Extensive crowdsourced dataset of in-situ evaluated binaural soundscapes of private dwellings containing subjective sound-related and situational ratings along with person factors to study time-varying influences on sound perception -- research data

Siegbert Versümer[1,2], Jochen Steffens[1,2], Fabian Rosenthal[1]

[1] University of Applied Sciences Düsseldorf, Germany
Institute of Sound and Vibration Engineering (ISAVE)

[2] Technische Universität Berlin, Germany
Audio Communication Group

siegbert.versuemer@hs-duesseldorf.de
isave@hs-duesseldorf.de

https://doi.org/10.5281/zenodo.7193938

March 2023

SPONSORED BY THE
Federal Ministry of Education and Research

I Abstract

The soundscape approach highlights the role of situational factors in sound evaluations; however, only a few studies have applied a multi-domain approach including sound-related, person-related, and time-varying situational variables. Therefore, we conducted a study based on the Experience Sampling Method to measure the relative contribution of a broad range of potentially relevant acoustic and non-auditory variables in predicting indoor soundscape evaluations. Here we present the comprehensive dataset for which 105 participants reported temporally (rather) stable trait variables such as noise sensitivity, trait affect, and quality of life. They rated 6.594 situations regarding the soundscape standard dimensions, perceived loudness, and the saliency of its sound components and evaluated situational variables such as state affect, perceived control, activity, and location. To complement these subject-centered data, we additionally crowdsourced object-centered data by having participants make binaural measurements of each indoor soundscape at their homes using a low-(self-)noise recorder. These recordings were used to compute (psycho-)acoustical indices such as the energetically averaged loudness level, the A-weighted energetically averaged equivalent continuous sound pressure level, and the A-weighted five-percent exceedance level. This complex hierarchical data can be used to investigate time-varying non-auditory influences on sound perception and to develop soundscape indicators based on the binaural recordings to predict soundscape evaluations.

II Description of the study

A Scope of the study

This study aims to investigate indoor soundscapes and their subjective perception determined by sound-related acoustic and perceptual predictors as well as non-auditory time-varying situational and rather temporally stable person-related ones.

B Participants

One-hundred-and-five participants from Düsseldorf/Germany and the surrounding area (57 women, 48 men, no diverse gender, *M* = 36 years, *SD* = 14 years, age range: 18-68 years) took part from April to October 2021. From these, 29 lived alone, 16 lived with children, and 99 had neighbors (frequency of neighbors: 52 above the participant's dwelling, 59 below, 82 next door – multiple responses were possible). Thirty-six participants had no hearing impairments (less than or equal to 20 dB HL; Age: *M* = 31, years, *SD* = 11 years), whereas 31 had mild impairments on at least one ear (over 20 and to 35 dB HL; Age: *M* = 40, years, *SD* = 13 years) and 11 lived with moderate impairments on at least one ear (over 35 dB HL; Age: *M* = 54, years, *SD* = 14 years).

Participants were invited through newspaper articles, social media posts, local radio and television broadcasts, and friends and acquaintances. Participants were excluded if they planned to be away from home for more than two days during the 10-day participation period if they were not expected to report five times a day and if they wore hearing aids. Two participants dropped out, one due to illness and one because the hourly assessments interfered too much with social relationships with other family members. One participant's results were excluded because the records could not be linked to the assessments (due to unsystematic time differences of more than 60 minutes).

C Design and Questionnaires

The field study is based on the Experience Sampling Method. Because participants were required to submit multiple reports, both a within- and between-subject design was used. There was no manipulation and no intervention, as the focus was on the natural perception of the participants. The predictors assessed stem from three different domains: the sound field (1), the non-auditory time-varying effects that change rather quickly from situation to situation (2), and the non-auditory, temporally rather stable person-related and socio-economic effects (3). Acoustical predictors were calculated based on the audio recordings of the sound pressure, the perceptual sound-related and situational predictors were assessed using a survey app on smartphones, and the person-related and socio-economic predictors were assessed using a tablet-and-pen questionnaire. For details, all German questionnaires are available together with a translation to English and the translated help texts.

