participants were given 60 Euro gift token for their participation. The experiment was approved by the ethics committee of the Hospital District of Helsinki and Uusimaa (166/13/03/00/2014 TMK12 §6, 16 July 2014).

C. Sound conditions

The descriptive names and abbreviations of the three *conditions* are quiet (Q, RTN was absent), low frequency RTN (LF), and high frequency RTN (HF). The A-weighted equivalent SPLs during the 8-hour-long night were 17.2 dB, 37.1 dB, and 37.2 dB L_{Aeq8h}, respectively. The spectra of the *conditions* are shown in **Fig. 1b**. C*ondition* Q involved only sound from the silent ventilation system and the level was below the hearing threshold for most frequency bands being nearly inaudible and carrying no specific frequency content. Thus, Quiet was a control *condition* which is expected to have no impact on sleep. *Conditions* LF and HF represent two alternative indoor RTNs, which exist inside a dwelling when the same outdoor RTN is transmitted through two different façade constructions having different sound insulation spectra (**Fig. 1a**).

The target equivalent SPL of *condition* LF, $L_{\mathrm{LF,eq}}$ [dB], was determined according to equation

$$L_{LF,eq} = L_{out,eq} - R_{LF}, \qquad (1)$$

where L_{out} [dB] is the SPL of RTN outdoors and R_{LF} [dB] represents the sound reduction index of the façade LF in **Fig. 1a**. Correspondingly, the target equivalent SPL of *condition* HF was determined according to equation

$$L_{HF,eq} = L_{out,eq} - R_{HF}, \qquad (2)$$

where R_{HF} [dB] is the sound reduction index of façade HF in **Fig. 1a**. Both R_{LF} and R_{LF} obtain the same SNQ value of $R_{\mathrm{w}} + C_{\mathrm{tr}} = 37$ dB according to ISO 717-1 standard.[22]

D. Laboratory

The experiment was conducted in the sleep laboratory of the Finnish Institute of Occupational Health in autumn 2014. Rooms LF, HF, and Q refer to the *conditions* LF, HF, and Q, respectively (**Fig. 2**). The participants changed the room according to the pre-defined order (**Table S1**[23]). The rooms were similar to each other, in regards to the floor area, room height, lighting, interior design, and temperature. The background noise level was 17 dB L_{Aeq} in each room.

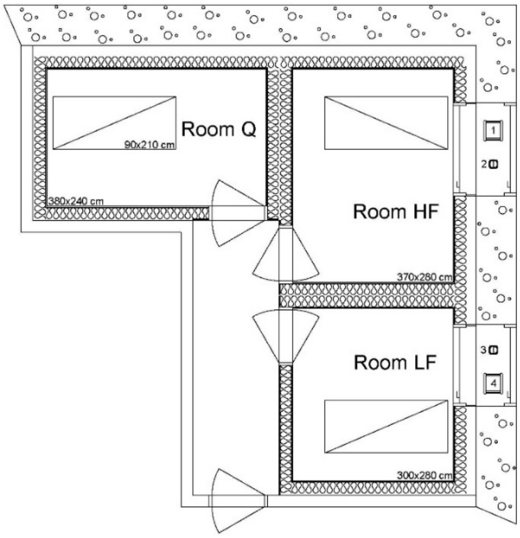

FIG. 2. Floor plan of the sleep laboratory. Four loudspeakers were used to produce the stimuli: two in room HF, and another two in room LF.

E. Creation of sound conditions

Road traffic noise was recorded in a junction of two busy streets of speed limit 50 km/h. **Figs. S1a–S1b** (See supplementary material[23]) describe the recording site. Recordings were conducted in May 2014 on two consecutive rainless weekday nights (recordings A1 and A2) between 22:00 and 07:30. Vehicles had normal tyres, outdoor temperature was above +8 °C, and wind was calm. Recordings were conducted using a digital recorder (Tascam HD-P2) and a free-field microphone (NTI Audio M2010). Both apparatus have flat frequency response within 20–20 000 Hz. System was calibrated by recording the sound of a level calibrator (B&K 4230) using the same input level settings as during the nocturnal recording.

All sounds slots that were not identified as road traffic sounds, such as emergency vehicles (8 min), church bells (8 min), birdsong (15 min), and human sounds (2 min), were removed from the recording A1. The removal was made for good measure although some of them might not be audible indoors. The removed slots were replaced either by silence (if the removal was short) or by a vehicle pass-by (if the removal was long and vehicle sounds were involved to the removed slot). The replacements were taken from recording A2 so that the night time noise level profile was not changed. Both removals and replacements extended always from silence to silence so that discontinuities were not remained in the recording. Only vehicle sounds were included in the outcome, which we call recording B.

After the replacements, the recording B was carefully listened by two researchers by headphones using much louder level than the level used in the sleep experiment. Implausible features such as discontinuities or periodicities were not observed. As a result, we had a recording of 9.5 hours, which consisted only of road traffic sound during 22:00–07:30. The time range from 22:45–07:00 was used for the sleep experiment. The spectrum and time profile of the recording B was determined (Sinus Harmonie Light, Samurai 1.5.12, Messtechnik GmbH, Germany). The unweighted 8-hour equivalent SPL spectrum (L1 in **Fig. 1a**) agreed well with the standard spectrum of urban road traffic noise.[22] The sound profile of **Fig. 1c** resembled typical urban nocturnal soundscape.

The experimental sounds were simultaneously produced to both LF and HF rooms during the sleep experiment (**Fig. 2** and **Fig. S1c** in Supplementary material[23]) by two pairs of loudspeakers (Genelec 1029A two-way speaker and Genelec 7050B subwoofer) to enable sound reproduction within 20–5000 Hz. Rooms LF and HF contained a fake window and a 1 m^3 chamber behind it. We wanted to avoid that the participants could see two loudspeakers in the room producing RTN. The loudspeakers were installed behind the fake window in both rooms (see **Fig. S1c** in Supplementary material[23]). Radiation of RTN through the fake window created a natural RTN environment to the room compared to the solution where the loudspeakers are located inside the room. The windows were covered by fixed curtains so that the subjects could not see the loudspeakers. The experimental sound produced to room LF or HF did not produce a sound level in the other two rooms that exceeded the background noise level 17 dB L_{Aeq}.

The recording B was filtered to obtain final recordings (C_LF for *condition* LF and C_HF for *condition* HF) as explained in **Sec. II.C** using digital third octave band filters (Adobe Audition 3.0) so that the measured SPLs in the position of the subject's head agreed with the target SPLs of **Fig. 1b**. It should be noted that because both final recordings C_LF and C_HF origin from the same recording B, they have exactly the same time profiles. A stereo playback file was created where C_LF was on the left channel and C_HF was on the right channel. This way, the playback was perfectly syncronized in rooms LF and HF regarding the occurrence time of every RTN event. Additional information about the temporal behavior of sound in conditions LF and HF are shown in Table S2 (See supplementary material[23]).

To verify the SPLs of the RTN exposure in the sleep laboratory, the final recordings C_LF and C_HF were measured in unoccupied rooms LF and HF by recording the full 8-hour-long presentation of *condition* simultaneously. The same measurement apparatus was used as in the original recordings. The recording lasted from 22:45 to 07:00. The equivalent $L_{\mathrm{Z,eq}}$ of the recordings agreed with the target $L_{\mathrm{Z,eq}}$ of **Fig. 1b** with an accuracy of ±2 dB in every 1/3-octave band.

Short calibration sounds (wide-band noise having the spectra of **Fig. 1b**) was added to the beginning of recordings C_LF and C_HF, respectively. The night nurse measured these calibration sounds in the pillow area of both rooms before the participant arrived in the room using a sound level meter (Cesva SC-15, Spain). No deviations from the target level (37 dB L_{Aeq}) were observed. Thus, each participant was exposed to the experimental sounds at the desired level and spectrum.

F. Experimental procedure

The experimental procedure is described in **Table S2** in Supplementary material.[23] The first night was for habituation to the laboratory conditions and the registration equipment. The following three nights (Tuesday–Friday) were the experimental nights. The participant was assigned to one of the three rooms, LF room, HF room, or Q room, according one of the pre-defined orders of the *conditions* according to **Table S1** in Supplementary material.[23]

On the first evening, participants were informed about the practices and restrictions during the research week and were told that the aim of this study was to assess the effects of RTN on sleep but that the RTN did not have any adverse health effects. No information about the sound conditions was given. During the first night, also airflow, snoring, limb movements, and movements of the rib cage and abdomen were recorded so that participants having sleeping problems or sleep disorders, such as sleep apnea or restless legs syndrome, could have been identified. Otherwise the first evening and night in the laboratory followed the same procedure as the other three nights.

Every evening, the participants were given a light snack and they also had some free time. Personal electronic devices were collected from the participants at 22.45 and the playback of the experimental sound was started. The lights were turned off at 23.00. The sound was stopped just before the subjects were woken up at 7.00. The participants were given a light breakfast every morning.

The participants were not allowed to take naps before the research nights. Consumption of alcohol and caffeinated drinks was prohibited after 3 p.m. before the research nights so that their effects on sleep could be minimized. During the days of the research week, the participants were allowed to continue their daily lives, for example, work, study, and exercise as usual. After the participants left the sleep laboratory on Friday morning, all restrictions related to the research were ceased.

G. Dependent variables

Sleep quality was measured both objectively and subjectively. Physiological changes induced by RTN were evaluated using polysomnography, PSG (Embla N7000). PSG signifies the simultaneous recording of electrical brain activity (electroencephalography, EEG), eye movements (electro-oculography, EOG), and muscle tone (electromyography, EMG), which are then used to classify sleep stages and to score arousals. The electrical activity of the heart (electrocardiography, ECG) was recorded to assess the functioning of the autonomic nervous system. The EEG electrodes were silver-silver chloride electrodes and disposable electrodes were used in the face area. The EEG electrodes were attached to the scalp and the EOG and EMG electrodes to the face of the subjects according to a well-established method.[24]

Two trained nurses evaluated the polysomnograms according to the American Academy of Sleep Medicine (AASM) 2.0 manual.[24] All the recordings of each individual subject were scored by the same nurse. The parameters that were derived from each recording were *total sleep time* (how long the subject slept during each night), *sleep efficiency* (the percentage of time spent asleep from the time spent in bed), *sleep latency* (time from wakefulness to first epoch of any sleep), and *WASO duration* (wake after sleep onset duration; the combined length of wake periods after initial sleep onset), and *number of arousals*. *Number of arousals* (abrupt shifts of EEG frequency that last at least 3 seconds), *arousal index* (the number of arousals per hour asleep), *number of awakenings*, *duration of awakenings*, and *awakening index* (the number of awakenings per hour asleep) were derived. The duration of each sleep stage was analyzed: *N1 duration* (light sleep), *N2 duration* (intermediate sleep), *N3 duration* (slow-wave sleep), and *REM duration* (rapid-eye-movement sleep) was analyzed. Sleep stages were scored in 30-second sequential epochs and a sleep stage was assigned for each epoch.