1 Sound-related aggregated technical Predictors

For calculating the three aggregated acoustic predictors provided with this dataset, the binaural recordings of 15-seconds duration each were calibrated and filtered first (see section 0.2 for more details). Second, the instantaneous loudness was calculated for each channel of each recording based on the ISO 532-1 (DIN ISO 532-1:2017) algorithm for time-varying sound that is provided with the *acousticLoudness* function of the MATLAB® *Audio Toolbox™* (The MathWorks, Inc., 2022). Each value of the time series derived by this procedure was then converted from Sone to Phon using the MATLAB® function *Sone2PhonTV2018* taken from ISO/DIS 532-3:2022(E) (ISO/DIS 532-3:2022). Then, the loudness level *LLz*(P) was calculated as suggested by Kuwano et al. (2013). In addition, the L_{Aeq} and L_{AF5} (ISO 1996-1:2003) were also calculated. All three physical aggregated measures are provided for both channels of each binaural measurement.[1]

2 Time series of the binaural measurements[2]

To protect participants' privacy, each binaural measurement is also provided as a two-channel time series of $L_{\mathrm{Aeq\_2\,ms}}$ and $L_{\mathrm{Zeq\_2\,ms}}$ (linear, i.e., un-weighted) values instead of real-time audio. These sound pressure levels for time windows of 2 ms together with spectrograms (Δf = 21.5 Hz) every 23 ms allow the generation of further acoustic indicators and further statistical analyses while protecting the privacy of the participants. To nevertheless provide values that are as close to the original as possible, only the frequency response of the recording devices was compensated, i.e., the signal of the measurement microphones (without free-field correction) was provided.

3 Sound-related perceptual Predictors[3]

The *Perceived loudness* of the sound environment was determined by a categorical loudness scale that allows an intuitive and verbally anchored assessment in the first place with the five scale levels very low-level (0), low-level (10), medium (20), loud (30), very loud (40). Second, due to the design capabilities of the smartphone app used for gathering the personal ratings of the sounds and situations, the categorical verbal scale was partitioned (i.e., subdivided (Heller, 1990; DIN ISO 16832:2007) by using a numerical scale ranging from one to ten that allowed a finer differentiation. Both scales were added to the interval scale *Perceived loudness* (ranging from 1 to 50).

Participants further rated their perception of the sound environment by evaluating *Soundscape pleasantness* and *Soundscape eventfulness*, the main components of the soundscape, according to the soundscape standard ISO 12913 (DIN ISO/TS 12913-3:2021) and the translation into German language (Aletta et al., 2020. Table 3).

[1] It cannot be ensured that channel 1 was always the left channel and channel 2 the right one, because the participants may have swapped the microphones.
[2] See the MAT and JSON files in the *TimeSeries_and_Spectrograms_xxxx.zip* and the *TimeSeries_and_Spectrograms_README.md*.
[3] For further descriptions of all variables together with their scale levels and value ranges see the *VariableDescriptions_EnglishPersonQuestionnaire.pdf*

Additionally, participants had to describe the sound environment as if it were a composition of sounds by assigning all audible sounds to one of the given eight sound categories and by indicating the saliency of each category using eleven-level Likert scales.

4 Situational Predictors[3]

The momentary affective state, consisting of *Valence* and *Arousal*, which can be described using the circumplex model of affect (Posner et al., 2005), was assessed directly after starting the questionnaire with one continuous slider each (which was in the middle at the beginning of the scale and had to be touched or moved to proceed to the next question).

In addition to *Arousal*, also *Wakefulness* was assessed to capture the perceived level of feeling tired—alert (in the style of Steyer, n.d; Steyer et al., 1997), although these were found to describe different—though not independent—dimensions (Hinz et al., 2012). This could be meaningful if you feel tired (low alertness level) and yet experience a high activation level because you need to complete a time-sensitive important task even though you are exhausted, or if, for example, you are very tired and yet relaxed (low arousal level) because of the relaxing upcoming free weekend after a busy work week.

To examine the possible dependence of loudness ratings on the task participants were engaged in before each poll, *Cognitive load* and *Physical load* were assessed using an adaption of the NASA Task Load Index, which was developed to examine the performance of individuals driving or operating machines, vehicles, or aircraft (NASA TLX. Hart, 2006).