The participants completed a questionnaire during the first evening concerning, for example, their sleeping habits (average sleeping time, nighttime use of earplugs). Noise sensitivity was assessed with seven sleep related items from The Noise Sensitivity Questionnaire, the answer options ranging from (1 = completely disagree to 7 = completely agree) and higher scores representing higher sensitivity to noise.[25]Questionnaires were used to assess how the participants themselves felt they had slept during the night and how they felt during the following day. These subjective variables are defined in **Table I**. The participants completed a morning questionnaire approximately 30 minutes after they had been woken up. The participants were asked to evaluate their sleep quality during the night by assessing their *sleep disturbance by various items, satisfaction with sleep, subjective sleep latency, subjective sleep difficulties, subjective recovery*, *morning sleepiness,* and *morning strain.* An evening questionnaire was completed after entering the experimental room to assess how the participants felt during the day and in the evening. The participants were asked to evaluate *evening sleepiness*, *evening strain,* and *relative daytime tiredness*.

TABLE I. The subjective variables, items and response scales used in morning and evening questionnaires.

Variable name. Item from the questionnaire.	**Response scale**
Sleep disturbance by item X. How much the following items disturbed your sleep during the last night? Items: Road traffic sounds, silence, cold, hot, apparatus on my body, bed quality, darkness, unfamiliar place, monitoring camera	1 Not at all, 2 Only a little, 3 To some extent, 4 Much, 5 Very much
Subjective sleep latency. Falling asleep took longer than usually.	Same as above
Subjective sleep difficulties. My sleep was discontinuous, and I had difficulties staying asleep.	Same as above
Subjective recovery. I recovered well from the previous day's strain during the night.	Same as above
Morning strain. Evening strain. How strained do you feel at the moment?	Same as above
Satisfaction with sleep. How satisfied are you with your sleep last night?	−2 Very dissatisfied, −1 Dissatisfied, 0 Neither satisfied nor dissatisfied, 1 Satisfied, 2 Very satisfied
Morning sleepiness. Evening sleepiness. How have you felt during last five minutes? (Karolinska Sleepiness Scale)	1 Extremely alert, 2 Very alert, 3 Alert, 4 Rather alert, 5 Neither alert nor sleepy, 6 Some signs of sleepiness, 7 Sleepy, no effort to stay awake, 8 Sleepy, some effort to stay awake, 9 Very sleepy, great effort to keep awake, fighting sleep
Relative daytime tiredness. How tired have you been during the day compared to an ordinary day?	

	1 Much more alert, 2 Somewhat more alert, 3 As alert or as tired as usually, 4 Somewhat more tired, 5 Much more tired

On the Friday evening (between 19:00–23:00) after leaving the sleep laboratory, the participants completed an evening questionnaire that was sent to them by email. This evening questionnaire included two additional questions regarding the overall ranking of the *conditions*: "1. On which night the sound environment disturbed your sleep the most (three alternatives)?" and "2. Point out the room (on a floor layout) where the sound environment was the most annoying." A sum variable was calculated as the mean of these two items so that the number of counts per *condition* can range from 0 to 21. This sum variable is called *retrospective sleep disturbance ranking*.

H. Statistical analyses

Tests for repeated measures were used to analyze the differences between the *conditions* regarding the effects of nocturnal traffic noise on sleep. The Shapiro-Wilk test[26] was used to assess the assumption of normal distribution. *Number of awakenings*, *awakening index*, *N2 duration*, *N2* percentage, *N1+N2 duration*, *N1+N2* percentage, *REM duration*, *REM percentage*, and *N3 duration* were normally distributed and analyzed with the parametric repeated measures analysis of variance (ANOVA). The Greenhouse-Geisser correction[27,28] was used for all the statistically significant differences. All the other PSG variables and all the questionnaire variables were analyzed with the non-parametric Friedman test.[29] Paired comparisons were done with paired samples t-test and Wilcoxon signed-rank test, for the repeated measures ANOVA and Friedman test, respectively. The results of the paired comparisons are reported using the p-value for the t-statistic and the approximate p-value for the Z-statistic.[29] For all the statistically significant results in the pairwise comparisons, we used the Benjamini-Hochberg procedure[30] using $p = 0.05$ as a critical value for the false discovery rate and report which results remain significant after the corrections. The χ^2 -test was used to compare the frequencies of the *retrospective sleep disturbance rankings* between the *conditions*. An alpha level of 0.05 was used throughout the analyses. The statistical analyses were performed using IBM SPSS Statistics for Windows, Version 23.0 (Armonk, NY: IBM Corp).

III. RESULTS

The study sample consisted of 21 healthy volunteers and the mean age of the sample was 24.7 years (standard deviation [SD] = 3.1). Most of the participants were university students (90%) and the rest were working full-time. On the background questionnaire, the participants reported usually sleeping on average 7.7 hours per night (SD = 0.8) during weekdays. None of the participants used earplugs regularly while sleeping. On average, the participants were not particularly sensitive to noise, as the mean score on the sleep items of the Noise Sensitivity Questionnaire was 23 (SD = 6.9) and the scores varied from 13 to 33 (maximum value is 49).