5 Person-related Predictors

Mean *Noise sensitivity* was measured using the German version (Eikmann et al., 2015. Table 2.9) of the NoiSeQ-R (Schütte et al., 2007) by averaging the three subscales for noise sensitivity regarding sleep, work, and habitation. Participants' *Hearing impairment* was measured for both ears across the octaves from 250 Hz to 8 kHz using the HEAD Audiometer with Sennheiser HDA-300 headphones, using the value of the greatest hearing loss for both ears in each frequency band.

The perceived general *Health* status was assessed using a single-item question: "How in general would you rate your health?". Single-item measures for the general health status showed to be a sufficient measure when different health aspects are not of special interest and a brief measure is needed (Idler and Benyamini, 1997; Bowling, 2005; Radun et al., 2019). Then, the participants averaged over the past two weeks their impression of the following items: Psychological *Wellbeing* was assessed using a German version of the WHO-5 questionnaire (Bech, 1999; Topp et al., 2015) which serves as a valid and internationally accepted time-efficient measure (Kulzer et al., 2006)). Also, participants assessed their state *Anxiety* using the GAD-7 questionnaire, which has been developed to be a reliable measure for the generalized anxiety disorder (Kroenke et al., 2007) and which was transferred into German (verified translation by back-translation).

The German Multidimensional Mood State Questionnaire (for the English translation of the German "Mehrdimensionale Befindlichkeitsfragebogen", MDBF, see Steyer, n.d.) served for measuring participant's three-dimensional mood (trait affect) over the past two weeks. It consists of the three scales *Trait mood* (good—bad), *Trait wakefulness* (awake—tired), and *Trait rest* (calm—nervous) (adapted from Steyer et al., 1997) which are highly correlated but still measure different dimensions (Hinz et al., 2012).

D Materials

1 *Survey Smartphones*

All perceptual sound-related and situational ratings reported by the participants were performed using a Nokia 4.2 smartphone and a survey app (movisensXS. movisens GmbH, Karlsruhe, Germany) that enables complex query procedures and makes the evaluated data immediately available to the study administration.

2 *Binaural recording Devices*

Some Facts in Brief

The binaural low-cost low self-noise audio recording devices were developed at the University of Applied Sciences Düsseldorf specifically for this study. Both electret microphones (Primo EM272)[4] were fitted into standard earbuds (with speakers and rubber plugs removed, see Figure 1 A and B; microphones pointing outward). Participants placed the earbuds loosely in their cavum conchae (approximately ±90-degree azimuth) so that the ear canal was not hermetically sealed. The recordings of 15 seconds duration, made at a sampling frequency of 44.1 kHz and an amplitude resolution of 16 Bit, were saved to a memory card and archived and removed after each participation. Reference recordings were made for calibration purposes. To do so, both microphones and a sound level meter were placed next to each other in a closed wooden box that includes a small loudspeaker (8 cm in diameter in a separate loudspeaker housing) to record a sine wave of 94 dB_{SPL} with a frequency of 100 Hz (to avoid acoustical modes). These two-channel recordings were used to calculate calibration factors that were applied to all of a device’s binaural recordings (i.e., all frequencies were adjusted equally). Additionally, all calibrated recordings were filtered to compensate for the implemented analog high-pass filter of the recording devices, to shift the reference plane from the microphone position outside the ear to the eardrum (convolving with the transfer function of the ear canal, derived with an artificial head with “independent of direction” equalization in an anechoic room), and to apply a free-field frequency correction (taken from Table 1 from (ISO 532-2:2017)).

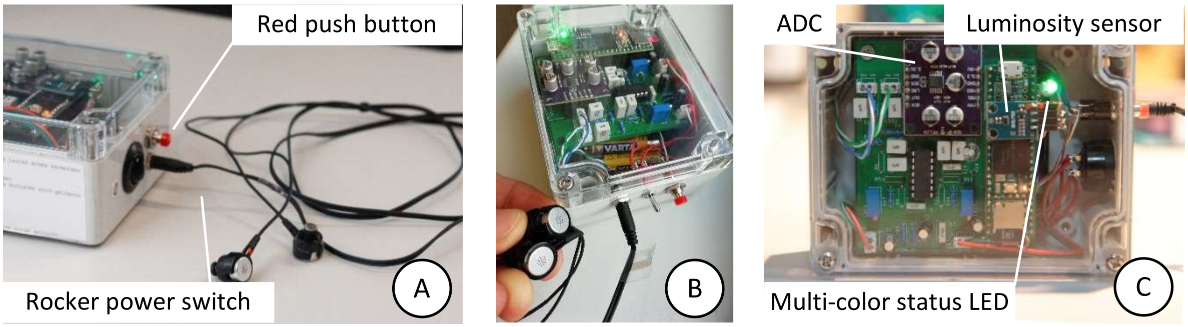

Figure 1: A: Binaural recorder with microphones built into the earbuds from which the speakers and rubber plugs were removed. The simple user interface allows to switch the device on or off completely (black rocker switch) and to start a recording (pressing the red push button shortly), to stop a running recording, and to delete the last recording (both by pressing the push button for more than two seconds). B and C: Pictures of the prototype. ADC: Analog to digital converter. Photos taken by Jörg Reich, 2020.

Low self-noise design

To achieve the low self-noise, the devices consist of two different electronic circuits, both powered by separate batteries: one of them is used for the most stable possible power supply of the analog part which includes the microphones, filtering, and preamplification. The second one supports the microcontroller (PRJC Teensy 3.6)[5] and can easily handle the current surges and the resulting voltage

---

[4] http://www.primomic.de/pdf/EM272Z1.pdf

[5] https://www.pjrc.com/store/teensy36.html

dips when storing the audio data on the memory card during recording - without affecting the analog part of the circuit unfavorably. In addition, the low-noise electret microphones were pre-amplified such that a sine wave of 94 dB_{SPL} reaches half the scale of the additional analog-digital converter (ADC, Texas Instruments PCM1808)[6], leaving 6 dB headroom for more impulsive or louder signals.

Environmental Sensors

The recording devices are also equipped with two environmental sensors: one measures temperature and relative humidity, as well as static air pressure (Bosch Sensortec BME280)[7] through a hole in the side of the case. Another measures brightness (ROHM Semiconductor BH1750)[8] - which is why the devices also have a transparent cover. – Unfortunately, environmental data was not collected in 786 situations due to a sporadic problem in the electronic circuitry (corresponding values are set to "NA"), which was evident after viewing the initial data and then gradually corrected.

User Interface

The devices provide a very simple user interface. Users can switch the device on or off completely (which is visible through the transparent lid and ensures that eavesdropping is not happening) using the black rocker power switch. By pressing the red push button briefly, participants can start the recording mode or confirm a low battery warning. Pressing it for two seconds aborts and deletes the current recording or deletes the last recording (since the device was switched on; can be not used several times in a row, hence, only the last recording and no recordings before it can be deleted). The multicolor LED makes the state of the device visible at any time. Possible states are:

Green, continuous:	The last action was completed successfully, and the next recording can be made.
Red, flashing slowly:	Indicates the five-second preparation time (after pressing the red push button) immediately before the recording begins.
Red, flashing fast:	The 15-second recording is running.
Purple, flashing fast:	The maximum input level is exceeded (overload) during a running recording.
Purple, continuous:	During the last recording, the maximum input level was exceeded (overload).
Yellow, flashing:	The running or the last recording is being deleted.
Yellow, continuous:	Signals a low battery voltage after switching on the device. – This warning can be acknowledged by pressing the red push button and further recordings can be made (until the microcontroller does not work anymore).

After activating the recording mode, the recording itself was delayed by five seconds, both so that participants could take deep breaths to remain still and motionless during the 15-second recording, and so that pressing the push-button would not become an audible part of the recording itself.

Frequency response

The frequency response of the analog high-pass filter (Figure 2) was measured with one of the ten recording devices and its two associated microphones in an acoustic free-field scenario. In addition, octave smoothing of the amplitude value was applied, and values below 21 Hz were overwritten with -14 dB to compensate the sound recordings for the characteristics of the recording device while ignoring possible microphone-specific frequency response deviations.