The means for the objective sleep parameters measured with PSG are shown in **Table II**. The only statistically significant differences between the *conditions* were found in the *N3 duration* (repeated measures ANOVA, $F(2) = 5.29$, $p = 0.010$) and the *N3 percentage* (Friedman test, $F_R = 6.95$, $p = 0.03$). The *N3 duration* was shorter both in *condition* LF ($t(20) = -3.0$, $p = 0.017$) and *condition* HF ($t(20) = -2.6$, $p = 0.033$) than in *condition* Q. The *N3 percentage* was lower in both in *condition* LF (Wilcoxon signed-rank test, $Z = -2.84$, $p = 0.010$) and *condition* HF ($Z = -2.37$, $p = 0.028$). These *p*-values remained significant after controlling with the Benjamini-Hochberg procedure. No difference was observed between *conditions* LF and HF in the *N3 duration* ($p = 0.52$) nor in the *N3 percentage* ($p = 0.53$). *Number of arousals* and the *arousal index* were slightly larger in *condition*

HF than in other *conditions*, but the differences between the *conditions* were not statistically significant ($p = 0.16$ for *number of arousals* and $p = 0.23$ for *arousal index*).

TABLE II. Means (M) and standard deviations (SD) of the PSG variables for the three *conditions*.

	Condition			
	LF	HF	Q	
PSG variable	M (SD)	M (SD)	M (SD)	*p* value[a]
Time in bed [min]	480	480	480	
Total sleep time [min]	448 (17)	449 (17)	453 (22)	0.433
Sleep efficiency	93 (4)	94 (4)	94 (5)	0.433
Sleep latency [min]	14 (13)	13 (10)	13 (11)	0.549
WASO duration [min]	19 (15)	18 (13)	15 (16)	0.554
No. of arousals	49 (23)	54 (24)	47 (24)	0.156
Arousal index	6.6 (3.0)	7.3 (3.1)	6.3 (3.1)	0.229
No. of awakenings	17 (9)	16 (9)	16 (6)	0.493[b]
Duration of awakenings [min]	1.1 (0.6)	1.0 (0.6)	1.0 (1.1)	0.101
Awakening index	2.5 (1.2)	2.2 (1.2)	2.1 (0.9)	0.349[b]
N1 duration [min]	29 (14)	32 (16)	27 (12)	0.795
N2 duration [min]	228 (28)	225 (28)	218 (42)	0.305[b]
N1+N2 duration [min]	258 (27)	256 (32)	245 (42)	0.197[b]
N3 duration [min]	86 (30)	89 (25)	97 (26)	0.010[b]
REM duration [min]	104 (24)	104 (23)	111 (27)	0.503[b]
N1 percentage [%]	6.6 (3.3)	7.0 (3.7)	6.0 (2.8)	0.795
N2 percentage [%]	51 (6)	50 (6)	48 (8)	0.094[b]
N1+N2 percentage [%]	58 (7)	57 (7)	54 (9)	0.064[b]
N3 percentage [%]	19 (7)	20 (5)	21 (6)	0.031
REM percentage [%]	23 (5)	23 (5)	24 (6)	0.561[b]

[a] The statistical significance of the main effect of the *condition*.
[b] Parametric test was used.

Results for *sleep disturbance caused by various items* are shown in **Table III**. The only statistically significant difference between the *conditions* was observed in the item *road traffic sounds*. The ratings for *sleep disturbance due to road traffic sounds* varied significantly between the *conditions* ($F_R = 25$, $p < 0.001$). Expectedly, the paired comparisons showed that the ratings were higher in *conditions* LF ($Z = -3.7$, $p < 0.001$) and HF ($Z = -3.7$, $p < 0.001$) than in *condition* Q. The *conditions* LF and HF did not differ from each other ($p = 0.29$). The higher rates during the habituation night in most of the other items than the *road traffic sounds* suggest that the effects found in our experiment were caused by the *condition* and not by other factors related to the laboratory environment.

TABLE III. Means (M) and standard deviations (SD) of *sleep disturbance by various items* in the three *conditions* and the habituation night. Habituation night was omitted from the analyses comparing the items in each *condition*.

	Habituation night	*Condition* LF	*Condition* HF	*Condition* *Q*
Item	M (SD)	M (SD)	M (SD)	M (SD)
*Road traffic sounds****	1.1 (0.3)	2.5 (1.2)	2.8 (1.1)	1.1 (0.2)
Silence	1.9 (1.3)	1.0 (0.0)	1.0 (0.0)	1.2 (0.5)
Cold	1.9 (0.9)	1.4 (0.7)	1.6 (0.9)	1.3 (0.7)
Hot	1.8 (1.0)	1.4 (0.7)	1.3 (0.6)	1.2 (0.5)
Bed quality	1.1 (0.3)	1.1 (0.3)	1.1 (0.5)	1.1 (0.5)
Apparatus on my body	2.8 (1.0)	1.1 (0.3)	1.1 (0.5)	1.1 (0.5)
Darkness	1.5 (1.2)	1.1 (0.4)	1.1 (0.2)	1.3 (0.6)
Unfamiliar place	2.6 (1.2)	1.3 (0.5)	1.3 (0.6)	1.4 (0.7)
Monitoring camera	1.5 (0.9)	1.3 (0.7)	1.1 (0.5)	1.2 (0.5)

Response scale: 1 Not at all, 5 Very much.
*** Significant difference between *conditions* LF, HF, and Q (p<0.001).

From the subjective evaluations of sleep and recovery, we observed a main effect of the *condition* on *satisfaction with sleep*, and *subjective sleep latency* (**Table IV**). Participants were significantly more satisfied with their sleep in *condition* Q than in *conditions* LF (Z = -2.6, p = 0.009) and HF (Z = -2.5, p =0.012) and these differences remained significant after controlling with the Benjamini-Hochberg procedure. However, no differences in *satisfaction with sleep* were observed between *conditions* LF and HF (p = 0.37). *Subjective sleep latency* was larger in *conditions* LF (Z = -2.8, p = 0.015) and HF (Z = -2.2, p = 0.046) than in *condition* Q, whereas the *conditions* LF and HF did not differ from each other significantly (Z = -1.3, p = 0.190). Paired comparison showed that *subjective sleep difficulties* were higher in *conditions* LF (Z = -2.3, p = 0.019) and HF (Z = -2.0, p = 0.043) than in *condition* Q. However, these differences in the *subjective sleep difficulties* did not remain significant after controlling for the false discovery rate with the Benjamini-Hochberg procedure.The c*onditions* LF and HF did not differ in relation to *subjective sleep difficulties* (Z = -0.66, p = 0.51).

The *condition* had no effect on *subjective recovery* ($p = 0.29$). Finally, the *condition* did not affect *morning sleepiness, morning strain*, *evening sleepiness, evening strain,* or the *relative daytime tiredness* (**Table S4** in Supplementary material[23]).

TABLE IV. Means (M) and standard deviations (SD) of subjective variables of the morning questionnaire. The main effect of *condition* is denoted by *p* value.

	Condition			
	LF	HF	Q	*p*
Variable	M (SD)	M (SD)	M (SD)	
Satisfaction with sleep[a]	0.1 (1.1)	0.3 (1.0)	1.0 (1.0)	0.019
Subjective sleep latency[b]	2.1 (1.0)	1.9 (1.0)	1.5 (0.8)	0.006
Subjective sleep difficulties[b]	2.2 (1.0)	2.1 (1.0)	1.6 (1.0)	0.010
Subjective recovery[b]	3.2 (1.0)	3.2 (0.9)	3.5 (1.1)	0.287

[a] Response scale: -2 Very dissatisfied, +2 Very satisfied.
[b] Response scale: 1 Not at all, 5 Very much.

Retrospective sleep disturbance ranking varied significantly between the *conditions* (χ^2 (2, $N = 21$) = 23, $p < 0.001$, **Fig. 3**). The largest number of ratings was obtained for *condition* HF. The paired comparisons showed statistically significant differences between all pairs of *conditions* (LF vs. HF: $p = 0.007$; LF vs. Q: $p = 0.046$; HF vs. Q: $p < 0.001$). The responses to the two retrospective items were perfectly associated (nights matched with rooms).

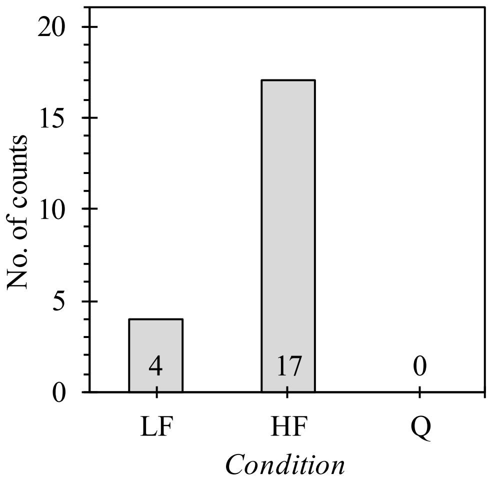

FIG. 3. *Retrospective sleep disturbance ranking* measured after the whole experiment.

IV. DISCUSSION

A. Main findings

We studied the effects of two spectrally different RTNs, i.e. low frequency RTN (*condition* LF, 37 dB L_{Aeq}) and high frequency RTN (*condition* HF, 37 dB L_{Aeq}), on sleep in a sleep laboratory. *Condition* Q (quiet, 17 dB L_{Aeq}), in which RTN was absent, was used as a control. *N3 duration* was significantly longer and *N3 percentage* higher in *condition* Q than in *conditions* LF and HF. No other differences were observed with objective measurements. *Satisfaction with sleep* was lower and *subjective sleep latency* was longer in *conditions* LF and HF compared to *condition* Q. Overall, the results suggest that the *conditions* involving RTN (LF and HF) were more adverse for sleep than *condition* Q. *Conditions* LF and HF did not differ from each other based on any objective or subjective measurement of sleep quality during the experiment.

Although we did not observe differences between the two *conditions* involving RTN in neither objective sleep quality nor participants' subjective evaluations during the experiment, the *condition* HF was evaluated to be less preferable than the *condition* LF after the whole experimental week.

This finding was perhaps not supported by the subjective ratings provided in the mornings after the experimental nights because the retrospective questions (see **Sec. II.G**) were measuring different experiences than those presented in the morning questionnaire (see **Table I**). The retrospective question, where only the most adverse option is forced to be selected, is more efficient for the ranking of the *conditions* than magnitude estimation (five-step response scale was used in subjective ratings). Two-alternative forced choice method is generally used in psychophysics instead of magnitude estimation when two acoustic stimuli, which do not have large differences, need to be reliably ranked. The retrospective questionnaire may have been more efficient in measuring the difference between the *conditions* LF and HF in respect to both general annoyance related to sound environment and sleep disturbance of the sound environment.