[6] https://www.ti.com/lit/gpn/pcm1808
[7] https://www.mouser.de/datasheet/2/783/bst_bme280_ds002-2238172.pdf
[8] https://www.mouser.com/datasheet/2/348/bh1750fvi-e-186247.pdf

Self-Noise

Although the microphones and other electronic components of the recorders have a minimal material cost of 110 EUR, a low-noise design has been achieved, resulting in an equivalent A-weighted sound pressure level of about 18 dBA (calibrated but no frequency response correction applied), allowing the recording of very quiet indoor soundscapes.

Figure 3 shows the A-weighted third-octave spectrograms of one 15-second binaural recording measured in the anechoic room of the University of Applied Sciences Düsseldorf, Germany, using one of the recording devices as well as a professional low-noise microphone[9] with which an equivalent sound pressure level of approximately 7 dBA was measured.

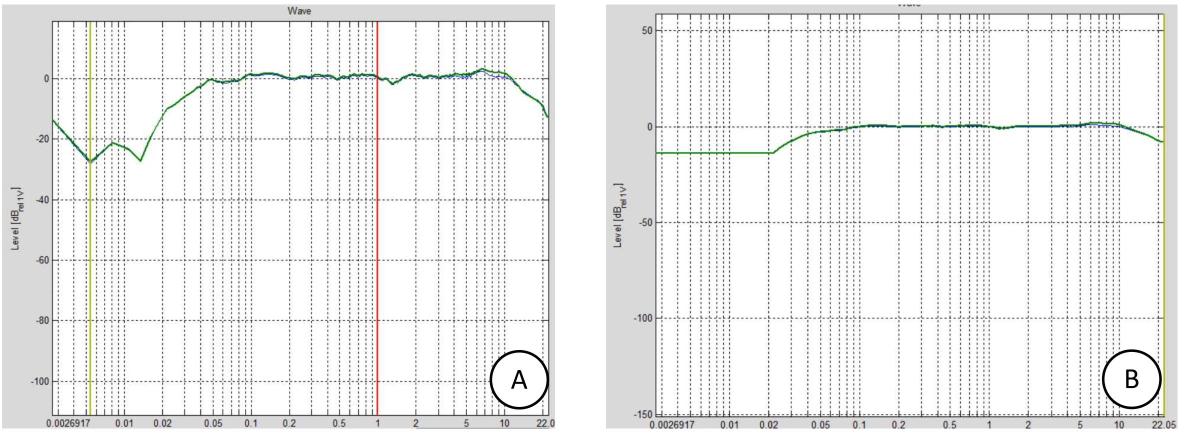

Figure 2: A: Frequency response of both microphones of one of the ten recording devices, measured in an anechoic room with a 1/3-octave magnitude smoothing applied. B: Same as A, but with octave smoothing and correction of frequencies below 20 Hz to -8 dB.

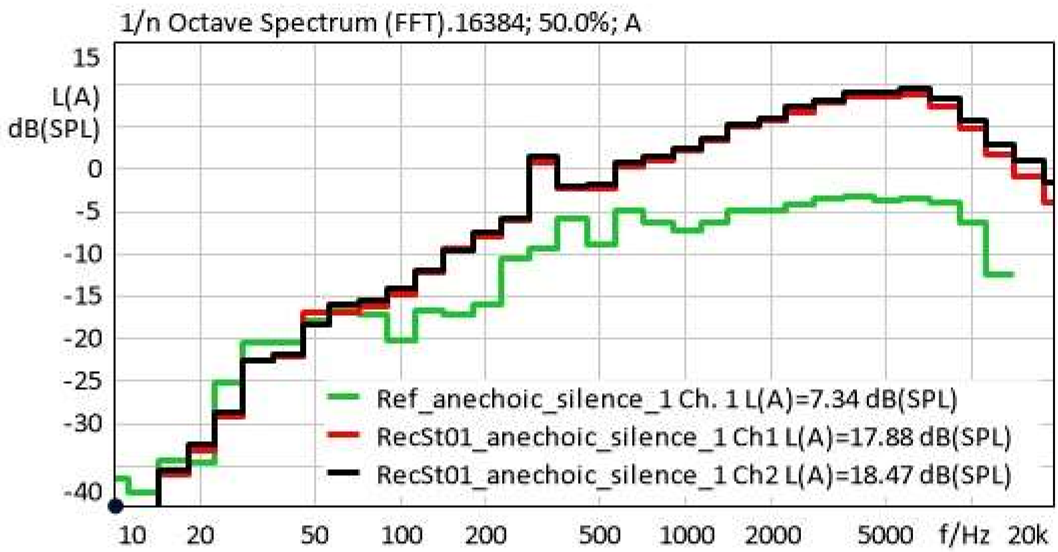

Figure 3: Third-octave spectrograms of the A-weighted equivalent sound pressure levels of the self-noise of one recording in an anechoic room (without additional excitation) using one binaural recording device and a professional ½" low-noise microphone (GRAS 47HC).