The retrospective finding was not supported by the objective findings either. RTN had a statistically significant effect on only one variable of objective sleep quality compared to *condition* Q (**Table II**). Because the sound levels of *conditions* LF and HF (37 dB L_{Aeq}) were much higher than the sound level of *condition* Q (17 dB L_{Aeq}), it seems afterwards improbable that objective sleep quality differences between *conditions* LF and HF would have appeared because they had equal sound levels. Differences in objective sleep quality might be possible to observe with much greater spectral differences. We chose to study two extreme spectra that can realistically occur along the same street due to two spectrally different but nominally equivalent (same $R_{\mathrm{w}}+C_{\mathrm{tr}}$) façade sound insulations. Spectral differences of RTN indoors, while keeping constant L_{Aeq}, are larger than in our study if both outdoor RTN spectrum and façade sound insulation spectrum are modified at the same time.[16]

The retrospective finding of our study is supported by several psychoacoustic studies suggesting that high-frequency sounds can be more annoying than low frequency sounds with the same A-weighted SPL.[15,31,32,33] Furthermore, our finding is supported by the sleep study of Smith et al.[13]

They studied the effect of single noise events from ground-borne railway tunnels that had two alternative spectra while we studied the effect of a full night RTN having two alternative spectra. Despite of methodologically different approaches, these two studies seemed to result in a similar suggestion that nocturnal exposure to high frequency spectrum might have more adverse effects than exposure to low frequency spectrum having the same A-weighted SPL. Further research in this field is strongly justified because A-weighting is a globally used SNQ for assessing the loudness and also the annoyance of sounds.

B. Rating of façade sound insulation

Façade construction is probably the most important means to provide noise control to residential dwellings in noisy environments. Requirements for the façade sound insulation are nearly always expressed using SNQs of sound insulation, such as those described in ISO 717-1 standard.[22] Therefore, it is extremely important that such SNQs are used, which are highly associated with noise annoyance[16] and also sleep disturbance. The *conditions* LF and HF (**Fig. 1b**) were created by simulating the sound insulation of two spectrally different façade constructions (**Fig. 1a**). Both façades were, on purpose, identical with respect to R_w+C_{tr}.[22] This SNQ is used in many countries to rate the sound insulation performance of façade constructions against urban RTN. However, the *condition* HF was ranked the most disturbing for sleep after the experiment. This finding gives reasons to suspect that R_w+C_{tr} might not be the most adequate SNQ for rating façade construction. Hongisto et al.[16] studied recently, which SNQ of the façade explained best the annoyance of RTN indoors. They found that $R_w+C_{50\text{-}3150}$ explained the annoyance of urban RTN transmitted through the façade better than the other standardized SNQs, such as R_w or R_w+C_{tr}. In the current study, the façade LF was 7 dB better than façade HF regarding their $R_w+C_{50\text{-}3150}$ value (**Fig. 1a**). Retrospective ranking of HF being more adverse than LF gives support to the findings of Hongisto et al.[16] according to which $R_w+C_{50\text{-}3150}$ could better rank the façade constructions against RTN than R_w+C_{tr}.

However, most of the results highlight the similarity of *conditions* LF and HF supporting the adequacy of $R_{w}+C_{tr}$. Further research on this issue is needed.

C. General effects of RTN

Some differences in the sleep quality outcomes were observed between *condition* Q and both conditions with RTN (*condition* LF and *condition* HF). *N3 duration* and *N3 percentage* were lower, and *satisfaction with sleep* and *subjective sleep latency* higher in *conditions* LF and HF than in *condition* Q. *N3 duration* was on average 10 minutes shorter in *condition* LF and 8 minutes shorter in *condition* HF than in *condition* Q. This finding of reduced *N3 duration* during RTN at 37 dB L_{Aeq} is in line with previous studies.[5,8,35] The reduction in *N3 duration* has thought to result from a general elevation of the organism's arousal level caused by the acute effects of noise on sleep, e.g. awakenings and body movements. Given the importance of N3 or slow wave sleep on health, the observation of decreased N3 sleep among healthy, normal sleepers with RTN levels of 37 dB L_{Aeq} is important. Furthermore, a trend towards a higher *number of arousals*, as well as longer *N2 duration* (and *N2 percentage*) and *N1+N2 duration* (and *N1+N2 percentage*) could be observed in *conditions* LF and HF than in *condition* Q. However, these differences were not statistically significant. It is possible that this study lacks the statistical power to detect differences between the RTN conditions and the quiet condition in relation to some of the objective sleep parameters.

Subjective experiences of sleepiness and strain during the morning and evening after being exposed to RTN did not differ between any of the *conditions*. These findings are somewhat discrepant with Öhrström[7] who found significant effects of RTN on tiredness despite of smaller L_{Aeq} values. However, the participants in that study were rather or very sensitive to noise, which may explain a greater impact of noise on tiredness in their study. We did not specifically recruit noise sensitive participants. Our study representsyoung and healthy adults living in urban environments because noise sensitivity has been found to affect subjective evaluations of sleep.[34]

D. Strengths and limitations

Our study is the first to explore the effects of RTN spectrum on sleep in a laboratory experiment. The main strengths of our study were the highly controlled experimental design and procedures. Firstly, the exposure order of the conditions was counter-balanced to eliminate possible order effects. Secondly, both conditions involving RTN resembled a natural sound environment because the loudspeakers could not be seen nor localized. Thirdly, one adaptation night prior to the experiment was sufficient for the adaptation to the paradigm. The adaptation night reduced possible bias due to sleeping in an unfamiliar place. Fourthly, the time profile of RTN corresponded to the natural nocturnal variation observed in normal city streets. Finally, the two *conditions* involving RTN were identical with respect to the overnight time profile, which made it possible to address all possible differences between the two *conditions* to the spectrum of RTN.