[9] GRAS 47HC ½" low-noise microphone: https://www.grasacoustics.com/products/special-microphone/low-noise-measuring-systems/product/781-47hc
Zodiac Heim DATaRec© 4 DIC6B interface: https://mh-gmbh.de/Downloadfile.php?get=DATaRec4/Technical%20Specification%20DIC6%202%20Hz_6.pdf

E Procedure and Participant Task

The participants received a standardized introduction during their individual appointments at the University of Applied Sciences Düsseldorf and completed an audiometry test (250 Hz – 8 kHz octaves). Answers to frequently asked questions and general help regarding the study were provided before the study and were also assessable via the survey smartphones throughout participation.[10]

The four trained instructors guided the participants through the questionnaire with the person-related and socio-economic predictors.[11] Participants received a survey smartphone and a binaural recording device. The devices were stored in a fanny bag so that the participants would have them at hand when the alarm reminded them to answer the questionnaire. A separate smartphone with a special training survey was used to familiarize participants with the use of their survey smartphone, the operation of the survey app, all possible rating scales and input types, and the use of the help section. Finally, participants were made familiar with the structure of the Experience Sampling Method questionnaire[12], comprehension of the questions and scales, and the definitions of the eight sound categories[13].

At home, participants paired the study smartphones with their private Wi-Fi so that survey results were uploaded immediately. After starting the survey app, participants set the daily time range in which they liked to be asked to answer the questionnaire. During their participation, which began on Friday and lasted ten consecutive days until the Monday after next, periodical alarms started approximately 15 minutes (±5 minutes) after the beginning of the daily time range. Further alarms then followed approximately every hour (±10 minutes). Participants were able to accept the alarm, delay it by five minutes twice, reject it altogether, or ignore it. Ignored alarms were repeated twice. Participants were asked not to mute the alarm sound, but to adjust the daily time range as often as necessary to tailor the study to their needs and times at home. In cases where participants spent only a few hours per day at home or were absent for a few workdays or a weekend, they could independently initialize the assessment to reach the target of 70 assessments during their time at home. Despite being tasked to conduct assessments on the hourly reminders whenever possible, participants were more likely to start assessments on their own (73% of the assessments were self-initiated). Thirty-five participants self-initiated more than 90 percent of their assessments, with fourteen of them doing so in all cases.

When an alarm was accepted, participants first indicated whether they could hear anything. If not, the questionnaire was canceled and otherwise continued. Once given, answers could not be changed later and answers on previous questionnaire pages could not be viewed again. Affective state questions were then answered to capture participants' emotional state preferably independent of a possible emotional impact of responding to the survey. Participants were then asked to make the audio recording.[14]

---

[10] See the *HelpTexts.pdf* file.
[11] See the *VariableDescriptions_EnglishPersonQuestionnaire.pdf* and *PersonQuestionnaire_OriginalGermanVersion.pdf* files.
[12] See the *ESM-Questionnaire.pdf* file.
[13] See the *VariableDescriptions_EnglishPersonQuestionnaire.pdf* file.
[14] In many cases, contrary to the request, the recording was made first and then the ESM assessment was started manually. When importing the data after the study, a matching of the audio recordings with the ESM assessments was performed. Only cases in which a clear assignment could be made were considered; all other recordings and ESM assessments were excluded.

The first of the three main sections followed: the evaluation of the most salient sound. In the second part, participants reported on the overall indoor sound environment. The third part dealt with the situational predictors, complementing the questions asked at the beginning about the affective state.

After their ten-day participation, the participants received a staggered compensation of up to 100 Euros. Twenty Euros were paid for participating in the introduction at the university, 30 Euros for evaluating 45 sound situations, and 2 Euros for each additional contribution, but not more than a total of 100 Euros for a total of 70 evaluations. Participants could report more beyond that without receiving further compensation.

III DATA AVAILABILITY

The data of this study is openly available at https://doi.org/10.5281/zenodo.7193938. Real-time audio recordings cannot be made publicly available for privacy reasons. Please contact the authors.

IV CONTRIBUTIONS

SV designed and conducted the study, performed the statistical analysis, interpreted the data, wrote the initial manuscript, and made the revisions. SV also preprocessed and published the dataset. JS contributed to the research questions, the design, and the statistical analysis. All authors reviewed the manuscript and approved the final version of the article.

V ACKNOWLEDGES

The author thanks Patrick Blättermann for advice and constructive discussion of the study design. Jenny Winter was responsible for the help texts, subject acquisition, and scheduling of subject introductions. The latter were carried out by Jan Roloff, JW, FR, and SV. Particular thanks go to Benjamin Müller, who co-developed the recording device and built the prototype as well as a 10-part small series. Likewise, special thanks go to Christian Epe for expert advice on the low-noise circuit design.

VI ETHICS STATEMENT AND DATA SECURITY

The studies involving human participants were reviewed and approved by the Ethics Committee of the Medical Faculty of the University of Duisburg-Essen, Germany. All participants provided digital written informed consent by confirming the declaration on data collection and processing before participating.

VII DECLARATION OF CONFLICT OF INTERESTS

The author declared no potential conflicts of interest concerning the research, authorship, and/or publication of this dataset.

VIII FUNDING

This study was sponsored by the German Federal Ministry of Education and Research. “FHprofUnt”-Funding Code: 13FH729IX6.

IX References

Aletta, F., Oberman, T., Axelsson, Ö., Xie, H., Zhang, Y., Lau, S.-K., Tang, S. K., Jambrošić, K., Coensel, B. de, van den Bosch, K. A.-M., Aumond, P., Guastavino, C., Lavandier, C., Fiebig, A., Schulte-Fortkamp, B., Sarwono, A., Astolfi, A., Nagahata, K., Jeon, J.-Y., Jo, H.-I., Chieng, J., Gan, W.-S., Hong,