Our study has also some limitations. Although the PSG was manually scored by trained nurses with a well-established method, the scoring may be subject to human errors. Furthermore, as the scoring was split between the two nurses, it was not possible to estimate the error resulting from inter-scorer variability. The scorers were blind to the *condition* and, thus, no systematic bias should be expected. Only healthy adults were enrolled in this study, which has been the case for most of the studies assessing the effects of noise on sleep.[10] Because our main purpose was to investigate the effect of RTN spectrum on sleep, we were not required to recruit participants which represents the whole population. If the spectrum has an effect with healthy adults who sleep well, it is probable that the effect would be found also among bad sleepers who are more sensitive to disturbances.

Healthy adults are hardly among the most vulnerable groups for sleep disturbances due to nocturnal traffic noise. This might partly explain why we found only few effects of RTN on sleep. Furthermore, as our study population consisted mainly of women (90%): we were not able to examine potential gender-differences on the effects of the spectrum on RTN on sleep, nor to account for gender possibly influencing our results. There have been some indications of men being more sensitive to traffic noise than women.[35] In addition, we only invited participants who are living close to noisy street so that the conditions involving RTN would not be too unusual. Stronger responses might be possible among participants not accustomed or habituated to RTN. Furthermore, the requirements of the participation were relatively demanding as the participants had to sleep for four consecutive nights in laboratory conditions. Bias may have occurred in such a way that only subjects who are highly interested on the topic or sleep research in general volunteered. Therefore, our results related to the differences in sleep quality between RTN conditions (L_{Aeq} = 37 dB) and quiet (17 dB) might not be generalizable to the entire population.

Although many methodological attempts were made to reduce the artificiality of the laboratory environment, the extrapolation of our findings to residential environments has to be made with caution. For example, fewer awakening reactions have been found in field settings than in laboratory studies with the same sound level.[36,37] On the other hand, Skånberg & Öhrström[38] exposed the same subjects to equal levels of RTN both in a laboratory and in their homes and did not found any significant differences in the sleep quality assessed with wrist-actigraphy and questionnaires.

V. CONCLUSIONS

The primary aim of our study was to determine whether the spectrum of road traffic noise (RTN) affects sleep quality. The effects of two different RTN spectra (low frequency prominent LF and high frequency prominent HF) with the same overall sound level (L_{Aeq} = 37 dB) were examined. Comparison was made to quiet (control, L_{Aeq} = 17 dB). Neither the objective variables nor the subjective evaluations of sleep quality revealed differences between the two RTN spectra. However, the retrospective ratings after the whole experiment showed that HF RTN was perceived as a significantly more disturbing sound environment for sleep and more annoying sound environment than LF RTN. Although the acute subjective findings did not show a similar difference, the retrospective finding provides a reason to further investigate the effect of RTN spectrum on sleep quality. This is supported by a recent sleep study of Smith et al.[13] that has showed objective effects of spectrum on sleep quality. Furthermore, our results may also have practical implications in sound insulation design of façades. It is justified to study further, which single-number quantity describing the sound insulation of façades at different frequencies would be most suitable in relation to sleep quality.

ACKNOWLEDGEMENTS

This study was part of ÄKK project (2012–2014) which was mainly funded by Business Finland (Tekes Grant no. 2296/31/2011). The other funders were the Ministry of Environment, Betoniteollisuus Assoc., Kestävä Kivitalo, Karelia-Upofloor Ltd., Wärtsilä Finland Ltd., Saint-Gobain Rakennustuotteet Ltd., STX Finland Cabins Ltd., Skaala Ltd., University of Turku, Tampere University (previously Tampere University of Technology), and Finnish Institute of Occupational Health. The final analyses and manuscript writing were funded by University of Turku and Turku University of Applied Sciences. Thanks are due to David Oliva (acoustic engineering), Jarkko Hakala (playback system and drawings), Dr Markku Sainio (medical aspects), Mrs. Nina Lapveteläinen and Mrs. Riitta Velin (sleep laboratory), and volunteers who participated in the experiment.

REFERENCES

1. World Health Organization. (**2009**). “Night Noise Guidelines for Europe,” Regional Office for Europe: Copenhagen, Denmark.

2. Pirrera, S., De Valck, E., Cluydts, R. (**2010**). “Nocturnal road traffic noise: A review on its assessment and consequences on sleep and health,” Environ. Int. **36**(5), 492–498.

3. Basner, M., McGuire, S. (**2018**).”WHO environmental noise guidelines for the European Region: a systematic review on environmental noise and effects on sleep,” Int. J. Environ. Res. Public Health **15**(3), 519.

4. Basner, M., Müller, U., Griefahn, B. (**2010**). “Practical guidance for risk assessment of traffic noise effects on sleep,” Appl. Acoust. **71**(6), 518–22.

5. Griefahn, B., Marks, A., Robens, S. (**2006**). “Noise emitted from road, rail and air traffic and their effects on sleep,” J Sound Vib. 295(1), 129–140.

6. Öhrström, E., Rylander, R. (**1982**). “Sleep disturbance effects of traffic noise—A laboratory study on after effects,” J. Sound Vib. **84**(1), 87–103.

7. Öhrström, E. (1995). “Effects of low levels of road traffic noise during the night: a laboratory study on number of events, maximum noise levels and noise sensitivity,” J. Sound Vib. **179**(4), 603–615.

8. Eberhardt, J. L., Stråle, L. O., Berlin, M. H. (1987). “The influence of continuous and intermittent traffic noise on sleep,” J. Sound Vib. **116**(3), 445–64.

9. World Health Organization Europe. (**2011**). “Burden of disease from environmental noise: Quantification of healthy life years lost in Europe,” Copenhagen, Denmark.