J.-Y., Lam, B., Ong, Z.-T., Kogan, P., Silva, E. S., Manzano, J. V., Yörükoğlu, P. N. D., Nguyen, T. L., and Kang, J. (**2020**). “Soundscape assessment: towards a validated translation of perceptual attributes in different languages,” inter.noise 2020.
Bech, P. (**1999**). “Health-related quality of life measurements in the assessment of pain clinic results,” Acta anaesthesiologica Scandinavica **43**, DOI: 10.1034/j.1399-6576.1999.430906.x, 893–896.
Bowling, A. (**2005**). “Just one question: If one question works, why ask several?,” Journal of epidemiology and community health **59**, DOI: 10.1136/jech.2004.021204, 342–345.
Eikmann, T., Nieden, A. zur, Ziedorn, D., Römer, K., Lengler, A., Harpel, S., Bürger, M., Pons-Kühnemann, J., Hudel, H., and Spilski, J. (**2015**), “Lärmwirkungsstudie NORAH, Band 5 - Endbericht - Blutdruckmontitoring” (“Noise Impact Study NORAH, Volume 5 - Final Report - Blood Pressure Monitoring.”). In NORAH: Study on Noise-Related Annoyance, Cognition and Health: A transportation noise effects monitoring program in Germany.
Hart, S. G. (**2006**). “Nasa-Task Load Index (NASA-TLX); 20 Years Later,” Proceedings of the Human Factors and Ergonomics Society Annual Meeting **50**, DOI: 10.1177/154193120605000909, 904–908.
Heller, O. (**1990**), “Scaling and orientation.” In F. Müller (Ed.). Fechner Day 90. Proceedings of the Sixth Annual Meeting of the International Society for Psychophysics (Würzburg University, Würzburg), 52–57.
Hinz, A., Daig, I., Petrowski, K., and Brähler, E. (**2012**). „Die Stimmung in der deutschen Bevölkerung: Referenzwerte für den Mehrdimensionalen Befindlichkeitsfragebogen MDBF“ („Mood in the German population: norms of the Multidimensional Mood Questionnaire MDBF“), Psychotherapie, Psychosomatik, medizinische Psychologie **62**, DOI: 10.1055/s-0031-1297960, 52–57.
Idler, E. L., and Benyamini, Y. (**1997**). “Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies,” Journal of Health and Social Behavior **38**, DOI: 10.2307/2955359, 21-37.
ISO (**2003).** ISO 1996-1(E), “Acoustics - Description, measurement and assessment of environmental noise: Part 1: Basic quantities and assessment procedures,” (International Organization for Standardization, Geneva, Switzerland).
ISO (**2007).** DIN ISO 16832, “Acoustics – Loudness scaling by means of categories,” (International Organization for Standardization, Geneva, Switzerland).
ISO (**2017).** DIN ISO 532-1, “Acoustics - Methods for calculating loudness - Part 1: Zwicker method,” (International Organization for Standardization, Geneva, Switzerland).
ISO (**2017).** ISO 532-2, “Acoustics - Methods for calculating loudness - Part 2: Moore-Glasberg method,” (International Organization for Standardization, Geneva, Switzerland).
ISO (**2021).** DIN ISO/TS 12913-3, “Acoustics - Soundscape: Part 3: Data analysis,” (International Organization for Standardization, Geneva, Switzerland).
ISO (**2022).** ISO/DIS 532-3(E), “Methods for calculating loudness — Part 3: Moore-Glasberg-Schlittenlacher method,” (International Organization for Standardization, Geneva, Switzerland).
Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., and Löwe, B. (**2007**). “Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection,” Annals of internal medicine **146**, DOI: 10.7326/0003-4819-146-5-200703060-00004, 317–325.
Kulzer, B., Hermanns, N., Kubiak, T., Krichbaum, M., and Haak, T. (**2006**). „Der WH0 5: Ein geeignetes Instrument zur Messung des Wohlbefindens und zum Depressionsscreening bei Diabetikern“ („The WH0 5: A Suitable Instrument for Measuring Well-Being and Screening for Depression in Diabetic Patients“), Diabetologie und Stoffwechsel **1**, DOI: 10.1055/s-2006-943822.
Kuwano, S., Hatoh, T., Kato, T., and Namba, S. (**2013**), “Evaluation of the loudness of stationary and non-stationary complex sounds.” In ICA 2013, 50188.

movisens GmbH, Karlsruhe, Germany, “movisensXS,” https://www.movisens.com/en/products/movisensXS/, accessed 06 October 2022.

Posner, J., Russell, J. A., and Peterson, B. S. (**2005**). “The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology,” Development and psychopathology **17**, DOI: 10.1017/S0954579405050340, 715–734.

Radun, J., Hongisto, V., and Suokas, M. (**2019**). “Variables associated with wind turbine noise annoyance and sleep disturbance,” Building and Environment **150**, DOI: 10.1016/j.buildenv.2018.12.039, 339–348.

Schütte, M., Marks, A., Wenning, E., and Griefahn, B. (**2007**). “The development of the noise sensitivity questionnaire,” Noise and Health **9**, DOI: 10.4103/1463-1741.34700, 15–24.

Steyer, R. (**n.d.**), “MDMQ questionnaire (English version of MDBF),” https://www.metheval.uni-jena.de/mdbf.php, accessed 08 August 2022.

Steyer, R., Schwenkmezger, P., Notz, P., and Eid, M. (**1997**). „Testtheoretische Analysen des Mehrdimensionalen Befindlichkeitsfragebogen (MDBF)“ („Test-theoretical analyses of the Multidimensional Mood State Questionnaire“), Diagnostica **40**, 320–328.

The MathWorks, Inc. (**2022**), “Audio Toolbox™ Reference Guide,” https://de.mathworks.com/help/pdf_doc/audio/audio_ref.pdf, accessed 25 May 2022.

Topp, C. W., Østergaard, S. D., Søndergaard, S., and Bech, P. (**2015**). “The WHO-5 Well-Being Index: a systematic review of the literature,” Psychotherapy and psychosomatics **84**, DOI: 10.1159/000376585, 167–176.