10. Muzet, A. (**2007**). “Environmental noise, sleep and health,” Sleep Med. Rev. **11**(2) 135−142.

11. Persson Waye, K., Clow, A., Edwards, S., Hucklebridge, F., Rylander, R. (**2003**). “Effects of nighttime low frequency noise on the cortisol response to awakening and subjective sleep quality,” Life Sciences **72**(8), 863−875.

12. Öhrström, E., Skånberg, A. (**2004**). “Sleep disturbances from road traffic and ventilation noise—laboratory and field experiments,” J. Sound Vib. **271**(1), 279−296.

13. Smith, M. G., Ögren, M., Morsing, J. A., Persson Waye, K. (**2019**). “Effects of ground-borne noise from railway tunnels on sleep: A polysomnographic study,” Build. Environ. **149**, 288−296.

14. Nilsson, M. E. (**2007**). “A-weighted sound pressure level as an indicator of short-term loudness or annoyance of road-traffic sound,” J. Sound Vib. **302**(1−2), 197−207.

15. Hongisto, V., Oliva, D., Rekola, L. (**2015**). ”Subjective and Objective Rating of Spectrally Different Pseudorandom Noises—Implications for Speech Masking Design,” J. Acoust. Soc. Am. **137**(3), 1344−1355.

16. Hongisto, V., Oliva, D., Rekola, L. (**2018**). ”Subjective and objective rating of the sound insulation of residential building façades against road traffic noise,” J. Acoust. Soc. Am. **144**(2), 1100−1112.

17. ISO (2003). “ISO 226:2003 Acoustics. Normal Equal-Loudness-Level Contours,” International Organization for Standardization, Geneva, Switzerland.

18. Keränen, J., Hakala, J., Hongisto, V. (**2019**). ”Sound insulation of various residential façades at frequencies from 5 to 5000 Hz,” Build. Environ. **156**, 12–20.

19. Gołebiewski, R., Makarewicz, R., Nowak, M., Preis, A. (**2003**). “Traffic noise reduction due to the porous road surface,” Appl. Acoust. **64**(5), 481−494.

20. Pedersen, S., Møller, H., Persson Waye, K. (**2007**). “Indoor measurements of noise at low frequencies—problems and solutions,” J. Low Freq. Noise. Vibr. Act. Con. **26**(4), 249−70.

21. Ministry of the Environment. (**2003**). “Dimensioning of the sound insulation of building façades,” Environment Guide 108-2003, Helsinki, Finland (Available in Finnish).

22. ISO (**2013**). “ISO 717-1:2013 Acoustics. Rating of sound insulation in buildings and of building elements. Part 1: Airborne sound insulation,” International Organization for Standardization, Geneva, Switzerland.

23. See supplementary material at ***URL*** for additional methods and results.

24. Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C. L., Vaughn, B.V. (**2012**). “The AASM manual for the scoring of sleep and associated events. Rules, Terminology and Technical Specifications,” American Academy of Sleep Medicine, Darien, Illinois, USA.

25. Schütte, M., Sandrock, S., Griefahn, B. (**2007**). “Factorial validity of the noise sensitivity questionnaire,” Noise Health. **9**(34), 15−24.

26. Barton, B., Peat, J. (**2014**). *Medical statistics: a guide to SPSS, data analysis, and critical appraisal* (2nd edition), Wiley Blackwell, Oxford, England.

27. Abdi, H. (**2010**). The Greenhouse-Geisser correction. In: N. J. Salkind (Ed.) *Encyclopedia of research design*, *1*, 544–548. SAGE Publications, Inc, Thousand Oaks, CA, USA.

28. Verma, J. (**2015**). *Repeated measures design for empirical researchers.* John Wiley & Sons, Incorporated, Hoboken, New Jersey, USA.

29. Corder, G., Foreman, D. (**2014**). *Nonparametric statistics : a step-by-step approach* (2nd edition), Wiley, Hoboken, New Jersey, USA.

30. Bhattacharjee, M., Dhar, S. K., & Subramanian, S. (2011). *Recent advances in biostatistics: false discovery rates, survival analysis, and related topics* (Vol. 4). World Scientific. https://doi.org/10.1142/8010

31. Torija, A. J., Flindell, I. H. (**2014**). “Differences in subjective loudness and annoyance depending on the road traffic noise spectrum,” J. Acoust. Soc. Am. **135**(1), 1−4.

32. Torija, A. J., Flindell, I. H. (**2015**). “The subjective effect of low frequency content in road traffic noise,” J. Acoust. Soc. Am. **137**(1), 189−198.

33. Oliva, D., Hongisto, V., Haapakangas, A. (**2017**). ”Annoyance of low-level tonal sounds– Factors affecting the penalty,” Build. Environ. **123**, 404−14.

34. Marks, A., Griefahn, B. (**2007**). “Associations between noise sensitivity and sleep, subjectively evaluated sleep quality, annoyance, and performance after exposure to nocturnal traffic noise,” Noise Health. **9**(34), 1−7.

35. Basner, M., Müller, U., Elmenhorst, E.-M. (**2011**). "Single and combined effects of air, road, and rail traffic noise on sleep and recuperation," Sleep **34**(1), 11−23.

36. Pearsons, K. S., Barber, D. S., Tabachnick, B. G., Fidell, S. (**1995**). “Predicting noise-induced sleep disturbance,” J Acoust Soc Am. **97**(1), 331−338.

37. Öhrström, E. (**2000**). “Sleep disturbances caused by road traffic noise − Studies in laboratory and field,” Noise Health **2**(8) 71−78.

38. Skånberg, A., Öhrström, E. (**2006**). “Sleep disturbances from road traffic noise: a comparison between laboratory and field settings,” J. Sound Vib. **290**(1−2), 3−16.